



**HAL**  
open science

# Caractérisations biochimiques et structurales d'une métacaspase et d'une enzyme branchante de glucane en alpha-(1,4) chez *Candida glabrata*

Léa Conchou

► **To cite this version:**

Léa Conchou. Caractérisations biochimiques et structurales d'une métacaspase et d'une enzyme branchante de glucane en alpha-(1,4) chez *Candida glabrata*. Biologie structurale [q-bio.BM]. Université de Lyon, 2022. Français. NNT : 2022LYSE1016 . tel-03968929

**HAL Id: tel-03968929**

**<https://theses.hal.science/tel-03968929>**

Submitted on 2 Feb 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



N°d'ordre NNT : 2022LYSE1016

## **THESE de DOCTORAT DE L'UNIVERSITE DE LYON**

opérée au sein de  
**l'Université Claude Bernard Lyon 1**

**Ecole Doctorale N° ED205**  
**(Ecole Doctorale Interdisciplinaire Science et Santé)**

**Spécialité de doctorat :**  
**Discipline :** Biochimie et Biologie Structurale

Soutenue publiquement le 17/02/2022, par :  
**Léa Pauline Conchou**

---

# **Caractérisations biochimiques et structurales d'une métacaspase et d'une enzyme branchante de glucane en $\alpha$ -(1,4) chez *Candida glabrata***

---

Devant le jury composé de :

BOMPARD, Coralie	CR	UGSF, Lille	Rapporteuse
ROUSSEL, Alain	DR	AFMB, Marseille	Rapporteur
CAMADRO, Jean-Michel	DR	Institut Jacques Monod, Paris	Rapporteur
GIRARD-EGROT, Agnès	PR	ICBMS, Lyon	Examinatrice
BALLUT, Lionel	MC	MMSB, Lyon	Directeur de thèse
AGHAJARI, Nushin	DR	MMSB, Lyon	Co-directrice de thèse
DOUMECHÉ, Bastien	MC	ICBMS, Lyon	Invité
GERARD-BARAGGIA, Francine	MC	MMSB, Lyon	Invitée





# Université Claude Bernard – LYON 1

Président de l'Université	M. Frédéric FLEURY
Président du Conseil Académique	M. Hamda BEN HADID
Vice-Président du Conseil d'Administration	M. Didier REVEL
Vice-Président du Conseil des Etudes et de la Vie Universitaire	M. Philippe CHEVALLIER
Vice-Président de la Commission de Recherche	M. Petru MIRONESCU
Directeur Général des Services	M. Pierre ROLLAND

## COMPOSANTES SANTE

Département de Formation et Centre de Recherche en Biologie Humaine	Directrice : Mme Anne-Marie SCHOTT
Faculté d'Odontologie	Doyenne : Mme Dominique SEUX
Faculté de Médecine et Maïeutique Lyon Sud Charles Mérieux	Doyenne : Mme Carole BURILLON
Faculté de Médecine Lyon-Est	Doyen : M. Gilles RODE
Institut des Sciences et Techniques de la Réadaptation (ISTR)	Directeur : M. Xavier PERROT
Institut des Sciences Pharmaceutiques et Biologiques (ISBP)	Directrice : Mme Christine VINCIGUERRA

## COMPOSANTES & DEPARTEMENTS DE SCIENCES & TECHNOLOGIE

Département Génie Electrique et des Procédés (GEP)	Directrice : Mme Rosaria FERRIGNO
Département Informatique	Directeur : M. Behzad SHARIAT
Département Mécanique	Directeur M. Marc BUFFAT
Ecole Supérieure de Chimie, Physique, Electronique (CPE Lyon)	Directeur : Gérard PIGNAULT
Institut de Science Financière et d'Assurances (ISFA)	Directeur : M. Nicolas LEBOISNE
Institut National du Professorat et de l'Education	Administrateur Provisoire : M. Pierre CHAREYRON
Institut Universitaire de Technologie de Lyon 1	Directeur : M. Christophe VITON
Observatoire de Lyon	Directrice : Mme Isabelle DANIEL
Polytechnique Lyon	Directeur : Emmanuel PERRIN
UFR Biosciences	Administratrice provisoire : Mme Kathrin GIESELER
UFR des Sciences et Techniques des Activités Physiques et Sportives (STAPS)	Directeur : M. Yannick VANPOULLE
UFR Faculté des Sciences	Directeur : M. Bruno ANDRIOLETTI



## RÉSUMÉ

Les infections causées par diverses espèces *Candida* ont considérablement augmenté ces dernières années notamment dans un cadre hospitalier où elles entraînent une forte mortalité. Parallèlement à *Candida albicans*, *Candida glabrata* est l'une des levures pathogènes ayant la plus forte prévalence chez l'humain. À ce jour, les composés antifongiques utilisés en traitement présentent de nombreux phénomènes de résistance nécessitant l'identification de nouvelles cibles thérapeutiques. Afin d'explorer de nouvelles voies, nous nous sommes intéressés à la métacaspase de *C. glabrata*. Ces protéases à cystéines sont des homologues structuraux des caspases et des effecteurs de la mort cellulaire régulée, un processus physiologique visant à maintenir l'homéostasie cellulaire chez les métazoaires et les levures. L'étude structurale, fonctionnelle et enzymatique a permis d'élucider le déroulement de la maturation de l'enzyme dans un mécanisme *trans* avec ses changements structuraux associés, ainsi que son activation, les deux processus étant dépendants de la présence d'ions calcium. Nous avons également étudié une enzyme branchante de glucane en  $\alpha$ -(1,4) de la sous famille des GH13\_8. Ces dernières sont impliquées dans le métabolisme énergétique du glycogène et plus particulièrement dans l'introduction de ramifications *via* la création de liaisons  $\alpha$ -(1,6). L'étude biochimique et structurale a permis la résolution de la structure tridimensionnelle de l'enzyme sous une forme libre et en complexe avec un maltotriose. Nous avons également mis en évidence des particularités structurales de l'enzyme propres aux levures et plus particulièrement au phylum des *Saccharomycotina*.

Mots clés : *Candida glabrata*, mort cellulaire régulée, métacaspase, glycoside hydrolase, enzyme branchante de glucane, cristallographie aux rayons X



## ABSTRACT

Infections caused by various *Candida* species have dramatically increased during the recent years, especially in hospitals where they are held accountable for a high mortality rate. Along with *Candida albicans*, *Candida glabrata* is one of the most prevailing pathogenic yeasts to humans. To date, antifungal compounds used in treatment present numerous resistance phenomena requiring the identification of new therapeutic targets. In order to explore new pathways, we focused on a metacaspase from *C. glabrata*. These cysteine proteases are structural homologues of caspases and effectors of regulated cell death, a physiological process aiming at maintaining cellular homeostasis in metazoans and yeasts. Structural, functional and enzymatic studies have allowed us to elucidate that the maturation process of the enzyme follows a *trans* mechanism, with the associated structural reorganization and its activation, both processes being dependent on the presence of calcium ions. We also studied an  $\alpha$ -(1,4)-glucan branching enzyme from the GH13\_8 subfamily. This latter is involved in the energy metabolism of glycogen and more precisely in the introduction of  $\alpha$ -(1,6) glucose branching points. Biochemical and structural studies allowed determining the resolution of the three-dimensional structure of the enzyme in a free form and in complex with a maltotriose. We moreover highlighted structural features of the enzyme only present in yeasts and more particularly in the *Saccharomycotina* phylum.

Key words: *Candida glabrata*, regulated cell death, metacaspase, glycoside hydrolase, glucane branching enzyme, X-ray crystallography



## REMERCIEMENTS

L'ensemble des travaux présentés dans cette thèse ont été réalisés dans l'équipe de « Biocristallographie et Biologie Structurale des Cibles Thérapeutiques » sous la direction scientifique du Dr. Lionel Ballut et du Dr. Nushin Aghajari.

J'adresse mes remerciements au Dr. Christophe Geourjon, Directeur de l'Institut de Biologie et Chimie des Protéines et au Dr. Christophe Grangeasse, Directeur de l'Unité « Molecular Microbiology and Structural Biochemistry » (UMR 5086).

Je remercie le Dr. Coralie Bompard, le Dr. Jean-Michel Camadro et le Dr. Alain Roussel d'avoir accepté de juger mon travail. Je remercie également le Dr. Agnes Girard-Egrot, le Dr. Francine Gerard-Barragia et le Dr. Bastien Doumèche de l'intérêt qu'ils ont porté à mon travail en acceptant de participer à mon jury de thèse.

Je souhaite remercier tous les membres de mon comité de suivi individuel, le Dr. Samuel Tranier, le Dr. Pascale Cotton et le Dr. Olivier Marcillat. Je remercie très sincèrement le Dr. Bastien Doumèche pour ses précieux conseils et toutes ces discussions scientifiques passionnantes qui ont abouti à notre collaboration par la rédaction d'un article.

Je remercie l'ensemble de nos collaborateurs, le Dr. Florence Guillière, le Dr. Isabelle Gonçalves, le Dr. Bastien Doumèche, le Dr. Anne-Lise Bienvenu et le Pr. Stéphane Picot. Merci beaucoup pour vos contributions scientifiques respectives à mon travail de thèse.

Je remercie également la plateforme *Protein Science Facility* (PSF, SFR Biosciences, UMS3444) qui m'a permis d'utiliser les équipements dont j'avais besoin pour mes travaux. Merci à Virginie et Eric de m'avoir aidé à dompter une Äkta. Merci également à Adeline et Frédéric pour vos conseils et votre expertise en spectrométrie de masse.

Je tiens à présenter mes remerciements au Dr. Nushin Aghajari. Merci de m'avoir accueillie dans ton laboratoire depuis mon stage de master et de m'avoir, par la suite, donné l'opportunité de réaliser cette thèse. Je suis très reconnaissante d'avoir pu travailler à tes côtés durant ces trois dernières années.

J'adresse ma profonde gratitude à mon directeur de thèse, le Dr Lionel Ballut. Merci pour la confiance que tu m'as accordée tout au long de mon parcours ; aussi bien lors de notre premier projet réalisé ensemble pendant la licence ou lors de mon stage de master, que jusqu'au dernier jour de cette thèse. Cette expérience n'aurait jamais été la même sans toi. Je garderai en mémoire la passion avec laquelle tu m'as transmis ton savoir pendant ces trois



années. Merci pour toutes ces discussions aussi bien animées que passionnantes. Cette thèse nous aura donné du fil à retordre, beaucoup d'encre de marqueur sur ma paillasse pour déchiffrer nos résultats, beaucoup de tablettes de chocolat pour garder la motivation, beaucoup de cristaux et de bonbons pour survivre aux longues heures sur les lignes de lumières, beaucoup de déceptions, mais surtout beaucoup de réussites. Merci d'avoir été un moteur dans cette thèse, de m'avoir encouragé, écouté, conseillé (et supporté ?). J'ai énormément appris à tes côtés et je ressors grandie de cette expérience.

Je souhaite également remercier tous les membres de l'équipe BBSCT. Seb, merci infiniment pour toute l'aide que tu m'as apportée. Merci pour ta bonne humeur quotidienne, ta bienveillance, tes conseils et toutes ces conversations scientifiques (ou non) que l'on a partagé. Grâce à toi, j'ai presque cru que j'étais une vraie princesse pendant trois ans. Fred, je te remercie chaleureusement pour tous les moments que nous avons passés ensemble. Merci de m'avoir fait rire presque tous les jours depuis la première retraite MMSB. Merci de m'avoir fait découvrir la fameuse crêperie de Gif-sur-Yvette qui restera gravée dans ma mémoire, c'est dommage qu'ils ne servent pas de polonium là-bas. Merci de m'avoir fait mesurer et comparer les empreintes de toutes les personnes du laboratoire pour trouver l'identité du corbeau qui m'envoyait des lettres anonymes. Je te remercie aussi pour toute ton aide scientifique et celle que tu m'as apportée pour accaparer tous les portoirs roses de ce laboratoire. Je te confie la garde de Marie Rose, je sais qu'elle est entre de bonnes mains, s'il te plaît veille à ce que le fond d'écran ait toujours une touche de rose.

Je tiens également à remercier les membres de l'équipe « Retroviruses and Structural Biochemistry » du Pr. Patrice Gouet. Merci à tous d'avoir été pour moi cette grande équipe du rez-de-chaussée Nord. Tof, merci d'avoir partagé chaque jour/midi/facebook une partie de l'immensité de ta culture générale. Merci pour toutes ces conversations qui finissent par dériver sur des sujets complètement loufoques. J'espère qu'au travers de mon dictionnaire des djeun's j'ai pu te transmettre une partie de mon savoir de la génération Y et Z. Merci aussi pour tous tes conseils scientifiques, d'avoir vérifié toutes les amorces de ma trentaine de mutants (seulement ?) et d'avoir partagé avec moi ta recette secrète de la mutagenèse (= le protocole ça ne sert à rien). Xav, merci mille fois d'encore accepter de me parler après le milliard de questions que je t'ai posé sur Esript. Merci aussi de ne pas m'avoir pris pour une folle quand j'avais préparé presque une centaine de fiches de questions/réponses pour mon oral de bourse de thèse. Plus sérieusement, je te remercie d'avoir été un des piliers de ma thèse. Merci pour ton soutien moral qui m'a précieusement aidé dans les périodes les plus difficiles. Merci aussi pour tous les moments que l'on a passés en dehors du labo, les rires et

les discussions où l'on a refait le monde pendant des heures. Fra, merci pour ton enthousiasme quotidien, tes rires et ta gentillesse. Merci d'avoir été un peu notre maman à tous et d'avoir fait saliver nos papilles gustatives avec tous tes gâteaux. Merci aussi pour nos craquages à inventer une histoire d'amour entre les tabourets du labo ... Prend bien soin du petit dernier !

Cette expérience aurait été bien différente sans le soutien et la présence de mes fidèles co-thésards. Mathieu, on a commencé notre thèse ensemble (non, je sais, tu as commencé 1 mois avant mais je suis malgré tout la cheffe des thésards ...) et on a partagé le même bureau durant trois ans. Même si tout a mal commencé avec une sombre histoire d'amidon séché ... ta présence et ton soutien auront été précieux pour moi. Merci de nous avoir fait écouter chaque jour la reine des neiges, non en fait pas merci pour ça. Mais merci pour Céline Dion ! Merci pour ton rire que l'on peut entendre jusqu'au troisième étage de l'IBCP. Merci pour tes gâteaux aux ganaches roses, vertes, arc-en-ciel ... Non pas merci pour ça non plus. Merci pour toutes les bêtises que l'on a faites ensemble, les moments où on s'imaginait thésards alors qu'on était en M2 puis les moments où on a déprimé parce qu'on était en thèse. Aujourd'hui, tu voles vers de nouvelles aventures suédoises et je te souhaite le meilleur. Cecilia ou comme je préfère t'appeler, Ceci, merci pour cette amitié que nous avons créée. Merci d'avoir partagé avec moi ta culture italienne (ton huile d'olive surtout) et ta passion dévorante pour la montagne (je vais encore avoir besoin de mon instructrice d'escalade préférée). Merci d'avoir été mon bulletin météorologique des conditions climatiques sur toutes les montagnes de France pendant ces dernières années. Merci pour les moments que l'on a vécus ensemble, j'ai hâte de tous ceux qui suivront (cet hiver, vite !!). Je te lègue, non sans émotion, le très prisé titre de cheffe des thésards et te souhaite bon courage pour la dernière ligne droite qui t'attend (forza !). Camille, notre bébé thésarde (je sais tu es en deuxième année maintenant mais je suis dans le déni), merci pour ton incroyable gentillesse et ta douceur. Merci d'avoir égayé cette dernière année, d'avoir été ma partenaire de laboratoire à 6h30 du matin pendant le couvre-feu et de m'avoir laissé t'entraîner dans l'aventure des représentants de l'école doc et de DocEdiss. Je te souhaite tout le meilleur pour la suite de ta thèse.

Je remercie également Vinesh et Minh-Chau, je vous souhaite toute la réussite dans vos nouvelles aventures scientifiques et je vous remercie pour les moments, notamment culinaires, que nous avons passés ensemble. Vinesh, merci encore pour ma couronne de princesse qui trône sur mon bureau lorsqu'elle n'est pas sur ma tête. Johan, merci pour ta bonne humeur pendant les six mois qu'on a partagé à l'intérieur et en dehors du labo. Je n'ai

pas oublié, chose promise, chose due, le J c'est le S ! Mes pioupioups, Sarah et Chloé, merci à vous d'avoir été les meilleures stagiaires. Sarah, ta bonne humeur et ton enthousiasme ont apporté un vent de fraîcheur au labo lors de tes deux stages. Chloé, c'était un plaisir de partager ces six derniers mois avec toi. J'espère t'avoir transmis mon amour de l'organisation et des agendas. Merci d'avoir été autant à l'écoute quand je te racontais mon blabla scientifique ou tout le reste. À toi la relève maintenant, je te transmets tous mes encouragements pour la suite de ta thèse et j'ai hâte de revenir assister à ta soutenance !

Je souhaite remercier toutes les personnes de l'IBCP (MMSB ET LGBTI) avec qui j'ai pu partager ces trois dernières années. D'abord les anciens, partis trop tôt, Halima (merci pour ton soutien pendant la dernière ligne droite de ma thèse), Julien (je voudrais venir goûter la bière de Belgique !), Alexis et Arthur (merci pour cette session karaoké mémorable !), Sylvain (merci d'avoir été le Ron de mon Harry). Merci à mes copains de Bioch avec lesquels j'ai pu vivre cette aventure, Marine (de notre colocation en L2 à nos thèses à un étage d'écart ... nos routes professionnelles se seront finalement séparées à l'IBCP mais pas notre amitié qui nous réserve encore de belles surprises), Adèle (nos partages d'expérience de thèse, ton écoute et ton enthousiasme ont été précieux pendant les moments de doute, nos rendez-vous de la dep me manqueront mais je ne doute pas que nous trouverons un moyen de les remplacer), Hugo (que nos soirées nous mènent devant un match de foot ou dans une fosse aux ours, ces moments ont tous participé à rendre ces dernières années plus belles), Benjamin (on aura fini par les faire ces trois années, merci d'avoir été mon collègue représentant des doctorants, prof de nanoDSF, combattant de prolongement !) et Thibaut (aka titi, quoi tu n'es pas en bioch ? merci pour les fou rires, les soirées KL, les tours en vélo et j'en passe). Je remercie aussi ceux qui sont encore là et particulièrement Kelly (partenaire des pauses clope, merci d'avoir tenté de développer ma culture des films d'horreur ... Il me reste encore beaucoup à faire, merci aussi pour ton soutien surtout pendant le confinement quand on paniquait à l'idée de rester enfermées chez nous pour le restant de nos jours) et Altan (merci pour toutes ces pauses à m'écouter me plaindre et à me redonner le sourire). Je remercie également toute la team After Work pour tous les bons souvenirs partagés.

Un grand merci à la team Barbra de m'avoir fait l'honneur de participer au meilleur événement qui puisse exister le lundi soir ; le pub quizz du Flaningan's. Margaux, ma thésarde en cheffe, merci de m'avoir guidée au début de ma thèse, merci de m'avoir transmis la passion des mots fléchés et du bullet journal. Merci aussi de m'avoir intégré à la team Barbra même si je n'ai probablement jamais du vous aider à trouver une seule réponse .. Stelle, on ne s'est jamais rencontré au labo mais les bonnes ondes que tu avais laissé sur ton

bureau m'ont aidées pendant ces trois années, j'en suis certaine. Tu avais raison, c'était bien le meilleur bureau (visibilité imprenable sur les visiteurs venant nous assaillir + le seul avec un bidon gris pouvant contenir une quantité infinie de rhum). Merci pour tous les moments qu'on a pu partager ensemble, merci de comprendre mes goûts musicaux de djeun's et d'être aussi emballée que moi à l'idée de chanter sur du Céline Dion en soirée. Marine, merci de m'avoir laissée t'initier progressivement à la culture djeun's .. il semblerait d'ailleurs que cela ait porté ses fruits (ipad, iphone, compte instagram .. Qu'elle sera la prochaine étape ?). J'ai découvert avec toi qu'il était possible d'avoir plus de plantes que de mètres carrés, que les plantes pouvaient vivre plus d'une année consécutive (comme tu l'as dit dans tes propres remerciements .. RIP mon terrarium) et que le plante sitting était une vraie activité que j'ai prise avec beaucoup de sérieux ! Beaucoup de plantes donc ... mais ... pas que, beaucoup de musique aussi ! Les lundis matin n'auront jamais été aussi difficiles que pendant notre période sucrée, fief de mon secteur géographique de prédilection : la presqu'île (Lyon = la presqu'île + le reste ... sauf quand il faut que je gare ma voiture dans ton garage). Heureusement que de temps en temps nos pas de danse nous ont traînés en cellule de dégrisement (l'IBCP). Merci pour ton aide, pas seulement quand tu as accepté de relire ma thèse (!! ) mais aussi tout au long de cette aventure. Merci aussi pour ta présence, ton écoute et tes conseils mais surtout ... merci pour le tiramisu !!

Je remercie aussi Elisa, Romane et Julie avec qui j'ai pu partager la passionnante, folle et épuisante aventure de DocEDISS et de l'organisation de la journée des doctorants. Merci à vous les filles, je suis heureuse d'avoir fait votre rencontre et vous souhaite à chacune le meilleur.

À mes chères amies, Lucie, Imen (aka Mimouche), Justine, Marie, Morgane, Laurène, Candice, vous faites toutes parties de mon équilibre au quotidien et je ne pourrais pas imaginer avoir traversé ces trois dernières années sans vous. Je souhaite à tout le monde d'avoir des amies aussi merveilleuses que vous. Quelques lignes ne suffiraient pas à exprimer toute ma gratitude. Lucie, dès les bancs de la fac on a trouvé en l'une et en l'autre une partenaire de laboratoire et maintenant de thèse mais au-delà de ça, une partenaire de vie. Je te souhaite tout le courage et la réussite pour ton doctorat et je tacherai de te soutenir autant que tu l'as fait pour moi. Mimouche, personne ne peut me faire rire autant que toi, tu es le clown de mes soirées et le grumpy cat de mes journées. Merci d'avoir mis des paillettes dans ma vie. Ju, on évolue ensemble depuis longtemps maintenant, merci pour toutes nos discussions sur le monde, nos remises en question, tu es une vraie source d'inspiration pour moi. Marie, depuis le lycée jusqu'au ... confinement (tu entends pétrir ?) et tous les jours au quotidien, tu es

l'épaule sur laquelle je peux me reposer, toujours là pour moi, merci. Morgane, ma sœur, merci d'être une si belle personne qui me pousse chaque jour à devenir la meilleure version de moi-même. Merci de partager cette passion complètement inexplicable pour les musiques des années 80 et surtout Céline. Laurène, on se connaît depuis l'âge de 6 ans, un déménagement et 12 ans plus tard on s'est retrouvées à Lyon. Je suis heureuse de pouvoir te compter parmi mes amies proches et la fin de cette thèse marque le début d'une jolie aventure pour nous deux. Candice, tu as été un modèle de travail et de persévérance, tu es l'exemple que quand on s'en donne les moyens on arrive à tout et regarde où tu en es aujourd'hui. Merci, pour tout ce que tu as su m'apporter. Chacune à votre manière ou toutes en même temps vous avez contribué à rendre plus douces ces dernières années et aussi à me rendre meilleure. Merci d'avoir toujours été là, de m'avoir écoutée parfois pendant des heures, soutenue, fait rire, danser et chanter. Je n'aurais jamais réussi sans vous, merci.

Je terminerai ces remerciements en remerciant du fond du cœur ma famille. Papa, Maman, merci de nous avoir toujours donné les clés pour réussir avec Thomas, je vous en serai éternellement reconnaissante. Merci d'avoir été et de continuer à être des parents si présents. Vous n'avez jamais douté de moi, pas même quand je vous ai annoncé que cinq ans d'études ça ne me suffisait pas. Merci à mon grand frère, Thomas, deux univers nous sépare professionnellement et pourtant, tu as toujours su trouver les mots pour m'aider à surmonter les épreuves et à aller jusqu'au bout. Merci d'avoir été un exemple depuis le plus jeune âge, même si à cette époque c'était l'amour vache entre nous, je suis heureuse de la proximité qu'on a aujourd'hui. Merci à ma grand-mère, que j'écoute depuis l'enfance me conter ses études de médecine et de biologie. Merci à toi de m'avoir appris que tout était possible et merci d'être une femme si inspirante.





## Contributions scientifiques issues de la thèse

### Articles dans des revues scientifiques à comité de lecture international :

L. Conchou, J. Martin, I. R. Gonçalves, F. Galisson, S. Violot, F. Guillière, N. Aghajari, L. Ballut. 2021. **The *Candida glabrata* glycogen branching enzyme structure reveals unique features of branching enzymes of the *Saccharomycetaceae* phylum.** *Glycobiology*. <https://doi.org/10.1093/glycob/cwab110>.

*Soumis en Juillet 2021 – Accepté en Octobre 2021*

L. Conchou, B. Doumèche, F. Galisson, S. Violot, C. Dugelay, E. Diesis, A. Page, N. Aghajari, L. Ballut. 2021. **Structural and molecular determinants of *Candida glabrata* metacaspase maturation and activation by calcium.**

*Soumis en Décembre 2021 à Communications Biology*

### Communication orale

**Étude d'une cible thérapeutique potentielle face aux infections à *Candida glabrata* : caractérisation biochimique et structurale de Métacaspase.**

Congrès national CAPF (Consortium Anti-Parasitaire et Anti-Fongique) – Mars 2020 (Montpellier)

**Study of a potential therapeutical target against *Candida glabrata* infections: structural and biochemical characterizations of a program cell death implicated protease.**

Séminaire interne de l'IBCP – Octobre 2020 et congrès de la journée de l'EDISS – Novembre 2020 (Lyon)

### Communication sous forme de poster

**Glycoside hydrolase as new therapeutic targets against *Candida glabrata* infections.**

Congrès national AFC (Association Française des Cristallographes) - Juillet 2018 (Lyon)

**Three-dimensional structure of a GH13 family enzyme from *Candida glabrata*.**

Congrès international ALAMY (Symposium on the Alpha-Amylase Family) – Octobre 2019 (Bratislava, Slovaquie)



**Three-dimensional structure of a GH13 family enzyme from *Candida glabrata* complexed to maltotriose.**

Congrès national CAPF (Consortium Anti-Parasitaire et Anti-Fongique) - Février 2020 (Montpellier)

**Contributions scientifiques indépendantes du sujet de thèse**

Article dans des revues scientifiques à comité de lecture international :

S. Violot, F. Galisson, L. Carrique, V. Jugnarain, L. Conchou, X. Robert, A. Thureau, W. Helbert, N. Aghajari, L. Ballut. **Exploring molecular determinants of polysaccharide lyase family 6-1 enzyme activity.** 2021. *Glycobiology*. <https://doi.org/10.093/glycob/cwab073>.

**« Avec un talent ordinaire et une persévérance  
extraordinaire, on peut tout obtenir »**

*Thomas Buxton*



# SOMMAIRE

## Liste des abréviations

## INTRODUCTION

### Liste des figures et des tableaux

<b>I.</b>	<b>CANDIDA GLABRATA, UN PATHOGENE OPPORTUNISTE DE L'HOMME</b> .....	<b>31</b>
A.	BIOLOGIE DE <i>CANDIDA GLABRATA</i> .....	31
1.	<i>Découverte et classification</i> .....	31
2.	<i>Morphologie</i> .....	32
3.	<i>Génome</i> .....	34
B.	PATHOGÉNÉCITÉ DE <i>CANDIDA GLABRATA</i> .....	34
1.	<i>Epidémiologie et manifestations cliniques</i> .....	34
2.	<i>Facteurs de risque</i> .....	35
3.	<i>Facteurs de virulence</i> .....	36
C.	LES ANTIFONGIQUES.....	39
1.	<i>Azols</i> .....	40
2.	<i>Echinocandines</i> .....	43
3.	<i>Polyènes</i> .....	43
4.	<i>Fluoropyrimidines</i> .....	44
<b>II.</b>	<b>PROJET DE THESE</b> .....	<b>46</b>
<b>III.</b>	<b>REFERENCES</b> .....	<b>47</b>

## CHAPITRE I - Caractérisation des relations structure-fonction-activité de la métacaspase chez l'organisme pathogène *Candida glabrata*

### Liste des figures et des tableaux

<b>I.</b>	<b>LE PHENOMENE DE MORT CELLULAIRE REGULEE</b> .....	<b>59</b>
A.	LA MORT CELLULAIRE REGULEE CHEZ LES METAZOAIRES .....	59
1.	<i>L'apoptose</i> .....	59
2.	<i>Les caspases</i> .....	64
B.	LA MORT CELLULAIRE CHEZ LA LEVURE.....	72
1.	<i>Mort cellulaire dépendante de la mitochondrie</i> .....	72
2.	<i>Les métacaspases</i> .....	75
<b>II.</b>	<b>RESULTATS</b> .....	<b>92</b>
A.	MATURATION DEPENDANTE DU CALCIUM DE CgMCA-I .....	92
1.	<i>Purification de CgMCA-I</i> .....	92
2.	<i>Effet des cations divalents sur la maturation</i> .....	94
B.	CARACTERISATION STRUCTURALE DE CgMCA-I .....	95
1.	<i>Structure de CgMCA-I<sup>Ca</sup></i> .....	98
2.	<i>Structure de CgMCA-I<sup>Mg</sup></i> .....	102
C.	CHANGEMENTS STRUCTURAUX DE CgMCA-I ASSOCIES A LA MATURATION DEPENDANTE DU CALCIUM .....	104
D.	PROCESSUS DE MATURATION .....	108
E.	CARACTERISATION ENZYMATIQUE DE CgMCA-I .....	111
<b>III.</b>	<b>DISCUSSION</b> .....	<b>116</b>
<b>IV.</b>	<b>CONCLUSION ET PERSPECTIVES</b> .....	<b>122</b>
<b>V.</b>	<b>MATERIEL ET METHODES</b> .....	<b>126</b>
A.	EXPRESSION ET PURIFICATION DE CgMCA-I .....	126
1.	<i>Optimisation des gènes</i> .....	126
1.	<i>Mutagenèse dirigée</i> .....	128

2.	<i>Expression et purification</i> .....	130
3.	<i>Etude du processus de maturation</i> .....	131
B.	ANALYSE PAR SPECTROMETRIE DE MASSE.....	132
1.	<i>Identification des sites de coupure de CgMCA-I par LC-MS</i> .....	132
C.	ANALYSE ENZYMATIQUE.....	133
D.	ANALYSE STRUCTURALE.....	137
1.	<i>Analyse par diffraction aux rayons X</i> .....	137
2.	<i>Analyse par diffusion des rayons X aux petits angles (SAXS)</i> .....	138
VI.	REFERENCES.....	140

## CHAPITRE II - Caractérisation biochimique et structurale de l'enzyme branchante de glucane en $\alpha$ -(1,4) chez l'organisme pathogène *Candida glabrata*

### Liste des figures et des tableaux

I.	L'ENZYME BRANCHANTE DE GLUCANE EN $\alpha$ -(1,4) DE <i>CANDIDA GLABRATA</i> .....	155
A.	LES GLYCOSIDE HYDROLASES .....	155
1.	<i>Classification</i> .....	155
B.	LES ENZYMES BRANCHANTES .....	160
1.	<i>Substrat</i> .....	160
2.	<i>Mécanisme catalytique</i> .....	163
3.	<i>Caractérisation structurale des BEs GH13_9</i> .....	166
4.	<i>Caractérisation structurale des BEs GH13_8</i> .....	175
II.	CARACTERISATION STRUCTURALE DE CGGBE.....	181
A.	ARTICLE .....	182
B.	FIGURES ET TABLEAUX .....	208
C.	FIGURES SUPPLEMENTAIRES.....	216
III.	CONCLUSION ET PERSPECTIVES.....	228
IV.	REFERENCES.....	231

## Liste des abréviations

5-FC : 5-fluorocytosine  
5-F(d)UMP : 5-fluorodesoxyuridine monophosphate  
5-FU : 5-fluorouracile  
5-FUTP : 5-fluorouracile triphosphate  
AA : « Auxiliary Activities », activité auxiliaire  
ABC : « ATP Binding Cassette », transporteurs à cassettes liant l'ATP  
ACD : « Accidental Cell Death », mort cellulaire accidentelle  
ACN : Acétonitrile  
AF2 : AlphaFold 2  
AIC : « Akaike Information Criteria », critère d'information Akaike  
AIF1 : « Apoptosis-induction factor 1 », facteur 1 d'induction de l'apoptose  
Als : « Agglutinin-like sequence », séquence de type agglutinine  
AmB : Amphotéricine B  
AMC : 7-amino-4-methyl coumarin  
APAF1 : « Apoptotic Peptidase Activating Factor 1 », facteur 1 d'activation de la peptidase apoptotique  
BE : « Branching Enzyme »  
CAI : « Codon Adaptation Index », indice d'adaptation des codons  
CARD : « CASpase Recruitment Domain », domaine de recrutement des caspases  
Caspase : « Cystein-dependant ASPartate-specific peptidASE », protéases cystéine dépendantes spécifiques de l'aspartate  
CAZymes : « Carbohydrate-active enzymes », enzymes actives sur les glucides  
CBM : « Carbohydrate Binding Module », module de liaisons aux glucides  
CDR : « *Candida* Drug Resistance », pompes de résistance des espèces *Candida*  
CE : « Carbohydrate Esterase »  
Co-NTA : Cobalt – acide nitriloacétique  
CWP : « Cell Wall Protein », protéine de la paroi fongique  
DD : « Death Domain », domaine de mort  
DED : « Death Effector Domain », domaine des effecteurs de mort  
DIABLO : « Direct IAP Binding protein with LOw pI », protéine de liaison aux IAP de faible point isoélectrique  
DISC : « Death Inducing Signaling Complex », complexe de signalisation induisant la mort  
dNTP : désoxyribonucléoside triphosphate  
DO : densité optique  
EC : « Enzyme Commission », Commission d'enzyme  
EMA : « European Medicines Agency », agence européenne des médicaments  
EPA : « Epithelial adhesins », adhésines épithéliales  
FADD : « Fas-associated Death Domain », facteur Fas associé au domaine de mort  
FDA : « US Food and Drug Administration », administration américaine des aliments et des médicaments  
FWHM : « Full Width at Half Maximum », largeur à mi-hauteur  
GBE : « Glycogen Branching Enzyme », enzyme de branchement du glycogène  
GH : « Glycoside Hydrolases »

GPI : Glycophosphatidylinositol

GT : Glycoside transferase

GS : « Glycogen synthase », glycogène synthase

Hwp1 : « Hyphae wall protein », protéine filamenteuse de la paroi

IAP : « Inhibitor of Apoptosis », inhibiteur de l'apoptose

ICE : « Interleukine-1 $\beta$  Converting Enzyme », enzyme de conversion de l'interleukine-1 $\beta$

IP3R : « Inositol 3 phosphate receptor type 1 », récepteur de type 1 de l'inositol 3 phosphate

IL-1 $\beta$  : « Interleukine-1 $\beta$  », interleukine 1

IPTG : isopropyl- $\beta$ -D-thiogalactopyranoside

IUB-MB : « International Union of Biochemistry et Molecular Biology », Union internationale de biochimie et de biologie moléculaire

LB : Luria-Bertani

MFS : « Major Facilitator Superfamily », superfamille des facilitateurs majeurs

MOMP : « Mitochondrial Outer Membrane Permeabilization », perméabilisation de la membrane externe mitochondriale

NCAC : « Non *Candida albicans* *Candida* », *Candida* non albicans

NCCD : « Nomenclature Committee on Cell Death », comité de la nomenclature de la mort cellulaire

Ni-NTA : Nickel – acide nitriloacétique

NUC1 : « endonuclease-G », endonucléase G

NNISS : « National Nosocomial Infections Surveillance », surveillance nationale des infections nosocomiales

PCD : « Programmed Cell Death », mort cellulaire programmée

PCR : « Polymerase Chain Reaction », réaction de polymérisation en chaîne

Pir : « Protein with internal repeats », protéine avec des répétitions internes

PL : « Polysaccharide Lyases »

PMCA : « Plasma-membrane Ca<sup>2+</sup>-ATPase », ATPase calcium dépendante de la membrane plasmique

PTP : « Permeability Transition Pore », pore de perméabilité membranaire

RMSD : « Root Mean Square Deviation », écart quadratique moyen

RCD : « Regulated Cell Death », mort cellulaire régulée

RE : Réticulum endoplasmique

ROS : « Reactive Oxygen Species », espèces réactives de l'oxygène

SAPs : « Secreted Aspartyl Proteinase », aspartate protéase sécrétée

SBE : « Starch Branching Enzyme », enzyme de branchement de l'amidon

SEC : « Steric Exclusion Chromatography », chromatographie d'exclusion stérique

SERCA : « Sarco(endo)plasmic reticulum Ca<sup>2+</sup>-ATPase », ATPase calcium dépendante du réticulum endoplasmique

SMAC : « Second Mitochondrial Activator of Caspase », second activateur mitochondriale des caspases

TFA : Acide trifluoroacétique

TNF : « Tumor Necrosis Factor », facteur de nécrose tumorale

TRAILR1/R2 : « TNF-related apoptosis inducing ligand-receptor 1/2 », récepteur 1/2 du ligand inducteur d'apoptose lié au TNF

UPRTase : uracile phosphoribosyl transferase

VIH : Virus de l'immunodéficience humaine

Ycal : « Yeast caspase-1 », caspase 1 de levure

YMUC : « Yeast mitochondrial Unselective Channel », canal mitochondrial non sélectif des levures





# **INTRODUCTION**



## Liste des figures et des tableaux

### FIGURES

1. Arbre phylogénétique des espèces du genre *Nakaseomyces* et d'autres levures p 32
2. Levures de *Candida glabrata* observées en microscopie électronique p 32
3. Image de la paroi fongique de *C. glabrata* par microscopie à transmission électronique p 33
4. Variation de la distribution géographique des espèces *Candida* impliquées dans des infections invasives p 35
5. Pathogénicité des infections invasives des espèces *Candida* p 36
6. Représentation schématique de l'évasion du système immunitaire de *C. albicans* et *C. glabrata* p 38
7. Cibles thérapeutiques des antifongiques Azolés, Echinocandines, 5'-fluorocytosine et Polyènes p 39
8. Représentation schématique de la voie de synthèse de l'ergostérol chez *C. glabrata* p 40
9. Vue schématique des différents mécanismes de résistance aux antifongiques azolés chez les levures *Candida* p 42
10. Représentation moléculaire de l'Amphotéricine B et de son assemblage dans les membranes plasmiques p 44
11. Mécanisme d'action de la 5-fluorocytosine p 45

### TABLEAUX

1. Comparaison des traits morphologiques et pathogéniques de *S. cerevisiae*, *C. albicans* et *C. glabrata* p 39



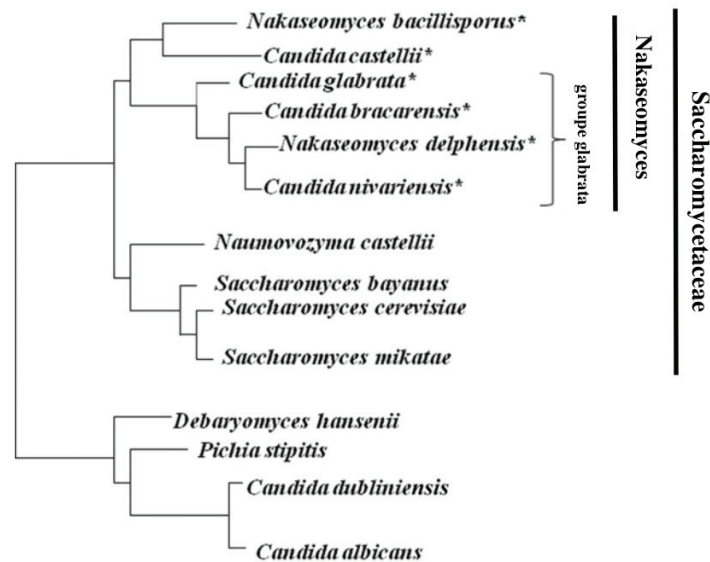
# I. *Candida glabrata*, un pathogène opportuniste de l'Homme

## A. Biologie de *Candida glabrata*

### 1. Découverte et classification

Historiquement, l'espèce *Candida glabrata* a été isolée pour la première fois en 1894 par Berlese puis identifiée chez l'Homme par Anderson en 1917 qui faisaient alors référence à l'espèce sous le nom de *Cryptococcus glabratus*. En 1894, ce champignon a été transféré au genre *Torulopsis*, dont les espèces sont caractérisées par l'absence de production de filaments mycéliens, et nommé *Torulopsis glabrata* (Fidel *et al.* 1999; Fidel *et al.* 2001). Ce n'est qu'en 1978, après que le développement sous la forme de filaments mycéliens ou pseudo-hyphes ne soit plus considéré comme un caractère distinctif des levures à *Candida* que les genres *Torulopsis* et *Candida* ont fusionné (Yarrow and Meyer 1978). Aujourd'hui, la communauté scientifique fait donc référence à l'espèce sous le terme de *Candida glabrata*. Pourtant, les récentes études en phylogénétique ont démontré une relation distante entre *C. glabrata* et la levure pathogène *C. albicans* et une plus grande proximité avec les espèces du genre *Saccharomyces* (Figure 1, Tableau 1). Depuis 2003, l'espèce a alors été reclassée au sein du genre *Nakaseomyces* par Kurtzman avec d'autres espèces dont il a été admis qu'elles possédaient toutes un ancêtre commun (Kurtzman 2003). Ce genre est divisé en deux lignées distinctes dont celle appelée « groupe *glabrata* » où on retrouve notamment les trois espèces pathogènes, *Candida bracarensis*, *Candida nivariensis* et *Candida glabrata* ainsi que *Nakaseomyces delphensis* (Figure 1).

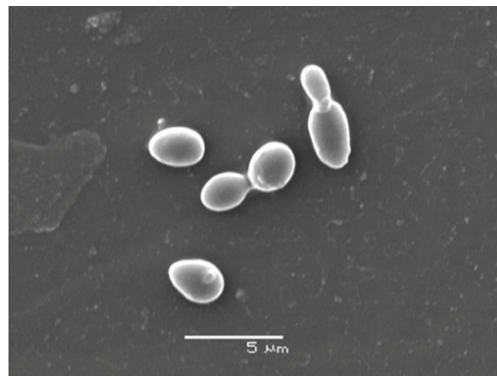
La taxonomie actuelle de *C. glabrata* est le règne des *Fungi*, sous-règne *Dikarya*, embranchement *Ascomycota*, sous-embranchement *Saccharomycotina*, classe *Saccharomycetes*, ordre *Saccharomycetales*, famille *Saccharomycetaceae*, genre *Nakaseomyces*, clade *Candida* et espèce *glabrata* (Kumar *et al.* 2019).



**Figure 1. Arbre phylogénétique des espèces du genre *Nakaseomyces* et d'autres levures d'après Ahmad *et al.* 2014.** La construction de l'arbre est basée sur l'alignement de 603 gènes orthologues, les espèces *Candida glabrata*, *Candida bracarensis*, *Nakaseomyces delphensis* et *Candida nivariensis* font parties du groupe glabrata.

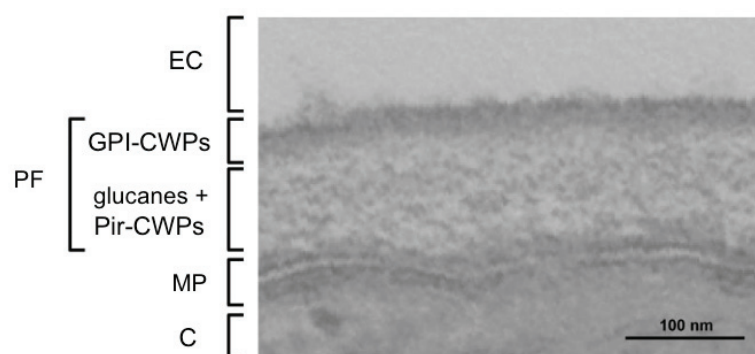
## 2. Morphologie

Sur un milieu standard, *C. glabrata* se présente sous forme de colonies brillantes, lisses et de couleur crème (Rodrigues *et al.* 2014). Microscopiquement, elle se présente sous la forme de levures allant de 1 à 4  $\mu\text{m}$  se multipliant par bourgeonnement (Figure 2). La différence majeure ayant provoqué tous les débats autour de la taxonomie de l'espèce est due à son incapacité à se développer sous la forme de filaments comme les autres espèces *Candida* (Tableau 1). Toutefois, en 2000, Csank et Haynes ont démontré que *C. glabrata* pouvait avoir une croissance en pseudo-filaments sur un milieu carencé en azote. Ce type de croissance n'a cependant jamais été rencontré dans des cas cliniques (Csank and Haynes 2000 ; Kaur *et al.* 2005).



**Figure 2. Levures de *Candida glabrata* observées en microscopie électronique d'après Ahmad *et al.* 2014.** L'image permet de distinguer deux cellules uniques et deux paires de cellules mère-fille pendant le processus de bourgeonnement.

Comme toutes les levures, *C. glabrata* est constituée d'une paroi fongique, une organelle dynamique et essentielle impliquée dans des processus importants. En effet, elle confère à la cellule une résistance mécanique lui permettant de maintenir sa forme mais également de résister aux variations de pressions osmotiques ou aux attaques imposées par l'environnement. La paroi conserve néanmoins une certaine plasticité permettant à la cellule de croître, se diviser et interagir avec son environnement. La paroi de *C. glabrata* a une épaisseur de 100 à 200 nm et ses constituants sont très semblables à ceux de *S. cerevisiae* et *C. albicans* (Figure 3) (de Groot *et al.* 2008). Parmi ces derniers on retrouve la chitine, un homopolymère linéaire de N-Acétyleglucosamine liés en  $\beta$ -(1,4), représentant seulement 1,1 à 1,3 % de la paroi (de Groot *et al.* 2008). On retrouve également des chaînes de glucanes et notamment le glucane  $\beta$ -(1,3), composant majeur de la paroi formant un réseau dans la partie inférieure. Le glucane  $\beta$ -(1,6) est présent en quantité bien moindre dans la partie extérieure de la paroi, il relie les composants entre eux et fait notamment le lien entre le glucane  $\beta$ -(1,3) et les mannoprotéines. Ces dernières sont aussi appelées CWP (Protéines de la paroi fongique ; « Cell Wall Proteins ») et sont réparties en deux classes : les Pir (Protéines avec des répétitions internes ; « Proteins with internal repeats ») qui sont liées au glucane  $\beta$ -(1,3) *via* des liaisons glycosidiques et les protéines à ancre glycosylphosphatidylinositol (GPI) dont font partie les adésines qui sont liées covalentement au glucane  $\beta$ -(1,6). Ces mannoprotéines ont un rôle essentiel dans les phénomènes d'adhérence, d'interactions hôte-pathogène et globalement dans la mise en place de la virulence de *C. glabrata* (De Groot *et al.* 2005 ; De Groot *et al.* 2008).



**Figure 3. Image de la paroi fongique de *C. glabrata* par microscopie à transmission électronique adapté de De Groot *et al.*, 2008.** La paroi est constituée de deux couches dont la partie la plus externe est composée de GPI-CWPs covalentement liées au glucane  $\beta$ -(1,6) et la partie la plus interne composée de Pir-CWPs liées au glucane  $\beta$ -(1,3). EC: domaine extra-cellulaire, PF: Paroi fongique, MP: membrane plasmique, C: Cytosol.



### 3. *Génome*

*C. glabrata* est une levure ayant un génome haploïde dont le génome de la souche CBS 138 a été intégralement séquencé par Dujon *et al.* en 2004 (Dujon *et al.* 2004). Ce dernier est composé de 13 chromosomes, nommés de A à M, pour un total de 12,3 Mb. Parmi les 5294 gènes de *C. glabrata*, 4870 sont orthologues à *S. cerevisiae* démontrant un haut degré de synténie (Kumar *et al.* 2019). De plus, les protéines orthologues aux deux espèces partagent en moyenne 65 % d'identité de séquence (Kaur *et al.* 2005). Cependant, en comparaison à levure non pathogène *S. cerevisiae*, *C. glabrata* a perdu de nombreux gènes impliqués dans l'assimilation de sucre, le métabolisme du phosphate, de l'azote et du sulfure (Kaur *et al.* 2005). Il a été suggéré que cette différence pourrait être associée à l'émergence du caractère pathogène de l'espèce et à son interaction avec l'hôte mammifère (Dujon *et al.*, 2004; Gabaldón *et al.* 2013).

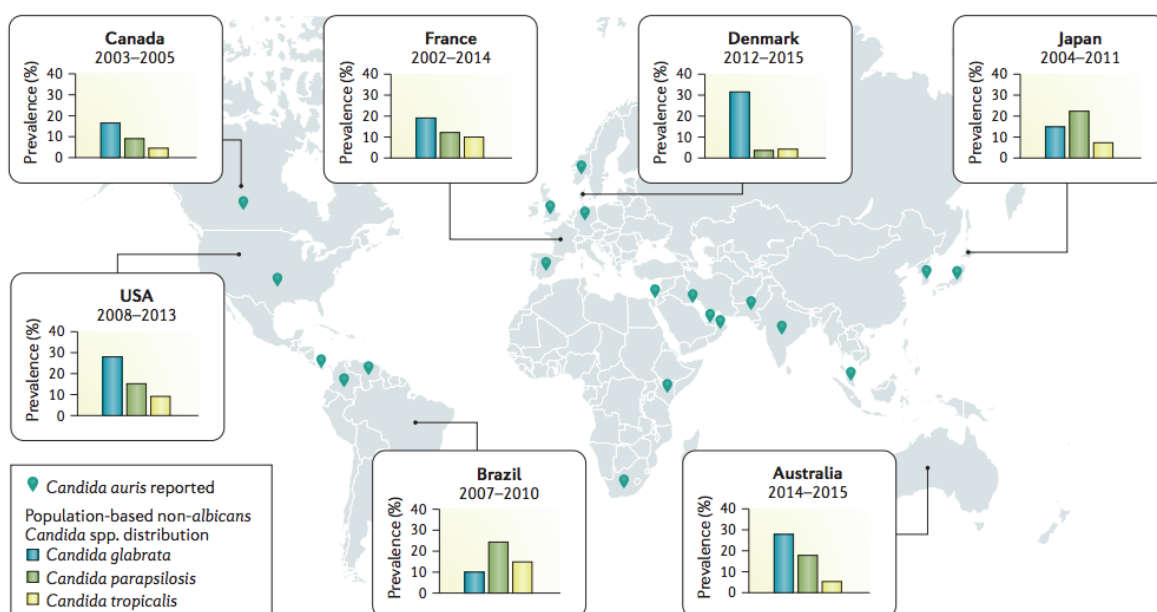
## B. Pathogénicité de *Candida glabrata*

Les levures *Candida glabrata* sont des levures majoritairement retrouvées à l'état commensal et peuplant la flore microbienne naturelle de l'Homme. Ce sont des saprophytes des muqueuses digestives, génitales et urinaires. *C. glabrata* est aussi un pathogène opportuniste n'exprimant son pouvoir pathogène qu'en présence de facteurs favorisants dits « facteurs de risque ». Aujourd'hui, *Candida glabrata* est considérée comme la seconde cause la plus fréquente d'infections nosocomiales juste après *Candida albicans* (Hassan *et al.* 2021).

### 1. *Epidémiologie et manifestations cliniques*

Les infections à *Candida* peuvent se manifester cliniquement sous la forme d'infections superficielles dites candidoses ou d'infections invasives dites candidémies. Les infections superficielles se retrouvent principalement au niveau des muqueuses et de la peau et peuvent engendrer des candidoses cutanées, oropharyngées, vulvo-vaginales ou encore urinaires. Bien que *C. albicans* reste le microorganisme majoritairement responsable de ce type d'infections, *C. glabrata* est la seconde cause la plus fréquente de candidoses (Pappas *et al.* 2018). En effet, il est mis en cause dans 20 % des cas pour les candidoses urinaires (Sobel *et al.* 2000 ; Lundstrom and Sobel 2001) et les candidoses vulvo-vaginales aux Etats-Unis, en Europe et en Australie (Gonçalves *et al.* 2016). En Afrique, Asie et en Inde, *Candida glabrata* est même isolée plus souvent que *C. albicans* dans des cas d'infections fongiques vaginales (Gonçalves *et al.* 2016).

Les infections invasives peuvent provoquer une fièvre modérée ou se propager jusque dans le sang et conduire à des septicémies. Ces candidémies sont associées aux infections nosocomiales du sang dans les établissements de santé notamment dans les unités de soin intensif où elles engendrent de forts taux de mortalité (Hassan *et al.* 2021). Les espèces *Candida* sont la quatrième cause d'infections nosocomiales du sang après *Staphylococcus aureus*, les espèces de staphylocoques à coagulase négative et les espèces *Enterococcus* (Pappas *et al.* 2018). De plus, elles sont responsables de 18 % des candidémies rapportées dans les centres de soin intensifs au niveau mondial (Pappas *et al.* 2018). Bien que *C. albicans* continue d'être l'espèce ayant la plus grande prévalence pour ces infections invasives, les dernières décennies ont vu apparaître des espèces *Candida* dites « *Candida non albicans* » ou NCAC (« *Non Candida albicans candida* ») avec une variabilité géographique (Figure 4). *Candida glabrata* est notamment la seconde espèce la plus retrouvée aux Etats-unis et en Europe de l'ouest (environ 20 % des candidémies en France entre 2002 et 2014) et est responsable d'infections invasives associées à une mortalité de 40 % (Pappas *et al.* 2018 ; Hassan *et al.* 2021).

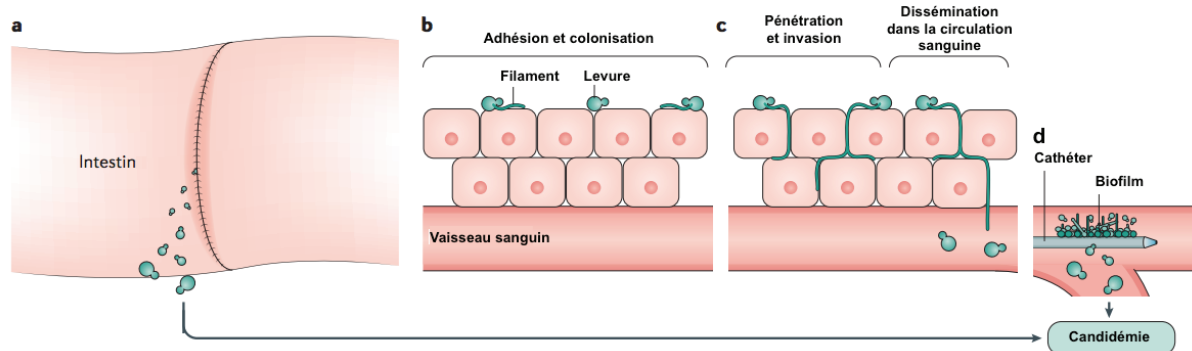


**Figure 4. Variation de la distribution géographique des espèces *Candida* impliquées dans des infections invasives d'après Pappas *et al.* 2018.** La prévalence (en %) des espèces *Candida glabrata*, *Candida parapsilosis* et *Candida tropicalis* est représentée pour le Canada, la France, le Danemark, le Japon, les États-Unis, le Brésil et l'Australie ainsi que la distribution des cas pour lesquels *Candida auris* était impliqué d'après les données du Centre de prévention et de contrôle des maladies (<https://www.cdc.gov/fungal/candida-auris/>).

## 2. Facteurs de risque

Lorsque des perturbations des muqueuses intestinales et/ou une baisse du système immunitaire apparaissent, les espèces *Candida* peuvent passer de l'état commensal à l'état de pathogène opportuniste. Parmi les facteurs de risques on retrouve l'utilisation au long terme

ou répétée d'antibiotiques à large spectre tels que l'ampicilline ou la tétracycline. En effet, la flore bactérienne étant le principal mécanisme de défense contre les espèces *Candida*, l'utilisation d'antibiotiques permet une augmentation de leur colonisation (Gonçalves *et al.* 2016 ; Pappas *et al.* 2018). Un autre facteur de risque est la rupture de la barrière gastro-intestinale due à une chirurgie, une perforation ou à l'utilisation de cathéters en intraveineuse permettant le passage du pathogène, des muqueuses intestinales à la circulation sanguine (Figure 5) (Galocha *et al.* 2019). La troisième cause est une diminution du système immunitaire dûe à des traitements de chimiothérapie, l'utilisation d'immuno-suppresseurs lors d'un transfert d'organe ou encore des maladies compromettant le système immunitaire comme le virus de l'immunodéficience humaine (VIH) (Figure 5 c) (Gonçalves *et al.* 2016 ; Pappas *et al.* 2018).



**Figure 5. Pathogénicité des infections invasives des espèces *Candida* adapté de Pappas *et al.* 2018.** Les infections invasives à *Candida* peuvent être dues à une rupture de la barrière gastro-intestinale (a), à une baisse du système immunitaire entraînant une invasion du pathogène au travers des barrières épithéliales et sa dissémination dans le sang (b, c) ou encore à l'utilisation de cathéters en intra-veineuse (d).

L'augmentation de la fréquence d'identification de *C. glabrata* serait aussi liée à un âge avancé du patient. En effet, *C. glabrata* serait majoritairement isolée chez des patients dont l'âge est supérieur à 70 ans (Hassan *et al.* 2021). Enfin, l'augmentation de l'incidence de *C. glabrata* et d'autres NCAC pourrait s'expliquer par leur faible susceptibilité aux antifongiques azolés dont l'utilisation prophylactique aurait entraîné une sélection des espèces moins sensibles (Goncalves *et al.*, 2016).

### 3. Facteurs de virulence

Lors d'une infection, les espèces *Candida* ont la possibilité d'employer des mécanismes d'agression pour favoriser leur colonisation et leur invasion, on parle de facteurs de virulence. Un des facteurs de virulence des espèces *Candida* est leur capacité à se développer sous une forme filamenteuse afin de traverser la barrière épithéliale et se disséminer dans la circulation

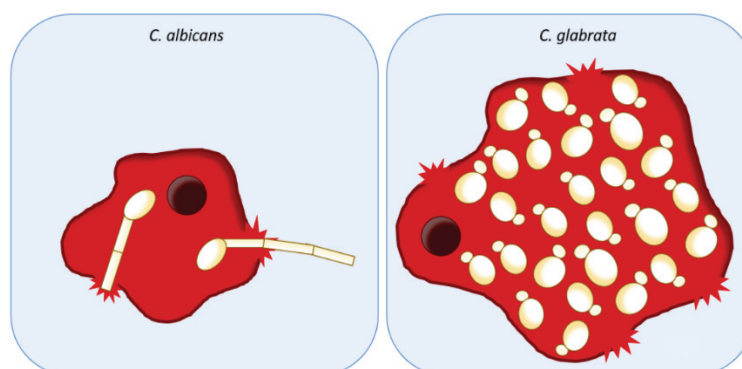
sanguine (Figure 5 b, c) (Galocha *et al.* 2019). Comme mentionné précédemment, il a été démontré que *C. glabrata* était capable de se développer sous forme pseudo filamenteuse dans des conditions de carence en azote *in vitro* (Csank and Haynes 2000). Bien que ce mode de croissance pourrait constituer un facteur de virulence dans l'invasion de l'hôte, des données expérimentales démontrant ce phénomène *in vivo* sont encore manquantes. Un autre élément permettant l'invasion des espèces *Candida* est leur capacité à sécréter des enzymes hydrolytiques détruisant les tissus de l'hôte (Galocha *et al.* 2019). Chez *C. albicans*, ces protéases sont en majorité des aspartates protéases sécrétées (SAPs ; « Secreted Aspartyl Proteinase »). En revanche, *C. glabrata* n'exprime pas de SAPs mais son génome contient 11 gènes (*CgYPSI-11*) codant des aspartates protéases non sécrétées à ancre GPI aussi appelées yapsine (Kaur *et al.* 2007). Ces protéases exposées à la surface sont requises à la mise en place de la virulence ainsi qu'au maintien de la paroi fongique et à la résistance du pathogène face à un environnement acide notamment lorsque ce dernier est internalisé par des macrophages de l'hôte (Kaur *et al.* 2007).

L'infection des levures pathogènes est initiée par l'adhérence et la colonisation du tissu épithélial hôte. L'adhérence est médiée par des protéines de surface appelées adhésines. Chez *C. albicans*, la famille de protéine Als (« Agglutinin-like sequence ») et la protéine Hwp1 (« Hyphae wall protein ») sont critiques pour l'adhérence et la virulence *in vivo* (Galocha *et al.* 2019). Chez *C. glabrata*, les adhésines, importantes pour l'adhérence et l'entrée dans les macrophages, font partie de la famille des adhésines épithéliales (EPA ; « Epithelial adhesins ») et sont codées par 17 à 23 gènes en fonction des souches (Galocha *et al.* 2019; Hassan *et al.* 2021).

Après leur adhésion, les espèces *Candida* sont capables de former un biofilm, une structure tridimensionnelle constituée de levures adhérant les unes aux autres et étant enchâssées dans une matrice extracellulaire (Hassan *et al.* 2021). Ce complexe peut adhérer à une surface telle que les tissus épithéliaux de l'hôte ou à une surface inerte telle qu'un matériel médical (ex. cathéter). La matrice extracellulaire confère des avantages écologiques à la communauté puisqu'elle les protège des macrophages de l'hôte et elle contribue à la pathogénicité en agissant comme une barrière diminuant la susceptibilité aux antifongiques et participant aux phénomènes de résistance (Araújo *et al.*, 2017). Dans le cas de *C. glabrata*, le biofilm est caractérisé par une monocouche ou multicouche dense de levures et d'une matrice plus riche en protéines et sucres que les autres espèces *Candida* (Silva *et al.*, 2009).

Lors de l'infection, les levures pathogènes sont confrontées aux cellules de l'immunité de l'hôte et leur capacité à s'en évader fait partie des facteurs de virulence. Chez un hôte

immunocompétent, les macrophages reconnaissent le glucane  $\beta$ -(1,3) ou la chitine des cellules invasives *via* des récepteurs à leur surface et induisant la phagocytose (Galocha *et al.* 2019). En revanche, des candidoses chroniques ou candidémies peuvent s'établir chez des individus immunocompromis. Afin de s'échapper du système immunitaire, les espèces *Candida* ont plusieurs stratégies à commencer par la modification de la composition de leur paroi fongique afin de masquer les composants reconnus par les cellules de l'immunité. Cependant, il semblerait que la reconnaissance par les macrophages fasse partie de la stratégie d'évasion de *C. glabrata* (Galocha *et al.* 2019). Une fois ingéré par les macrophages, le phagosome fusionne avec un lysosome et c'est dans ce dernier que les levures doivent survivre à des conditions environnementales drastiques telles que la limitation de source de carbone, la production d'espèces réactives de l'oxygène et l'acidification du compartiment (Galocha *et al.* 2019). Toutefois, *C. albicans* et *C. glabrata* ont démontré une capacité à limiter la maturation du phagosome et à prévenir son acidification notamment en utilisant des acides aminés comme source de carbone induisant une production d'ammonium et l'alkalynisation de l'environnement (Galocha *et al.* 2019). *C. albicans*, étant capable de se développer sous une forme filamenteuse, peut s'évader des macrophages en perçant leur membrane. En revanche, *C. glabrata*, qui ne possède pas cette plasticité morphologique, adopte une autre stratégie. La levure survit aux conditions extrêmes des macrophages et se réplique, menant à la lyse du macrophage après plusieurs jours, du à une surcharge de cellules pathogènes (Figure 6) (Galocha *et al.* 2019; Hassan *et al.* 2021).



**Figure 6. Représentation schématique de l'évasion du système immunitaire de *C. albicans* et *C. glabrata* d'après Galocha *et al.* 2019.** *C. albicans* s'échappe des macrophages par sa croissance sous la forme de filaments, *C. glabrata* persiste dans le macrophage, se réplique et mène à la lyse du macrophage due à une surcharge de cellules pathogènes

	<i>C. glabrata</i>	<i>C. albicans</i>	<i>S. cerevisiae</i>
<b>Ploïdie</b>	Haploïde	Diploïde	Diploïde
<b>Morphologie</b>	Levure, pseudo-filaments	Levure, pseudo-filaments, filaments	Levure, pseudo-filaments
<b>Sites majeurs d'infection</b>	Oral, vaginal, disséminé	Oral, vaginal, disséminé	Non infectieux
<b>Gènes de reproduction</b>	Présents	Présents	Présents
<b>Formation de biofilm</b>	Présent	Présent	Présent
<b>Adhésines majeures</b>	Lectines (Codées par les gènes <i>EPA</i> )	Lectines (Als et Hwp1)	Lectines (codées par les gènes <i>FLO</i> )
<b>Résistance aux azolés</b>	Résistance innée	Susceptible	Susceptible

Tableau 1. Comparaison des traits morphologiques et pathogéniques de *S. cerevisiae*, *C. albicans* et *C. glabrata* adapté de Kaur *et al.*, 2005 et Kumar *et al.*, 2019.

### C. Les antifongiques

Afin de traiter les infections fongiques, plusieurs classes d'antifongiques ont été développées et approuvées par la FDA (« US Food and Drug Administration ») et l'EMA (« European Medicines Agency »). Parmi elles, on retrouve quatre classes majoritaires qui vont agir sur différentes voies métaboliques des pathogènes : les azolés, les échinocandines, les polyènes et les analogues de pyrimidine (Figure 7).

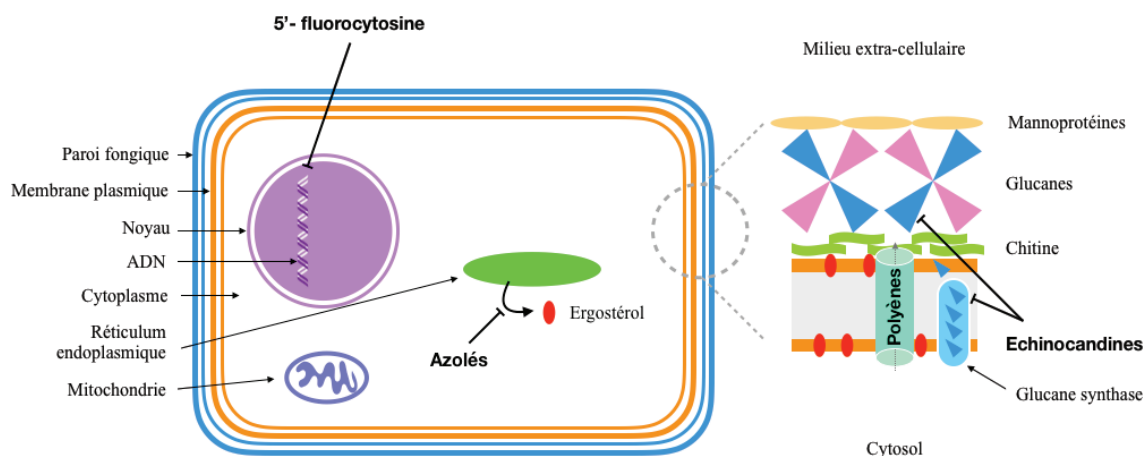
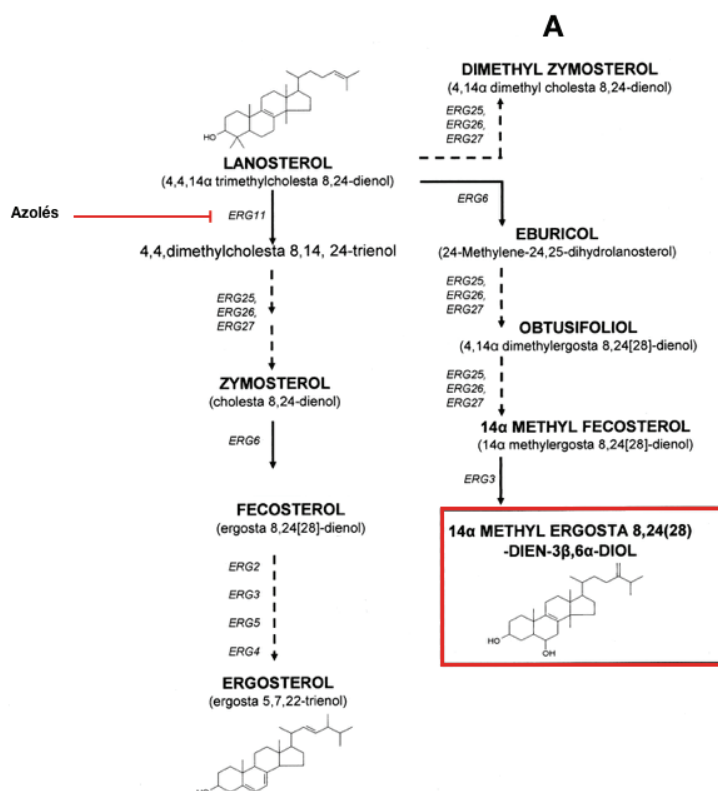


Figure 7. Cibles thérapeutiques des antifongiques Azolés, Echinocandines, 5'-fluorocytosine et Polyènes, schéma adapté de Kathiravan *et al.* 2012. L'analogue de pyrimidine, 5'-fluorocytosine, cible la synthèse de l'ADN, les azolés ciblent la synthèse de l'ergostérol, les polyènes créent des pores dans la membrane plasmique et les échinocandines ciblent la synthèse de glucanes.



## 1. Azolés

Les azolés sont la plus large famille d'antifongiques qui existe à ce jour. Ce sont des inhibiteurs de la 14 $\alpha$  déméthylase, une enzyme impliquée dans la synthèse de l'ergostérol qui est un analogue du cholestérol présent chez les cellules animales et un composant majeur des membranes plasmiques (Spampinato and Leonardi 2013). L'inhibition de la 14 $\alpha$  déméthylase, codée par le gène *ERG11*, empêche la déméthylation du lanostérol et donc la formation d'ergostérol. Cela aura pour conséquence, une accumulation de stérols 14 $\alpha$  méthylés intermédiaires pouvant être toxiques pour la cellule, ainsi qu'une modification de la fluidité membranaire et d'autres processus vitaux tels que la signalisation, le transport, l'endocytose et l'exocytose (Figure 8) (Carrillo-Muñoz *et al.* 2006 ; Rodrigues *et al.* 2014). Parmi la famille des azolés, on retrouve des imidazoles comme le miconazole, mais également des triazolés comme le fluconazole. Dès l'apparition de ce dernier sur le marché en 1994, il a été le plus employé en traitement curatif ou en prophylaxie. Cependant, sa commercialisation coïncide avec l'apparition de phénomènes de résistance et l'augmentation de la prévalence d'espèces moins sensibles à cet antifongique comme *C. glabrata* (Rodrigues *et al.* 2014).



**Figure 8. Représentation schématique de la voie de synthèse de l'ergostérol chez *C. glabrata* adaptée de Hull *et al.* 2012.** Accumulation de stérols 14- $\alpha$  méthylés intermédiaires suite à l'inhibition de la 14 $\alpha$  déméthylase codée par *ERG11*. Les gènes *ERG2*, *ERG3*, *ERG4*, *ERG5*, *ERG6*, *ERG25*, *ERG26* et *ERG27* codent respectivement pour la C5-desaturase, C24-reductase, C22-desaturase, C24-méthyl transférase, C4-méthyloxidase, C4-decarboxylase, C3-ketoreductase.

### a) *Résistance aux azolés*

Quatre mécanismes de résistance ont été décrits chez les espèces *Candida*, à commencer par la diminution des azolés dans le milieu intracellulaire dûe à une surexpression de pompes à efflux (Figure 9 A). Deux familles de protéines sont concernées, les pompes de résistance CDR (« *Candida Drug Resistance* ») qui sont des transporteurs ABC (« *ATP Binding Cassette* ») codées par *CgCDR1* et *CgCDR2* et les pompes MFS (*Major Facilitator Superfamily*) codées par *CgMDR1* (Sanglard *et al.* 1999). Le deuxième mécanisme est la diminution de l'affinité de la cible pour les azolés, causée par des mutations du gène *ERG11* codant pour la 14 $\alpha$  déméthylase et entraînant un défaut de fixation de la drogue (Marichal *et al.* 1999) (Figure 9 B). Environ 140 altérations de ce gène ont été identifiées (Hassan *et al.* 2021). Un autre phénomène de résistance est la surexpression de la 14 $\alpha$  déméthylase, impliquant donc une augmentation de la cible des azolés dans le milieu intracellulaire (Figure 9 C). L'antifongique n'est pas en quantité suffisante pour inhiber toutes les cibles et ainsi bloquer la voie de biosynthèse de l'ergostérol (Marichal *et al.* 1999). Enfin, le dernier mécanisme de résistance ayant été mis en évidence est une altération des dernières étapes de la voie de biosynthèse de l'ergostérol *via* l'inhibition de la protéine codée *ERG3* et responsable de la synthèse des dérivés de stérols 14 $\alpha$  méthylés toxiques (Figure 9 D). La mutation de *ERG3* permet une inhibition de la synthèse de ces dérivés toxiques au profit d'un dérivé stérol non toxique (Miyazaki *et al.* 1999 ; Spampinato and Leonardi 2013).



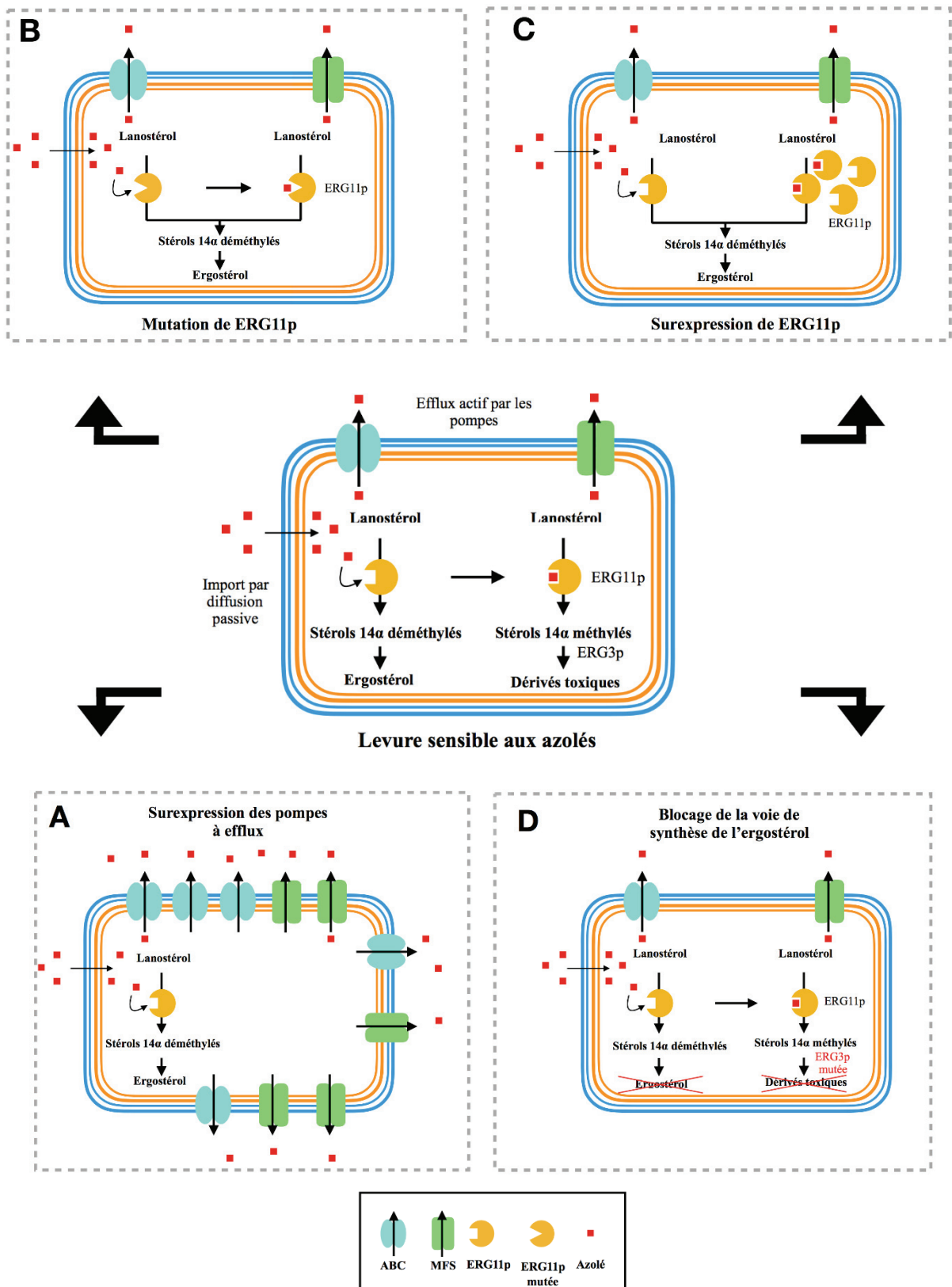


Figure 9. Vue schématique des différents mécanismes de résistance aux antifongiques azolés chez les levures *Candida* adapté de Sanglard 2003. A. Surexpression des pompes à efflux ABC et MFS prenant en charge les composés azolés. B. La mutation de la protéine codée par *ERG11*, ERG11p, diminue son affinité pour les azolés. C. La surexpression de ERG11p augmente le nombre de cible pour les azolés qui ne sont pas en quantité suffisante. D. La mutation de la protéine codée par *ERG3*, ERG3p, empêche la synthèse des dérivés 14 $\alpha$ -méthylés toxiques.

## 2. *Echinocandines*

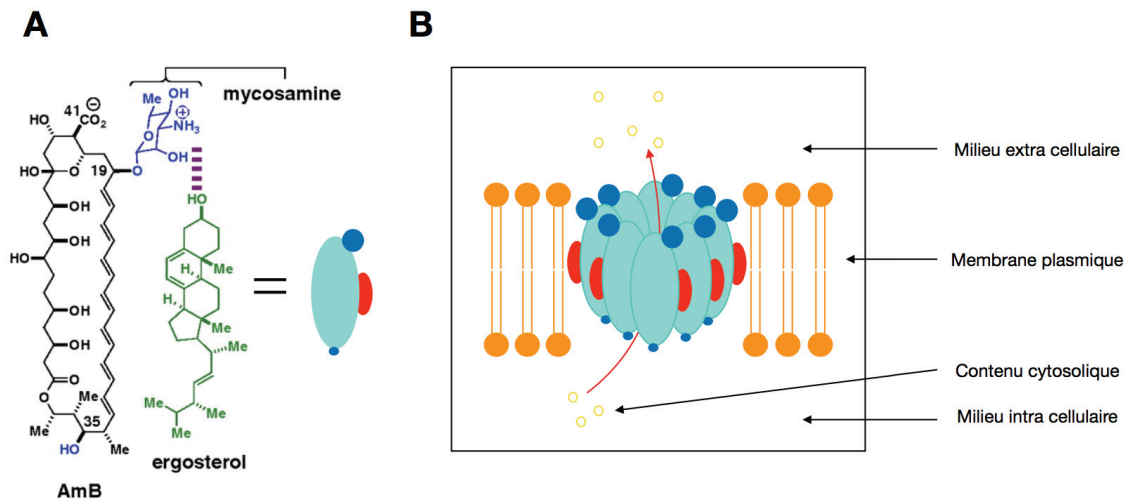
Les échinocandines sont la dernière classe d'antifongiques découverte comprenant la caspofongine, la micafongine et l'anidulafongine. Leur mise sur le marché date de 2002 et ils sont aujourd'hui la première stratégie thérapeutique employée contre des infections invasives de *C. albicans* et *C. glabrata* (Denning 2002; Hassan *et al.* 2021). Ce sont des dérivés synthétiques de lipopeptides qui inhibent, de manière non compétitive, les sous-unités de la  $\beta$ -(1,3)-D-glucane synthétase présente dans la membrane plasmique (Figure 7). L'inhibition de cette enzyme conduit à une diminution d'incorporation de glucose dans les chaînes de glucanes  $\beta$ -(1,3) et aura pour conséquence l'altération de l'intégrité structurale de la paroi fongique rendant la cellule vulnérable à la lyse (Kathiravan *et al.* 2012 ; Spampinato and Leonardi 2013).

### a) *Résistance aux échinocandines*

Malgré une faible prévalence de souches intrinsèquement résistantes aux échinocandines, 4 % chez *C. glabrata* et 1 % chez *C. albicans*, un traitement prolongé conduirait à l'apparition de résistance et à la diminution de la sensibilité de la cible notamment pour *C. glabrata* (Hassan *et al.* 2021). Cette résistance est acquise par la mutation des gènes *FKS1* et *FKS2* de *C. glabrata* codant des protéines constituant les sous-unités de  $\beta$ -(1,3)-D-glucane synthétase (Thompson *et al.* 2008 ; Alexander *et al.* 2013). Plus inquiétant encore, depuis quelques années, des phénomènes de co-résistance aux échinocandines et aux azolés ont été mis en avant chez *C. glabrata* (Pfaller *et al.* 2011).

## 3. *Polyènes*

Les polyènes sont des composés chimiques ayant été isolés chez les bactéries du genre *Streptomyces* et possédant un pouvoir antifongique notamment l'amphotéricine B (AmB) et la nystatine (Kathiravan *et al.* 2012). Le caractère amphotère des polyènes leur permet de s'intégrer dans la bicouche lipidique de la membrane plasmique, de se lier à l'ergostérol et de créer des pores (Spampinato and Leonardi 2013) (Figure 10). Cela entraîne non seulement une perturbation de la fluidité membranaire mais également une fuite du contenu cytosolique conduisant à la mort de la cellule. AmB possède un spectre d'action fongicide très large parmi les espèces fongiques, mais a également une affinité non négligeable pour le cholestérol présent dans les cellules de mammifères et cause de nombreux effets secondaires (Kathiravan *et al.* 2012).



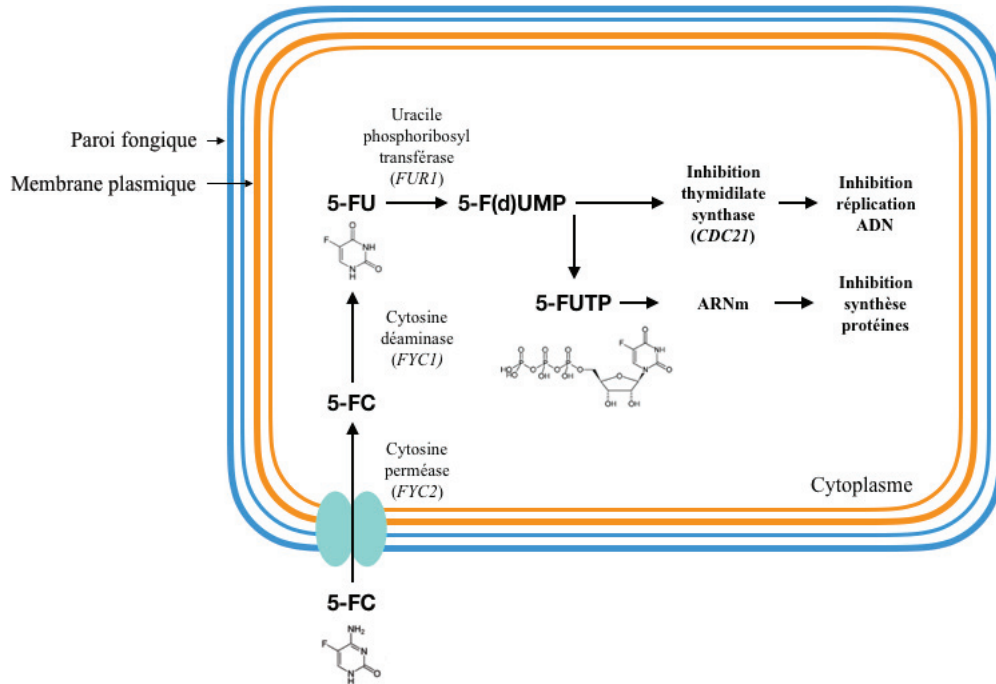
**Figure 10. Représentation moléculaire de l'Amphotéricine B et de son assemblage dans les membranes plasmiques d'après Gray *et al.* 2012. A. Représentation moléculaire de l'Amphotéricine B (AmB), de l'ergostérol et de leur interaction. B. Représentation du mécanisme de perméabilisation membranaire lié à l'activité antifongique de AmB.**

#### a) Résistance aux polyènes

Même s'ils sont moins fréquemment observés que pour les azolés ou les échinocandines, des phénomènes de résistance à l'amphotéricine B ont été remarqués dans des isolats de *C. glabrata* (Pfaller and Diekema 2004; Kanafani and Perfect 2008). Cette résistance se traduirait par des mutations des gènes *ERG2*, *ERG3*, *ERG5*, *ERG6* et *ERG11* codant des protéines qui interviennent dans la synthèse d'ergostérol (Vandeputte *et al.* 2008; Hassan *et al.* 2021). La conséquence est une modification du profil des stérols composant la membrane plasmique des levures et une diminution de leur susceptibilité aux polyènes.

#### 4. Fluoropyrimidines

Cette classe d'antifongiques correspond à des analogues structuraux synthétiques d'un acide nucléique entrant dans la composition de l'ADN et l'ARN : la cytosine. La 5-fluorocytosine (5-FC), une fois dans les cellules fongiques est convertie en 5-fluorouracile (5-FU) par la cytosine déaminase. 5-FU est transformé par l'uracile phosphoribosyl transférase (UPRTase) en 5-fluorodesoxyuridine monophosphate (5-F(d)UMP), un inhibiteur de la thymidilate synthase qui est une enzyme de la biosynthèse d'ADN (Kathiravan *et al.* 2012 ; Spampinato and Leonardi 2013). Ce même composé peut aussi être phosphorylé en 5-fluorouracile triphosphate (5-FUTP) qui s'intègre à l'ARNm et bloque la synthèse protéique (Figure 11).



**Figure 11. Mécanisme d'action de la 5-fluorocytosine adapté de Accoceberry et Noël 2006.** 5-FC : 5-fluorocytosine, 5-FU : 5-fluorouracile, 5-F(d)UMP : 5-fluorodesoxyuridine monophosphate, 5-FUTP : 5-fluorouracile triphosphate.

#### a) Résistance aux fluoropyrimidines

Les phénomènes de résistance associés à l'utilisation de la 5-fluorocytosine sont assez fréquents. Ceci explique notamment que cet antifongique est rarement utilisé en mono thérapie et plus souvent combiné à un autre antifongique. Plusieurs mécanismes de résistance ont été identifiés chez *C. glabrata* et le premier consiste à réduire le taux intracellulaire de la drogue par une mutation du gène *FCY2* codant la cytosine perméase (Espinel-Ingroff 2008 ; Vandeputte *et al.* 2011) (Figure 11). La résistance peut également provenir de mutations du gène *FYCI* codant la cytosine désaminase ou du gène *FURI* codant l'uracile phosphoribosyl transférase, des enzymes de la conversion de la 5-FC (Francis and Walsh 1992; Vandeputte et al. 2011). Enfin, le dernier mécanisme de résistance identifié serait une augmentation des cibles de la 5-FC par la surexpression du gène *CDC21* codant la thymidilate synthétase, l'enzyme ciblée par cet antifongique (Vandeputte *et al.* 2011).

## II. Projet de thèse

Dans le cadre de l'étude structurale et biochimique de protéines d'organismes pathogènes au sein du laboratoire de biocristallographie et biologie structurale de cibles thérapeutiques (BBSCT), ce projet de thèse vise à caractériser les relations structure-fonction-activité de potentielles cibles thérapeutiques appartenant au pathogène *C. glabrata*. Ainsi, deux cibles, impliquées dans deux processus métaboliques distincts, ont été sélectionnées : la métacaspase intervenant dans le processus de mort cellulaire programmée et l'enzyme branchante de glucane participant au métabolisme énergétique du glycogène. L'étude structurale, fonctionnelle et enzymatique de ces deux protéines a pour but d'identifier leur mécanisme d'action avec l'objectif au long terme de réaliser de la conception rationnelle d'inhibiteurs afin d'identifier des alternatives aux antifongiques actuels faisant face à l'augmentation des phénomènes de résistance. Afin de réaliser ce travail, les deux cibles protéiques ont été exprimées en système bactérien, cultivées en milieu liquide et purifiées de manière suffisante pour les caractériser par des techniques de biologie moléculaire, de biochimie, de biophysique et de biologie structurale. Concernant cette dernière, plusieurs méthodes ont été employées de manière à obtenir des informations structurales du repliement de la protéine par diffusion des rayons X aux petits angles (SAXS) et pour obtenir de l'information structurale à l'échelle atomique par cristallographie aux rayons X. Des analyses enzymatiques ont également été réalisées afin de caractériser l'activité des deux protéines. Enfin, l'enzyme branchante de glucane a fait l'objet d'une étude phylogénétique.

Afin de faciliter la lecture de ce manuscrit, les deux protéines, impliquées dans deux processus métaboliques distincts, feront l'objet de deux chapitres distincts. Le chapitre portant sur la caractérisation des relations structure-fonction-activité de la métacaspase sera développé en introduction, résultats, discussion, conclusion et perspectives puis matériels et méthodes. Le chapitre portant sur la caractérisation structurale et enzymatique de l'enzyme branchante de glucane sera rédigé sous la forme d'une thèse article avec une introduction, l'article accepté dans un journal à comité de lecture international ainsi que la conclusion et les perspectives de ce projet de recherche.

### III. Références

Accoceberry I, Noël T. 2006. Antifongiques : cibles cellulaires et mécanismes de résistance. *Thérapie*. 61(3):195–199. doi:10.2515/therapie:2006048.

Ahmad KM, Kokošar J, Guo X, Gu Z, Ishchuk OP, Piškur J. 2014. Genome structure and dynamics of the yeast pathogen *Candida glabrata*. *FEMS Yeast Research*. 14(4):529–535. doi:10.1111/1567-1364.12145.

Alexander BD, Johnson MD, Pfeiffer CD, Jiménez-Ortigosa C, Catania J, Booker R, Castanheira M, Messer SA, Perlin DS, Pfaller MA. 2013. Increasing echinocandin resistance in *Candida glabrata*: clinical failure correlates with presence of FKS mutations and elevated minimum inhibitory concentrations. *Clin Infect Dis*. 56(12):1724–1732. doi:10.1093/cid/cit136.

Araújo D, Henriques M, Silva S. 2017. Portrait of *Candida* Species Biofilm Regulatory Network Genes. *Trends Microbiol*. 25(1):62–75. doi:10.1016/j.tim.2016.09.004.

Carrillo-Muñoz AJ, Giusiano G, Ezkurra PA, Quindós G. 2006. Antifungal agents: mode of action in yeast cells. *Rev Esp Quimioter*. 19(2):130–139.

Csank C, Haynes K. 2000. *Candida glabrata* displays pseudohyphal growth. *FEMS Microbiol Lett*. 189(1):115–120. doi:10.1111/j.1574-6968.2000.tb09216.x.

De Groot PWJ, Ram AF, Klis FM. 2005. Features and functions of covalently linked proteins in fungal cell walls. *Fungal Genet Biol*. 42(8):657–675. doi:10.1016/j.fgb.2005.04.002.

Denning DW. 2002. Echinocandins: a new class of antifungal. *Journal of Antimicrobial Chemotherapy*. 49(6):889–891. doi:10.1093/jac/dkf045.

Dujon B, Sherman D, Fischer G, Durrens P, Casaregola S, Lafontaine I, de Montigny J, Marck C, Neuvéglise C, Talla E, et al. 2004. Genome evolution in yeasts. *Nature*. 430(6995):35–44. doi:10.1038/nature02579.

Espinel-Ingroff A. 2008. Mechanisms of resistance to antifungal agents: yeasts and filamentous fungi. *Rev Iberoam Micol*. 25(2):101–106. doi:10.1016/s1130-1406(08)70027-5.

Fidel PL, Vazquez JA, Sobel JD. 1999. *Candida glabrata*: review of epidemiology, pathogenesis, and clinical disease with comparison to *C. albicans*. *Clin Microbiol Rev*. 12(1):80–96. doi:10.1128/CMR.12.1.80.

Fidel PL, Vazquez JA, Sobel JD. 2001. *Candida glabrata*: an important fungal pathogen for the 21st century. *Clinical Microbiology Newsletter*. 23(22):171–176. doi:10.1016/S0196-4399(01)80058-1.

Francis P, Walsh TJ. 1992. Evolving role of flucytosine in immunocompromised patients: new insights into safety, pharmacokinetics, and antifungal therapy. *Clin Infect Dis*. 15(6):1003–1018. doi:10.1093/clind/15.6.1003.

Gabaldón T, Martin T, Marcet-Houben M, Durrens P, Bolotin-Fukuhara M, Lespinet O, Arnais S, Boissard S, Aguilera G, Atanasova R, et al. 2013. Comparative genomics of emerging pathogens in the *Candida glabrata* clade. *BMC Genomics*. 14(1):623. doi:10.1186/1471-2164-14-623.

Galocha M, Pais P, Cavalheiro M, Pereira D, Viana R, Teixeira MC. 2019. Divergent Approaches to Virulence in *C. albicans* and *C. glabrata*: Two Sides of the Same Coin. *Int J Mol Sci*. 20(9):2345. doi:10.3390/ijms20092345.

Gonçalves IR, Brouillet S, Soulié M-C, Gribaldo S, Sirven C, Charron N, Boccara M, Choquer M. 2016. Genome-wide analyses of chitin synthases identify horizontal gene transfers towards bacteria and allow a robust and unifying classification into fungi. *BMC Evol Biol*. 16. doi:10.1186/s12862-016-0815-9.

Gray KC, Palacios DS, Dailey I, Endo MM, Uno BE, Wilcock BC, Burke MD. 2012. Amphotericin primarily kills yeast by simply binding ergosterol. *Proceedings of the National Academy of Sciences*. 109(7):2234–2239. doi:10.1073/pnas.1117280109.

de Groot PWJ, Kraneveld EA, Yin QY, Dekker HL, Groß U, Crielaard W, de Koster CG, Bader O, Klis FM, Weig M. 2008. The Cell Wall of the Human Pathogen *Candida glabrata*: Differential Incorporation of Novel Adhesin-Like Wall Proteins. *Eukaryot Cell*. 7(11):1951–1964. doi:10.1128/EC.00284-08.

Hassan Y, Chew SY, Than LTL. 2021. *Candida glabrata*: Pathogenicity and Resistance Mechanisms for Adaptation and Survival. *J Fungi (Basel)*. 7(8):667. doi:10.3390/jof7080667.

Hull CM, Bader O, Parker JE, Weig M, Gross U, Warrillow AGS, Kelly DE, Kelly SL. 2012. Two Clinical Isolates of *Candida glabrata* Exhibiting Reduced Sensitivity to Amphotericin B Both Harbor Mutations in *ERG2*. *Antimicrob Agents Chemother*. 56(12):6417–6421. doi:10.1128/AAC.01145-12.



- Kanafani ZA, Perfect JR. 2008. Antimicrobial resistance: resistance to antifungal agents: mechanisms and clinical impact. *Clin Infect Dis.* 46(1):120–128. doi:10.1086/524071.
- Kathiravan MK, Salake AB, Chothe AS, Dudhe PB, Watode RP, Mukta MS, Gadhwe S. 2012. The biology and chemistry of antifungal agents: a review. *Bioorg Med Chem.* 20(19):5678–5698. doi:10.1016/j.bmc.2012.04.045.
- Kaur R, Domergue R, Zupancic ML, Cormack BP. 2005. A yeast by any other name: *Candida glabrata* and its interaction with the host. *Current Opinion in Microbiology.* 8(4):378–384. doi:10.1016/j.mib.2005.06.012.
- Kaur R, Ma B, Cormack BP. 2007. A family of glycosylphosphatidylinositol-linked aspartyl proteases is required for virulence of *Candida glabrata*. *PNAS.* 104(18):7628–7633. doi:10.1073/pnas.0611195104.
- Kumar K, Askari F, Sahu MS, Kaur R. 2019. *Candida glabrata*: A Lot More Than Meets the Eye. *Microorganisms.* 7(2):39. doi:10.3390/microorganisms7020039.
- Kurtzman CP. 2003. Phylogenetic circumscription of *Saccharomyces*, *Kluyveromyces* and other members of the *Saccharomycetaceae*, and the proposal of the new genera *Lachancea*, *Nakaseomyces*, *Naumovia*, *Vanderwaltozyma* and *Zygorulaspora*. *FEMS Yeast Res.* 4(3):233–245. doi:10.1016/S1567-1356(03)00175-2.
- Lundstrom T, Sobel J. 2001. Nosocomial candiduria: a review. *Clin Infect Dis.* 32(11):1602–1607. doi:10.1086/320531.
- Marichal P, Koymans L, Willemsens S, Bellens D, Verhasselt P, Luyten W, Borgers M, Ramaekers FCS, Odds FC, Vanden Bossche H. 1999. Contribution of mutations in the cytochrome P450 14 $\alpha$ -demethylase (Erg11p, Cyp51p) to azole resistance in *Candida albicans*. *Microbiology.* 145 ( Pt 10):2701–2713. doi:10.1099/00221287-145-10-2701.
- Miyazaki Y, Geber A, Miyazaki H, Falconer D, Parkinson T, Hitchcock C, Grimberg B, Nyswaner K, Bennett JE. 1999. Cloning, sequencing, expression and allelic sequence diversity of ERG3 (C-5 sterol desaturase gene) in *Candida albicans*. *Gene.* 236(1):43–51. doi:10.1016/s0378-1119(99)00263-2.
- Pappas PG, Lionakis MS, Arendrup MC, Ostrosky-Zeichner L, Kullberg BJ. 2018. Invasive candidiasis. *Nat Rev Dis Primers.* 4(1):18026. doi:10.1038/nrdp.2018.26.



- Pfaller MA, Diekema DJ. 2004. Rare and Emerging Opportunistic Fungal Pathogens: Concern for Resistance beyond *Candida albicans* and *Aspergillus fumigatus*. *J Clin Microbiol.* 42(10):4419–4431. doi:10.1128/JCM.42.10.4419-4431.2004.
- Pfaller MA, Moet GJ, Messer SA, Jones RN, Castanheira M. 2011. Geographic Variations in Species Distribution and Echinocandin and Azole Antifungal Resistance Rates among *Candida* Bloodstream Infection Isolates: Report from the SENTRY Antimicrobial Surveillance Program (2008 to 2009). *J Clin Microbiol.* 49(1):396–399. doi:10.1128/JCM.01398-10.
- Rodrigues CF, Silva S, Henriques M. 2014. *Candida glabrata*: a review of its features and resistance. *Eur J Clin Microbiol Infect Dis.* 33(5):673–688. doi:10.1007/s10096-013-2009-3.
- Sanglard D. 2003. Resistance and tolerance mechanisms to antifungal drugs in fungal pathogens. *Mycologist.* 17(2):74–78. doi:10.1017/S0269-915X(03)00207-6.
- Sanglard D, Ischer F, Calabrese D, Majcherczyk PA, Bille J. 1999. The ATP Binding Cassette Transporter Gene CgCDR1 from *Candida glabrata* Is Involved in the Resistance of Clinical Isolates to Azole Antifungal Agents. *Antimicrob Agents Chemother.* 43(11):2753–2765.
- Silva S, Henriques M, Martins A, Oliveira R, Williams D, Azeredo J. 2009. Biofilms of non-*Candida albicans* *Candida* species: quantification, structure and matrix composition. *Med Mycol.* 47(7):681–689. doi:10.3109/13693780802549594.
- Sobel JD, Kauffman CA, McKinsey D, Zervos M, Vazquez JA, Karchmer AW, Lee J, Thomas C, Panzer H, Dismukes WE. 2000. Candiduria: a randomized, double-blind study of treatment with fluconazole and placebo. The National Institute of Allergy and Infectious Diseases (NIAID) Mycoses Study Group. *Clin Infect Dis.* 30(1):19–24. doi:10.1086/313580.
- Spampinato C, Leonardi D. 2013. *Candida* Infections, Causes, Targets, and Resistance Mechanisms: Traditional and Alternative Antifungal Agents. *BioMed Research International.* 2013:1–13. doi:10.1155/2013/204237.
- Thompson GR, Wiederhold NP, Vallor AC, Villareal NC, Lewis JS, Patterson TF. 2008. Development of Caspofungin Resistance following Prolonged Therapy for Invasive Candidiasis Secondary to *Candida glabrata* Infection. *Antimicrob Agents Chemother.* 52(10):3783–3785. doi:10.1128/AAC.00473-08.

Vandeputte P, Pineau L, Larcher G, Noel T, Brèthes D, Chabasse D, Bouchara J-P. 2011. Molecular mechanisms of resistance to 5-fluorocytosine in laboratory mutants of *Candida glabrata*. *Mycopathologia*. 171(1):11–21. doi:10.1007/s11046-010-9342-1.

Vandeputte P, Tronchin G, Larcher G, Ernoult E, Bergès T, Chabasse D, Bouchara J-P. 2008. A nonsense mutation in the ERG6 gene leads to reduced susceptibility to polyenes in a clinical isolate of *Candida glabrata*. *Antimicrob Agents Chemother*. 52(10):3701–3709. doi:10.1128/AAC.00423-08.

Yarrow D, Meyer SA. 1978. Proposal for amendment of the diagnosis of the genus *Candida* Berkhout nom. cons. *International journal of systematic bacteriology*. 28(4):611–615. doi:10.1099/00207713-28-4-611.



## **CHAPITRE I**

**Caractérisation des relations structure – fonction – activité de la métacaspase chez l’organisme pathogène *Candida glabrata***



## Liste des figures et des tableaux

### FIGURES

1. **Représentation schématique simplifiée des voies d'activation intrinsèque et extrinsèque de la mort cellulaire** p 62
2. **Représentation schématique des flux intracellulaires de calcium dans un contexte de vie ou de mort cellulaire** p 64
3. **Représentation schématique des domaines des caspases** p 65
4. **Topologie et structure de la caspase-7** p 67
5. **Représentation schématique du mécanisme d'activation des caspases initiatrices et effectrices** p 68
6. **Comparaison structurale du site actif de la procaspase-7 zymogène et de la caspase-7 sous forme active** p 69
7. **Représentation schématique du mécanisme d'activation de la caspase-7** p 70
8. **Voies intra-cellulaires dépendantes du calcium dans la mort cellulaire régulée chez les mammifères et chez les levures** p 75
9. **Représentation des protéases de la famille C14 dans différents groupes taxonomiques** p 77
10. **Représentation de la structure tridimensionnelle de *TbMCA-Ib-C213G*** p 79
11. **Comparaison structurale et topologique des structures de *TbMCA-Ib* et de la caspase-7** p 81
12. **Représentation de la structure tridimensionnelle de *ScMCA-I*** p 82
13. **Représentation de la structure tridimensionnelle et de la topologie de *AtMCA-IIId-C139A*** p 84
14. **Fixation de l'ion  $\text{Sm}^{3+}$  et changements conformationnels de *TbMCA-Ib-C213A*** p 87
15. **Comparaison de la poche catalytique de la caspase-3 et de *TbMCA-Id*** p 88
16. **Représentation du potentiel électrostatique de surface de *TbMCA-Id*** p 89
17. **Purification caractéristique de *CgMCA-I* par chromatographie d'affinité au cobalt** p 93
18. **Purification par chromatographie d'exclusion stérique de *CgMCA-I*** p 93
19. **Maturation de *CgMCA-I* en présence ou absence de  $\text{Ca}^{2+}$  ou d'EGTA** p 94
20. **Maturation de *CgMCA-I* en présence de  $\text{Mn}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Zn}^{2+}$**  p 95

21.	<b>Maturation des mutants C238A, H182A et C238A/H182A en présence ou en absence de <math>\text{Ca}^{2+}</math> ou d'EGTA</b>	p 96
22.	<b>Organisation structurale de <i>CgMCA-I</i><sup>Ca</sup></b>	p 99
23.	<b>Organisation structurale du site catalytique de <i>CgMCA-I</i><sup>Ca</sup> et de <i>ScMCA-I</i></b>	p 99
24.	<b>Etude de la diffusion des rayons X aux petits angles (SAXS) de <i>CgMCA-I</i></b>	p 100
25.	<b>Identification des sites de clivages et des séquences peptidiques associées par spectrométrie de masse</b>	p 101
26.	<b>Organisation structurale de <i>CgMCA-I</i><sup>Mg</sup></b>	p 103
27.	<b>Organisation structurale du site catalytique de <i>CgMCA-I</i><sup>Mg</sup></b>	p 103
28.	<b>Organisation structurale du site de liaison aux cations de <i>CgMCA-I</i><sup>Ca</sup> et <i>CgMCA-I</i><sup>Mg</sup></b>	p 105
29.	<b>Organisation structurale du site de liaison au calcium de <i>CgMCA-I</i><sup>Ca</sup></b>	p 105
30.	<b>Réorganisation du site catalytique de <i>CgMCA-I</i><sup>Ca</sup> et de <i>CgMCA-I</i><sup>Mg</sup></b>	p 107
31.	<b>Organisation structurale des sites de liaisons aux ions <math>\text{Mg}^{2+}</math> de <i>CgMCA-I</i><sup>Mg</sup></b>	p 107
32.	<b>Maturation des constructions <sup>His</sup><i>CgMCA-I</i> et <i>CgMCA-I</i><sup>His</sup> en absence de <math>\text{Ca}^{2+}</math></b>	p 108
33.	<b>Maturation du mutant R54A de <i>CgMCA-I</i></b>	p 109
34.	<b>Maturation des mutants K263A, K263D et K263F de <i>CgMCA-I</i></b>	p 110
35.	<b>Description du mécanisme de maturation de <i>CgMCA-I</i></b>	p 111
36.	<b>Représentation de Michaelis-Menten de <i>CgMCA-I</i> non maturée (jour 0) et maturée (jour 8) en utilisant le substrat Z-GGR-AMC (0-50 <math>\mu\text{M}</math>) sans correction de l'hydrolyse chimique</b>	p 111
37.	<b>Transition de la population enzymatique <math>E_{\text{low}}</math> à la population <math>E_{\text{high}}</math> mis en évidence par la mesure de l'activité hydrolytique pendant 10 jours</b>	p 113
38.	<b>Maturation du mutant C238A/H182A après mélange avec <i>CgMCA-I</i> sauvage maturée au jour 8</b>	p 115
39.	<b>Alignement des séquences et des structures secondaires de <i>CgMCA-I</i><sup>Ca</sup>, <i>CgMCA-I</i><sup>Mg</sup>, <i>TbMCA-Ib</i> et <i>ScMCA-I</i></b>	p 117
40.	<b>Réorganisation structurale de la boucle L7 chez <i>CgMCA-I</i><sup>Ca</sup> et <i>CgMCA-I</i><sup>Mg</sup></b>	p 119
41.	<b>Superposition des structures <i>CgMCA-I</i><sup>Ca</sup> et <i>CgMCA-I</i><sup>Mg</sup> avec la modélisation de <i>CgMCA-I</i> par AlphaFold</b>	p 121

42.	Comparaison de la résolution structurale des boucles L6 et L7 pour les structures de <i>CgMCA-I<sup>Ca</sup></i> , <i>CgMCA-I<sup>Mg</sup></i> et la structure <i>CgMCA-I</i> modélisée par AlphaFold	p 121
43.	Carte du vecteur pEt15b	p 127
44.	Carte du vecteur pEt52b	p 127

## TABLEAUX

1.	Comparatif des dénominations des boucles entre les structures de la caspase-7, <i>TbMCA-Ib</i> , <i>ScMCA-I</i> et <i>AtMCA-IIId</i>	p 78
2.	Fonctions étudiées des métacaspases chez <i>C. albicans</i> et <i>S. cerevisiae</i>	p 90
3.	Statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de <i>CgMCA-I</i>	p 97
4.	Paramètres cinétiques de <i>CgMCA-I</i> aux jours de maturation 0 (J0) et 8 (J8)	p 112
5.	Paramètres optimisés du modèle <i>Cis-Trans</i> (meilleur modèle)	p 115
6.	Amorces sens et anti-sens employées pour la mutagenèse dirigée de C238A, H182A, C238A/H182A, R54A, K263A, K263D et K263F de <i>CgGBE</i> et pour la construction <i>CgMCA-I<sup>His</sup></i>	p 129
7.	Liste des tampons employés pour l'analyse du processus de maturation de <i>CgMCA-I</i> et de ses mutants, par BioSprint	p 131
8.	Liste des tampons employés pour l'analyse du processus de maturation de <i>CgMCA-I</i> en présence des ions divalents $Mn^{2+}$ , $Mg^{2+}$ et $Zn^{2+}$ , par BioSprint	p 132
9.	Modèles proposés décrivant la conversion de la forme $E_{low}$ à la forme $E_{high}$	p 135
10.	Jeux de paramètres initiaux pour la discrimination du modèle	p 137





# I. Le phénomène de mort cellulaire régulée

## A. La mort cellulaire régulée chez les métazoaires

Le développement et la survie des organismes pluricellulaires sont accompagnés d'un contrôle précis du nombre de cellules grâce à un équilibre entre phénomènes de mort et de prolifération cellulaire. D'après la NCCD (« Nomenclature Committee on Cell Death »), deux types de morts cellulaires sont à différencier : la mort cellulaire accidentelle (ACD : « Accidental Cell Death ») et la mort cellulaire régulée (RCD : « Regulated Cell Death ») (Galluzzi *et al.*, 2018). Alors que l'ACD serait instantanée, incontrôlée et causée par l'exposition à des contraintes physicochimiques et mécaniques sévères, la RCD repose sur une machinerie cellulaire très régulée. Cette dernière peut intervenir de manière physiologique et on fait alors référence à la mort cellulaire programmée (PCD : « Programmed Cell Death »), ou bien être induite par des perturbations environnementales intra- ou extracellulaires. Dans les deux cas, ces phénomènes constituent une stratégie de maintien de l'homéostasie cellulaire pour les organismes en éliminant les cellules dangereuses ou en libérant des molécules alertant les cellules environnantes d'une potentielle menace. Historiquement, trois types de RCD ont été décrites, la mort cellulaire de type I ou apoptose, la mort cellulaire de type II ou autophagie et enfin la mort cellulaire de type III ou nécrose. L'autophagie est un processus de dégradation et de recyclage d'éléments cytoplasmiques se manifestant par la formation de vacuoles dont le contenu sera éliminé après une fusion lysosomale. La nécrose est caractérisée par une modification de la perméabilité membranaire entraînant un gonflement du cytoplasme et des organelles face à l'augmentation de la pression osmotique intracellulaire. Ceci conduira à la rupture des membranes cellulaires et à la libération du contenu cytosolique induisant une réponse inflammatoire. Nous insisterons sur la dernière forme de mort cellulaire qui concerne plus directement les travaux qui ont été menés au cours de cette thèse.

### 1. *L'apoptose*

L'apoptose est un processus biologique très régulé permettant d'éliminer les cellules surnuméraires ou dysfonctionnelles d'un organisme. Le mot apoptose est introduit en 1972 par John Kerr, Andrew Wyllie et Alastair Currie pour faire opposition au modèle de la nécrose. Ils décrivent ce processus comme actif et associé à des changements morphologiques, biochimiques et structuraux conduisant à la destruction de la cellule « sans traces » (Kerr, Wyllie and Currie, 1972). Ce phénomène se traduit notamment par une rétractation progressive de la cellule s'accompagnant d'une condensation de la chromatine et

du cytoplasme, d'une fragmentation de l'ADN et d'une invagination des membranes plasmiques et nucléaires conduisant ainsi à la formation de corps apoptotiques. Ces vésicules seront phagocytées par les macrophages ou d'autres cellules de manière rapide, en évitant le relargage du contenu intracellulaire et donc sans induction d'une réaction inflammatoire. Le processus apoptotique pouvant être activé par deux voies différentes, on parle d'apoptose intrinsèque ou d'apoptose extrinsèque (Figure 1).

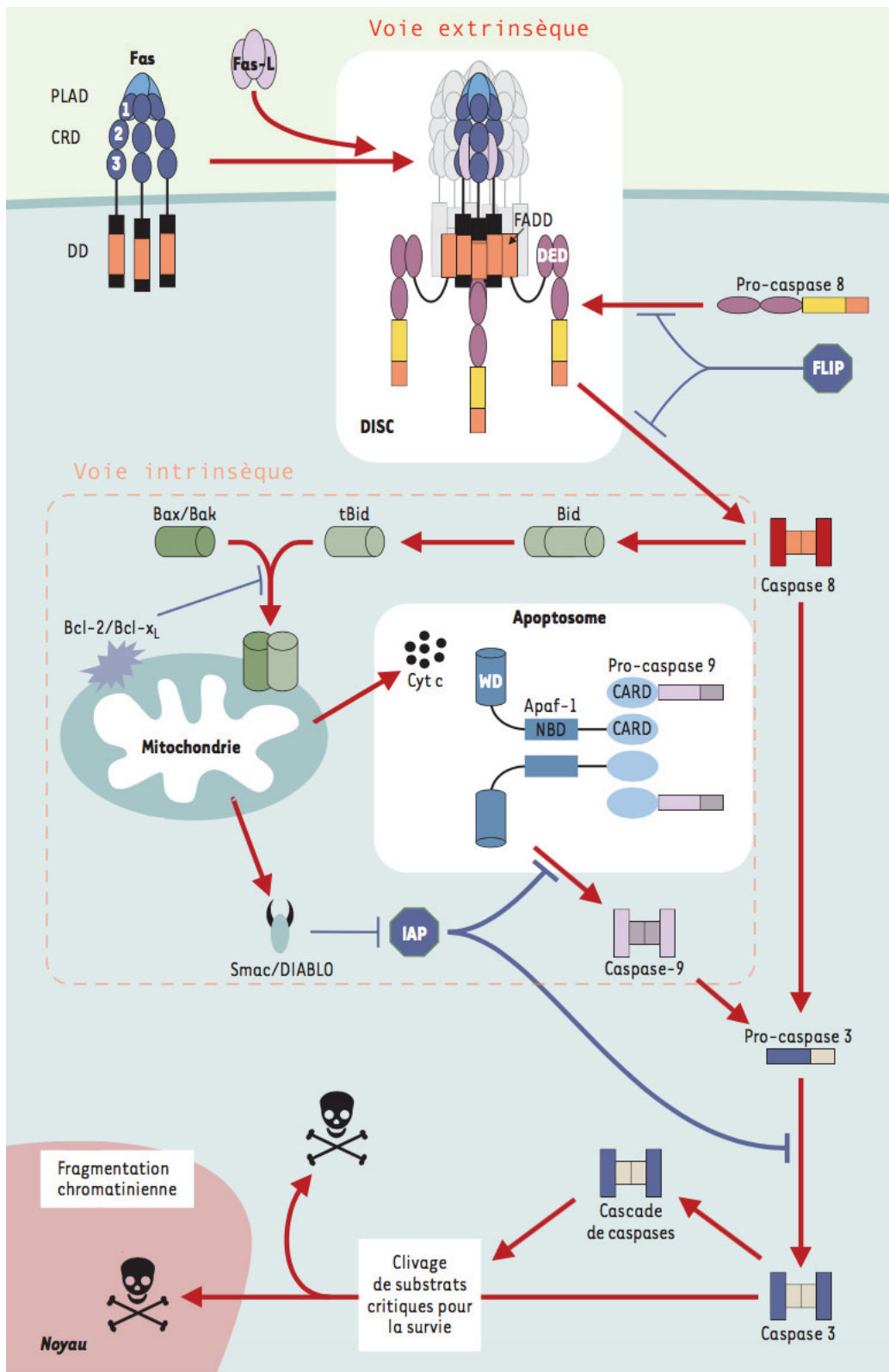
#### a) *Apoptose intrinsèque*

L'apoptose intrinsèque est, par définition, un mode d'apoptose induit par des perturbations intracellulaires telles que des dommages à l'ADN, un stress au niveau du réticulum endoplasmique (RE) ou une présence trop importante d'espèces réactives de l'oxygène (« Reactive Oxygen Species » ou ROS) résultant de la perte d'intégrité membranaire mitochondriale. Cette voie est notamment régulée par les protéines de la famille Bcl-2 composée de trois sous-familles : les facteurs pro-apoptotiques BH3 (tels que BID et BIM), les effecteurs pro-apoptotiques (BAX et BAK) et les protéines anti-apoptotiques (telles que BCL-2 et BCL-xL) (Czabotar *et al.*, 2014). Dans une cellule saine, les effecteurs BAX et BAK sont inhibés par les facteurs anti-apoptotiques. En revanche, en présence d'un stimulus de mort, les facteurs pro-apoptotiques BH3 sont régulés positivement, inhibant l'action des facteurs anti-apoptotiques et activant les effecteurs BAX et BAK localisés dans la membrane externe mitochondriale (Figure 1). Leur activation se traduit par leur oligomérisation, créant des pores dans la membrane mitochondriale et induisant une chute du potentiel transmembranaire. Ce phénomène est appelé perméabilisation de la membrane externe mitochondriale ou MOMP (« Mitochondrial Outer Membrane Permeabilization ») (Galluzzi *et al.*, 2018). Cette perméabilisation entraîne la libération de facteurs mitochondriaux résidant dans l'espace intermembranaire mitochondriale tel que le cytochrome *c*, intervenant dans le flux d'électrons de la chaîne respiratoire mitochondriale, et le second activateur mitochondrial de caspase dit SMAC (« Second Mitochondrial Activator of Caspase »). Le pool cytosolique de cytochrome *c* se lie au facteur 1 d'activation de la peptidase apoptotique (APAF1, « Apoptotic Peptidase Activating Factor 1 ») de manière desoxy-ATP dépendante pour former un complexe supramoléculaire appelé apoptosome (Li *et al.*, 1997). Ce complexe heptamérique se lie à la procaspase-9 *via* son domaine de recrutement CARD (« CASpase Recruitment Domain ») et permet son activation par homodimérisation. Son autoprotéolyse lui permet d'être sous forme mature (caspase-9) et d'activer par protéolyse des caspases dites exécutrices, la procaspase-7 et majoritairement la procaspase-3 (Yuan and Akey, 2013; Nagata, 2018). En parallèle, SMAC aussi appelé DIABLO (« Direct IAP Binding protein with

LOW pI », protéine de liaison aux IAP de faible point isoélectrique), précipite le processus apoptotique en s'associant avec les protéines IAP pour bloquer leur action. Les IAP sont des inhibiteurs de l'apoptose (« Inhibitor of Apoptosis ») bloquant l'activité des caspases (Galluzzi *et al.*, 2018).

#### b) *L'apoptose extrinsèque*

La voie de l'apoptose extrinsèque implique l'engagement de récepteurs de mort membranaires de la famille TNF (« Tumor Necrosis Factor »), le récepteur FAS, le récepteur 1 (TNFR1) et les récepteurs TRAILR1 et TRAILR2 (« TNF-related apoptosis inducing ligand-receptor 1/2») (Figure 1) (Nagata, 2018). Ce sont des protéines membranaires trimériques dont la partie cytosolique est appelée domaine de mort (« Death domain » ; DD). Lorsque ces récepteurs fixent les facteurs de mort (FasL, TNF- $\alpha$ , TRAIL), la protéine FADD (« Fas-associated Death Domain ») dont l'extrémité C-terminale est un domaine DD, est recrutée par une interaction homotypique entre les domaines de mort. FADD est également composée d'un domaine effecteur de mort (DED ; « Death Effector Domain ») par lequel elle fixe les procaspases initiatrices -8 ou -10 *via* leur domaine DED en tandem. L'ensemble multi protéique constitue le DISC (« Death Inducing Signaling Complex ») (Schleich *et al.*, 2012). Les procaspases-8 et -10 seront activées par homodimérisation et maturées par autoprotéolyse. Elles activent à leur tour la procaspase-3 exécutive qui propage le signal de mort cellulaire (Galluzzi *et al.*, 2018). Il a également été démontré que, dans les cellules de type II (hépatocytes et fibroblastes), la caspase-8 protéolyse le facteur pro-apoptotique BID cytosolique (tBID), induisant ainsi sa translocation à la membrane mitochondriale et son interaction avec BAX et BAK déclenchant alors la voie intrinsèque de l'apoptose (Strasser, Jost and Nagata, 2009).



**Figure 1. Représentation schématique simplifiée des voies d'activation intrinsèque et extrinsèque de la mort cellulaire régulée chez les metazoaires d'après Couzinet, Hérics et Hueber, 2002.** L'apoptose peut être déclenchée par des récepteurs de mort membranaires (voie extrinsèque) ou induite par des perturbations intracellulaires (voie intrinsèque). Les deux voies conduisent à l'activation de caspases initiateurs (caspase-8 et -9) qui activeront les caspases exécutrices (caspase-3), responsables de la propagation du signal de mort cellulaire.

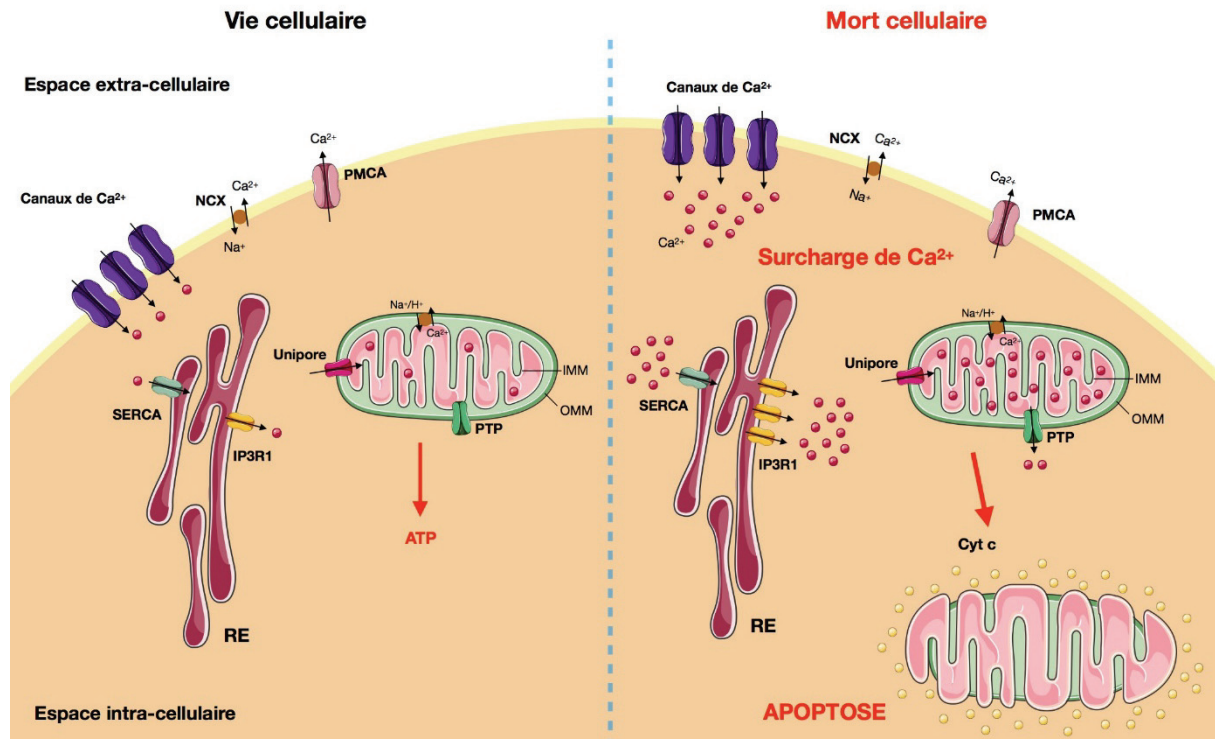
### c) *Activation par le calcium*

La régulation des flux ioniques intra et intercellulaires est très importante pour le maintien de l'homéostasie cellulaire. Un défaut de ce flux d'ions, résultant du dysfonctionnement d'un canal ou d'une pompe ionique, peut avoir des conséquences fatales menant jusqu'à la mort cellulaire (Orrenius, Zhivotovsky and Nicotera, 2003). Le calcium intracellulaire est importé par différents canaux ioniques présents à la membrane plasmique cellulaire (Figure 2). A l'intérieur des cellules il peut soit se coupler à des protéines liant le calcium, soit être séquestré dans le RE ou dans la mitochondrie (Orrenius, Zhivotovsky and Nicotera, 2003). La concentration intracellulaire de calcium est maintenue constante à environ 100 nM par une balance entre absorption par le RE *via* la pompe ATPase SERCA (« Sarco(endo)plasmic reticulum  $\text{Ca}^{2+}$ -ATPase ») et son extrusion du RE et de la cellule respectivement par les pompes IP3R (« Inositol 3 phosphate receptor type 1 ») et la pompe plasmique ATPase PMCA (« Plasma-membrane  $\text{Ca}^{2+}$ -ATPase »). L'absorption de calcium dans la mitochondrie se fait quant à elle par un transporteur unipore. Son relargage se fait par un transporteur  $\text{Na}^+/\text{H}^+$  dépendant du calcium ou par le pore de perméabilité membranaire mitochondriale dépendant du calcium (PTP ; « Permeability Transition Pore ») se trouvant dans la membrane mitochondriale (Figure 2) (Orrenius, Zhivotovsky and Nicotera, 2003).

Dans les années 1960, il a été mis en évidence qu'une augmentation intracellulaire de calcium était toxique pour la cellule et avait pour conséquence l'activation massive de protéases (Pinton *et al.*, 2008). L'identification des protéines de la famille Bcl-2 au niveau des organelles de stockage du calcium (mitochondrie et RE) et l'augmentation de calcium intracellulaire lors des phases précoces et matures de l'apoptose n'ont fait que renforcer le lien entre homéostasie du calcium et mort cellulaire programmée (Kruman, Guo and Mattson, 1998). Des études ont montré qu'une surexpression de la protéine anti-apoptotique BCL-2 induisait une réduction de la concentration en calcium du RE (Pinton *et al.*, 2000). A l'inverse, une surexpression de la protéine pro-apoptotique BAX induirait une augmentation de la concentration en calcium du RE (Pinton *et al.*, 2000).

Bien que le transporteur du calcium mitochondrial ait une faible affinité pour le calcium, la localisation stratégique de la mitochondrie, proche du RE, induit des concentrations locales de calcium très élevées et une absorption rapide et importante (Pressman, 1976; Pinton *et al.*, 2008). Ce mouvement de calcium du RE à la mitochondrie semble être un processus clé dans l'activation de l'apoptose (Scorrano *et al.*, 2003). L'absorption massive de calcium induirait un second mécanisme de perméabilisation membranaire, autre que le MOMP, *via* l'activation du pore PTP induisant un gonflement de la membrane mitochondriale interne, la rupture de la

membrane externe mitochondriale et un relargage du cytochrome *c* dans le cytosol (Figure 2) (Pinton *et al.*, 2001).



**Figure 2. Représentation schématique des flux intracellulaires de calcium dans un contexte de vie ou de mort cellulaire adapté de Pinton *et al.*, 2008.** Dans un contexte de vie cellulaire le calcium est importé dans les cellules *via* des canaux calciques, il est stocké majoritairement dans le réticulum endoplasmique (RE) *via* la pompe SERCA puis est libéré par le récepteur IP3R1. La concentration intracellulaire est maintenue grâce à la balance entre stockage au RE et externalisation cellulaire par la pompe PMCA. Le calcium est internalisé dans la matrice mitochondriale *via* un unipore et externalisé *via* l'activation d'un pore de perméabilité membranaire (PTP). Dans un contexte de mort cellulaire, la concentration intracellulaire de calcium augmente ce qui entraîne une surcharge de la concentration de calcium du RE et de la mitochondrie. La conséquence est le gonflement de la mitochondrie et la rupture de sa membrane intra- et extramitochondriale (respectivement IMM et OMM) libérant le cytochrome *c* et activant l'apoptose.

## 2. Les caspases

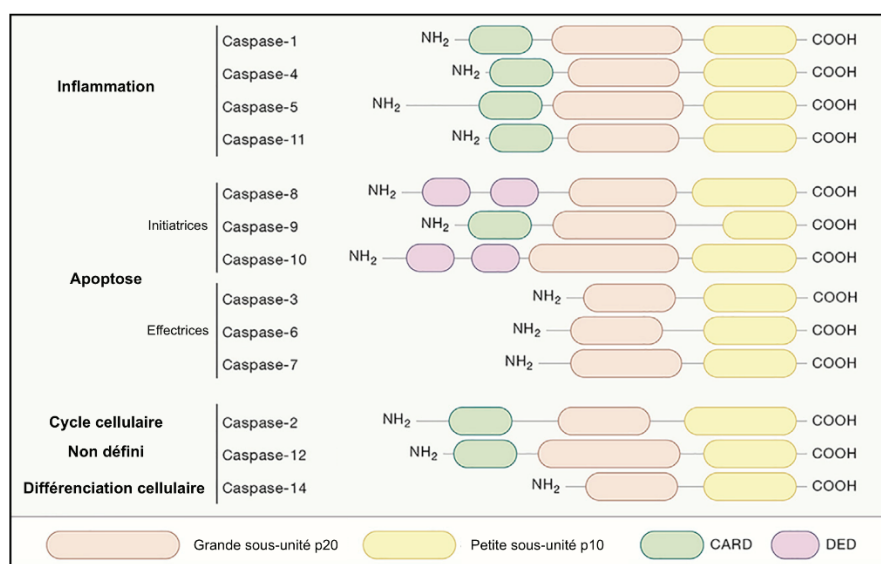
Les caspases (« cystein-dépendant **asp**artate-specific peptidase ») sont une famille de protéases intracellulaires cystéine dépendantes spécifiques de l'aspartate. Ces enzymes protéolytiques hautement régulées contrôlent des processus biologiques comme la mort cellulaire programmée chez les métazoaires mais également l'inflammation, la prolifération et la différenciation cellulaire (Li and Yuan, 2008).



### a) Classification

Les caspases font parties de la famille des cystéine-peptidases C14 et du clan CD selon la classification de la base de données MEROPS (<https://www.ebi.ac.uk/merops>) (Rawlings *et al.*, 2018). La première caspase identifiée comme participant à la mort cellulaire programmée est la cystéine protéase CED-3 découverte chez le nématode *Caenorhabditis elegans* (Yuan *et al.*, 1993). Depuis, 14 protéases homologues ont été identifiées chez les mammifères et peuvent être réparties selon trois grands groupes en fonction de leur structure et de leur fonction (Pinton *et al.*, 2001) :

- Les caspases pro-apoptotiques sous classées en caspases initiatrices (caspase-8, -9, -10) et caspases effectrices (caspase-3, -6, -7)
- Les caspases inflammatoires (caspase-1, -4, -5, -11)
- Les trois autres caspases dont les fonctions diffèrent (caspases-2, -12, -14)



**Figure 3. Représentation schématique des domaines des caspases d'après Van Opdenbosch et Lamkanfi, 2019.** Classification des caspases impliquées dans l'inflammation (-1, -4, -5, -11), dans l'apoptose avec les caspases initiatrices (-8, -9, -10) et effectrices (-3, -6, -7), dans le cycle cellulaire (-2), dans la différenciation cellulaire (-14) et dont le rôle reste à définir (-12).

Les caspases sont synthétisées sous forme de précurseurs inactifs ou dits zymogènes (procaspase) et vont nécessiter des étapes protéolytiques et d'oligomérisation afin d'être sous leur forme active dimérique. Sous leur forme zymogène, les caspases peuvent être constituées d'un pro-domaine ainsi que d'une grande sous unité p20 (environ 20kDa), contenant la diade catalytique His/Cys, et d'une petite sous unité p10 (environ 10kDa). Le domaine N-terminale peut être de taille variable. Pour les caspases initiatrices c'est un large pro-domaine

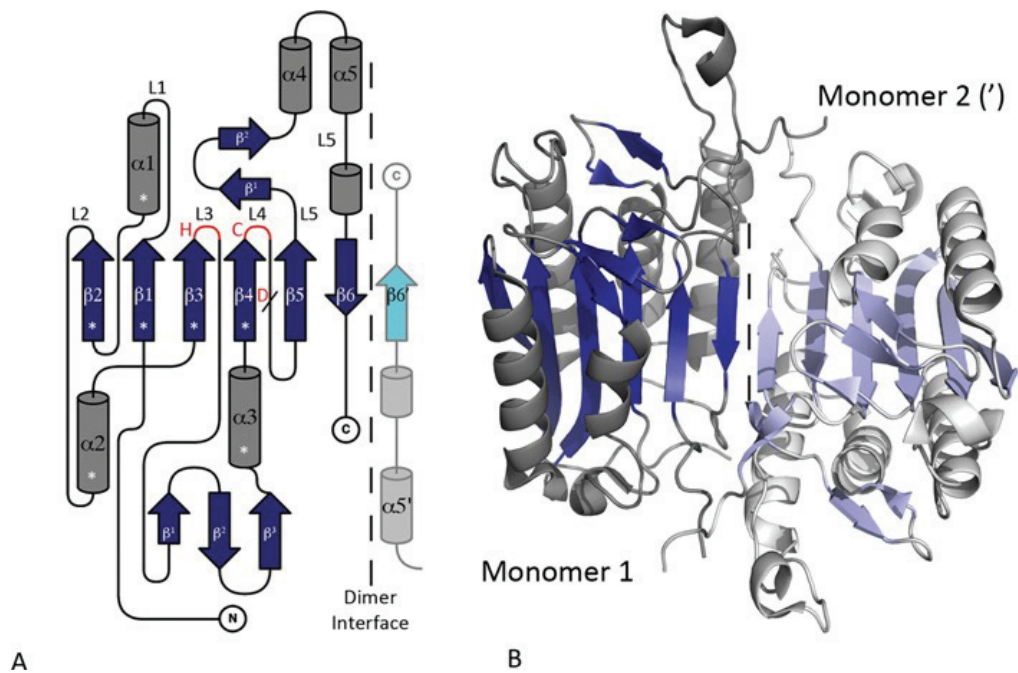


comprenant soit deux domaines DED pour les caspases-8 et -10, soit un domaine CARD pour la caspase -9. En revanche, les caspases effectrices sont constituées d'un pro-domaine court.

### b) *Repliement structural*

La première structure résolue de caspase fut celle de la caspase-1 en 1994 par deux équipes distinctes de chercheurs (Walker *et al.*, 1994 ; Wilson *et al.*, 1994). Les caspases ont d'une manière générale un repliement structural identique que l'on nomme repliement hémoglobinase (caractéristique structurale du clan CD) (Aravind and Koonin, 2002). On utilisera dans ce manuscrit la structure de la caspase-7 pour décrire cette famille de protéines (Figure 4). Comme indiqué précédemment, il est admis que les caspases sont actives sous une forme dimérique. Un monomère est composé d'un feuillet de 5 brins  $\beta$  parallèles et un brin  $\beta$  antiparallèle ( $\beta_1$ - $\beta_6$ ) selon la topologie  $2\uparrow 1\uparrow 3\uparrow 4\uparrow 5\uparrow 6\downarrow$  (numérotation pour la caspase-7) (McLuskey and Mottram, 2015). Les deux domaines s'associent de manière antiparallèle par le brin  $\beta_6$  à l'extrémité C-terminale générant ainsi un long feuillet continu de 12 brins  $\beta$  (Figure 4 B). Le feuillet  $\beta$  de chacun des monomères est entouré de 5 hélices  $\alpha$  ( $\alpha_1$ - $\alpha_5$ ) et contient également un feuillet de 3 brins  $\beta$  courts ( $\beta^1$ - $\beta^3$ ) entre le brin  $\beta_3$  et l'hélice  $\alpha_3$ . Les brins  $\beta_3$  et  $\beta_4$  portent respectivement les résidus catalytiques conservés His et Cys à leur extrémité C-terminale au niveau des boucles L3 et L4. Le résidu Asp, très conservé, représentant le site de protéolyse entre la sous-unité p20 ( $\beta_1$ - $\beta_4$ ) et p10 ( $\beta_5$ - $\beta_6$ ), se situe sur la boucle L4 entre les brins  $\beta_4$  et  $\beta_5$ .

Les domaines de recrutement DED et CARD, malgré une faible identité de séquences, possèdent un repliement structural commun. Ce sont des domaines globulaires composés de 6 hélices  $\alpha$ , dont les hélices  $\alpha_1$ - $\alpha_5$  ont un repliement en clé grecque (Fuentes-Prior and Salvesen, 2004).



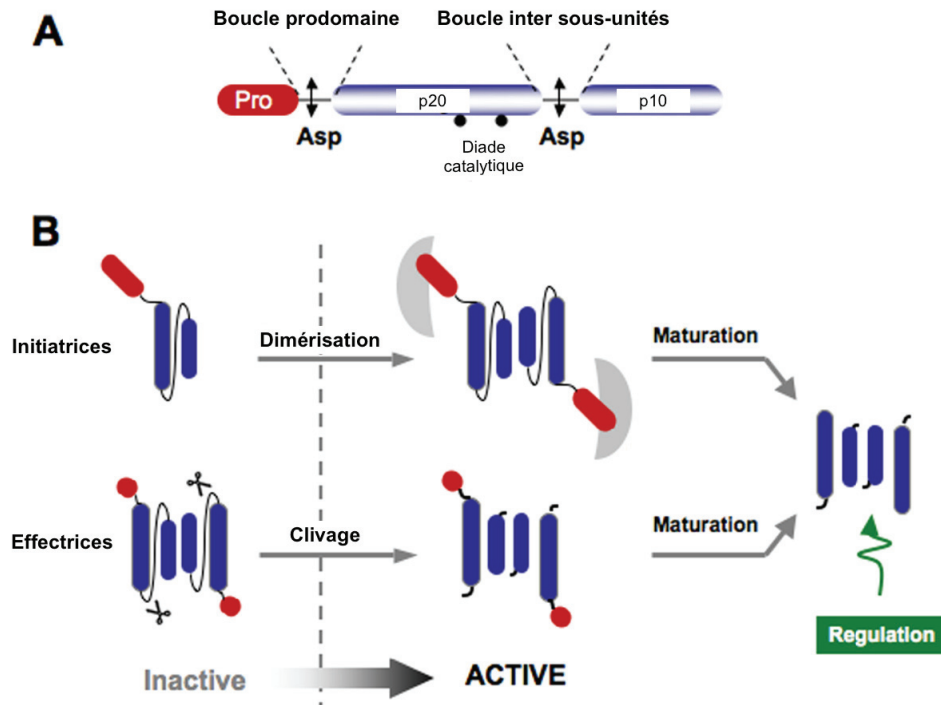
**Figure 4. Topologie et structure de la caspase-7 d'après McLuskey et Mottram, 2015.** **A.** La caspase-7 (PDB : 1F1J) est représentée de manière topologique et tridimensionnelle avec les hélices  $\alpha$  en gris, les brins  $\beta$  en bleu. Les éléments structuraux du dimère sont colorés en gris clair et bleu turquoise. Les hélices  $\alpha$  et les brins du feuillet  $\beta$  central sont numérotés à partir de l'extrémité N-terminale de  $\alpha 1$  à  $\alpha 5$  et de  $\beta 1$  à  $\beta 6$ . Les boucles sont numérotées en fonction du brin  $\beta$  qu'elles suivent (L1-L5). Les courtes sections de brins  $\beta$  suivant les boucles L3 et L5 sont numérotés respectivement  $\beta^1$ - $\beta^3$  et  $\beta^1$ - $\beta^2$ . Les résidus catalytiques His et Cys sont indiqués en rouge sur les boucles L3 et L4 ainsi que le résidu Asp conservé sur la boucle L4. Une ligne en pointillée verticale représente l'interface de dimérisation de la caspase-7. **B.** Image de la structure tridimensionnelle en ruban du dimère de la caspase-7.

### c) Mécanisme d'activation

Les caspases initiatrices et effectrices possèdent deux modes d'activation différents. Les premières sont synthétisées sous une forme zymogène monomérique et sont activées par dimérisation (Figure 5). Ce processus dépend du recrutement des caspases initiatrices vers des plateformes moléculaires (DISC ou apoptosome) *via* leur prodomaine (DED ou CARD) et l'on parle d'une dimérisation induite par proximité (Lamkanfi *et al.*, 2002). Puis, les caspases vont subir un processus de maturation ne semblant pas nécessaire à la mise en place de leur activité dans certains processus cellulaires mais étant indispensable à leur implication dans la mort cellulaire (Pop and Salvesen, 2009). La maturation des caspases initiatrices est une autoprotéolyse de leur pro-domaine ainsi que de la boucle inter sous-unités (p20 et p10) au résidu Asp conservé de la boucle L4 (Figure 4).

Les caspases effectrices quant à elles, existent dans le cytosol sous une forme zymogène dimérique nécessitant un clivage protéolytique de la boucle inter sous-unités afin d'être actives (Figure 5) (Fuentes-Prior and Salvesen, 2004 ; Pop and Salvesen, 2009). Ce clivage

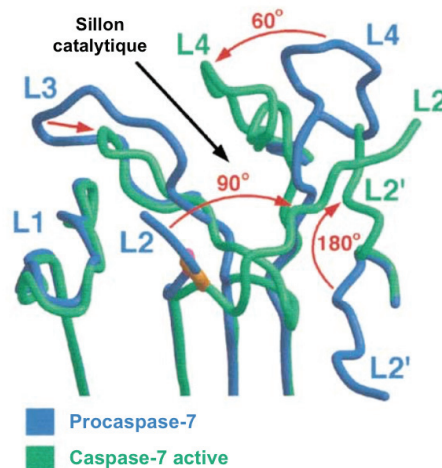
est effectué par l'activité catalytique des caspases initiatrices actives au niveau du résidu Asp conservé de la boucle L4 (Shi, 2002). La maturation des caspases effectrices, réalisée à nouveau par les caspases initiatrices, permet l'élimination par protéolyse du pro-domaine à l'extrémité N-terminale des monomères composant le dimère actif (Pop and Salvesen, 2009).



**Figure 5. Représentation schématique du mécanisme d'activation des caspases initiatrices et effectrices d'après Pop et Salvesen, 2009.** A. Organisation structurale des domaines de caspase. B. Mécanisme d'activation des caspases initiatrices et effectrices. Les initiatrices sont activées par la dimérisation des monomères inactifs alors que les effectrices sont activées par le clivage de la boucle de liaison entre les sous unités p20 et p10. L'activation est suivie d'une maturation correspondant à des événements protéolytiques permettant de stabiliser la caspase active.

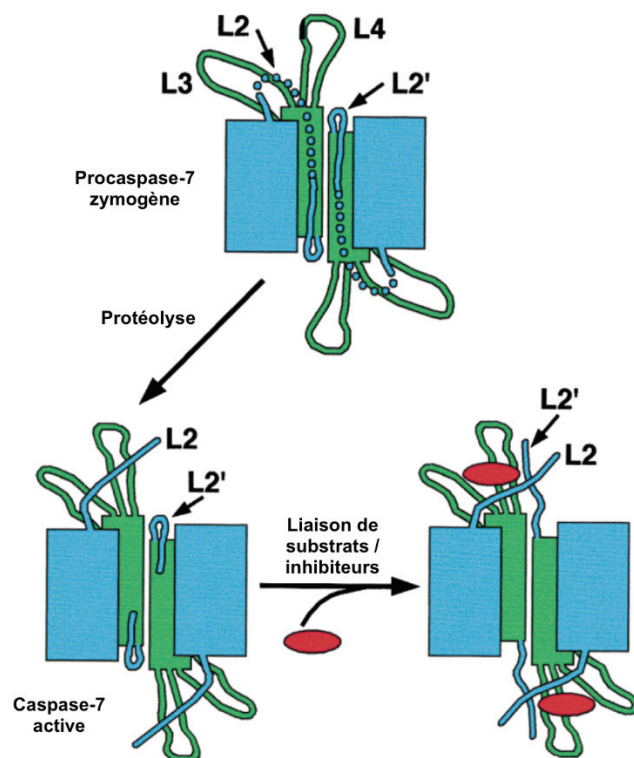
La résolution de la structure de la caspase-7 effectrice sous sa forme zymogène (procaspase-7), grâce à la mutation de la cystéine catalytique (Cys<sup>186</sup>) en Ala, et sous sa forme active a permis d'identifier la conservation de l'organisation spatiale des sous unités p20 et p10 (Jijie Chai *et al.*, 2001). En effet, les auteurs ont pu déterminer un RMSD (« Root Mean Square Deviation ») de 0,8 Å en superposant les chaînes carbonées des deux formes de caspase-7. En revanche, 3 boucles parmi les 4 composant le site actif adoptent des conformations différentes. En effet, bien que la boucle L5 suivant l'hélice  $\alpha 5$  (Figure 4), aussi appelée L4 (Figure 6), ne soit pas intégralement visible pour la procaspase-7, les auteurs font l'hypothèse que cette dernière se loge dans le sillon catalytique (Jijie Chai *et al.*, 2001). De plus, il est possible de constater qu'après protéolyse de la boucle L4 (Figure 4), nommée L2 par les auteurs (Figure 6), au niveau de l'Asp conservé (Asp<sup>198</sup>), celle-ci effectue une bascule de 90°

rendant alors la Cys<sup>186</sup> accessible au solvant (Figure 6) (Jijie Chai *et al.*, 2001). Dans le second monomère, la boucle L4' (nommée L2' par les auteurs) bascule quant à elle de 180°. Enfin, la boucle L5 $\alpha$  effectue une rotation de 60° et se rapproche de la boucle L5 (nommée L3 par les auteurs) stabilisant ainsi le sillon catalytique de l'enzyme (Jijie Chai *et al.*, 2001).



**Figure 6. Comparaison structurale du site actif de la procaspase-7 zymogène et de la caspase-7 sous forme active adapté de Jijie Chai *et al.*, 2001.** Les deux structures, procaspase-7 et caspase-7, sont respectivement représentées en *bleu* et en *vert*, le mouvement des boucles est indiqué par des flèches rouges. Les boucles L1, L2, L3 et L4 sont nommées d'après la publication est correspondent respectivement aux boucles L1, L4, L5 et L5 $\alpha$  d'après la numérotation issue de McLuskey and Mottram, 2015 utilisé dans ce manuscrit. Le résidu catalytique Cys<sup>186</sup> est coloré en or et rose respectivement pour l'enzyme active et inactive.

La proximité des boucles L4 (L2), L4' (L2') et L5 $\alpha$  (L4) semble importante pour la stabilisation et le repliement du site actif de l'enzyme en conformation active. Ce regroupement de boucles a d'ailleurs été identifié comme lieu d'interaction avec des inhibiteurs de caspase pour la structure de la caspase-7 en complexe avec un inhibiteur de la famille des IAP (J. Chai *et al.*, 2001). A l'inverse, la position des boucles dans l'enzyme inactive ne permet pas un accès et une stabilisation du site actif suffisante rendant la procaspase-7 inactive. Ces observations ont permis aux auteurs d'émettre l'hypothèse d'un mécanisme d'activation et de changements structuraux associés (Figure 7).



**Figure 7. Représentation schématique du mécanisme d'activation de la caspase-7 adapté de Jijie Chai *et al.*, 2001.** Les sous unités p20 et p10 sont respectivement représentées en *bleu* et *vert*. La protéolyse de la boucle L4 (appelée L2) et L4' (appelée L2'), du second monomère, au résidu Asp<sup>198</sup> induit l'accessibilité au site actif et l'activation de l'enzyme pouvant alors lier un substrat/inhibiteur.

#### d) Spécificité de clivage

Les caspases sont des protéases dont la spécificité de clivage dépend de la reconnaissance de 4 résidus consécutifs respectivement P4-P3-P2-P1. Le clivage se produit à l'extrémité C-terminale, soit au niveau du résidu P1 étant généralement un résidu aspartate (Shi, 2002). Chez les caspases de mammifère, le résidu P3 est un résidu Glu (Thornberry *et al.*, 1997). Ainsi la spécificité de clivage est décrite comme étant X-Glu-X-Asp. La nature du résidu P4 permet la séparation des caspases en trois groupes, les caspases -1, -4 et -5 préfèrent la séquence WEHD alors que les caspases -2, -3 et -7 préfèrent la séquence DEXD et la caspase-9 la séquence (L/V)EXD (Lavrik, Golks and Krammer, 2005). Comme indiqué précédemment le site catalytique des caspases se compose des résidus His et Cys. Ces derniers se trouvent au sein d'une poche « basique » due à la conservation de résidus chargés positivement expliquant la préférence d'un résidu Asp chargé négativement à la position P1 (cf. Figure 15, p88) (Klemenčič and Funk, 2018).

#### e) Rôle des caspases

Comme indiqué précédemment, le rôle des caspases initiatrices (-8 et -9) réside dans l'activation des caspases effectrices (-3, -6 et -7) n'étant pas capables de s'autoprotéolyser. En

revanche, la caspase-10, outre son implication dans l'induction de l'apoptose, a également été associée à l'autophagie, l'activation du facteur NF- $\kappa$ B et la survie cellulaire (Lamy *et al.*, 2013; Horn *et al.*, 2017).

Les caspases effectrices (-3, -6, -7) sont responsables des changements morphologiques typiques de l'apoptose, la rétractation de la cellule, la fragmentation de l'ADN, le bourgeonnement et la formation de corps apoptotiques contenant les constituants cellulaires qui seront dégradés par les macrophages (Ramirez and Salvesen, 2018). Plus d'une centaine de substrats des caspases effectrices ont pu être mis en évidence, parmi lesquels peu d'entre eux ont montré un rôle dans la dégradation des cellules apoptotiques ou l'externalisation de signaux menant à leur phagocytose (Julien and Wells, 2017).

Parmi les caspases inflammatoires, la caspase-1 est la plus étudiée. Elle fut identifiée lors de recherches de l'enzyme responsable de la maturation de la cytokine inflammatoire IL-1 $\beta$  (« Interleukine-1 $\beta$  ») et initialement nommée ICE pour « Interleukine-1 $\beta$  Converting Enzyme » (Thornberry *et al.*, 1992). Bien que son implication dans l'apoptose semble non essentielle, il a été montré qu'elle participait à la mise en place de la pyroptose, une forme de mort cellulaire programmée par nécrose induisant une inflammation. Les caspases inflammatoires -4, -5 et la caspase murine orthologue, caspase-11, ne sont pas capables de cliver IL-1 $\beta$  mais la protéine gasdermin D (« GSDMD ») participant également à la mise en place de la pyroptose (Shi *et al.*, 2015).

Les fonctions des caspases -2 et -14 ne sont pas encore complètement comprises mais ces dernières semblent respectivement participer à l'arrêt du cycle cellulaire et à la différenciation des kératinocytes (Van Opend Bosch and Lamkanfi, 2019).



## B. La mort cellulaire chez la levure

L'existence d'un phénomène de mort cellulaire régulé a longtemps été remise en question car considérée comme non nécessaire aux organismes unicellulaires. En effet, la mort cellulaire régulée d'un organisme unicellulaire induit, par définition, la mort de l'organisme en lui-même. De plus, aucune protéine homologue aux régulateurs de l'apoptose n'avait été identifiée. Cependant, en 1997, Frank Madeo et son équipe mettent en évidence un certain nombre de changements morphologiques caractéristiques de l'apoptose (condensation de la chromatine, fragmentation de l'ADN, externalisation de phosphatidyl-sérines) chez la levure *Saccharomyces cerevisiae* lorsqu'ils mutent le gène *CDC48* (Madeo, Fröhlich and Fröhlich, 1997). La protéine codée par ce gène intervient dans l'ubiquitylation des protéines et sa mutation arrête le cycle cellulaire, déclenchant ainsi un signal de mort cellulaire. La même équipe de recherche a également démontré ces mêmes changements morphologiques ainsi que la production de ROS lorsque la cellule est traitée au peroxyde d'hydrogène (Madeo *et al.*, 1999). Les mêmes observations ont également été faites chez l'organisme *Candida albicans*, lorsque les levures étaient soumises au peroxyde d'hydrogène ou à l'acide acétique (Phillips, Sudbery and Ramsdale, 2003 ; Phillips, Crowe and Ramsdale, 2006). En identifiant ces changements morphologiques et biochimiques, il a aussi été possible d'observer que l'apoptose est un mécanisme activé lors de la sénescence cellulaire et donc participant au vieillissement des levures. L'évidence d'une mort cellulaire régulée chez les levures est expliquée par l'avantage évolutif que ce processus représente. En effet, la mort cellulaire des organismes unicellulaires est considérée comme un suicide altruiste permettant à des cellules âgées ou ayant subies des dommages fatales, de stopper leur prolifération et de ne plus consommer des ressources du milieu pouvant bénéficier à la survie des cellules voisines (Herker *et al.*, 2004). C'est le mode de vie communautaire de ces organismes unicellulaires qui leur confère un comportement d'entité multicellulaire.

### 1. *Mort cellulaire dépendante de la mitochondrie*

Bien que la majeure partie de la production des ROS se déroule à la mitochondrie, la première mise en évidence d'une voie de mort cellulaire dépendante de la mitochondrie, similaire à l'apoptose intrinsèque des mammifères, provient du traitement de cellules de *S. cerevisiae* à l'acide acétique. En effet, ce stimulus induirait la libération du cytochrome *c* dans le cytosol ainsi qu'une dépolarisation et perméabilisation de la membrane mitochondriale (Ludovico *et al.*, 2001, 2002). Une protéine homologue aux protéines pro-apoptotiques BH3 régulant le phénomène de MOMP a été identifiée chez *S. cerevisiae*, il s'agit de la protéine YBH3 (Büttner *et al.*, 2011). En cas de stimuli de mort cellulaire, cette dernière est transloquée à la

membrane mitochondriale. Son interaction avec un transporteur de phosphate de la membrane mitochondriale (MIR1) et la sous-unité du complexe III (COR1) de la chaîne respiratoire, induit une chute du potentiel transmembranaire et une externalisation du cytochrome *c* (Büttner *et al.*, 2011). La levure *S. cerevisiae* présente également la capacité de former un canal dans la membrane mitochondriale interne, appelé YMUC (« Yeast Mitochondrial Unselective Channel », canal mitochondriale non sélectif des levures), semblable au pore de perméabilisation PTP.

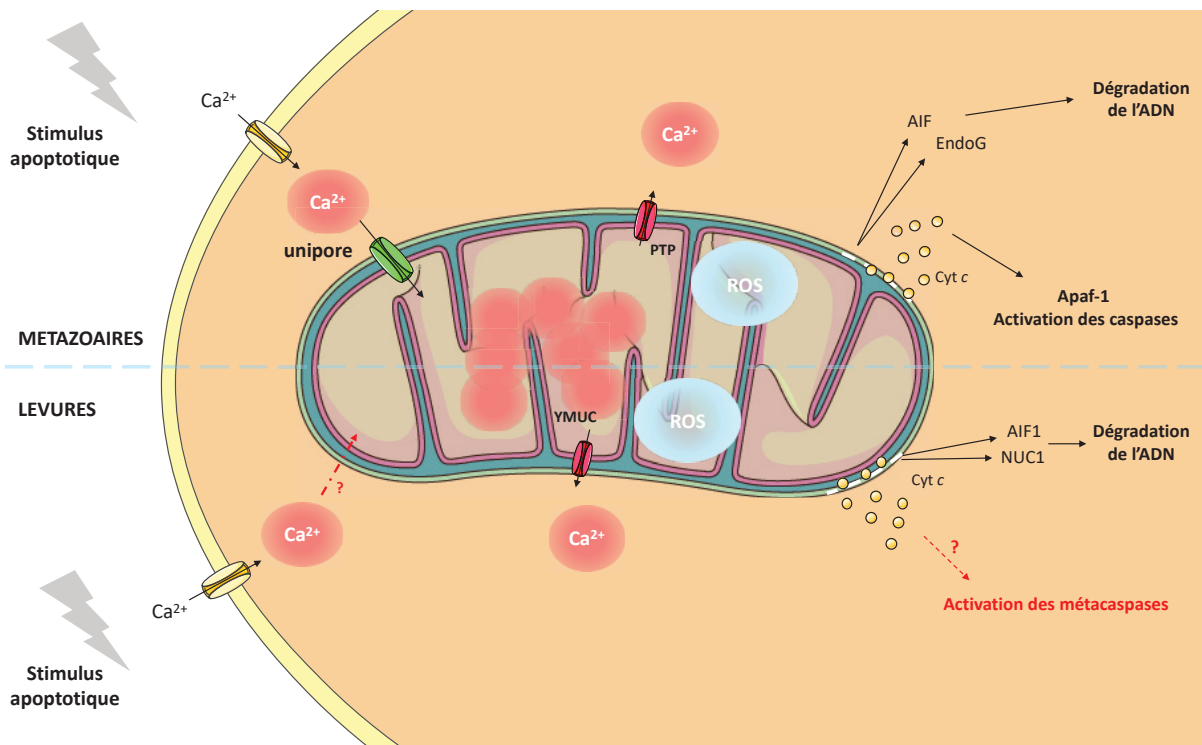
La perméabilisation mitochondriale entraîne aussi la libération de protéines mitochondriales, propageant le signal apoptotique, telles que le facteur AIF1 (« Apoptosis-induction factor 1 », le facteur 1 d'induction de l'apoptose) et le facteur NUC1 (« endonuclease-G », l'endonucléase 1), respectivement homologues à AIF et EndoG (endonucléase G) participant à la condensation de la chromatine et la fragmentation de l'ADN (Falcone and Mazzoni, 2016; Carmona-Gutierrez *et al.*, 2018) (Figure 8). Ces effecteurs de la mort cellulaire ne sont pas dépendants de la voie d'activation de mort *via* l'action des métacaspases et sont dits « métacaspase-indépendants ». En effet, chez les levures et les plantes, des homologues structuraux aux caspases ont pu être mis en évidence et ont été nommés métacaspases (Uren *et al.*, 2000). En revanche, l'existence d'un homologue fonctionnel à l'apoptosome n'ayant pas été identifié, un lien entre cytochrome *c* et activation des métacaspases reste à définir. L'activation de ces voies dépendantes ou non des métacaspases indiquent l'existence de multiples voies régulant la mort cellulaire (Guaragnella *et al.*, 2012).

Un phénomène de type apoptose extrinsèque n'est pas décrit chez les levures ce qui est notamment dû à l'absence de récepteurs de mort. En revanche, des molécules du milieu extracellulaire, comme des toxines virales sécrétées ou des molécules pharmacologiques, ont été identifiées comme induisant la mort cellulaire « métacaspase-dépendante » des levures (Reiter *et al.*, 2005 ; Shirazi and Kontoyiannis, 2015 ; Scariot *et al.*, 2016). Un manque de données ne permet pas encore d'expliquer si ces facteurs agissent comme des facteurs intracellulaires ou sont reconnus par des récepteurs plasmiques.



a) *Activation par le calcium*

Comme chez les mammifères, la concentration cytosolique de calcium est finement régulée chez les levures et maintenue à de faibles concentrations (50-200 nM) (Carraro and Bernardi, 2016). Son absorption depuis le milieu extracellulaire est médiée par des canaux voltage dépendants présents dans la membrane plasmique. Au niveau mitochondrial, la production de ROS et l'augmentation de la concentration en calcium sont des événements clés dans la mise en place du phénomène de mort cellulaire. Cependant, chez les levures, le lien entre l'augmentation de calcium intracellulaire et le phénomène de perméabilisation membranaire reste encore énigmatique étant donné qu'elles ne possèdent pas de transporteur mitochondriale de calcium (Carafoli and Lehninger, 1971 ; Carraro and Bernardi, 2016). Pourtant, en réponse à des stimuli apoptotiques, la concentration intramitochondriale de calcium augmenterait de manière similaire aux mammifères jusqu'à environ 100 mM (Carraro and Bernardi, 2016). Ces résultats suggèrent l'existence d'un système de transport spécifique du calcium n'ayant pas encore été identifié mais permettant une accumulation de calcium, induisant la perméabilisation membranaire et l'ouverture du pore YMUC (Carafoli and Lehninger, 1971) (Figure 8). Cette ouverture est suivie par le gonflement de la matrice mitochondriale, des dommages à la membrane externe et la libération de protéines mitochondriales (Pereira *et al.*, 2008; Carraro and Bernardi, 2016).



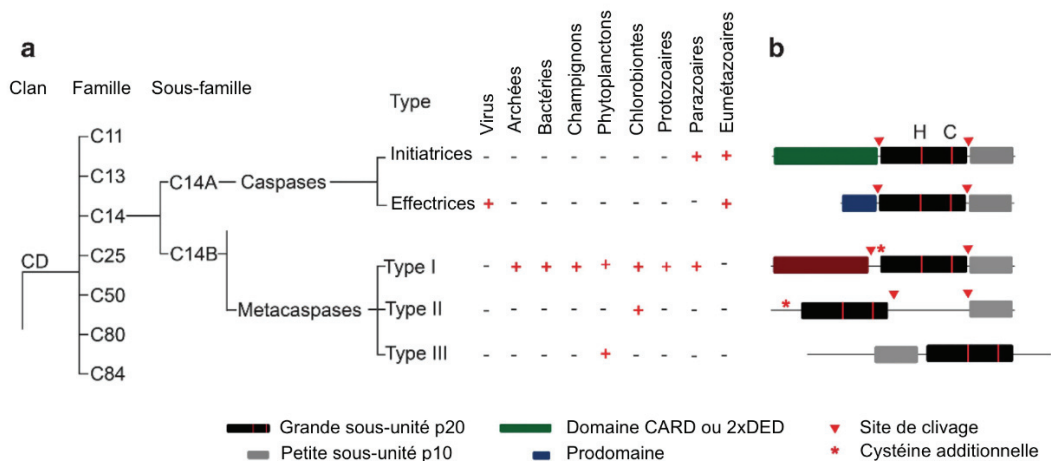
**Figure 8. Voies intracellulaires dépendantes du calcium dans la mort cellulaire régulée chez les mammifères et chez les levures, adapté de Carraro et Benardi, 2016.** Dans la partie supérieure de l'image, l'augmentation de calcium intracellulaire due à un stimulus apoptotique chez les métazoaires entraîne l'activation de l'unipore transportant le calcium dans la matrice mitochondriale. La concentration de calcium augmente dans la mitochondrie déclenchant l'augmentation de la production de ROS, l'activation du pore de perméabilisation mitochondriale (PTP) suivi par la rupture de la membrane mitochondriale externe et de la libération du cytochrome *c* (Cyt *c*) et des facteurs AIF et EndoG. Alors que Cyt *c* interagit avec le complexe Apaf-1 et induit l'activation des caspases, AIF et EndoG exercent leur action de condensation de la chromatine et de dégradation de l'ADN. Dans la partie inférieure, l'augmentation de calcium intracellulaire, chez les levures, induit une absorption massive de l'ion calcium par un mécanisme ou un transporteur encore inconnu. Cette augmentation de concentration aura le même effet que chez les métazoaires. En revanche, il n'existe pas de lien direct d'activation entre le cytochrome *c* et les métacaspases.

## 2. Les métacaspases

Bien que l'identification de marqueurs cellulaires typiques de l'apoptose ait permis de mettre en évidence la mort cellulaire régulée chez les organismes unicellulaires, des protéines homologues aux caspases tardaient à être identifiées. Ce n'est que grâce à des alignements de séquence et à des prédictions de structures secondaires, que des auteurs ont pu identifier des homologues structuraux possédant un faible pourcentage d'identité de séquence. Deux familles de type caspase ont ainsi été découvertes et nommées paracaspase et métacaspase (Uren *et al.*, 2000). Alors que la première famille est retrouvée chez les métazoaires, les myxomycètes et les bactéries, les métacaspases sont retrouvées dans de nombreux organismes pro- et eucaryotes dont les levures mais sont absentes chez les virus et les eumétazoaires (Minina *et al.*, 2017) (Figure 9).

### a) *Classification*

La première métacaspase de levure a été identifiée chez *S. cerevisiae* et nommée Yca1 pour « Yeast caspase-1 » ou caspase-1 de levure (Uren *et al.*, 2000). Cette découverte a été suivie de l'identification de la métacaspase de *C. albicans* nommée Métacaspase 1 (Cao *et al.*, 2009). D'après la classification MEROPS, les métacaspases comme les caspases et paracaspases appartiennent à la famille C14 des protéases à cystéine et le clan CD. Elles sont également sous-classées avec les paracaspases dans la famille C14B (Minina *et al.*, 2020). Cette classification est effectuée sur la base de similarités significatives de séquences en acides aminés du domaine catalytique pour la classification en famille et sur la similarité des structures tertiaires pour la classification en clans, sans prendre en considération les propriétés biochimiques des protéines (Rawlings *et al.*, 2018). La famille C14 regroupe des protéases possédant une spécificité de clivage d'un résidu Asp en position P1. Or, ni les métacaspases ni les paracaspases ne possèdent de spécificité de clivage après un résidu Asp. En revanche, elles préfèrent cliver après un résidu Arg ou Lys (Vercammen *et al.*, 2004 ; Rebeaud *et al.*, 2008). Ceci remet également en question leur dénomination issue du mot caspase correspondant initialement à des protéases cystéine dépendantes et spécifiques de l'aspartate. Cependant, les termes ayant été utilisés depuis de nombreuses années, la communauté s'est accordée sur le fait de conserver cette dénomination (Minina *et al.*, 2020). Récemment, une nomenclature commune a été proposée, l'abréviation latine de l'organisme suivi du terme MCA pour métacaspase, du type de la protéine écrit en chiffre latin et éventuellement d'une lettre (dans l'ordre alphabétique) permettant de différencier les homologues de la protéine au sein d'un même organisme (Minina *et al.*, 2020). En effet, bien que les levures ne possèdent qu'un seul gène codant pour une métacaspase, les plantes et les protozoaires possèdent plusieurs isoformes de métacaspases (Madeo *et al.*, 2002 ; Vercammen *et al.*, 2004 ; McLuskey *et al.*, 2012).



**Figure 9. Représentation des protéases de la famille C14 dans différents groupes taxonomiques adapté de Minina *et al.*, 2017.** A. Les caspases sont classées dans la sous famille C14A et sont retrouvées chez les virus, les parazoaires et les métazoaires. Les métacaspases sont classées dans la sous famille C14B, le type II et III sont respectivement retrouvés chez les chlorobiontes et les phytoplanctons. Le type I est retrouvé chez une large gamme d’organismes, les archées, bactéries, champignons, phytoplanctons, protozoaires et parazoaires mais sont absents chez les virus et les eumétazoaires.

Les métacaspases sont divisées en trois types en fonction de leurs différences structurales (Figure 9) (Madeo *et al.*, 2002 ; Vercammen *et al.*, 2004 ; McLuskey *et al.*, 2012). Le type I est caractérisé par une extension riche en proline et la présence d’un domaine en doigt de zinc à l’extrémité N-terminale qui est absente chez les types II et III (Vercammen *et al.*, 2004). Les métacaspases, comme les caspases, sont composées des deux sous-unités p20, contenant la diade catalytique His/Cys, et p10 (McLuskey and Mottram, 2015). Bien que les trois types de métacaspase possèdent un domaine de liaison entre les deux sous-unités, le type II se caractérise par la présence d’un domaine de liaison étendu. Pour le type III, les deux sous-unités sont inversées avec le domaine p10 puis p20 en partant de l’extrémité N-terminale (Minina *et al.*, 2020). Le type I est retrouvé chez un large nombre d’organismes comme les protistes, les champignons, les plantes, les algues, les bactéries et les archées (Minina *et al.*, 2017). En revanche, les types II et III n’ont été identifiés, respectivement, que dans les chlorobiontes (algues vertes et plantes terrestres) et les algues brunes (en particulier le phytoplancton) (Minina *et al.*, 2017).

## b) Repliement structural

À ce jour, seulement trois structures tridimensionnelles ont été déterminées par la technique de cristallographie aux rayons X. Il s'agit des métacaspases de type I de l'organisme *Trypanosoma brucei* (*T. brucei*) et *S. cerevisiae* en 2012 puis la métacaspase-4 de type II de *Arabidopsis thaliana* (*A. thaliana*) en 2020. Ces dernières ont permis d'évaluer les ressemblances et les différences structurales qui les séparent notamment des caspases effectrices. En raison de la grande diversité de dénomination des boucles entre les différentes caspases et les métacaspases, ces dernières seront renommées d'après la caspase-7 décrite par Mottram et McLuskey (McLuskey and Mottram, 2015) où elles portent le numéro du brin  $\beta$  ou de l'hélice  $\alpha$  les précédents. Un tableau récapitulatif de leur dénomination se trouve ci-dessous (Tableau 1).

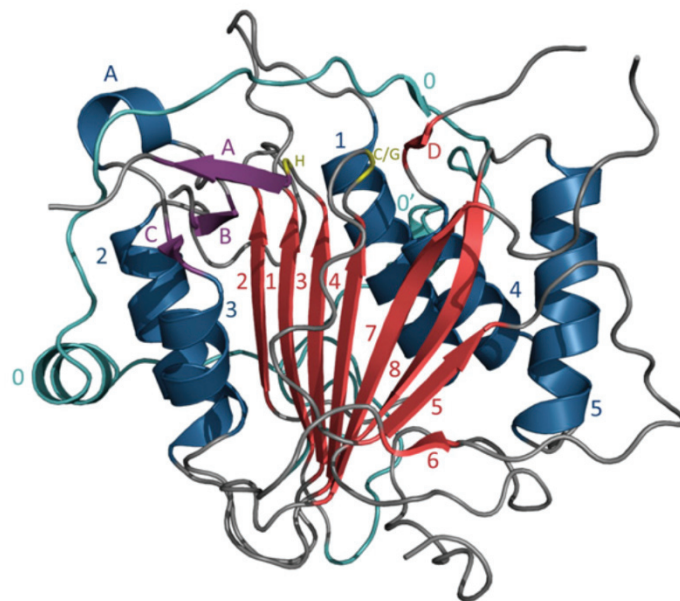
Caspase											
Jijie Chai <i>et al.</i> , 2001	Caspase-7	L1				L2	L3	L4			
Mottram and McLuskey 2015		L1	L2	L3	L4	L5	L5 $\alpha$				
Metacaspase type I											
McLuskey <i>et al.</i> , 2012	<i>Tb</i> MCA-Ib						Boucle 280				
Wong, Yan and Shi, 2012	<i>Sc</i> MCA-I	L1				L2	L3	L4			
Metacaspase type II											
Zhu <i>et al.</i> , 2020	<i>At</i> MCA-IIa	L1				L2	L3	L4	L5	L6	L7

**Tableau 1. Comparatif des dénominations des boucles entre les structures de la caspase-7, *Tb*MCA-Ib, *Sc*MCA-I et *At*MCA-IIa.**

### (1) *Tb*MCA-Ib

La première structure de métacaspase résolue a été celle du parasite protozoaire *T. brucei* responsable de la trypanosomiase humaine africaine. Le génome de ce parasite code 5 métacaspases que nous nommerons *Tb*MCA-Ia à *Tb*MCA-Ie, par respect de la nomenclature actuelle. Les auteurs n'ayant pu déterminer la structure de la forme active de l'enzyme, à cause de son autoprotéolyse *in vitro* et à la présence de produits de dégradation, ils ont résolu la structure tridimensionnelle de l'enzyme sous une forme mutée (McLuskey *et al.*, 2012). En effet, des expériences de mutagenèse dirigée de la cystéine catalytique (C213) et la cristallisation de l'enzyme sous cette forme mutée ont permis de résoudre la structure de *Tb*MCA-Ib-C213G et de *Tb*MCA-Ib-C213A respectivement à 1,4 Å (PDB : 4AFV) et 1,6 Å (PDB : 4AFR) de résolution. Les deux structures étant identiques, c'est *Tb*MCA-Ib-C213G qui est décrite. La qualité de la carte de densité électronique leur a permis de reconstruire la structure de l'enzyme du résidu 3 à 347 à l'exception de deux boucles constituées des résidus 166-170 et 269-275 (cette dernière étant nommée boucle 280).

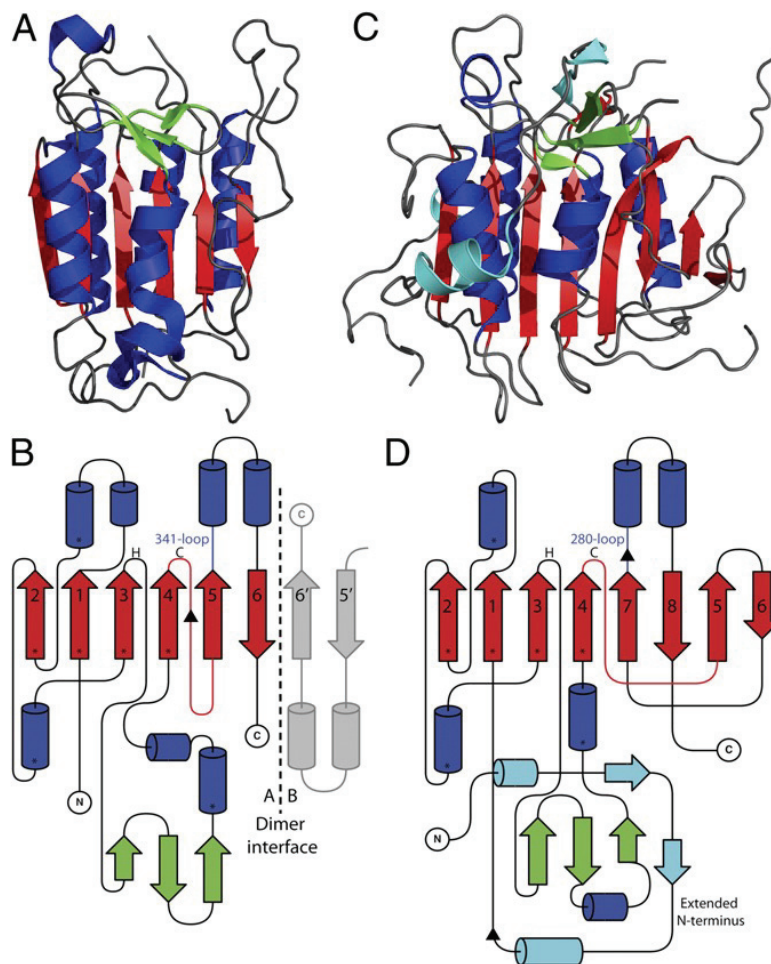
*TbMCA2-C213G* est constituée d'un feuillet de 8 brins  $\beta$  ( $\beta 1$ - $\beta 8$ ) dont 6 parallèles et 2 antiparallèles selon la typologie :  $2\uparrow 1\uparrow 3\uparrow 4\uparrow 7\uparrow 8\downarrow 5\uparrow 6\downarrow$  (Figure 10). Le feuillet est entouré par 5 hélices  $\alpha$  ( $\alpha 1$ - $\alpha 5$ ), avec les hélices  $\alpha 1$ ,  $\alpha 4$  et  $\alpha 5$  se trouvant d'un côté du feuillet et les hélices  $\alpha 2$  et  $\alpha 3$  du côté opposé. A l'extrémité C-terminale du brin  $\beta 3$ , se trouve une extension de trois courts brins  $\beta$  formant un feuillet ( $\beta A$ - $\beta C$ ) avec les brins  $\beta B$  et  $\beta C$  connectés par une boucle dont une petite section est repliée en hélice  $\alpha$  ( $\alpha A$ ). Les résidus catalytiques, His<sup>158</sup> et Cys<sup>213</sup> (ce dernier étant muté en Gly dans ce cas), se situent respectivement à l'extrémité C-terminale des brins  $\beta 3$  et  $\beta 4$ . La boucle portant le résidu Cys entre le brin  $\beta 4$  et  $\beta 5$ , appelée boucle catalytique, est constituée de 11 résidus et présente une grande flexibilité. A l'extrémité N-terminale du brin  $\beta 1$ , se trouve également une extension de 70 résidus étant constituée d'une hélice  $3_{10}$  ( $\alpha 0'$ ), 2 brins  $\beta$  ( $\beta 0'$  et  $\beta 0$ ) et une hélice  $\alpha$  ( $\alpha 0$ ). Cette boucle N-terminale, encercle l'enzyme *via* 129 résidus impliqués dans l'interaction avec le cœur de la protéine. Elle traverse cette dernière au-dessus du site actif enfouissant ainsi les deux résidus catalytiques (McLuskey *et al.*, 2012).



**Figure 10. Représentation de la structure tridimensionnelle de *TbMCA-Ib-C213G* d'après McLuskey *et al.*, 2012.** La boucle N-terminale flexible et entourant le cœur de l'enzyme est représentée en *turquoise*, les hélices  $\alpha$  sont représentées en *bleu*, les brins  $\beta$  du feuillet  $\beta$  central sont représentés en *rouge*, les brins  $\beta$  de l'extension en C-terminale du brin  $\beta 3$  sont représentés en *violet* et les autres boucles sont représentées en *gris*.

Des analyses comparatives ont permis de mettre en évidence que la structure de *TbMCA-Ib-C213G* possède une similarité de structure secondaire de 65 % avec la caspase-7 effectrice humaine (Figure 11 A, C) (McLuskey *et al.*, 2012). De manière intéressante, le feuillet  $\beta$  central de *TbMCA-Ib-C213G* est constitué de deux brins  $\beta$  supplémentaires ( $\beta 5$  et  $\beta 6$ ) par rapport à la caspase-7 (Figure 11 B, D). L'enzyme du parasite existe sous une forme monomérique active en solution et la présence de nombreuses boucles flexibles en surface n'est pas favorable à une dimérisation par les brins  $\beta$  du feuillet principal (McLuskey *et al.*, 2012). Une superposition des deux structures a permis d'observer que le brin  $\beta 5$  de *T. brucei* se superpose au brin  $\beta 6'$  du second monomère de la caspase-7. Ainsi, il semblerait que les deux brins  $\beta$  additionnels miment l'interface de dimérisation existante chez les caspases et stabilisent l'enzyme sous une forme monomérique. De manière similaire aux caspases, il est possible de distinguer les deux sous-unités, la p10 étant composée des brins  $\beta 7$  et  $\beta 8$  et la p20 étant constituée des autres brins  $\beta$  du feuillet principal (Figure 11 D). La superposition des résidus catalytiques a permis d'observer que leur position était conservée comme chez les autres enzymes du clan CD à la différence que pour *TbMCA-Ib*, les deux résidus ne se trouvent pas en surface dû à la position de l'extrémité N-terminale. Les auteurs émettent l'hypothèse que cette extension N-terminale bloque l'accès des substrats au site actif et qu'un déplacement de celle-ci est nécessaire à la mise en place de l'activité catalytique (McLuskey *et al.*, 2012).





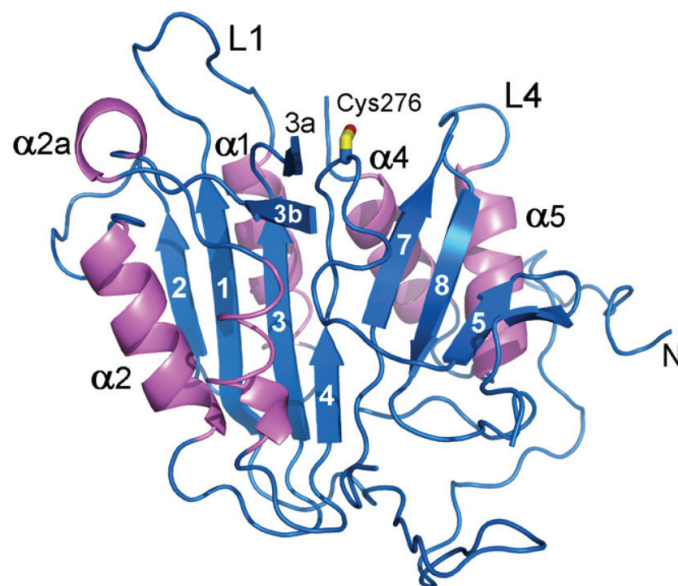
**Figure 11. Comparaison structurale et topologique des structures de *TbMCA-1b* et de la caspase-7 d'après de McLuskey *et al.*, 2012. (A, C) Représentation de la structure tridimensionnelle de la caspase-7 (A) et de *TbMCA2* (C) avec les boucles en *gris*, les hélices  $\alpha$  en *bleu*, les brins  $\beta$  du feuillet central en *rouge* et l'extension de brins  $\beta$  entre le brin  $\beta_3$  et  $\beta_4$  en *vert*. Les structures secondaires de l'extension N-terminale de *TbMCA2* sont représentées en *turquoise*. (B, D) Représentation de la typologie de la caspase-7 (B) et de *TbMCA2* (D) avec les mêmes couleurs que A et C. L'interface dimérique (« dimer interface ») de la caspase-7 est représentée en pointillés et la structure partielle du dimère en *gris*. Les résidus catalytiques His (H) et Cys (C) sont respectivement localisés à l'extrémité C-terminale des brins  $\beta_3$  et  $\beta_4$  et la boucle catalytique portant C est représentée en *rouge*. Les boucles suivant le brin  $\beta_5$  pour la caspase-7 et le brin  $\beta_7$  sont respectivement nommées boucle 341 (« 341 loop ») et boucle 280 (« 280 loop »), elles correspondent à la boucle L5 et sont identifiées en *bleu*. Les triangles *noirs* représentent les zones de clivage identifiées.**

## (2) ScMCA-I

La découverte de l'existence des homologues des caspases a été initiée par l'identification du gène codant pour la métacaspase de *S. cerevisiae* appelée Yca1 et que nous renommerons ScMCA-I par respect de la nomenclature actuelle (Uren *et al.*, 2000). Afin de résoudre la structure de l'enzyme, les auteurs ont réalisé une étape de protéolyse limitée visant à éliminer l'extrémité N-terminale flexible. Ainsi, la structure tridimensionnelle de ScMCA-I, tronquée des 83 premiers acides aminés, a été déterminée à 1,68 Å de résolution (PDB : 4F6O) (Wong, Yan and Shi, 2012).



La structure de *ScMCA-I* est composée d'un feuillet de 8 brins  $\beta$  ( $\beta 1$ - $\beta 8$ ) entouré de 3 hélices  $\alpha$  ( $\alpha 1$ ,  $\alpha 4$ ,  $\alpha 5$ ) d'un côté et 2 hélices  $\alpha$  ( $\alpha 2$  et  $\alpha 3$ ) du côté opposé (Figure 12) (Wong, Yan and Shi, 2012). Entre le brin  $\beta 3$  et l'hélice  $\alpha 3$  se trouvent une épingle à cheveux  $\beta$  constituée du brin  $\beta 3a$  et  $\beta 3b$  ainsi qu'une courte hélice  $\alpha$ ,  $\alpha 2a$ . La structure est également composée de 3 boucles ordonnées nommées L1, L2 et L4 de manière conventionnelle par rapport à la caspase-3 (PDB : 1CP3) et correspondant aux boucles L1, L4 et L5 $\alpha$  selon la nomenclature de la caspase-7 (PDB : 1F1J) (Figure 4, Tableau 1). La première se situe en C-terminale du brin  $\beta 1$ , la seconde entre les brins  $\beta 4$  et  $\beta 5$  et enfin la dernière entre l'hélice  $\alpha 5$  et le brin  $\beta 8$ . Les résidus catalytiques His<sup>220</sup> et Cys<sup>276</sup>, sont portés respectivement par une boucle à l'extrémité C-terminale du brin  $\beta 3$  et par la boucle nommée L2 par les auteurs (Wong, Yan and Shi, 2012). La boucle suivant le brin  $\beta 7$ , nommée L3, est décrite comme servant de base au sillon catalytique mais n'est pas visible dans la structure.



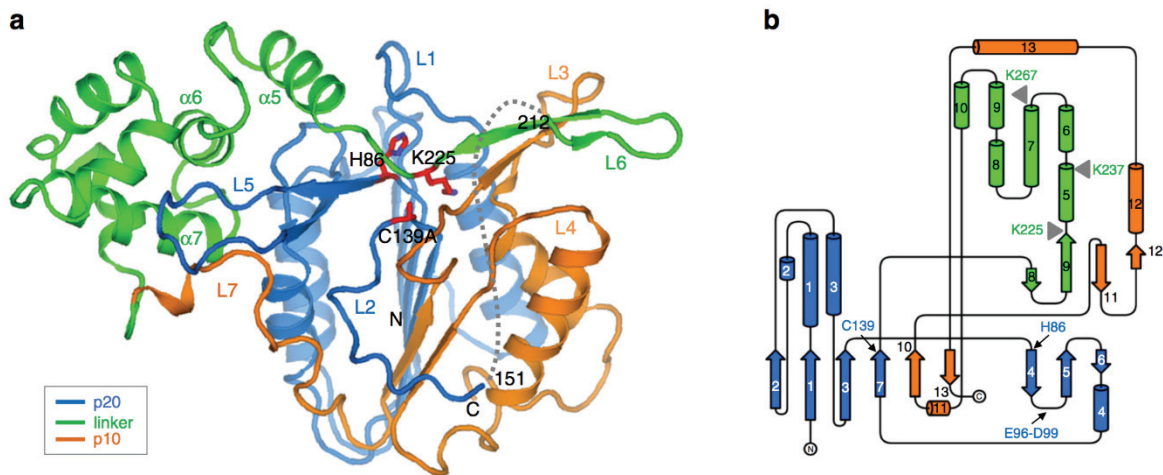
**Figure 12. Représentation de la structure tridimensionnelle de *ScMCA-I* d'après Wong, Yan et Shi, 2012.** Les hélices  $\alpha$  sont représentées en rose, les brins  $\beta$  en bleu et le résidu catalytique Cys<sup>276</sup> est représenté en bâtonnet. Les boucles L1 et L4 ont été nommées selon la nomenclature de la caspase-7 (Jijie Chai *et al.*, 2001) et correspondent aux boucles L1 et L5 $\alpha$ .

Comme pour *TbMCA-Ib*, le feuillet  $\beta$  central de *ScMCA-I* est constitué de 2 brins  $\beta$  additionnels ( $\beta 5$  et  $\beta 6$ ) en comparaison aux caspases (Wong, Yan and Shi, 2012). Les auteurs, de la même manière que pour *TbMCA-Ib*, observent que ces brins sont localisés à l'interface dimérique des caspases, expliquant ainsi l'incapacité des métacaspases à former des dimères. Ces différences structurales importantes sont notamment expliquées par une très faible

homologie de séquence avec les caspases ; *ScMCA-I* possède seulement 10,2 % et 11,4 % respectivement avec les caspase-3 et -9. La superposition des structures a permis la mesure d'un RMSD important de 7,45 Å et 7,61 Å respectivement pour la caspase-3 et -9 (Wong, Yan and Shi, 2012). En revanche, *ScMCA-I* partage une homologie de séquence de 24,9 % et un RMSD de 2,3 Å avec *TbMCA-Ib*.

### (3) *AtMCA-II*d

La troisième structure de métacaspase résolue est celle de l'organisme *Arabidopsis thaliana*. Il s'agit de la première structure d'une métacaspase de type II. L'enzyme native et l'enzyme mutée au résidu catalytique Cys<sup>139</sup> en Ala, ont été respectivement déterminées à 3,5 Å (*AtMCA-II*d-WT) et 2,8 Å (*AtMCA-II*d-C139A) de résolution (Zhu *et al.*, 2020). La structure est constituée des deux sous-unités p20 et p10 ainsi que d'une large région de liaison entre les deux, caractéristique de ce type de métacaspase. Cette région forme un domaine structural s'étendant du brin  $\beta$ 7 à l'hélice  $\alpha$ 11 et se composant d'une épingle à cheveux  $\beta$  ( $\beta$ 8 et  $\beta$ 9) et d'un ensemble d'hélices  $\alpha$  ( $\alpha$ 5- $\alpha$ 10) (Figure 13 A, B). Le cœur de la protéine est constitué par un feuillet  $\beta$  central avec 2 brins  $\beta$  antiparallèles du domaine p10 et 6 brins  $\beta$  du domaine p20. Deux brins supplémentaires du domaine p20, identifiés par analyse topologique, ne sont pas résolus dans la structure de *AtMCA-II*d. De plus, la boucle nommée L2 (positionnée entre les brins  $\beta$ 3 et  $\beta$ 4) en amont de la zone non résolue, semble se positionner au niveau du potentiel site de dimérisation expliquant que l'enzyme soit monomérique. Il est pris en sandwich par plusieurs hélices de part et d'autre du cœur de l'enzyme ( $\alpha$ 1- $\alpha$ 4,  $\alpha$ 11- $\alpha$ 13). Les résidus catalytiques, His<sup>86</sup> et Cys<sup>139</sup>, se trouvent respectivement en N-terminale du brin  $\beta$ 4 et en C-terminale du brin  $\beta$ 7. La partie N-terminale de la région de liaison traverse la protéine notamment au niveau du sillon catalytique où le résidu Lys<sup>225</sup>, identifié comme un site de clivage, est positionné dans la poche catalytique négativement chargée (Zhu *et al.*, 2020). En comparaison à la caspase-7, *AtMCA-II*d possède une boucle L5 plus courte (Figure 13), correspondant à la boucle entre les brins  $\beta$ <sup>1</sup> et  $\beta$ <sup>2</sup> du feuillet  $\beta$  de la caspase-7 (Figure 4). Cette boucle semble participer au regroupement de boucles au site actif comme c'est le cas pour la boucle L5 $\alpha$  (Figure 4), aussi appelée boucle L4 (Figure 6), pour la caspase-7.



**Figure 13. Représentation de la structure tridimensionnelle et de la topologie de *AtMCA-IIc-C139A* d'après Zhu *et al.*, 2020. A.** Le domaine p20, p10 et le domaine de liaison entre les deux sont respectivement représentés en *bleu*, *orange* et *vert*. Les résidus catalytiques (His<sup>86</sup>, Cys<sup>139</sup>) ainsi que le résidu identifié comme site de clivage (Lys<sup>225</sup>) sont représentés en *rouge*. **B.** La représentation topologique de *AtMCA-IIc* respecte le code couleur en **A**.

### c) Mécanisme d'activation

Comme indiqué précédemment, les caspases, afin d'être actives, nécessitent une étape de dimérisation suivie par une maturation correspondant à la protéolyse de la région de liaison entre les sous-unités p20 et p10. L'identification des métacaspases a conduit à leur étude biochimique et enzymatique afin d'observer si le même phénomène se produisait. De nombreuses études témoignent d'un clivage autoprotéolytique pour les métacaspases de type II (*AtMCA-IIb*, *AtMCA-IIc*, *AtMCA-IIf* précédemment nommées *AtMC2b*, *AtMC2d*, *AtMC2f* respectivement) (Vercammen *et al.*, 2004 ; Watanabe and Lam, 2005) mais également pour les métacaspases de type I (*ScMCA-I* et *TbMCA-Ib*) (Moss *et al.*, 2007; Wong, Yan and Shi, 2012). Des études de séquençage ont permis d'identifier un ou plusieurs sites de clivage majeurs ainsi que des sites de clivages secondaires associés correspondant toujours à des résidus Arg ou Lys présents dans des boucles flexibles et entre les sous-unités p10 et p20 (Vercammen *et al.*, 2004 ; Watanabe and Lam, 2005, 2011 ; Moss *et al.*, 2007 ; Wong, Yan and Shi, 2012). L'étude structurale de ces enzymes a permis d'observer que le clivage aux sites identifiés, se situant sur des boucles en amont ou en aval des brins  $\beta$  du feuillet central, n'induisait pas une séparation des deux sous-unités p10 et p20 en conditions non-dénaturantes, puisque le feuillet  $\beta$  n'est pas déstructuré (Figure 10) (McLuskey *et al.*, 2012).

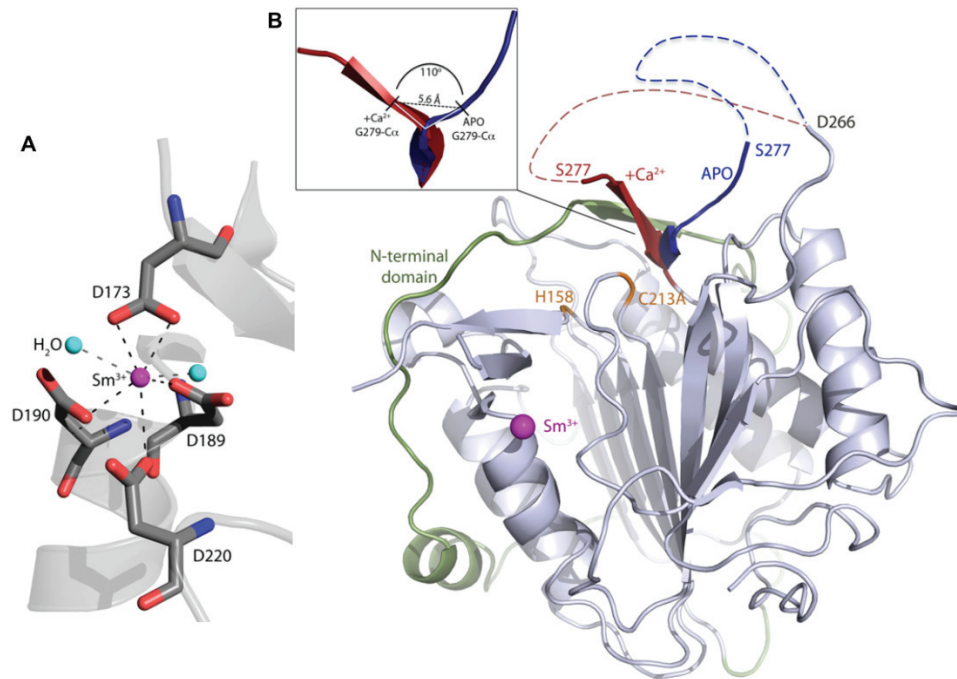
Alors que pour *AtMCA-IIc*, la mutation d'un résidu catalytique (Cys<sup>139</sup>) inhibe l'autoprotéolyse de l'enzyme et la rend inactive sur des substrats peptidiques constitués de Arg et/ou Lys (Vercammen *et al.*, 2004 ; Watanabe and Lam, 2011) ce n'est pas le cas pour

*ScMCA-I*. En effet, la mutation des résidus catalytiques (Cys<sup>276</sup> et His<sup>220</sup>) n'inhibe pas entièrement l'autoprotéolyse puisqu'un large fragment est toujours observable par SDS-PAGE (Wong, Yan and Shi, 2012). Les auteurs suggèrent que ce clivage serait attribué à l'action d'une autre protéase durant la phase de croissance cellulaire ou à l'utilisation de nombreuses protéases inhibitrices lors de la purification de *ScMCA-I*. Cependant chez *TbMCA-Ib*, bien que la mutation de deux sites de clivage empêche l'autoprotéolyse, elle n'altère pas l'activité enzymatique de l'enzyme permettant aux auteurs de conclure que le processus autoprotéolytique de cette enzyme n'est pas requis pour la mise en place de son activité (Moss *et al.*, 2007). Il a également été observé que *TbMCA-Ib* ne subissait pas d'autoprotéolyse *in vivo* et que ce phénomène visible *in vitro* serait une conséquence d'une forte concentration de l'enzyme et de la position des résidus basiques sur des boucles exposées au solvant (McLuskey *et al.*, 2012). Cependant, plus récemment, une autre équipe a identifié que le mutant ne serait actif que sur des substrats synthétiques et que seul l'enzyme native autoprotéolysée serait capable d'être active sur des larges substrats (Gilio *et al.*, 2017).

L'observation d'un mécanisme autoprotéolytique similaire aux caspases s'est également accompagné d'études permettant de discriminer entre un mécanisme d'activation en *cis* ou en *trans*. Une activation en *cis* impliquerait que l'enzyme soit capable de se maturer elle-même. Une activation en *trans* impliquerait l'intervention d'une seconde métacaspase active qui procéderait à la protéolyse de la première métacaspase. Chez *ScMCA-I*, après évaluation des distances entre les sites de clivage et les résidus catalytiques, il apparaît que le clivage aux résidus Lys<sup>331</sup> et Lys<sup>334</sup> pourrait se dérouler en *cis* alors que le clivage aux résidus Arg<sup>72</sup> et Lys<sup>86</sup>, plus lointain du site actif, serait en *trans* (Wong, Yan and Shi, 2012). Pour *AtMCA-IIId*, il a été observé que l'enzyme native était capable de protéolyser le mutant catalytique ne subissant pas d'autoprotéolyse, suggérant ainsi un mécanisme en *trans* (Watanabe and Lam, 2011).

Les études catalytiques des métacaspases ont également démontré un lien direct de dépendance au calcium pour la mise en place de leur activité protéolytique. Notamment, l'étude de *AtMCA-IIId* a montré que le calcium était spécifiquement requis *in vitro* pour que l'enzyme soit active puisqu'en absence de calcium aucune activité n'a été détectée sur un substrat spécifique des métacaspases (Boc-GRR-MCA) (Watanabe and Lam, 2011). De plus, la présence de calcium pour *AtMCA-IIb*, *AtMCA-IIId*, *ScMCA-I* et *TbMCA-Ib* accélère le processus d'autoprotéolyse témoignant d'une dépendance au calcium pour l'activation et l'autoprotéolyse de l'enzyme (Watanabe and Lam, 2005, 2011 ; McLuskey *et al.*, 2012 ; Wong, Yan and Shi, 2012). Chez les métacaspases de type II, la présence de calcium active

transitoirement l'enzyme en clivant la région de liaison entre les sous-unités p10 et p20 libérant ainsi l'accès au site actif (Zhu *et al.*, 2020). En revanche, une concentration élevée de calcium semblerait conduire à une protéolyse complète de l'enzyme et par conséquent à son inactivation (Watanabe and Lam, 2011). Cependant, les concentrations de calcium ayant permis de démontrer cette dépendance sont de l'ordre du millimolaire et sont largement supérieures aux concentrations physiologiques (Watanabe and Lam, 2005, 2011). L'étude de *TbMCA-Id* a également révélé que la présence de calcium était requise pour l'activité de l'enzyme et la liaison de substrat (McLuskey *et al.*, 2012). En effet, le trempage de cristaux de la protéine dans du calcium, bien que n'ayant pas permis de déterminer une structure en complexe avec cet ion, a permis d'identifier un mouvement structural de la boucle suivant le brin  $\beta 7$  (boucle 280) au site actif stabilisant l'interaction enzyme-substrat (Figure 14 B) (McLuskey *et al.*, 2012). En revanche, la détermination de la structure tridimensionnelle de *TbMCA-Ib* en complexe avec du Samarium ( $\text{Sm}^{3+}$ ), un substitut du calcium, a permis d'identifier les 4 résidus ( $\text{Asp}^{189}$ ,  $\text{Asp}^{190}$ ,  $\text{Asp}^{220}$  et  $\text{Asp}^{173}$ ) potentiellement impliqués dans la fixation de calcium (Figure 14 A). Ces résidus présents dans la sous-unité p20 sont conservés chez la métacaspase *ScMCA-I*. Ce site, à distance du site actif, serait considéré comme le site de haute affinité pour le calcium, fixant celui-ci de manière constante mais n'ayant pas un effet direct sur la conformation de *TbMCA-Ib* (Machado *et al.*, 2013). Il serait complété par un second site de faible affinité constitué par deux régions négativement chargées se trouvant dans la sous-unité p20 au niveau de la boucle précédent le brin  $\beta 7$  (Figure 10) et dans la sous-unité p10 (Machado *et al.*, 2013; Acosta-Maspons *et al.*, 2014; Fortin and Lam, 2018; Klemenčič and Funk, 2018). En comparaison, *AtMCA-II* n'ayant démontré aucune dépendance au calcium dans la mise en place de son activité catalytique, ne possède pas les acides aminés chargés négativement conservés dans la boucle flexible de la sous-unité p20 (Zhang and Lam, 2011; Zhu *et al.*, 2020).

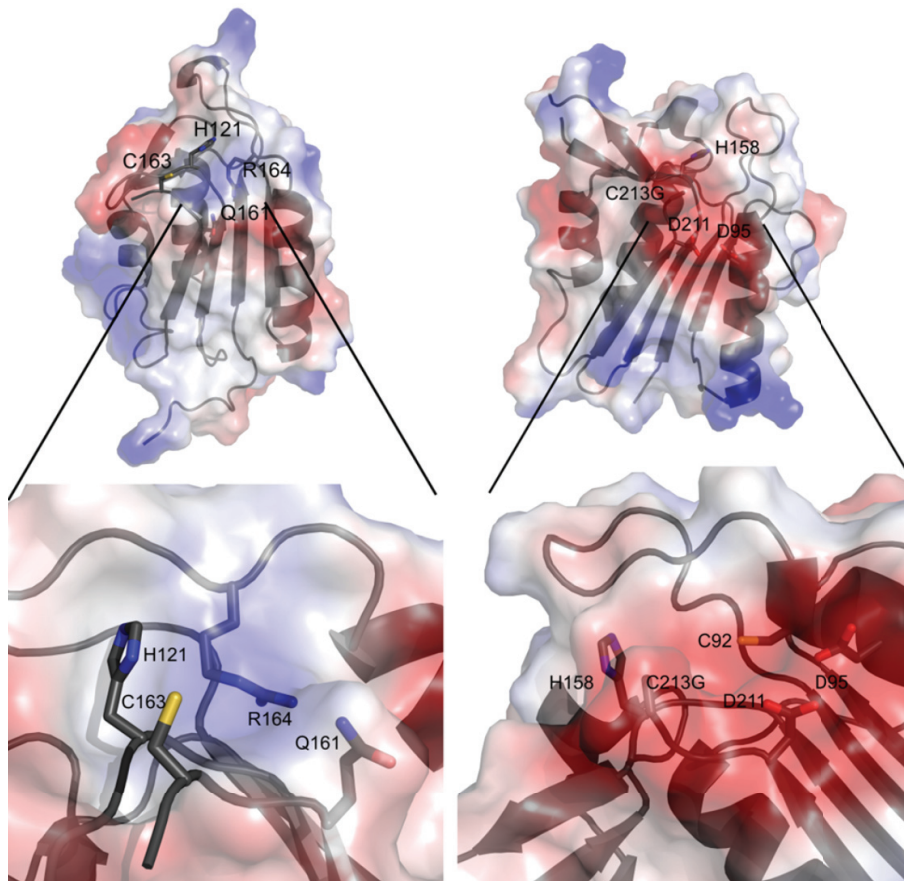


**Figure 14. Fixation de l'ion  $\text{Sm}^{3+}$  et changements de l'organisation structurale de *Tbmca-Ib-C213A* d'après McLuskey *et al.*, 2012. A. L'ion  $\text{Sm}^{3+}$  se lie à la surface de *Tbmca-Ib-C213A* via l'interaction de quatre résidus (Asp<sup>173</sup>, Asp<sup>189</sup>, Asp<sup>190</sup>, Asp<sup>220</sup>) et deux molécules d'eau (turquoise). B. La structure tridimensionnelle de *Tbmca-Ib-C213A* est représentée en *bleu clair* et le domaine N-terminale en *vert*. Les résidus catalytiques sont indiqués en *orange* et la boucle 280 est représentée en *bleu* pour la structure de l'enzyme Apo et en *rouge* pour la structure obtenue après trempage des cristaux dans du calcium. La rotation de la boucle 280, de  $110^\circ$ , est montrée dans l'encadré.**

#### d) Spécificité de clivage

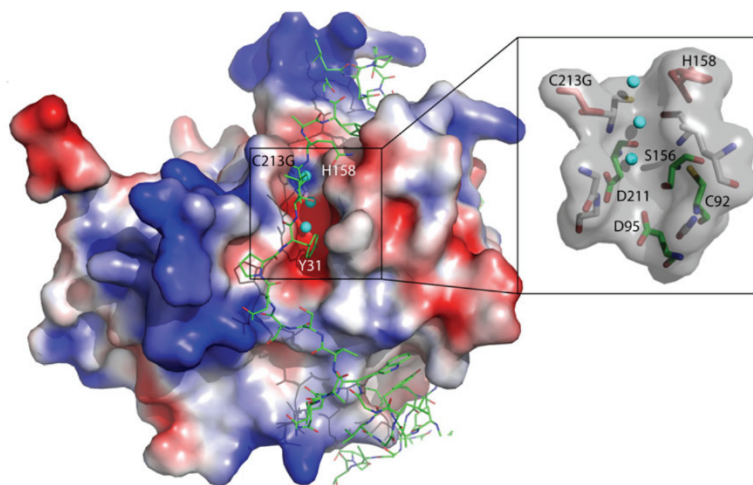
L'étude par séquençage N-terminale des produits de protéolyse de la métacaspase *AtMCIIf* a permis d'identifier les sites de coupures correspondant au résidu Arg<sup>183</sup>, au niveau du domaine C-terminale p10, ainsi qu'un résidu Lys libérant l'étiquette histidine de la protéine recombinante (Vercammen *et al.*, 2004). Ces résultats, au-delà de valider un processus protéolytique similaire aux caspases, ont permis de mettre en évidence la spécificité de résidus chargés positivement Arg/Lys en position P1 pour les métacaspases. De plus, *AtMCIId* et *AtMCIIf* n'ont démontré aucune activité protéolytique contre des substrats présentant la spécificité Asp de clivage des caspases (Ac-DEVD-AMC, Ac-YVAD-AMC, Z-VAD-AMC) (Vercammen *et al.*, 2004). La détermination de structures tridimensionnelles a notamment permis d'observer que les résidus catalytiques se trouvent dans une poche acide créant un environnement favorable à la liaison d'un résidu chargé positivement (Figure 15) (Klemenčič and Funk, 2018).





**Figure 15. Comparaison de la poche catalytique de la caspase-3 et de *TbMCA-Ib* d'après Klemenčič et Funk, 2018.** Représentation des potentiels électrostatiques de surface de la caspase-3 à gauche (PDB : 3GJT) et de *TbMCA-Ib* à droite (PDB : 4AF8). Le *bleu* indique les acides aminés basiques et le *rouge* indique les acides aminés acides.

Pour *TbMCA-Ib*, les résidus catalytiques (Cys<sup>213</sup> et His<sup>158</sup>) sont enfouis sous le domaine N-terminale qui prévient l'autoprotéolyse de l'enzyme ainsi que la fixation de substrat. Les résidus impliqués dans la fixation au site P1 sont les résidus Cys<sup>92</sup>, Asp<sup>95</sup>, Ser<sup>156</sup>, Asp<sup>211</sup> et Tyr<sup>31</sup>, et ils sont hautement conservés notamment chez d'autres métacaspases de *T. brucei* ainsi que chez *ScMCA-I* et *AtMCA-IIe* (McLuskey *et al.*, 2012). Les résidus Asp<sup>95</sup> et Ser<sup>156</sup> forment une liaison hydrogène avec le résidu Tyr<sup>31</sup> du domaine N-terminale et la mutation de ce dernier (Y31A) a montré un haut degré d'autoprotéolyse suggérant son implication dans l'accès au site actif (Figure 16). Un mécanisme suggéré est que Tyr<sup>31</sup> agirait comme le gardien du site actif, lors de l'activation de l'enzyme par le calcium, l'interaction avec les résidus du site P1 est rompue par l'autoclivage du domaine N-terminale entraînant un changement structural et l'accès de substrats au site actif (McLuskey *et al.*, 2012).



**Figure 16. Représentation du potentiel électrostatique de surface de *TbMCA-Id* d'après Klemenčič et Funk, 2018.** Les zones bleues et rouges représentent respectivement les zones positivement et négativement chargées. Le domaine N-terminale de l'enzyme est représenté en bâtonnets *verts* et un zoom du site P1 du site actif est présent sur la droite de la figure. Les acides aminés du site P1 sont représentés en bâtonnets, les molécules d'eau sont représentées par des sphères *bleu clair*.

e) *Rôle*

(1) *Mort cellulaire programmée*

Les métacaspases, en tant qu'homologues structuraux des caspases, ont d'abord été identifiées comme des effecteurs de la mort cellulaire chez les levures. En effet, Frank Madeo et son équipe avaient été les premiers à identifier un mécanisme de mort cellulaire dépendant d'une métacaspase en soumettant des souches, contenant des mutations du gène codant la protéine *ScMCA-I*, à un stress oxydatif (Madeo *et al.*, 2002). Lorsque la protéine est surproduite, la mort cellulaire est augmentée. En revanche, lorsque le gène de *ScMCA-I* est invalidé, les cellules ayant été soumises à un stress ont une survie plus élevée. D'autres expériences ont également démontré l'implication de *ScMCA-I* dans des mécanismes de mort cellulaire induits en réponses à des toxines virales (Ivanovska and Hardwick, 2005), au vieillissement cellulaire (Madeo *et al.*, 2002 ; Herker *et al.*, 2004) ainsi qu'à des défauts de processus tels que la réplication de l'ADN (Weinberger, 2005), la stabilité des ARN messagers (Mazzoni *et al.*, 2005) et l'ubiquitinylation des protéines (Walter, Matter and Fahrenkrog, 2010) (Tableau 2). Chez *C. albicans*, il a été démontré que *CaMCA-I* était également impliquée dans la mise en place de la mort cellulaire lorsque les levures étaient soumises à un stress oxydant (Cao *et al.*, 2009). De plus, le farnésol, un composé sécrété par les levures lorsque la densité cellulaire et la compétition pour les nutriments augmentent, a été identifié comme induisant la mort cellulaire notamment en activant l'expression de *CaMCA-I* (Shirtliff *et al.*, 2009 ; Léger *et al.*, 2015).



<i>Organismes</i>	<i>Nb. De gènes</i>	<i>Protéine(s)</i>	<i>Méthode</i>	<i>Fonction</i>	<i>Références</i>
<i>Candida albicans</i>	1	CaMCA-I	KO	Requise pour l'induction de la mort cellulaire suite à un stress oxydatif	Cao <i>et al.</i> , 2009
<i>Saccharomyces cerevisiae</i>	1	ScMCA-I	KO, OE	Requise pour l'induction de la mort cellulaire suite à un stress oxydatif	Madeo <i>et al.</i> , 2002
			KO, OE	Requise pour la mort cellulaire associée à des défauts d'ubiquitylation, de réplication d'ADN, de stabilité d'ARNm	Walter <i>et al.</i> , 2010 ; Weinberger <i>et al.</i> , 2005 ; Mazzoni <i>et al.</i> , 2005
			KO	Requise pour la mort cellulaire induite par des toxines virales	Ivanovska et Hardwick, 2005
			KO	Requise pour la mort cellulaire associée au vieillissement cellulaire	Madeo <i>et al.</i> , 2002 ; Herker <i>et al.</i> , 2004
			KO, IM	Requise pour l'élimination d'agrégats protéiques	Lee <i>et al.</i> , 2010
			KO, IM	Requise pour la régulation du cycle cellulaire	Lee <i>et al.</i> , 2008

**Tableau 2. Fonctions étudiées des métacaspases chez *C. albicans* et *S. cerevisiae* d'après Tsiatiani *et al.*, 2011.** L'identification des fonctions exercées par CaMCA-I et ScMCA-I est indiquée ainsi que la méthode expérimentale utilisée, l'inactivation du gène codant la protéine (KO ; knock out), la surexpression du gène (OE ; over expression) ou un encore un mutant inactif de la protéine (IM ; inactive mutant).

Bien que l'implication des métacaspases de levure dans la mort cellulaire ne semble plus à prouver, ScMCA-I ne serait responsable de la mise en place de ce processus que dans 40 % des cas (Madeo *et al.*, 2009). En effet, la protéine ne serait pas impliquée lorsque la mort cellulaire est induite par exposition des cellules à l'acide formique (Du *et al.*, 2008) ou le cuivre (Liang and Zhou, 2007). De plus, comme mentionné précédemment les effecteurs AIF1 et NUC1 peuvent propager le signal de mort cellulaire indépendamment de l'action de la métacaspase (Madeo *et al.*, 2009).

Il est aussi important de mentionner que de nombreuses études d'activité catalytique des métacaspases ont été réalisées avec des substrats synthétiques spécifiques des caspases (YVAD-AMC ou DEVD-AMC). Ces derniers étant composé d'un Asp, en position P1, ne correspondant pas à la spécificité Arg/Lys de substrat des métacaspases (Tsiatiani *et al.*,

2011). Il est donc envisageable que l'implication de ces protéases dans les processus de mort cellulaire soit plus importante que ce qui a été rapporté dans la littérature.

## (2) Multifonctionnalité

Outre les phénomènes de mort cellulaire, *ScMCA-I* a été identifiée comme participant à d'autres phénomènes tel que la régulation du cycle cellulaire en intervenant lors de la transition entre la phase G1 et S (Lee *et al.*, 2008) (Tableau 2). *ScMCA-I* serait également impliquée dans le recrutement d'agrégats protéiques au protéasome et interviendrait donc dans le contrôle qualité des protéines (Lee *et al.*, 2010). Il a été suggéré que l'enzyme pourrait exercer son activité protéolytique directement sur les agrégats protéiques afin de les pré-cliver ou stabiliserait l'interaction entre les agrégats et les protéines impliquées dans leur désassemblage (Lee *et al.*, 2010; Hill and Nyström, 2015). D'après Hill et Nyström, la mort cellulaire dépendante des métacaspases ne serait déclenchée que lorsque cela représenterait un bénéfice pour la population de levure et cette transition serait opérée par le calcium agissant comme un interrupteur moléculaire. En effet, le calcium étant un activateur de l'autoclivage et de l'activité protéolytique des métacaspases, il pourrait, en présence de signaux comme une densité cellulaire trop forte ou un manque de nutriments, activer la mort cellulaire de manière semblable au quorum sensing chez les bactéries (Hill and Nyström, 2015). Ce rôle multifonctionnel chez les levures pourrait notamment être attribué à la présence d'un seul gène codant pour une métacaspase comme chez *S. cerevisiae* et *C. glabrata*.

### f) Substrats

Bien que la connaissance de substrats naturels *in vivo* de métacaspase soit encore méconnu, il a été démontré que *ScMCA-I* était capable de cliver *in vitro* la GAPDH (Glyceraldéhyde-3-P-déshydrogénase), une enzyme de la glycolyse (Silva *et al.*, 2011). Il a également été démontré *in vitro* que des substrats de *CaMCA-I* étaient les protéines « heat shock », responsables de l'adressage des agrégats protéiques au protéasome (Léger *et al.*, 2015). De manière intéressante, il semblerait que *ScMCA-I* ne clive pas ces protéines mais possède un rôle dans le recrutement ou l'interaction avec la machinerie cellulaire du contrôle qualité des protéines (Lee *et al.*, 2010). Dans le cas de *CaMCA-I*, les résultats suggèreraient que la protéine réduirait la capacité des cellules stressées à répondre aux altérations protéiques. Lors de cette même étude, les auteurs ont également montré que la protéine *cdc48p* était un substrat potentiel de *CaMCA-I*. La mutation du gène codant cette protéine avait initialement permis d'identifier des changements morphologiques typiques d'un processus apoptotique chez *S. cerevisiae* (Léger *et al.*, 2015).

## II. Résultats

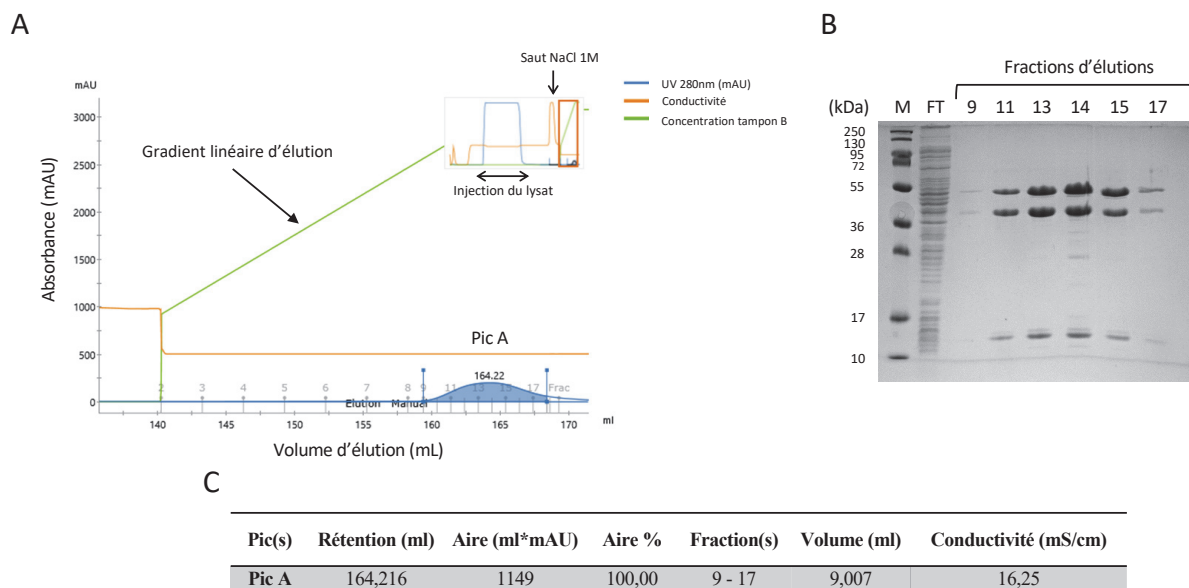
### A. Maturation dépendante du calcium de CgMCA-I

#### 1. *Purification de CgMCA-I*

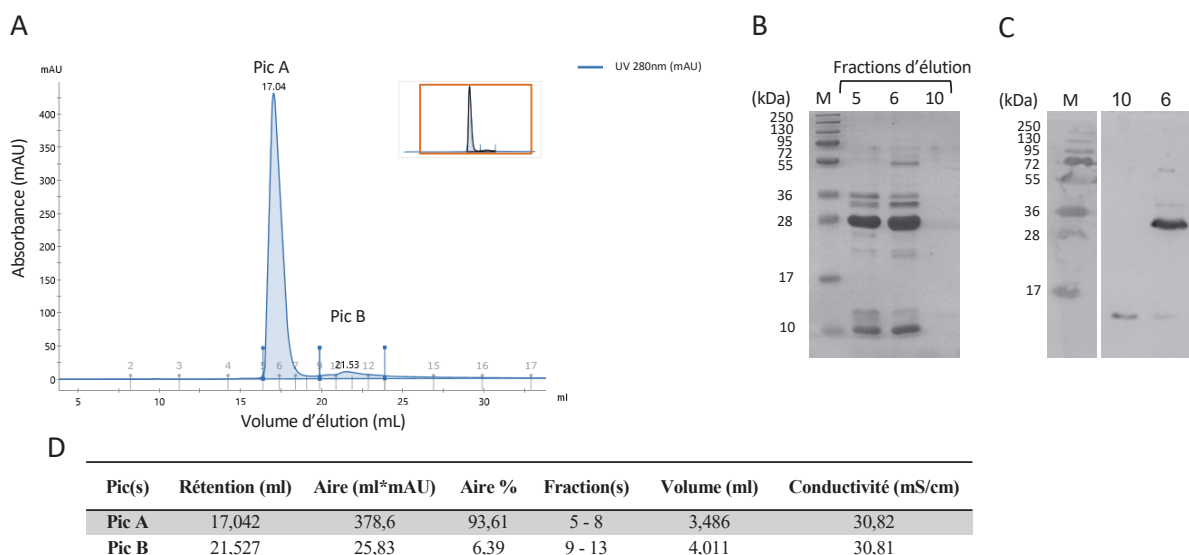
Après suproduction de la protéine sauvage, CgMCA-I WT fusionnée à une étiquette histidine (48 kDa), des mutants ou des différentes constructions génétiques (cf. Matériels et méthodes p 126-129), les bactéries sont lysées et les protéines purifiées par une première étape de chromatographie d'affinité pour le cobalt (Figure 17).

Les fractions d'élution sont analysées par SDS-PAGE afin de confirmer la présence de la protéine d'intérêt (Figure 17 B). On observe la présence d'une bande d'environ 48 kDa correspondant à la masse théorique de CgMCA-I ainsi que de nombreuses autres bandes de poids moléculaires inférieurs. Afin d'obtenir un échantillon très pur pour les prochaines étapes d'étude biochimique et de biologie structurale, les fractions sont collectées et concentrées afin d'être injectées sur une colonne de chromatographie d'exclusion stérique (Figure 18).

Le chromatogramme d'élution permet d'observer la présence de deux pics, un pic A dont le volume d'élution correspond à une masse apparente d'environ 38 kDa et un second pic B correspondant à une masse apparente d'environ 10 kDa. Les fractions d'élution 5 et 6 du pic A et la fraction 10 du pic B sont analysés par SDS-PAGE (Figure 18 B). À nouveau, on observe de nombreuses bandes mais également une absence de la bande à 48 kDa suggérant que CgMCA-I pourrait être soumise à un processus de maturation. La fraction 10, quant à elle, ne montre aucune bande apparente. En revanche, l'analyse de cette fraction par western blot, à l'aide d'un anticorps dirigé contre l'étiquette histidine, révèle la présence d'une faible bande d'environ 10 kDa, correspondant à un fragment contenant l'extrémité N-terminale, issu de la protéolyse de CgMCA-I.



**Figure 17. Purification caractéristique de CgMCA-I par chromatographie d'affinité au cobalt. A.** Chromatogramme d'élution de CgMCA-I indiquant l'absorbance à 280 nm ( $UV_{280\text{ nm}}$ ) en fonction du volume d'élution (mL). L'encadré en haut à droite montre le chromatogramme entier où l'on peut observer l'injection du lysat sur la colonne suivie par un lavage à forte concentration ionique (saut NaCl 1M). L'élution est réalisée par un gradient linéaire de 0 à 100 % du tampon d'élution et suivi par une étape à 100 % de tampon d'élution. **B.** Les fractions 9 à 17, du pic A d'élution (A, C) ont été déposées sur un gel SDS-PAGE 15 %. **C.** Caractéristiques du pic A d'élution.

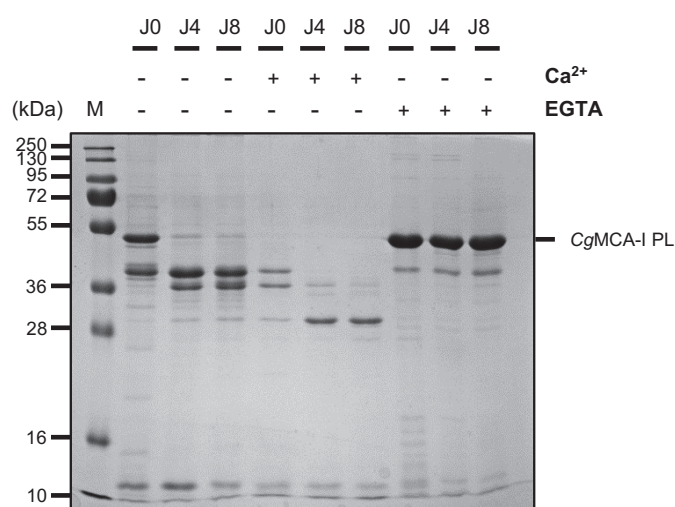


**Figure 18. Purification par chromatographie d'exclusion stérique de CgMCA-I. A.** Chromatogramme d'élution de CgMCA-I indiquant l'absorbance à 280 nm ( $UV_{280\text{ nm}}$ ) en fonction du volume d'élution (mL). L'élution est réalisée de manière isocratique avec le tampon de dessalage **B.** Les fractions d'élution 5, 6 et 10 ont été analysées par SDS-PAGE 15 %. **C.** Western-blot des fractions 6 et 10 avec un anticorps anti-histidine dirigé contre l'étiquette N-terminale de CgMCA-I. **D.** Caractéristiques du pic A et B d'élution.

## 2. Effet des cations divalents sur la maturation

Le calcium ayant été montré comme étant requis ou accélérant le processus de maturation des métacaspases (Watanabe and Lam, 2011 ; McLuskey *et al.*, 2012 ; Wong, Yan and Shi, 2012 ; Zhu *et al.*, 2020), nous avons réalisé des purifications de CgMCA-I en présence ou absence de calcium ( $\text{Ca}^{2+}$ ) ou d'EGTA, un agent chélatant les ions calcium, et nous avons prélevé des aliquots aux jours 0, 4 et 8 après purification (Figure 19).

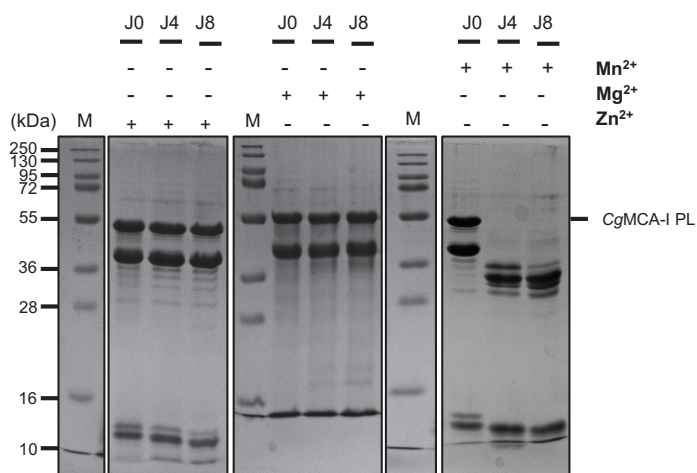
En absence de calcium ou d'EGTA, CgMCA-I semble subir un clivage protéolytique se poursuivant sur plusieurs jours *in vitro*, sans atteindre une forme de maturation complète au jour 8. Ce clivage protéolytique s'observe par l'apparition de nombreuses bandes intermédiaires entre la bande à 48 kDa et celle à 10 kDa. La maturation en présence de calcium est largement accélérée puisque la protéine pleine longueur (CgMCA-I PL) n'est plus visible au jour 0 de maturation. De plus, au jour 8, une fine bande d'environ 23 kDa apparaît, pouvant correspondre au domaine p20 caractéristique. En présence d'EGTA, le processus semble être arrêté dès le jour 0, confirmant l'importance du rôle du calcium. En revanche, une seconde bande est observée à environ 38 kDa indiquant que la maturation aurait pu débiter dans les cellules avant l'ajout d'EGTA dans les tampons de lyse et de purification.



**Figure 19. Maturation de CgMCA-I en présence ou absence de  $\text{Ca}^{2+}$  ou d'EGTA.** SDS-PAGE des aliquots de CgMCA-I aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en présence ou absence de  $\text{Ca}^{2+}$  (10 mM) ou d'EGTA (1 mM). La bande à 48 kDa correspond à la protéine pleine longueur (PL) et les bandes inférieures ont été attribuées à des séquences de la protéine.

Afin d'évaluer, si le calcium est le seul ion divalent capable d'agir sur la maturation de CgMCA-I, la protéine a également été purifiée en présence de  $\text{Mg}^{2+}$ ,  $\text{Mn}^{2+}$  et  $\text{Zn}^{2+}$  (Figure 20).

En présence de  $Mn^{2+}$ , la maturation de CgMCA-I est similaire à celle observée en absence d'ion divalent ou d'EGTA. Cependant, en présence de  $Mg^{2+}$  ou  $Zn^{2+}$ , la maturation est arrêtée après le premier clivage, de la même manière qu'en présence d'EGTA, suggérant un rôle inhibiteur de ces ions divalents.

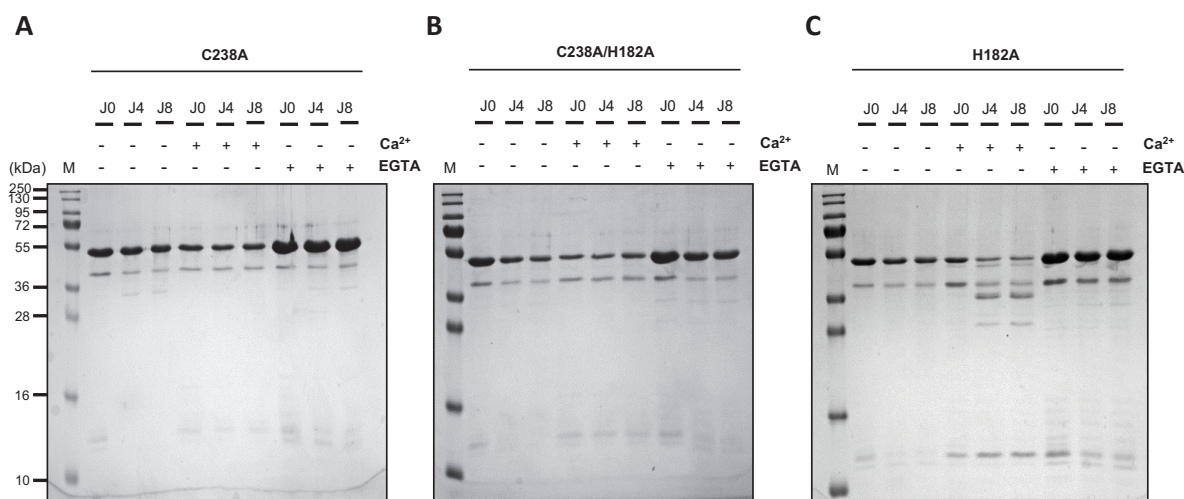


**Figure 20. Maturation de CgMCA-I en présence de  $Mn^{2+}$ ,  $Mg^{2+}$ ,  $Zn^{2+}$ .** SDS-PAGE des aliquots de CgMCA-I aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en présence de  $Mn^{2+}$  (10 mM),  $Mg^{2+}$  (10 mM) ou  $Zn^{2+}$  (10 mM).

## B. Caractérisation structurale de CgMCA-I

Afin d'identifier l'effet du calcium sur l'activation de CgMCA-I, des essais de cristallogénèse ont été entrepris avec la protéine purifiée en absence ou en présence d'additifs (10 mM  $Ca^{2+}$  ou 1 mM EGTA). Ceci dans le but de résoudre les structures cristallines de l'enzyme sous une forme non maturée, de maturation intermédiaire et totalement maturée.

En parallèle, des expériences de mutagenèse dirigée ont été réalisées afin d'obtenir les simples et doubles mutants des résidus catalytiques, C238A, H182A et C238A/H182A. Ces mutants ont été purifiés, analysés par SDS-PAGE (Figure 21) et ont été employés pour des expériences de cristallogénèse afin de résoudre la structure de l'enzyme non maturée. En revanche, de manière surprenante, aucun de ces mutants n'a permis une inhibition complète de la maturation de CgMCA-I. De plus, en présence de l'ion divalent calcium, le mutant H182A semble pouvoir se maturer partiellement, suggérant qu'un autre accepteur de proton dans l'environnement du site actif puisse remplacer son rôle de donneur de protons.



**Figure 21. Maturation des mutants C238A, C238A/H182A et H182A en présence ou en absence de  $\text{Ca}^{2+}$  ou d'EGTA.** SDS-PAGE des aliquots des mutants C238A (A), C238A/H182A (B) et H182A (C) de CgMCA-I aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en présence ou absence de  $\text{Ca}^{2+}$  (10 mM) ou d'EGTA (1 mM).

Environ 3000 conditions de cristallisation ont été testées (voir Matériel et Méthodes) avec des concentrations différentes de protéines et seules deux conditions ont permis d'obtenir des cristaux de l'enzyme CgMCA-I WT purifiée en présence de  $\text{Ca}^{2+}$  et employée pour la cristallogénèse au jour de maturation 0 et 4. Les cristaux de la protéine au jour de maturation 0 ont été obtenus dans la condition  $(\text{NH}_4)_2\text{SO}_4$  0,2 M, PEG 4000 30 % (*m/v*), et ceux de la protéine au jour de maturation 4 ont été obtenus dans la condition Acétate de magnésium 0,2 M, PEG 3350 20 % (*m/v*).

Deux structures tridimensionnelles ont pu être résolues, la structure de CgMCA-I<sup>Ca</sup> montrant une densité électronique pour un ion  $\text{Ca}^{2+}$  et la structure de CgMCA-I<sup>Mg</sup> montrant une densité électronique pour un ion  $\text{Mg}^{2+}$ . La résolution des structures a été réalisée en employant la technique du remplacement moléculaire avec comme modèle guide, pour le jeu de données de CgMCA-I<sup>Ca</sup>, la structure de la métacaspase de *S. cerevisiae* (ScMCA-I, code PDB : 4F6O) partageant un pourcentage d'identité de séquence de 66,5 %. La structure de CgMCA-I<sup>Mg</sup>, quant à elle, a été résolue par remplacement moléculaire en employant la structure de CgMCA-I<sup>Ca</sup> comme modèle guide. Les statistiques des deux jeux de données sont résumées dans le Tableau 3.



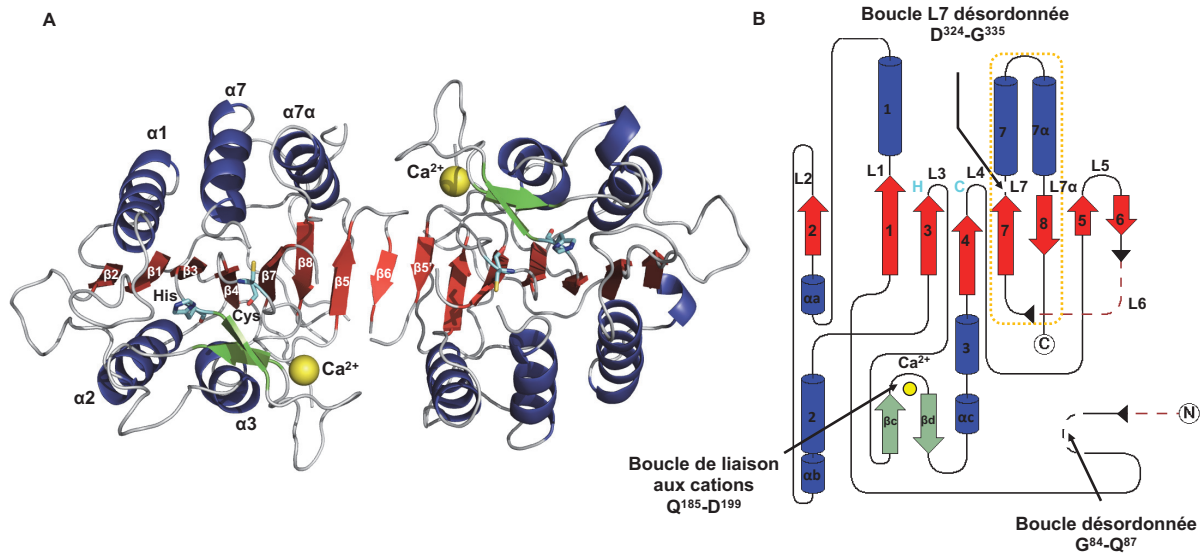
Structure	<i>Cg</i> MCA-I <sup>Ca</sup>	<i>Cg</i> MCA-I <sup>Mg</sup>
Collection des données		
Ligne de lumière	PXIII	PROXIMA1
Longueur d'onde (Å)	0,99999	0,97857
Groupe d'espace	<i>P32</i>	<i>C2</i>
Dimensions		
a, b, c (Å)	97,37 97,37 54,60	957 83,6 87,83
α, β, γ (°)	90,0 90,0 120,0	90,0 106,6 90,0
Gamme de résolution (Å)	40,0 - 3,0	30,0 - 1,6
Nb. de réflexions totales	123226	303345
Nb. de réflexions uniques	11590	87056
R <sub>meas</sub> (%)	9,9 (101,1)	10,3 (16,9)
CC <sub>1/2</sub> (%)	99,9 (82,3)	99,7 (99,2)
I/σ(I)	20,5 (2,5)	8,44(6,44)
Multiplicité	10,6 (10,7)	3,5 (3,5)
Complétude (%)	100 (100)	99,6 (99,7)
Nb. mol. /unit. asymm.	2	2
Affinement		
R <sub>work</sub> /R <sub>free</sub> (%)	24,94 / 27,59	17,53 / 20,23
Nb. atomes		
Protéine	4014	4111
Ligand/ion	10	36
Eau	16	627
Facteur B moyen (Å <sup>2</sup> )		
Protéine	83,9	25,6
Ligand/ion	65,5	50,2
Eau	83	39,6
r.m.s.d.		
longueur des liaisons (Å)	0,002	0,017
Angles (°)	0,654	1,618
Ramachandran		
Favorisé (%)	93,3	98,8
Autorisé (%)	5,7	1,2
Abbérations (%)	1	0

Tableau 3. Statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de *Cg*MCA-I.

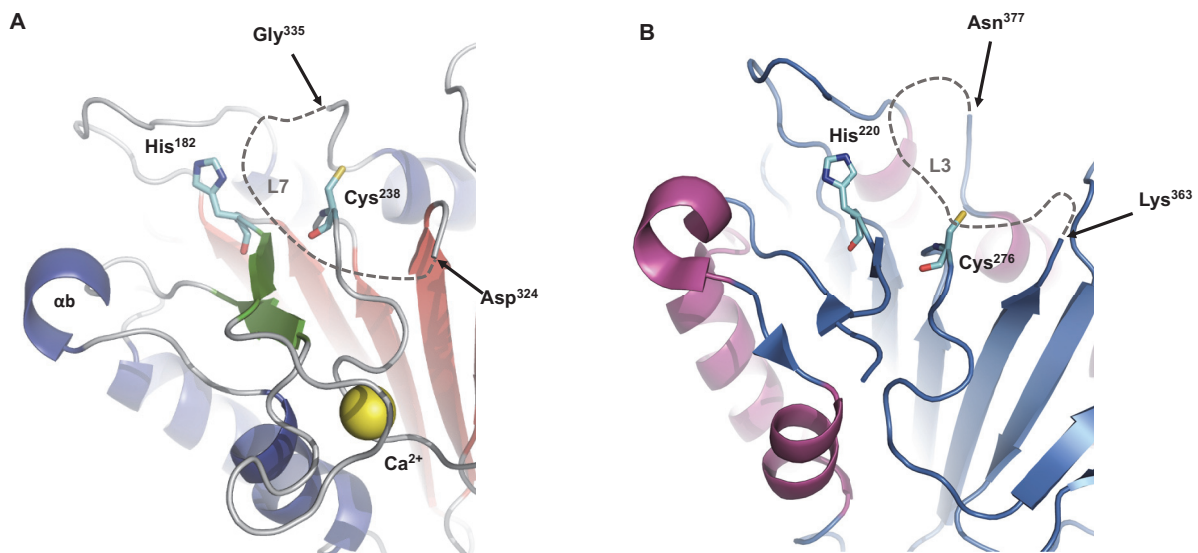


## 1. Structure de CgMCA-I<sup>Ca</sup>

La structure de CgMCA-I<sup>Ca</sup> a été déterminée à partir de protéines ayant cristallisé dans le groupe d'espace P32 avec deux molécules dans l'unité asymétrique (Figure 22 A) et a été déterminée à 3,0 Å de résolution. Les deux molécules sont composées d'un domaine p20 et d'un domaine p10 caractéristiques des caspases, avec deux brins  $\beta$  antiparallèles formant la sous-unité p10 et complétant les six brins  $\beta$  parallèles formant le feuillet  $\beta$  de la sous-unité p20 (Figure 22 A, B). Le feuillet central est entouré par, d'un côté trois hélices  $\alpha$  ( $\alpha 1$ ,  $\alpha 7$  et  $\alpha 7\alpha$ ) et de l'autre côté deux hélices  $\alpha$  ( $\alpha 2$  et  $\alpha 3$ ). De plus, une épingle  $\beta$ , composée des brins  $\beta c$  et  $\beta d$ , est présente entre les brins  $\beta 3$  et  $\beta 4$  ainsi que trois courtes hélices  $\alpha$  ( $\alpha a$ ,  $\alpha b$ ,  $\alpha c$ ) complétant le domaine p20. Les deux résidus catalytiques de la diade sont portés par les boucles CgL4 et CgL3 (Figure 22 B). Comme observé chez ScMCA-I (Wong, Yan and Shi, 2012), la boucle CgL7 (nommée L3 pour ScMCA-I), qui vient se positionner au-dessus du site actif, est désordonnée du résidu Asp<sup>324</sup> au résidu Gly<sup>335</sup> (Figure 23 A, B). Cependant, contrairement à ScMCA-I, où cette zone forme une courte boucle, la région Gly<sup>84</sup>-Gln<sup>87</sup> est désordonnée (Figure 22 B).

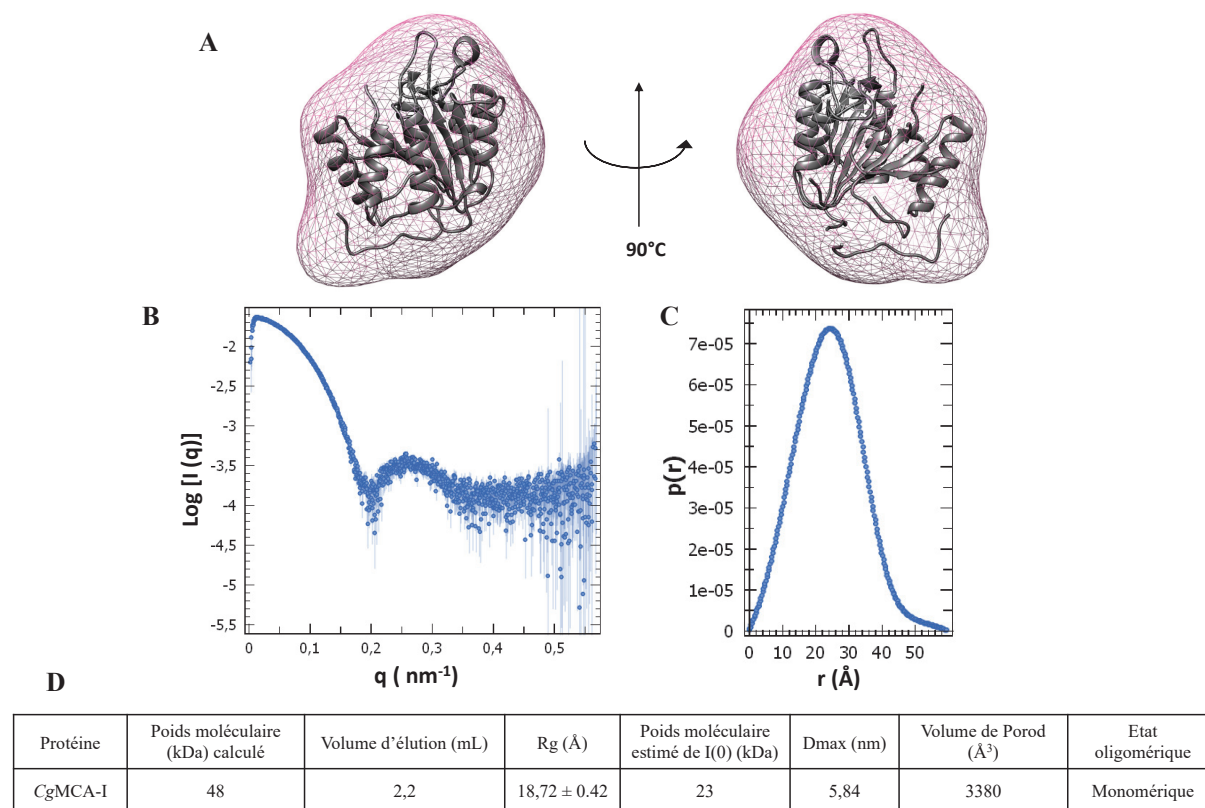


**Figure 22. Organisation structurale de *CgMCA-I<sup>Ca</sup>*.** A. Structure tridimensionnelle des deux monomères de *CgMCA-I<sup>Ca</sup>*, présents dans l'unité asymétrique cristalline, en complexe avec un ion calcium ( $\text{Ca}^{2+}$ ) représenté par une sphère *jaune*. Les hélices (*bleues*) et les brins (*rouges*) sont respectivement numérotés en *noir* et en *blanc*. Les deux brins  $\beta$  courts ( $\beta\text{c}$  et  $\beta\text{d}$ ) sont représentés en *vert*. Les résidus catalytiques, His<sup>182</sup> et Cys<sup>238</sup>, sont représentés en *cyan*. B. Représentation topologique de *CgMCA-I<sup>Ca</sup>*, les boucles sont représentées en *noir* et numérotées (L1 à L7/L7 $\alpha$ ) et les zones désordonnées sont en pointillés *noir* et les zones manquantes dues au processus de maturation sont en *rouge foncé*. Les sites de clivage sont indiqués par des flèches *noires* et la sous-unité p10 est représentée en pointillés *orange*.



**Figure 23. Organisation structurale du site catalytique de *CgMCA-I<sup>Ca</sup>* et de *ScMCA-I*.** A, B. Les résidus catalytiques de *CgMCA-I* (Cys<sup>238</sup> et His<sup>182</sup>) (A) et de *ScMCA-I* (Cys<sup>276</sup> et His<sup>220</sup>) (B) sont représentés en *cyan* et l'ion calcium ( $\text{Ca}^{2+}$ ) de *CgMCA-I* est représenté en sphère *jaune* (A). Les boucles désordonnées L7 (Gly<sup>335</sup>-Asp<sup>324</sup>) de *CgMCA-I* (A) et L3 (Lys<sup>363</sup>-Asn<sup>377</sup>) de *ScMCA-I* (B) sont indiquées en pointillés *gris*.

L'organisation du dimère cristallin ressemblant à l'organisation dimérique des caspases, des expériences de SAXS ont été réalisées afin de déterminer l'organisation oligomérique de CgMCA-I en solution. Il a été ainsi possible de confirmer son état monomérique (Figure 24).



**Figure 24. Etude de la diffusion des rayons X aux petits angles (SAXS) de CgMCA-I.** A. Le modèle *ab initio* de CgMCA-I a été généré par SEC-SAXS en utilisant le programme DAMMIF (Franke and Svergun, 2009). La structure cristalline de CgMCA-I a été superposée à l'enveloppe SAXS visualisée (18 Å) avec le logiciel CHIMERA (University of California, San Francisco) et confirme l'état oligomérique de la protéine. B. La courbe de diffusion, Log I(q) en fonction de q, a été générée en utilisant PRIMUS (Konarev *et al.*, 2003). C. La courbe P(r) en fonction de r, a été générée à partir des données de (B) en utilisant CRY SOL (Svergun, Barberato and Koch, 1995). D. Données SAXS de l'analyse de CgMCA-I.

D'après la structure de CgMCA-I<sup>Ca</sup>, la purification de la protéine en présence de calcium, a permis la maturation de cette dernière. En effet, le domaine N-terminale (Met<sup>1</sup>-Gly<sup>68</sup>) ainsi que la boucle CgL6 (Asn<sup>260</sup>-Ile<sup>309</sup>), correspondant au domaine de liaison entre la sous-unité p20 et p10, ne montrent aucune densité électronique (Figure 22 B). L'identification précise des résidus ciblés par le clivage n'est pas possible structuralement, ce qui est dû au fait que ces zones de coupures sont situées dans des boucles flexibles. Ainsi, afin de déterminer ces résidus, des études de spectrométrie de masse ont été réalisées en collaboration avec la plateforme scientifique PSF.

La protéine maturée en présence de calcium a été injectée sur une colonne C4 en conditions dénaturantes, permettant la séparation des fragments clivés en fonction de leur hydrophobicité. Cette colonne est couplée à un spectromètre de masse qui permet, après déconvolution des spectres MS, l'identification des extrémités de chaque fragment avec une précision inférieure au Dalton (Figure 25). Les résultats, ont permis l'identification de trois sites de clivages situés aux résidus Arg<sup>307</sup>, Lys<sup>263</sup> et Arg<sup>54</sup> permettant, respectivement, le clivage à l'extrémité C-terminale du domaine de liaison, de la sous-unité p20 et du domaine N-terminale (Figure 25 B).

**A**

Masse monoisotopique (Da)	Masse théorique (Da)	$\Delta$ ppm	Séquence peptidique	Position
6470,721	6470,724	0,003	(R)GSHMYPGSGNYSYNNRPSMPPPGFNGDGQGYRQYEGNQYGGGYQQQ QYQDQYQGENR	17 - 74
9364,673	9364,685	1,2	(R)QIKFSPADVIMLSGSKDNQTSADTFADGQNIAMSHAFISVMTRQPQQSY LSQLNLRNELAGKYSQKPLSASHPIDVNLQFIM	328 - 412
23590,42	23590,43	0,2	(R)GQYQGQYQDQPEYGRPPSGMVRPPSSIQQNGGQQFQYSQMTGRRKAL LIGINYIGSKNALRGCINDAHNIFNYLITTYCGYRPEDIVMLTDDQREMVKIPLK ENIIRAMQWLKDAQPNDALFFHYSGHGGQTKDLDGDEEDGMDDVIYPVD FESVGPLIDDTMHDIMVKSPLPQGARLTALFDSCHSGTVLDPYTYSTKGVKEP NMWK	75 - 283

**B**

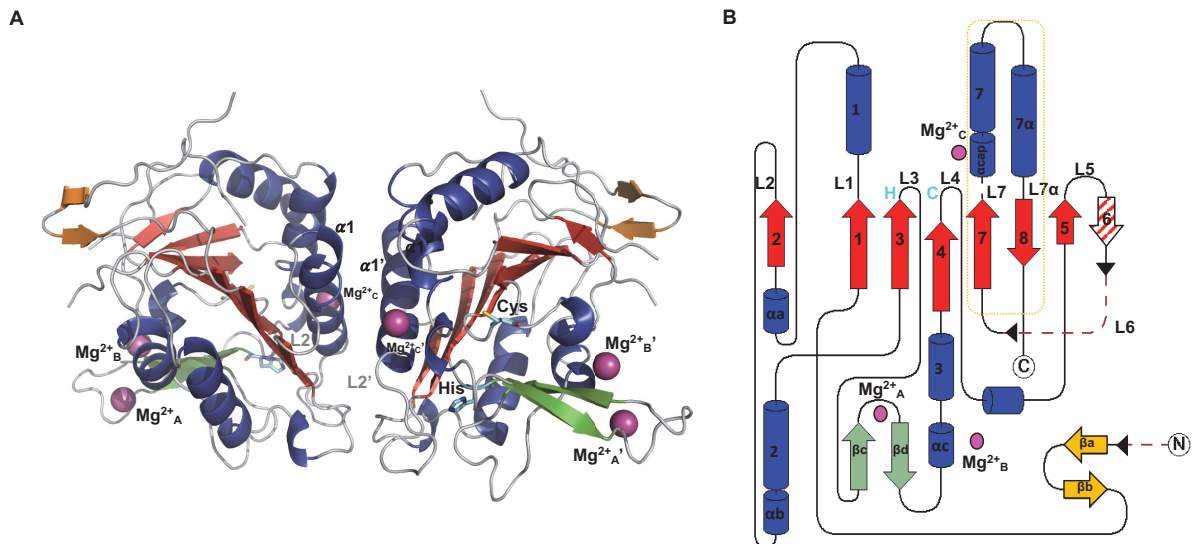
10	20	30	40	50	60	Nter – M1 – R54 (6,2 kDa)
MYPGSGNYSY	NNRPSMPPPG	FNGDGQGYRQ	EYGNQYGGGY	QQQYQDQYQ	GEYRQYQGGY	
70	80	90	100	110	120	p20 - G75 – K263 (23,6 kDa)
YQDQPEYGRP	PSGMVRPPSS	IQQNGGQQFQ	YSQMTGRRKA	LLIGINYIGS	KNALRGCIND	
130	140	150	160	170	180	
AHNIFNYLTT	YCGYRPEDIV	MLTDDQREMV	KIPLKENIIR	AMQWLKDAQ	PNDALFFHYS	
190	200	210	220	230	240	D264 - R307 (5,87 kDa)
GHGGQTKDLD	GDEEDGMDDV	IYPVDFESVG	PLIDDTMHD	MVKSPLPQGAR	LTALFDSCHS	
250	260	270	280	290	300	(p10) - Q308 – M392 (9,37kDa)
GTVLDPYTY	STKGVKEPN	MVKIVGSDGI	QAAMAYATGN	RSALFSSIGN	MVSSVTKKQN	
310	320	330	340	350	360	
VDRERVRIK	FSPADVIMLS	GSKDNQTSAD	TFADGQNIAM	MSHAFISVMT	RQPQQSYLSL	
370	380	390				
LQNLRLNELAG	KYSQKPLSA	SHPIDVNLQF	IM			

**Figure 25. Identification des sites de clivages et des séquences peptidiques associées par spectrométrie de masse.** **A.** Les séquences peptidiques majoritaires identifiées en spectrométrie de masse ont permis de mettre en évidence trois sites de clivage. Les masses théoriques des peptides identifiés ont également été calculées en utilisant l'outil Findpept (Expasy) ainsi que l'écart entre les valeurs expérimentales et théoriques ( $\Delta$ ppm). La numérotation des résidus prend en compte les 20 résidus de l'étiquette histidine. **B.** La région N-terminale et les sous-unités p20 et p10 sont respectivement indiquées en *bleu*, *jaune* et *vert*. Les sites de clivage sont mis en évidence par des ronds *rouges*.

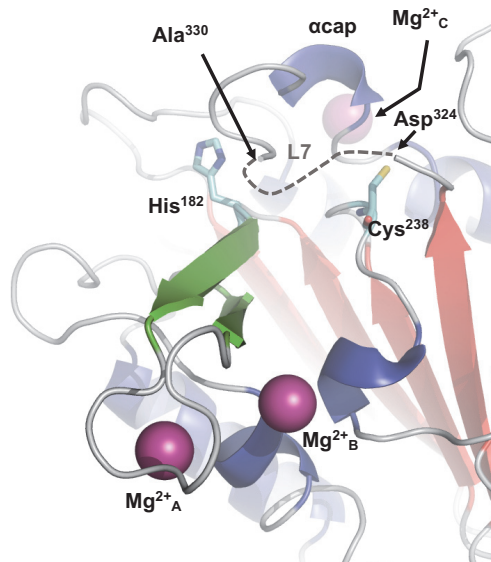
## 2. Structure de CgMCA-I<sup>Mg</sup>

La structure de CgMCA-I<sup>Mg</sup> a été déterminée à partir de protéines ayant cristallisé dans le groupe d'espace C2, à 1,6 Å de résolution et avec également deux molécules dans l'unité asymétrique (Figure 26 A). En comparaison à CgMCA-I<sup>Ca</sup>, il n'y a pas de complémentation du brin  $\beta_6$  pour stabiliser le dimère cristallin. Le contact se fait à la place, entre l'hélice  $\alpha_1$  et la boucle CgL2.

Globalement, la structure de CgMCA-I<sup>Mg</sup> est très similaire à celle de CgMCA-I<sup>Ca</sup>, avec un RMSD de 0,823 Å, calculé à partir des distances entre carbones  $\alpha$ . De manière similaire, la protéine semble être complètement maturée, on retrouve l'absence du domaine N-terminale (Met<sup>1</sup>-Ser<sup>72</sup>), et du domaine de liaison entre la sous-unité p20 et p10 (Asn<sup>360</sup>-Ile<sup>309</sup>) (Figure 26 B, 27). En revanche, la zone Gly<sup>84</sup> à Gln<sup>87</sup> est résolue et structurée en une épingle à cheveux composée de deux courts brins  $\beta$  ( $\beta_a$  et  $\beta_c$ ) (Figure 26 B). De plus, la boucle désordonnée CgL7 chez CgMCA-I<sup>Ca</sup>, a pu être mieux reconstruite pour le segment Ala<sup>329</sup> à Gly<sup>335</sup> chez CgMCA-I<sup>Mg</sup>. Ce segment se structure en un tour d'hélice  $\alpha$ , nommée  $\alpha_{cap}$ , venant également coiffer le site actif de l'enzyme et semble prévenir l'accès au site catalytique (Figure 27).



**Figure 26. Organisation structurale de CgMCA-I<sup>Mg</sup>.** A. Structure tridimensionnelle des deux monomères de CgMCA-I<sup>Mg</sup> présents dans l'unité asymétrique en complexe avec trois ions magnésium (Mg<sup>2+</sup><sub>A</sub>, Mg<sup>2+</sup><sub>B</sub>, Mg<sup>2+</sup><sub>C</sub>) représentés en sphères violettes. Les hélices sont représentées en bleu, les brins du feuillet principal sont en rouge et les brins βa-βb et βc-βd sont respectivement représentés en orange et en vert. Les résidus catalytiques, His<sup>182</sup> et Cys<sup>238</sup>, sont représentés en cyan. B. Représentation topologique de CgMCA-I<sup>Ca</sup>, les boucles sont représentées en noir et numérotées (L1 à L7/L7α), les zones désordonnées sont en pointillés noirs et les zones manquantes dues au processus de maturation sont en rouge foncé. Le brin β8 manquant est représenté par une flèche rayée rouge et blanche. Les sites de clivage sont indiqués par des flèches noires et la sous-unité p10 est représentée en pointillés oranges.



**Figure 27. Organisation structurale du site catalytique de CgMCA-I<sup>Mg</sup>.** Les résidus catalytiques Cys<sup>238</sup> et His<sup>182</sup> sont représentés en cyan et les ions magnésium (Mg<sup>2+</sup><sub>A</sub>, Mg<sup>2+</sup><sub>B</sub>, Mg<sup>2+</sup><sub>C</sub>) sont représentés en sphères violettes. La boucle désordonnée L7 (Ala<sup>330</sup>-Asp<sup>324</sup>) est indiquée en pointillés gris.

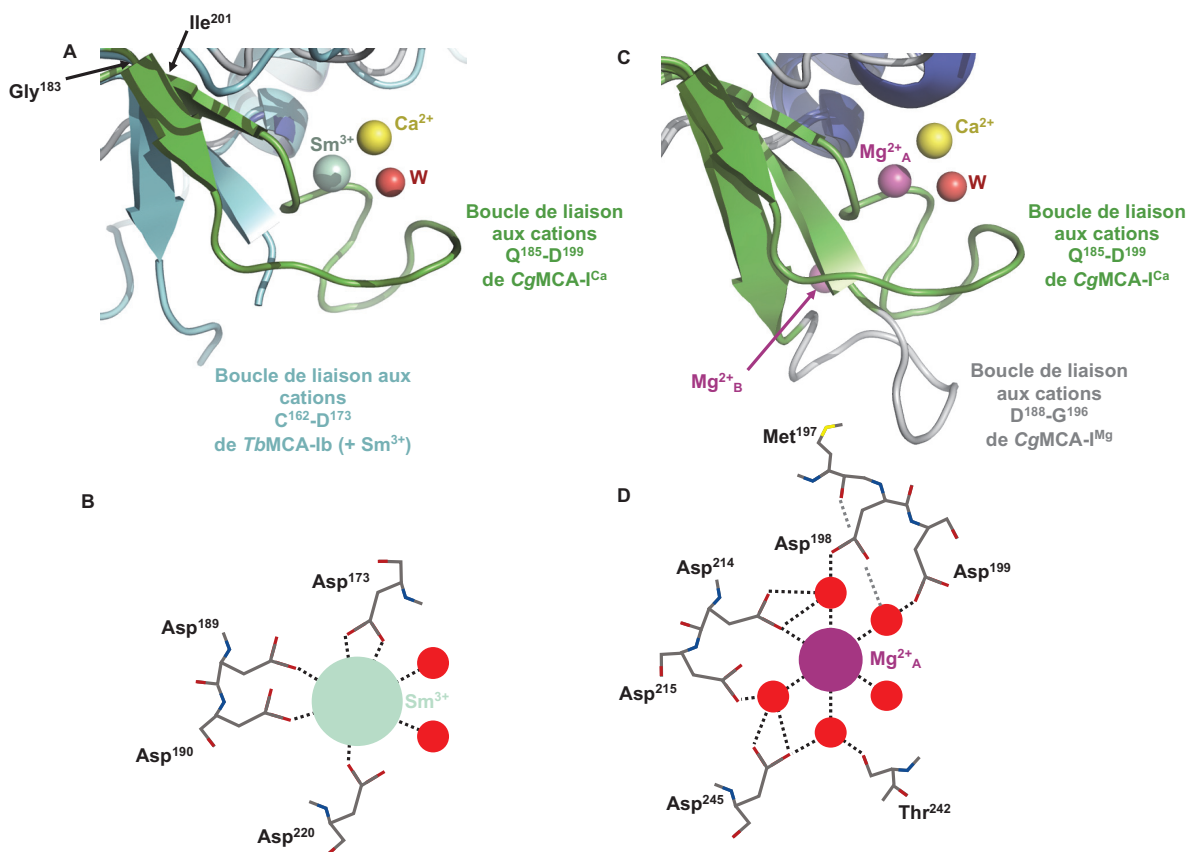


### C. Changements structuraux de CgMCA-I associés à la maturation dépendante du calcium

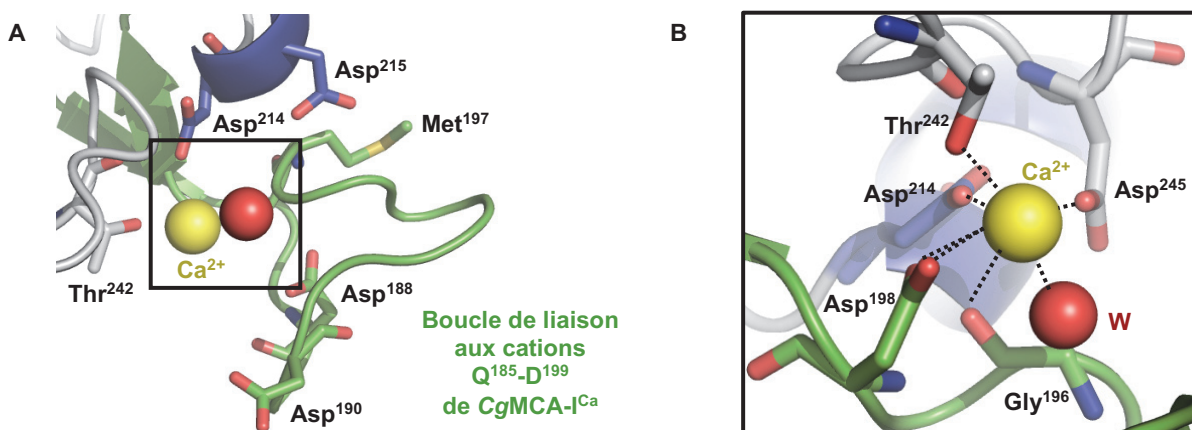
Comme indiqué précédemment, les cristaux ayant conduit à la résolution de la structure cristalline de CgMCA-I<sup>Ca</sup> ont poussés en employant l'enzyme partiellement maturée, purifiée en présence de calcium au jour 0 de maturation. Les cristaux sont apparus dans la condition de cristallisation après trois mois de stockage à 18°C. Pourtant, la structure montre une protéine dont la maturation est complète (Figure 22 A, B). Ceci semble indiquer que le processus de maturation se serait poursuivi lors des phases de cristallisation de la protéine.

Une densité électronique pour l'ion calcium a été observée dans la structure de CgMCA-I<sup>Ca</sup> et est proche de celle ayant été identifiée pour l'ion samarium dans la structure de TbmCA-Ib (Figure 22 A, 23 A, 28 A). Chez CgMCA-I<sup>Ca</sup>, l'ion est coordonné par 7 liaisons avec une molécule d'eau, les résidus Gly<sup>196</sup>, Thr<sup>242</sup> et les résidus Asp<sup>198</sup>, Asp<sup>214</sup> et Asp<sup>245</sup> correspondant respectivement aux résidus Asp<sup>173</sup>, Asp<sup>189</sup> et Asp<sup>220</sup> de TbmCA-Ib (Figure 28 B, 29 A, B). L'ion Sm<sup>3+</sup>, présent dans la structure de la métacapsase de *T. brucei*, est également coordonné par 7 liaisons avec les trois résidus Asp conservés ainsi qu'avec deux molécules d'eau et le résidu Asp<sup>190</sup> (Figure 28 B). Le mode de fixation des deux ions est donc relativement différent entre les deux métacaspases et la fixation du calcium chez CgMCA-I semble notamment induire une réorganisation complète du segment Gly<sup>183</sup>-Ile<sup>201</sup> (Figure 28 A).

En effet, l'ion Ca<sup>2+</sup> est déplacé de 3,4 Å par rapport à l'ion Sm<sup>3+</sup> (Figure 28 A). De plus, la boucle désordonnée (Cys<sup>162</sup>-Asp<sup>173</sup>) chez TbmCA-Ib est stabilisée chez CgMCA-I<sup>Ca</sup> (Gln<sup>185</sup>-Asp<sup>199</sup>), dû à la coordination du Ca<sup>2+</sup> par les résidus Gly<sup>196</sup> et Asp<sup>198</sup> (Figure 29 B). Dans cette conformation, cette boucle coiffe l'ion calcium et on l'appellera boucle de liaison aux cations. Elle fait partie d'un petit domaine constitué des deux brins β antiparallèles (βc et βd) qui fait directement suite à la boucle CgL3 et qui porte l'histidine catalytique (Figure 22 B). Enfin, le segment Asp<sup>236</sup>-Asp<sup>245</sup> (Asp<sup>211</sup>-Asp<sup>220</sup> chez TbmCA-I) contenant le résidu catalytique Cys<sup>238</sup> est quant à lui stabilisé *via* la coordination du Ca<sup>2+</sup> par les résidus Thr<sup>242</sup> et Asp<sup>245</sup> (Figure 29 B, 30). De manière intéressante, la boucle de liaison aux cations n'est pas résolue chez ScMCA-I (Gln<sup>223</sup>-Asp<sup>237</sup>) et TbmCA-Ib (Cys<sup>162</sup>-Asp<sup>173</sup>). Ceci suggère que le calcium possède un rôle de stabilisation à la fois de la boucle de liaison aux cations et du segment Asp<sup>236</sup>-Asp<sup>245</sup> résultant d'une stabilisation des deux résidus catalytiques au site actif.



**Figure 28. Organisation structurale du site de liaison aux cations de  $\text{CgMCA-I}^{\text{Ca}}$  et  $\text{CgMCA-I}^{\text{Mg}}$ .** **A**, **C**. Position de l'ion  $\text{Ca}^{2+}$  en jaune ( $\text{CgMCA-I}^{\text{Ca}}$ ) et de sa molécule d'eau (W), en rouge, coordonnée respectivement avec l'ion  $\text{Sm}^{3+}$  en vert clair ( $\text{TbmCA-I}$ , code PDB : 4AFP) (**A**) et l'ion magnésium en violet ( $\text{CgMCA-I}^{\text{Mg}}$ ) (**B**). La boucle de liaison aux cations (Gln<sup>185</sup>-Asp<sup>199</sup>) entourant l'ion  $\text{Ca}^{2+}$  ( $\text{CgMCA-I}^{\text{Ca}}$ ) est représentée en vert (**A**, **C**). La boucle de liaison aux cations désordonnée (Cys<sup>162</sup>-Asp<sup>173</sup>) chez  $\text{TbmCA-Ib}$  est indiquée en cyan (**A**) et cette même boucle chez  $\text{CgMCA-I}^{\text{Mg}}$  (Asp<sup>188</sup>-Gly<sup>196</sup>) est indiquée en grise (**C**). **B**, **D**. Représentation schématique de la coordination de l'ion  $\text{Sm}^{3+}$  (**B**) et de l'ion  $\text{Mg}^{2+}_{\text{A}}$  (**D**) observés respectivement dans les structures de  $\text{TbmCA-Ib}$  et  $\text{CgMCA-I}^{\text{Mg}}$ .

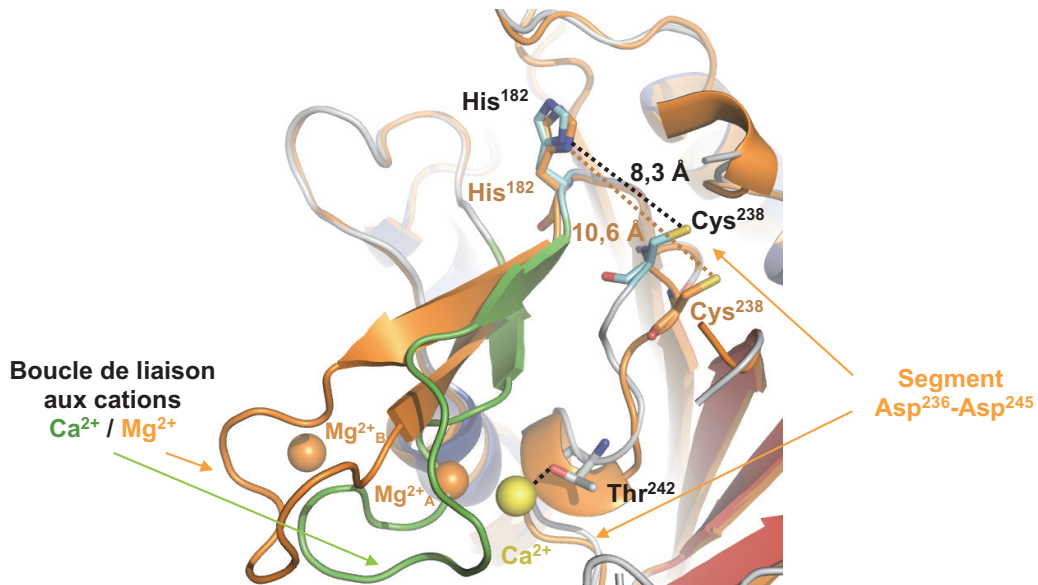


**Figure 29. Organisation structurale du site de liaison au calcium de  $\text{CgMCA-I}^{\text{Ca}}$ .** **A**, **B**. L'ion calcium en jaune ( $\text{CgMCA-I}^{\text{Ca}}$ ) est coordonné par une molécule d'eau (W) en rouge et les résidus de la boucle de liaison aux cations (Gln<sup>185</sup>-Asp<sup>199</sup>) en vert.

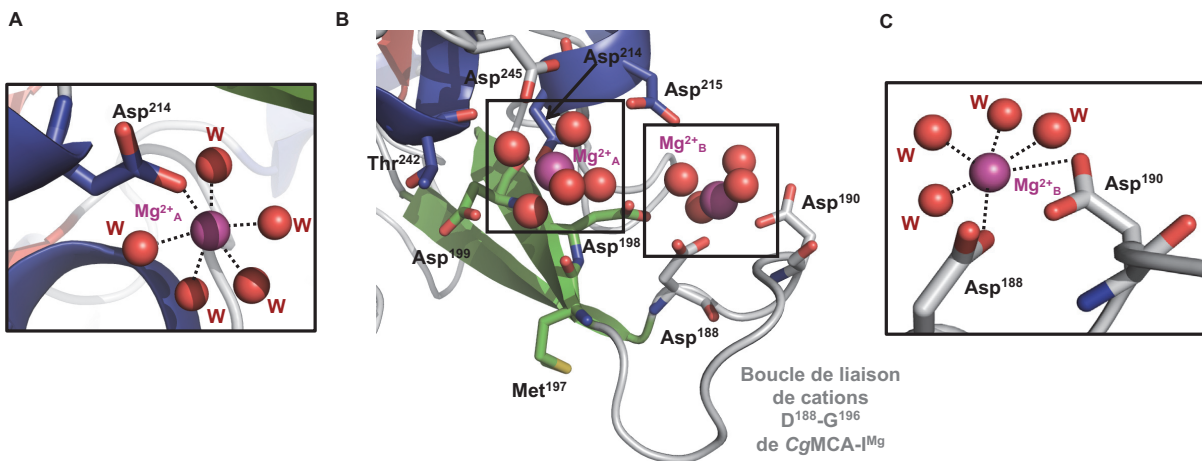


Les cristaux ayant conduit à la résolution de la structure cristalline *CgMCA-I<sup>Mg</sup>* ont poussés en employant l'enzyme purifiée en présence de calcium et au jour 4 de maturation. L'enzyme était donc dans une forme de maturation avancée et a cristallisé dans une condition contenant 200 mM de  $Mg^{2+}$ .

Trois zones de densité électronique ont été observées pour cet ion ( $Mg^{2+}_A, B, C$ ), chacun d'entre eux étant coordonné par 6 liaisons (Figure 31 A, B, C). Un des ions  $Mg^{2+}$  est positionné exactement à la place de l'ion  $Sm^{3+}$  observé chez *TbMCA-Ib* mais est coordonné différemment, par cinq molécules d'eau et le résidu Asp<sup>214</sup> (Asp<sup>189</sup> chez *TbMCA-Ib*) (Figure 28 C, D, 31 A). De manière intéressante, la plupart des résidus impliqués dans la fixation du  $Sm^{3+}$  sont conservés pour la fixation du  $Mg^{2+}_A$  et interagissent avec trois molécules d'eau qui coordonnent l'ion, l'Asp<sup>198</sup> (Asp<sup>173</sup> chez *TbMCA-Ib*), l'Asp<sup>215</sup> (Asp<sup>190</sup> chez *TbMCA-Ib*) et l'Asp<sup>245</sup> (Asp<sup>220</sup> chez *TbMCA-Ib*) (Figure 28 B, D). En plus de ces trois résidus, l'Asp<sup>199</sup> et la chaîne principale de la Thr<sup>242</sup> et de la Met<sup>197</sup> interagissent avec les deux autres molécules d'eau (Figure 28 D). Ainsi, il semblerait que l'ion  $Sm^{3+}$  mime le  $Mg^{2+}$  plutôt que l'ion  $Ca^{2+}$ . Le fait que cet ion ait été identifié pour une forme non maturée de *TbMCA-Ib* pourrait expliquer ces organisations différentes. En effet, bien que *CgMCA-I<sup>Mg</sup>* corresponde à une forme complète de maturation de l'enzyme, l'ion  $Mg^{2+}$  a précédemment été démontré comme empêchant le processus de maturation. De plus, pour *CgMCA-I<sup>Mg</sup>*, les deux résidus catalytiques His<sup>182</sup> et Cys<sup>238</sup> sont à une distance de 10,6 Å en comparaison à une distance de 8,3 Å pour *CgMCA-I<sup>Ca</sup>* suggérant que l'ion calcium entraîne une réorganisation structurale du site actif permettant à l'enzyme d'adopter une conformation dans laquelle elle est active (Figure 30). Par ailleurs, un second magnésium ( $Mg^{2+}_B$ ), coordonné par deux résidus Asp (Asp<sup>188</sup> et Asp<sup>190</sup>), se trouve aussi au niveau de la boucle de liaison aux cations et pourrait contribuer à sa stabilisation et sa réorientation (Figure 31 B, C).



**Figure 30.** Réorganisation du site catalytique de *CgMCA-I<sup>Ca</sup>* et de *CgMCA-I<sup>Mg</sup>*. *CgMCA-I<sup>Ca</sup>* est représentée en *bleu* (hélices  $\alpha$ ), *rouge* (brins  $\beta$ ), *vert* ( $\beta$ c et  $\beta$ d) et *gris* (boucles) et *CgMCA-I<sup>Mg</sup>* est représentée en *orange*. La distance a été calculée entre le groupement thiol de la cystéine catalytique Cys<sup>238</sup> et l'azote du groupement imidazole de l'histidine catalytique His<sup>182</sup>.

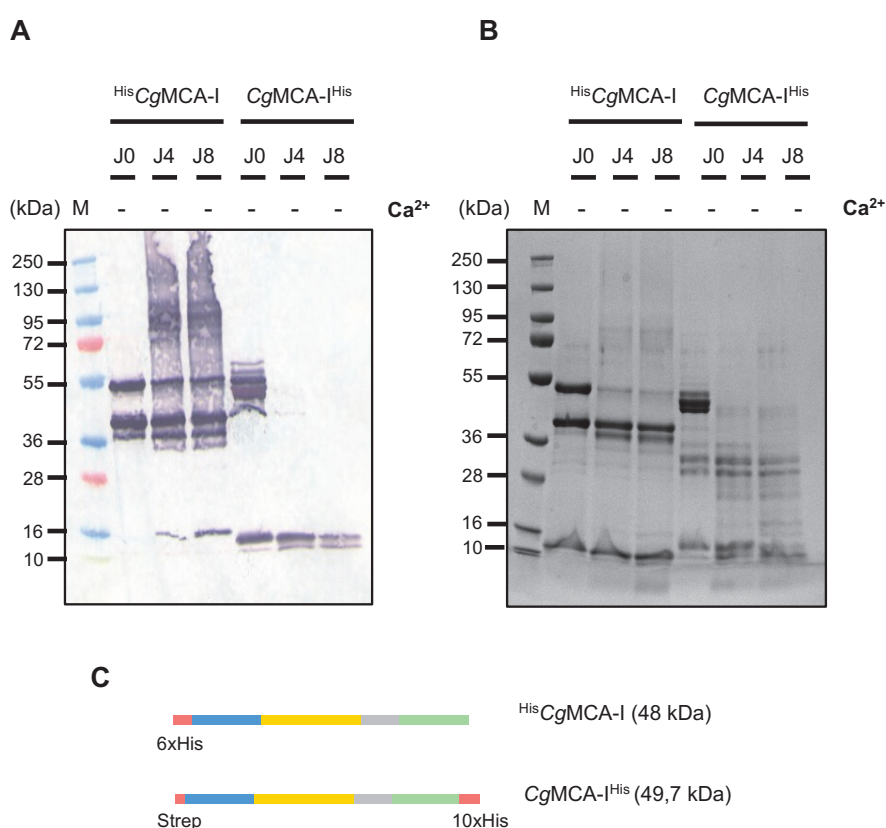


**Figure 31.** Organisation structurale des sites de liaisons aux ions  $Mg^{2+}$  de *CgMCA-I<sup>Mg</sup>*. A, B, C. Les ions magnésium,  $Mg^{2+}_A$  (A, B) et  $Mg^{2+}_B$  (B, C) sont représentés par des sphères *violettes*. Ils sont majoritairement coordonnés par des molécules d'eau (W) en *rouge*. *CgMCA-I<sup>Mg</sup>* est représentée en *bleu* (hélices  $\alpha$ ), *rouge* (brins  $\beta$ ), *vert* ( $\beta$ c et  $\beta$ d) et *gris* (boucles) (A, B).

## D. Processus de maturation

En complément des résultats de spectrométrie de masse, des constructions génétiques et des mutants de CgMCA-I ont été réalisées afin de déchiffrer le processus de maturation de l'enzyme.

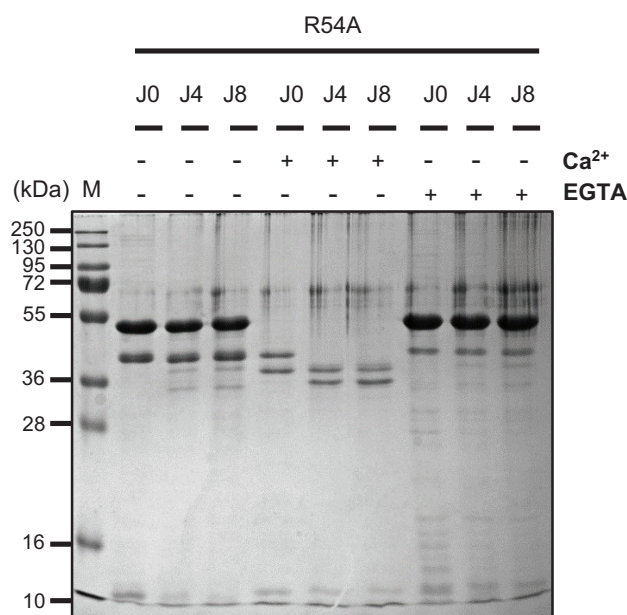
Pour ce faire, CgMCA-I a été cloné dans le vecteur pET52b résultant d'une enzyme recombinante et fusionnée à une étiquette 10xHis à l'extrémité C-terminale (CgMCA-I<sup>His</sup>) (Figure 32 C). De manière similaire à l'étude de la construction initiale, <sup>His</sup>CgMCA-I, des aliquots de la protéine purifiée en absence de Ca<sup>2+</sup> et d'EGTA ont été prélevés aux jours 0, 4 et 8 de maturation (Figure 32 A). En complément, un western-blot a été réalisé avec un anticorps dirigé contre l'étiquette histidine (Figure 32 B).



**Figure 32. Maturation des constructions <sup>His</sup>CgMCA-I et CgMCA-I<sup>His</sup> en absence de Ca<sup>2+</sup>.** Western-blot (A) et SDS-PAGE (B) des aliquots des constructions de <sup>His</sup>CgMCA-I et CgMCA-I<sup>His</sup> aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en absence de Ca<sup>2+</sup> (10 mM) ou d'EGTA (1 mM).

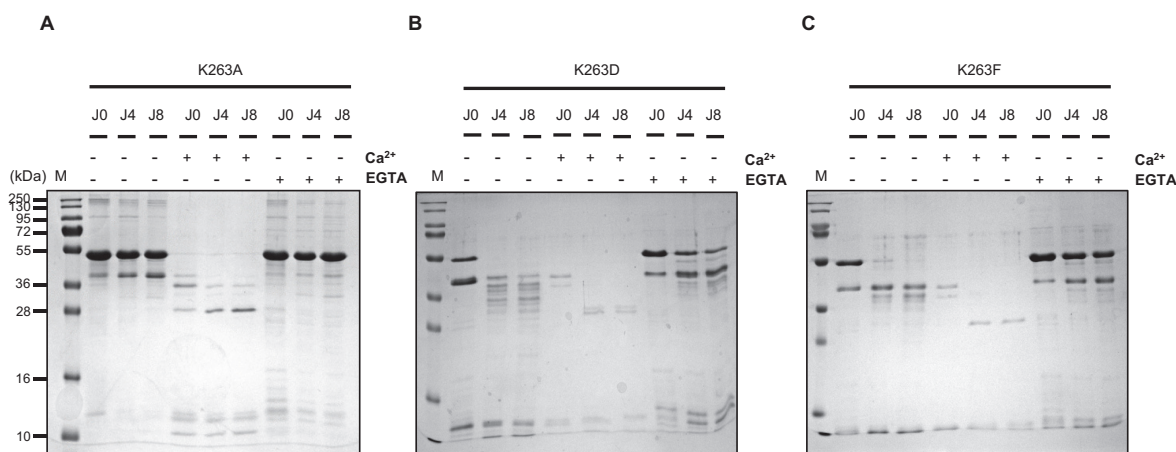
Alors qu'une bande d'environ 10 kDa apparaît au jour 0 pour CgMCA-I<sup>His</sup>, elle n'apparaît qu'au jour 4 pour <sup>His</sup>CgMCA-I, indiquant que le clivage de l'extrémité C-terminale, correspondant à la sous unité p10 clivée au résidu Arg<sup>307</sup>, interviendrait en premier dans le processus de maturation. De plus, l'absence de bande à 48 kDa, correspondant à la masse de

l'enzyme pleine longueur, au jour 4 de maturation pour *CgMCA-I<sup>His</sup>* indique que le départ de l'extrémité C-terminale a lieu dans les premières étapes. A l'inverse, pour la construction *HisCgMCA-I*, la présence d'une bande à 48 kDa jusqu'au jour 4 de maturation, indique que le clivage de l'extrémité N-terminale se déroule dans les dernières étapes. Pour confirmer cela, le mutant du site de clivage identifié R54A a été réalisé (Figure 33). L'analyse de sa maturation a révélé que même en présence de calcium, le processus de maturation est arrêté avant d'atteindre une forme mature complète.



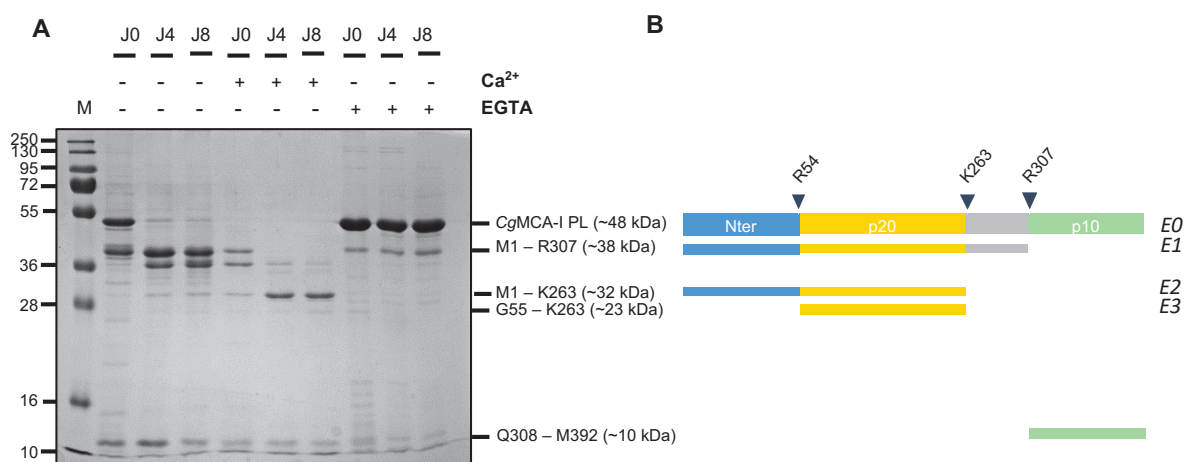
**Figure 33. Maturation du mutant R54A de *CgMCA-I*.** SDS-PAGE des aliquots du mutant R54A de *CgMCA-I* aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en présence ou absence de Ca<sup>2+</sup> (10 mM) ou d'EGTA (1 mM).

Afin d'étudier le second site de clivage qui semble se dérouler au résidu Lys<sup>263</sup> d'après les résultats de spectrométrie de masse, le mutant K263A a été étudié (Figure 34 A). De manière surprenante, en présence de calcium, la maturation de l'enzyme mutée se déroule de manière similaire à l'enzyme sauvage. Deux autres mutants ont été réalisés pour lesquels la Lys<sup>263</sup> était mutée en Asp et en Phe afin d'inverser la charge du résidu (K263D) et l'encombrement stérique (K263F). Pour les deux mutants, en absence ou en présence de Ca<sup>2+</sup>, une maturation aberrante est observée à partir du second clivage (Figure 34 B, C). En effet de nombreuses bandes non identifiables sont présentes à partir du jour 4 de maturation, résultant d'une disparition presque totale de l'enzyme en présence de Ca<sup>2+</sup>. Ces résultats suggèrent que la présence d'un acide aminé basique ne semble pas requise à cette position à condition que les résidus remplaçant la lysine ne changent pas drastiquement la fonction assumée par sa chaîne latérale.



**Figure 34. Maturation des mutants K263A, K263D et K263F de *CgMCA-I*.** SDS-PAGE des aliquots des mutants K263A (A), K263D (B) et K263F (C) de *CgMCA-I* aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en présence ou absence de  $\text{Ca}^{2+}$  (10 mM) ou d'EGTA (1 mM).

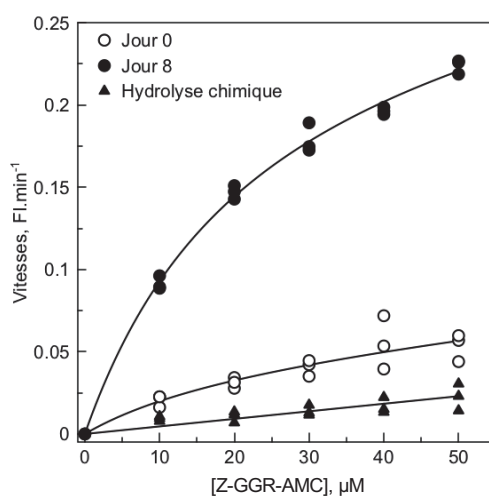
D'après les résultats obtenus, voici l'hypothèse émise sur le processus de maturation de *CgMCA-I* : la première étape de protéolyse de la forme non mature de *CgMCA-I* (48 kDa), nommée forme E0, permet de libérer le domaine C-terminale (Gln<sup>308</sup>-Met<sup>392</sup>) après un premier clivage au résidu Arg<sup>307</sup>. La disparition de ce fragment protéique, d'environ 10 kDa, conduit à la forme E1 associée à l'apparition d'une bande d'environ 38 kDa sur un gel d'électrophorèse (Met<sup>1</sup>-Arg<sup>307</sup>) (Figure 35 A, B). Un second clivage, au résidu Lys<sup>263</sup> induit la libération du domaine de liaison entre la sous-unité p20 et p10 d'environ 6 kDa et à l'apparition d'une bande d'environ 32 kDa sur gel d'électrophorèse correspondant à la forme E2 (Met<sup>1</sup>-Lys<sup>263</sup>). Enfin, la dernière étape correspond au clivage au résidu Arg<sup>54</sup> et à la libération du domaine N-terminale d'environ 6 kDa (Met<sup>1</sup>-Arg<sup>54</sup>). Cette étape correspond également à l'apparition de la forme mature de *CgMCA-I* (Gly<sup>55</sup>-Lys<sup>264</sup>), d'environ 23 kDa, nommée forme E3. Le clivage de *CgMCA-I*, à ces résidus, coïncide avec la taille des fragments identifiables par SDS-PAGE (Figure 35 B).



**Figure 35. Description du mécanisme de maturation de CgMCA-I.** A. SDS-PAGE des aliquots de CgMCA-I aux jours 0 (J0), 4 (J4) ou 8 (J8) de maturation en présence ou absence de Ca<sup>2+</sup> (10 mM) ou d'EGTA (1 mM). CgMCA-I PL correspond à la protéine pleine longueur de 48 kDa, les autres fragments protéiques identifiés correspondent aux différentes formes de maturations identifiées de CgMCA-I. B. Schéma de maturation de CgMCA-I (forme E0) après clivage aux résidus Arg<sup>307</sup>, Lys<sup>263</sup> et Arg<sup>54</sup> conduisant respectivement à l'apparition des formes E1, E2 et E3.

## E. Caractérisation enzymatique de CgMCA-I

Après des premières expériences préliminaires, les constantes de cinétiques de CgMCA-I ont été déterminées en présence de 10 mM de Ca<sup>2+</sup> aux jours 0 et 8 de maturation en utilisant le substrat Z-GGR-AMC (Figure 36). Un modèle cinétique de l'hydrolyse enzymatique du Z-GGR-AMC incluant l'hydrolyse chimique du substrat, a été utilisé pour déterminer les paramètres cinétiques (Tableau 3).



**Figure 36. Représentation de Michaelis-Menten de CgMCA-I non maturée (jour 0) et maturée (jour 8) en utilisant le substrat Z-GGR-AMC (0-50 µM) sans correction de l'hydrolyse chimique.**

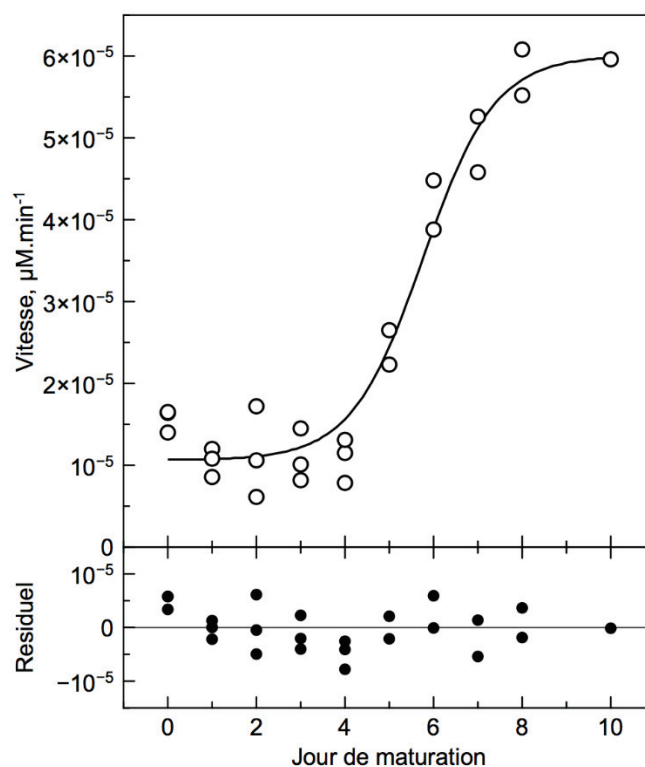


Paramètres	Valeur	Déviatiion standard	CV (%)
J0			
$K_m$	18,0 $\mu\text{M}$	12,0	64,7
$k_{\text{cat}}$	2,8 $\text{min}^{-1}$	0,7	24,9
$k_{\text{chem}}$	0,107 $\text{min}^{-1}$	0,012	10,8
J8			
$K_m$	18,2 $\mu\text{M}$	1,7	9,1
$k_{\text{cat}}$	15,8 $\text{min}^{-1}$	0,6	3,6
$k_{\text{chem}}$	0,1183 $\text{min}^{-1}$	0,0092	7,8

**Tableau 4. Paramètres cinétiques de CgMCA-I aux jours de maturation 0 (J0) et 8 (J8).** Les données ont été obtenues en utilisant le script 1 avec le logiciel DynaFit (cf. Matériels et méthodes, p 134). CV représente le coefficient de variabilité (en %).

La constante cinétique de premier ordre pour la réaction non catalysée ( $k_{\text{chem}}$ ) est comprise entre 0,107 +/- 0,012  $\text{min}^{-1}$  et 0,1183 +/- 0,0092  $\text{min}^{-1}$ . Comme attendu, la réaction enzymatique possède un comportement Michaelien dans la gamme de concentration de substrat testée. Au jour 0 de maturation, la valeur de  $k_{\text{cat}}$  est de 2,8 +/- 0,7  $\text{min}^{-1}$  et atteint 15,8 +/- 0,6  $\text{min}^{-1}$  au jour 8 de maturation indiquant que CgMCA-I mature est 5 à 6 fois plus active que sa forme non maturée. En revanche, la valeur du  $K_m$  pour Z-GGR-AMC (18 +/- 2  $\mu\text{M}$ ) ne change pas lors du processus de maturation. Par conséquent, le site de fixation du substrat est peu affecté pendant la maturation et il est plus probable que ce processus affecte la réorganisation spatiale des acides aminés catalytiques ce qui est en accord avec ce qui a été précédemment démontré dans l'étude structurale du site actif de CgMCA-I<sup>Ca</sup> et CgMCA-I<sup>Mg</sup>.

Afin de suivre le processus de maturation, l'enzyme purifiée a été stockée à 22 °C en absence de calcium et l'activité est mesurée quotidiennement en présence de 10 mM de  $\text{Ca}^{2+}$  pendant 10 jours (Figure 37). La courbe d'activité est une sigmoïde qui permet d'élucider le mécanisme de maturation.



**Figure 37. Transition de la population enzymatique  $E_{low}$  à la population  $E_{high}$  mise en évidence par la mesure de l'activité hydrolytique pendant 10 jours.**

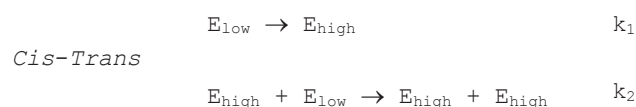
De 0 à 4 jours, l'activité est légèrement supérieure au bruit de fond ( $\sim 10 \cdot 10^{-6} \mu\text{M} \cdot \text{min}^{-1}$ ). Les principales formes de CgMCA-I (E0, E1, E2) identifiées par électrophorèse sur gel au stade précoce de la maturation ne sont pas totalement actives catalytiquement (Figure 35, 37). Au jour 5, l'activité augmente fortement pour atteindre un maximum de  $\sim 60 \cdot 10^{-6} \mu\text{M} \cdot \text{min}^{-1}$  au jour 8. Ceci est cohérent avec la multiplication par 6 de la constante  $k_{cat}$  observée dans l'expérience précédente. Le jour 8 de maturation coïncide avec l'apparition de la forme E3 identifiée par électrophorèse et semble correspondre à l'état enzymatique pleinement actif de CgMCA-I. Après le 8ème jour, toutes les molécules enzymatiques sont pour la plupart dans leur forme catalytiquement active (mature).

Du point de vue de l'activité, la forme sigmoïdale de cette courbe suggère l'existence de deux populations d'enzymes : une première population (nommée  $E_{low}$ ) dont l'activité spécifique est la plus faible et qui comprend les états E0, E1 et E2 ; et une seconde population (nommée  $E_{high}$ ) dont l'activité spécifique est la plus élevée et qui est composée de l'état E3. Au sein de la population  $E_{low}$  la distinction entre les enzymes de même activité, est impossible.

La forme sigmoïdale de la courbe suggère un mécanisme d'action en *trans*, c'est-à-dire que de la même manière que les caspases initiatrices mûrissent les caspases effectrices, les



enzymes de la population  $E_{high}$  sont capables de maturer les enzymes de la population  $E_{low}$  pour donner des enzymes  $E_{high}$ . Plusieurs modèles ont été évalués, à l'aide du logiciel Dynafit®, pour décrire le mécanisme cinétique de la transition de la population  $E_{low}$  à la population  $E_{high}$  (cf. Matériels et méthodes, Tableau 9). Le modèle le plus acceptable est un modèle à deux étapes où la première étape est la transition de la population  $E_{low}$  à la population  $E_{high}$  par un mécanisme *cis* tandis que la seconde étape est un mécanisme *trans* où les enzymes de la population  $E_{high}$  catalysent la maturation des enzymes  $E_{low}$  en population d'enzymes  $E_{high}$ , similaire à un mécanisme autocatalytique :



Dans les différents modèles on suppose que la vitesse de réaction est directement liée à la concentration de  $E_{low}$  et  $E_{high}$ , chaque forme ayant sa propre réponse spécifique :

$$\text{Vitesse } (\mu\text{M}\cdot\text{min}^{-1}) = r_{low} * [E_{low}] + r_{high} * [E_{high}]$$

$r_{low}$  et  $r_{high}$  sont respectivement les réponses spécifiques de  $E_{low}$  et  $E_{high}$ , et  $[E_{low}]$  et  $[E_{high}]$  sont respectivement les concentrations molaires de  $E_{low}$  et  $E_{high}$ . Les valeurs de  $k_1$  et  $k_2$  ainsi que les réponses spécifiques sont ajustées par le logiciel pour chaque modèle et les modèles sont comparés entre eux afin d'en définir le plus probable (cf. Matériels et méthodes, Tableau 10).

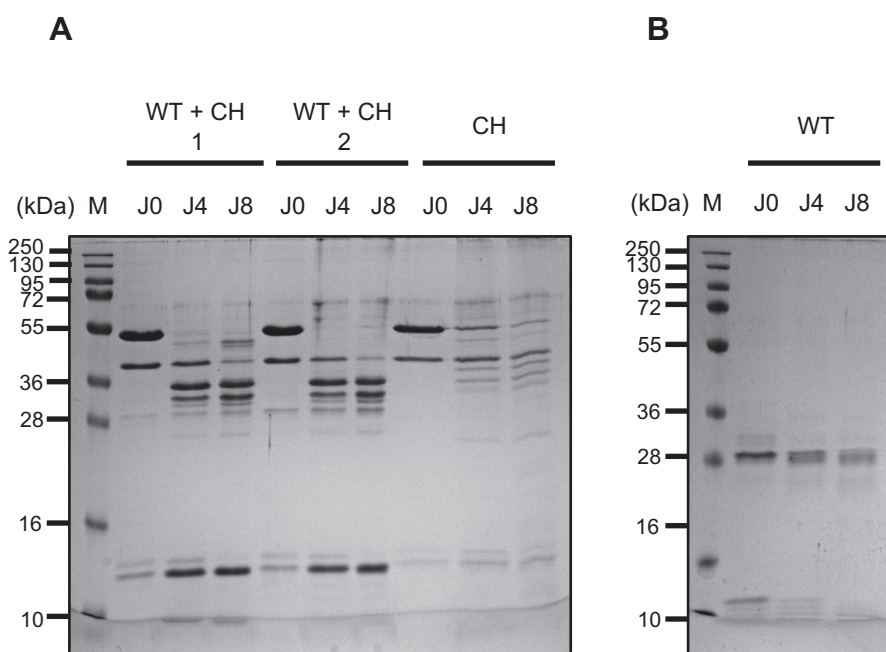
Les constantes cinétiques du meilleur modèle *Cis-Trans* ont été déterminées pour le jeu de paramètre où la constante  $k_1$  est fixée à 0,001 (Tableau 4) et il apparaît que la constante apparente de la première étape est environ 600 fois plus faible que celle de la seconde étape. Ceci explique que, dans nos conditions expérimentales, l'apparition de l'enzyme à haute activité (E3) suit quasiment un comportement OFF-ON.

Paramètre	Valeur	CV (%)
$k_1$ (jour <sup>-1</sup> )	0,001	n.a.
$k_2$ (μM <sup>-1</sup> .jour <sup>-1</sup> )	0,617 ± 0,025	4,0
$r_{low}$ (μM.min <sup>-1</sup> .μM <sup>-1</sup> )	(5,3 ± 0,6) × 10 <sup>-6</sup>	10,4
$r_{high}$ (μM.min <sup>-1</sup> .μM <sup>-1</sup> )	(30,0 ± 1,4) × 10 <sup>-6</sup>	4,5

**Tableau 5. Paramètres optimisés du modèle *Cis-Trans* (meilleur modèle).** CV est le coefficient de variabilité (en %), n.a : non applicable.

Lorsque la réaction est effectuée en présence d'EGTA (1 mM) ou en absence de calcium, aucune augmentation significative de l'activité n'est observée en huit jours, ce qui prouve que le calcium est nécessaire à la fois pour l'activité et la maturation.

Afin de valider cette hypothèse, nous avons mélangé une enzyme complètement maturée au jour 8 (1 ou 2 μM) avec un mutant catalytique incapable de se maturer (C238A/H182A ; 9 ou 8 μM) (Figure 38). Nous avons observé après une période de 8 jours que le mutant catalytique était maturé par l'enzyme active présente à faible concentration permettant de confirmer l'action *Trans* de *CgMCA-I*.



**Figure 38. Maturation du mutant C238A/H182A après mélange avec *CgMCA-I* sauvage maturée au jour 8.** A. SDS-PAGE des aliquots du mélange de *CgMCA-I* sauvage (WT) au jour 8 de maturation avec le mutant C238A/H182A (CH) n'étant pas capable de se maturer, aux jours 0 (J0), 4 (J4) et 8 (J8). Des mélanges à différent ratio ont été réalisés, le ratio 1 μM WT / 9 μM CH et le ratio 2 μM WT / 8 μM. Le mutant CH a été déposé sur gel à une concentration de 9 μM. B. SDS-PAGE des aliquots de *CgMCA-I* sauvage (WT), précédemment maturée 8 jours, après 0 (J0), 4 (J4) et 8 (J8) jours de maturation correspondant donc respectivement aux jours 8, 12 et 16 de maturation.

### III. Discussion

Dans cette étude, nous fournissons des preuves fonctionnelles et structurales de la dépendance au  $\text{Ca}^{2+}$  des métacaspases pour leur activité autocatalytique. L'activation complète de CgMCA-I requiert au moins trois clivages conduisant à une enzyme complètement mature avec ses domaines p20 et p10.

Le premier clivage a lieu à l'extrémité C-terminale de la protéine, libérant le domaine p10. Les deuxième et troisième clivages ont eu lieu respectivement en position 263 et 54, libérant la boucle centrale CgL6 (Asp<sup>264</sup>-Arg<sup>307</sup>) et, finalement, l'extrémité N-terminale de la protéine (Figure 35). Dans la structure de TbmCA-Ib, la dyade catalytique est enterrée sous le N-terminale qui encercle le corps principal de l'enzyme, cette conformation indiquant que le clivage du domaine N-terminale pourrait être un prérequis à l'activation complète de l'enzyme (Figure 10). En ce qui concerne la structure de ScMCA-I ou CgMCA-I, les domaines N-terminaux ne sont pas visibles, les structures étant obtenues avec des enzymes matures. Il a été suggéré que le domaine N-terminale de ScMCA-I n'avait aucun rôle d'auto-inhibition (Wong, Yan and Shi, 2012). Le fait que le domaine N-terminale de CgMCA-I soit la dernière partie de l'enzyme à être clivée, confirme que son rôle n'est pas de contrôler ou de restreindre l'activation autocatalytique de ce type d'enzyme.

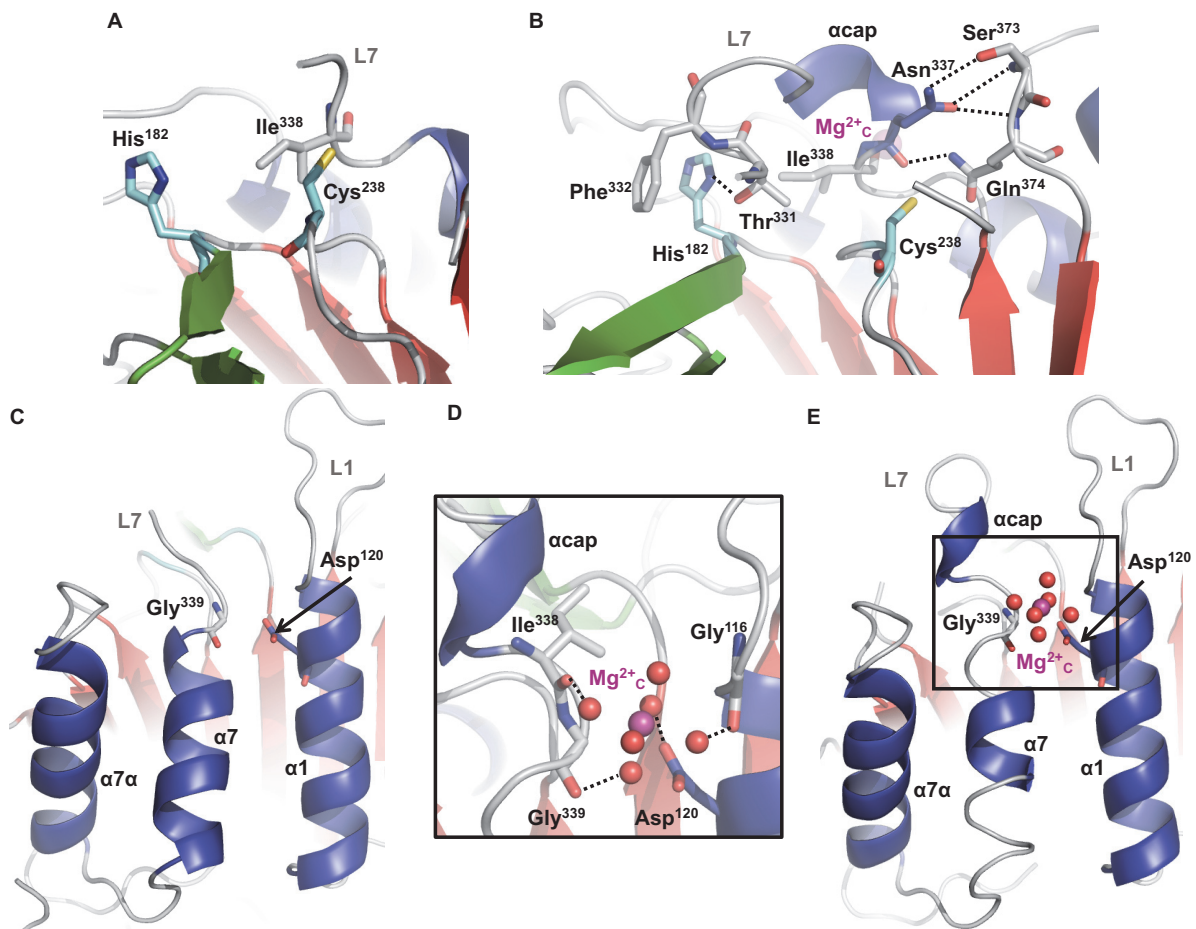
Sur la base de la structure et des sites de clivage identifiés, nous nous sommes interrogés sur la possibilité d'un clivage autocatalytique en *cis* versus un clivage autocatalytique en *trans*. Comme souligné précédemment pour la structure de ScMCA-I, les distances entre le centre catalytique et les résidus clivés sont difficilement compatibles avec un clivage en *cis* (Wong, Yan and Shi, 2012). Nous avons pu confirmer que l'activation de CgMCA-I pouvait reposer sur une activation *trans*. En revanche, la spécificité de clivage ne semblerait pas dépendre de la séquence, compte tenu du manque de conservation entre les différents sites de clivage (Figure 39), mais plutôt de la conformation des boucles cibles. Cela pourrait également indiquer que pour éviter un clivage systématique, le site actif n'est accessible ou restreint qu'à cette partie étendue des protéines cibles, l'accessibilité reposant finalement sur la présence de résidus basiques.



Afin d'accommoder leurs substrats, les différentes structures disponibles de métacaspase montrent la même organisation du site actif avec un regroupement de boucles composé de CgL3 et CgL4 portant respectivement l'histidine et la cystéine catalytiques (la boucle portant la cystéine catalytique étant appelée L2 pour la caspase-7 exécutive et ScMCA-I ; Tableau 1), la boucle CgL7 coiffant le site catalytique (nommée L5 ou L3 pour la caspase-7, L3 pour ScMCA-I et boucle 280 pour TbmCA-I) et les boucles CgL1 et CgL7 $\alpha$  (nommées L4 pour la caspase-7 et ScMCA-I) de part et d'autre du site catalytique (Figure 39). Dans cet assemblage, la boucle L1 (CgL1), observée dans les structures de métacaspase disponibles, présente exactement la même taille et orientation (McLuskey *et al.*, 2012; Wong, Yan and Shi, 2012; McLuskey and Mottram, 2015). D'autre part, CgL7 et en particulier la boucle  $\alpha$ cap montrent une séquence, une taille ou une organisation de structure très variable (Figure 39). Chez la métacaspase de type II de *A. thaliana* (AtMCA-IIId), le site catalytique est coiffé par le domaine de liaison entre les sous-unités p20 et p10. Cette conformation unique force une lysine (Lys<sup>225</sup>) à entrer dans le site catalytique, ce qui entraîne l'auto-inhibition de AtMCA-IIId (Zhu *et al.*, 2020). Pour TbmCA-Ib, la structure obtenue à partir d'un cristal trempé dans du calcium montre une réorientation complète de la boucle 280 (CgL7) qui forme un feuillet à deux brins avec l'extrémité N-terminale intégrée dans le site catalytique. Dans CgmCA-I, deux organisations différentes sont également observées. En présence de calcium, la boucle CgL7 est en grande partie déstabilisée, cependant, en présence de magnésium, la boucle forme une hélice à un tour qui est stabilisée par une liaison hydrogène entre Asn<sup>337</sup> et la chaîne latérale de Ser<sup>373</sup> et Gln<sup>374</sup> (Figure 40 A, B). CgL7 est également stabilisée par la présence d'un Mg<sup>2+</sup> (Mg<sup>2+</sup><sub>C</sub>) qui est coordonné par 6 molécules d'eau interagissant avec la chaîne latérale de Asp<sup>120</sup> et la chaîne principale de Gly<sup>116</sup>, Ile<sup>338</sup> et Gly<sup>339</sup> (Figure 40 C, D, E). Sous cette conformation, CgL7 entre dans le site catalytique en formant une liaison hydrogène entre le His<sup>182</sup> catalytique et Thr<sup>331</sup> (Figure 40 B). De plus, nous avons montré que la présence des deux Mg<sup>2+</sup> dans la boucle Gln<sup>185</sup>-Asp<sup>199</sup> entraîne une augmentation de la distance entre les deux résidus catalytiques (Figure 30). En remplaçant le magnésium, le Ca<sup>2+</sup> pourrait forcer l'enzyme à passer à une forme active et la verrouiller dans cette configuration. Concernant TbmCA-Ib, des auteurs ont suggéré qu'un second site de faible affinité pourrait exister pour le Ca<sup>2+</sup> (Machado *et al.*, 2013; Acosta-Maspons *et al.*, 2014; Klemenčič and Funk, 2018). Ayant identifié un ion Mg<sup>2+</sup> partageant le même emplacement que le site liaison principal du Ca<sup>2+</sup>, nous émettons l'hypothèse que le site de liaison du Mg<sup>2+</sup><sub>C</sub> au niveau de la boucle CgL7 pourrait représenter le site de liaison de faible affinité du Ca<sup>2+</sup>. Ce dernier ne pourrait fixer un ion que lorsque les concentrations intracellulaires de calcium seraient élevées et induirait un mouvement de la boucle CgL7, rendant accessible le site actif de l'enzyme. Ainsi, la



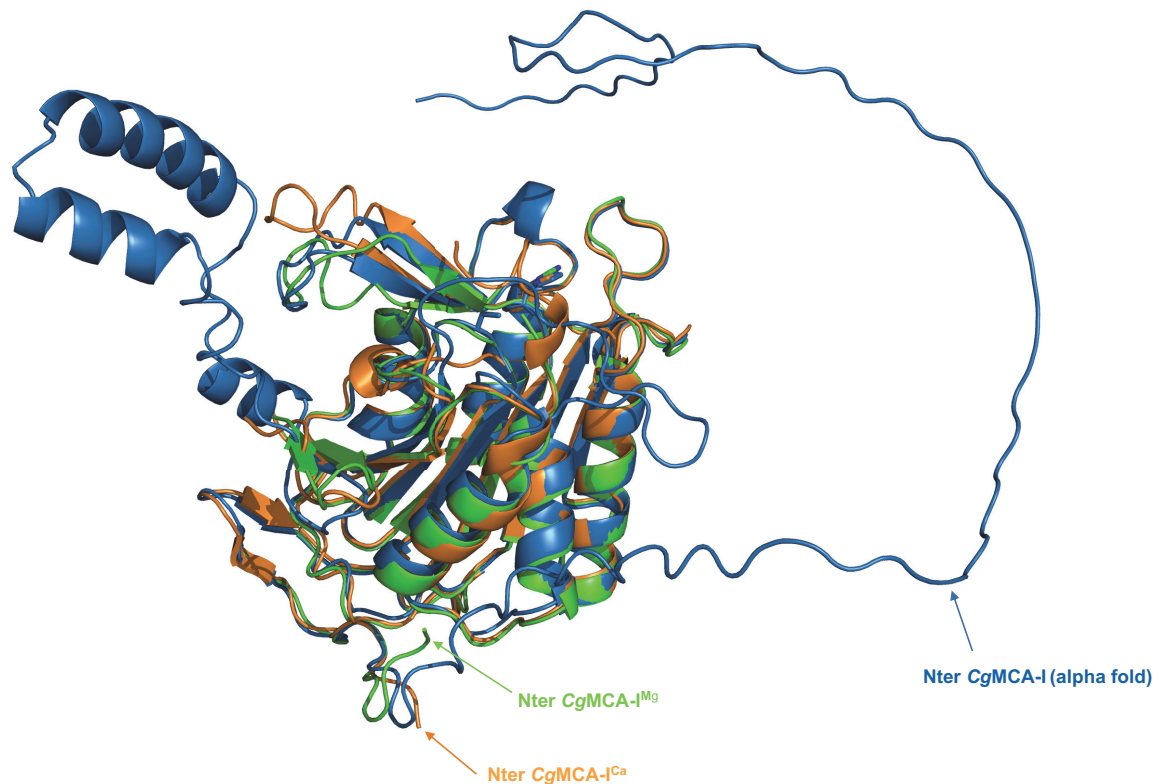
concentration intracellulaire de calcium serait un moyen de contrôler l'état physiologique des métacaspases.



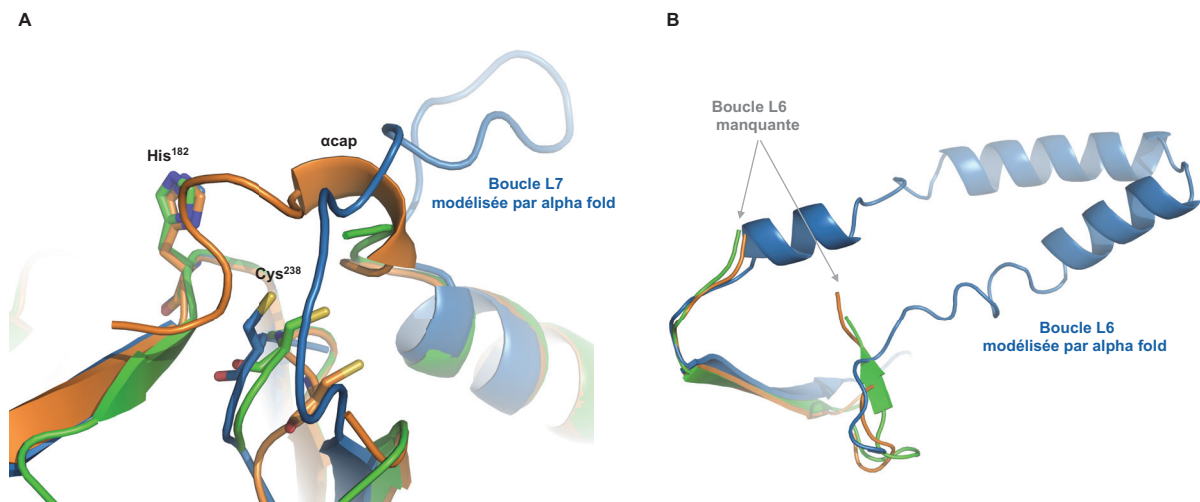
**Figure 40. Réorganisation structurale de la boucle L7 chez *CgMCA-I*<sup>Ca</sup> et *CgMCA-I*<sup>Mg</sup>.** A, B, C, D, E. Les structures de *CgMCA-I*<sup>Ca</sup> (A, C) et *CgMCA-I*<sup>Mg</sup> (B, D, E) sont représentées en *bleu* (hélices  $\alpha$ ), *rouge* (brins  $\beta$ ), *vert* ( $\beta$ c et  $\beta$ d), *gris* (boucles) et les résidus catalytiques sont indiqués en *cyan* (A, B). L'ion  $Mg^{2+}_c$  est représenté par une sphère *violette* (B, D, E). La boucle de *CgMCA-I*<sup>Mg</sup> qui coiffe le site actif est nommée  $\alpha$ cap (B, D, E).

N'ayant pu étudier l'enzyme non maturée d'un point de vue fonctionnel et structural, nous avons entrepris de modéliser la structure de *CgMCA-I* à l'aide du logiciel d'intelligence artificielle AlphaFold 2 (AF2). Ce dernier est capable, à partir de la séquence en acides aminés, de prédire la structure tridimensionnelle d'une protéine (Jumper *et al.*, 2021). Ainsi, nous avons pu observer le modèle AF2 de la structure tridimensionnelle de *CgMCA-I* non maturée et confirmer que, selon ce modèle, comme pour la structure de *TbMCA-Ib*, l'extrémité N-terminale de l'enzyme n'est pas structurée (Figure 41). Toutefois, nous émettons l'hypothèse, qu'à la différence de *TbMCA-I*, l'extrémité N-terminale ne viendrait pas se positionner au niveau du site catalytique afin de restreindre l'accès au site actif en raison de la présence de la boucle *Cg*L7 (Figure 10, Figure 41, Figure 42 A). Concernant cette boucle, nous observons par ailleurs, qu'Alphafold la modélise avec une orientation différente

de celle observée dans la structure de *CgMCA-I<sup>Mg</sup>*, confirmant ainsi la flexibilité de ce segment et la possibilité que le mouvement de celui-ci puisse limiter l'accès au site actif en enfouissant les deux résidus catalytiques (Figure 42 A). Enfin, de manière surprenante, il apparaît que la boucle de liaison entre les sous-unités p10 et p20 (*CgL6*), absente des structures tridimensionnelles de *CgMCA-I<sup>Ca</sup>* et *CgMCA-I<sup>Mg</sup>* dû à son clivage lors de la maturation de l'enzyme, serait susceptible de se structurer en hélices  $\alpha$  (Figure 42 B). Cette observation soulève de nombreuses interrogations, cette structuration a-t-elle un rôle de stabilisation de l'enzyme avant que celle-ci soit maturée et que ce domaine de liaison entre sous-unités soit clivé ? Représenterait-elle d'être un motif structural reconnu par un partenaire protéique ? Cette reconnaissance pourrait-elle intervenir dans un contexte où l'enzyme est impliquée dans des phénomènes de survie cellulaire et n'exerce pas nécessairement une activité catalytique ? Une étude plus approfondie de ce motif structural permettrait de comprendre, à terme, son implication au niveau fonctionnel.



**Figure 41.** Superposition des structures  $CgMCA-I^{Ca}$  et  $CgMCA-I^{Mg}$  avec la modélisation de  $CgMCA-I$  par AlphaFold 2. La structure de  $CgMCA-I^{Ca}$ ,  $CgMCA-I^{Mg}$  et la modélisation de la structure de  $CgMCA-I$  par alpha fold sont respectivement représentées en vert, orange et bleu.



**Figure 42.** Comparaison des boucles L6 et L7 pour les structures de  $CgMCA-I^{Ca}$ ,  $CgMCA-I^{Mg}$  et la structure  $CgMCA-I$  modélisée par AlphaFold 2. La structure de  $CgMCA-I^{Ca}$ ,  $CgMCA-I^{Mg}$  et le modèle AF2 de  $CgMCA-I$  par alpha fold sont respectivement représentées en vert, orange et bleu. La boucle L6, manquante chez  $CgMCA-I^{Ca}$  et  $CgMCA-I^{Mg}$ , est modélisée par AF2 en trois hélices  $\alpha$  (A). La boucle L7 manquante chez  $CgMCA-I^{Ca}$  et structurée sous la forme d'une courte hélice  $\alpha$  ( $\alpha cap$ ) chez  $CgMCA-I^{Mg}$  est sous la forme d'une boucle flexible possédant une orientation différente d'après le modèle AF2 (B).



## IV. Conclusion et perspectives

Au cours de cette thèse, nous avons pu étudier les relations structurales, fonctionnelles et enzymatiques de la métacaspase de *C. glabrata*. Après une première stratégie infructueuse visant à inhiber l'autoprotéolyse de l'enzyme en mutant les acides aminés catalytiques ou les résidus ciblés par le clivage, nous avons réorienté notre axe de recherche. En effet, notre deuxième stratégie a été d'étudier l'effet du calcium sur la structure de l'enzyme et notamment sur sa maturation et son activation. Nos recherches nous ont non seulement permis de comprendre le mécanisme séquentiel de maturation de l'enzyme mais nous ont également permis d'obtenir, pour la première fois, des informations importantes sur le rôle de l'ion  $\text{Ca}^{2+}$  dans la réorganisation structurale du site actif et sur la maturation de l'enzyme. Par l'étude enzymatique de l'enzyme, nous avons observé, qu'au-delà du rôle structural et fonctionnel de cet ion, il était également requis pour la mise en place de l'activité de CgMCA-I. Ainsi, l'activité et la maturation de cette enzyme pourrait être dépendante des flux de calcium intracellulaire. Comme cela a été mentionné en introduction, une conséquence de la perception du signal de mort pour une cellule est l'augmentation de la concentration intracellulaire et intramitochondriale de calcium. Ce flux d'ions, en plus de provoquer le phénomène de MOMP pourrait également être responsable de l'activation et de la maturation de la métacaspase, constitutivement exprimée sous une forme zymogène. La nécessité d'un phénomène de maturation dépendant du calcium pourrait constituer une sécurité à l'échelle de la cellule en empêchant l'enzyme d'exercer son activité protéolytique dans des conditions physiologiques. Bien que l'activité protéolytique des métacaspases de levure n'ait pas encore été démontrée *in vivo*, des expériences ont permis d'identifier des substrats naturels chez *Candida albicans*, tel que la GAPDH (Silva *et al.*, 2011), une enzyme de la glycolyse possédant un homologue chez *C. glabrata* (codé par le gène *TDH3*). Il est tentant de postuler que ces enzymes participeraient à la propagation du signal de mort cellulaire en protéolysant des cibles protéiques essentielles à la survie de l'organisme.

De nombreuses questions restent malgré tout en suspens concernant notamment la distinction entre l'implication de ces enzymes dans des phénomènes de mort et de survie cellulaire, d'autant plus maintenant qu'il a été démontré que leur activation est dépendante d'une concentration importante de calcium. Il serait envisageable qu'en conditions physiologiques, les métacaspases soient exprimées sous leur forme zymogène et ne puissent donc pas exercer leur activité protéolytique. Dans cette configuration, elles pourraient alors exercer un rôle de protéine chaperonne ou interagir avec d'autres protéines participant à des phénomènes de

survie cellulaire comme cela a été suggéré pour leur implication dans l'adressage des agrégats protéiques au protéasome (Lee *et al.*, 2010 ; Hill and Nyström, 2015). Le motif structural en hélices  $\alpha$  du domaine de liaison des sous-unités p10 et p20, suggéré par modélisation *via* le logiciel AlphaFold 2, pourrait constituer une zone d'interaction avec des partenaires protéiques. Le calcium pourrait ainsi agir comme un interrupteur moléculaire permettant de passer d'une fonction de survie à une fonction de mort cellulaire, en induisant le processus de maturation et la mise en place de l'activité catalytique de l'enzyme.

De nombreuses autres expérimentations pourraient être envisagées afin d'obtenir une meilleure compréhension du mécanisme moléculaire des métacaspases ainsi que du chemin métabolique de mort cellulaire dans lequel elles interviennent. L'obtention de l'enzyme sous sa conformation non mature permettrait notamment de consolider les hypothèses émises lors de ces travaux. En effet, il serait intéressant d'observer la configuration structurale du site actif de l'enzyme, que nous supposons similaire à celle en présence de  $Mg^{2+}$ , et de la mettre en perspective avec celle du site actif en présence de  $Ca^{2+}$ . En complément, il serait pertinent d'envisager une approche bioinformatique visant à comprendre les aspects dynamiques. Les résultats obtenus par modélisation de la structure non mature *via* AlphaFold 2 semblent indiquer que le positionnement des résidus catalytiques est similaire à ce qui a été observé chez *CgMCA-I<sup>Ca</sup>* (Figure 42 A). Bien qu'AlphaFold soit une avancée technique majeure, le logiciel montre ici sa limite à donner une information plus complète que la biologie structurale expérimentale sur cette question précise. Il faudrait alors envisager des travaux de dynamique moléculaire sur de grandes échelles de temps afin de modéliser les mouvements qui peuvent s'opérer au sein de la protéine. Ces techniques de modélisation, aujourd'hui considérées comme fiables, nécessitent néanmoins une expertise dédiée et de très gros moyens de calculs. La résolution de la structure tridimensionnelle de l'enzyme non mutée permettrait également de confirmer le positionnement de la boucle *CgL7* lorsque l'enzyme est sous une forme inactive. En parallèle, il serait aussi intéressant d'obtenir de l'information structurale d'un second (ou plus) site de liaison au calcium afin de valider l'hypothèse de sa position au niveau de la boucle *CgL7*. L'étude enzymatique de ces deux sites permettrait de confirmer leur différence d'affinité pour le calcium tel que cela a été démontré précédemment chez *TbMCA-Ib* (Machado *et al.*, 2013). D'un point de vue enzymatique, il serait aussi envisageable d'étudier l'influence des autres ions et notamment du magnésium sur l'activité de l'enzyme. En effet, cela permettrait d'observer si, de la même manière qu'il inhibe la maturation de l'enzyme, il empêcherait la mise en place de l'activité catalytique de *CgMCA-I*. Chez *S. cerevisiae* et *C. glabrata*, un seul gène codant la métacaspase a été identifié.

Toutefois, l'activation en *trans* de ces enzymes semble suggérer qu'en présence de signaux d'activation de la mort cellulaire, comme l'augmentation intracellulaire de calcium, un nombre de copies de l'enzyme relativement important serait nécessaire à ce mode d'activation. Ainsi, il conviendrait d'évaluer le taux d'expression de l'enzyme dans des conditions d'activation du signal de mort cellulaire dépendant des métacaspases. Une approche cellulaire *in vivo* pourrait également mettre en évidence la co-localisation de ces copies dans la cellule de levure. Enfin, il apparaît essentiel de mener des études d'interaction de l'enzyme avec des partenaires protéiques pouvant être des cibles ou des activateurs de la protéase dans un contexte de mort cellulaire ou des partenaires protéiques s'inscrivant dans un contexte de survie cellulaire. En lien avec les résultats structuraux obtenus par modélisation à l'aide d'AlphaFold 2, il conviendrait d'étudier la boucle CgL6, qui semble se structurer en hélices  $\alpha$ , et évaluer si ce segment pourrait représenter un motif de reconnaissance par des partenaires. Par ailleurs, l'identification de substrats naturels de l'enzyme permettrait d'obtenir de plus amples informations concernant son implication dans les voies métaboliques dans lesquelles elle est impliquée ou éventuellement, d'en découvrir de nouvelles.

L'étude de CgMCA-I et plus généralement de la voie métabolique de la mort cellulaire régulée des levures, impliquant les métacaspases, pourrait s'inscrire dans la recherche d'alternatives thérapeutiques aux antifongiques présentant des résistances dans le traitement des infections à *Candida*. En effet, bien qu'inhiber une protéase impliquée dans la mort cellulaire d'un pathogène puisse sembler être une stratégie thérapeutique antagoniste, l'implication de cette enzyme dans de nombreux processus essentiels à la survie et son absence chez l'Homme peuvent faire d'elle une cible thérapeutique de choix. La miltéfosine, un traitement utilisé contre les infections à *Leishmania*, a notamment montré son efficacité en stimulant le signal de mort cellulaire du parasite en ciblant la voie de biosynthèse des métacaspases (Kumar, Pandey and Samant, 2018 ; Vandana *et al.*, 2019). D'autres auteurs ont pu montrer que la miltéfosine aurait un effet activateur de la métacaspase de *Leishmania major* (*LmMCA*) (Basmaciyani, Azas and Casanova, 2018). Ils ont également pu mettre en évidence que l'utilisation couplée d'amphotéricine B, de curcumin et de peroxyde d'hydrogène déclencherait un phénomène de mort cellulaire où *LmMCA* serait inhibée. Face à ce dernier phénomène, deux hypothèses peuvent être faites, soit l'inhibition de la métacaspase induit directement un phénomène « apoptotique » puisque cette dernière est aussi impliquée dans des phénomènes de survie, soit la mort cellulaire serait le résultat de l'activation d'une voie de mort cellulaire indépendante des métacaspases.

Outre leur intérêt d'un point de vue thérapeutique, la compréhension du mécanisme d'action de ces enzymes à l'échelle moléculaire est cruciale pour déchiffrer leur fonctionnement ainsi que le processus de mort cellulaire auxquelles elles participent. Depuis la mise en évidence de ces homologues des caspases, il y a une vingtaine d'années, des structures de métacaspases ont été résolues chez seulement trois organismes. De plus, la distinction entre métacaspases et caspases a longtemps été assez floue, notamment du point de vue de leur spécificité enzymatique. En effet, de nombreuses caractérisations de l'activité des métacaspases ont été réalisées à l'aide de substrat correspondant à la spécificité de clivage des caspases, clivant après des résidus Asp. Enfin, il faut noter l'importance d'une approche structurale – fonctionnelle – enzymatique, permettant une compréhension globale des mécanismes de maturation et d'activité de ces enzymes.

## V. Matériel et méthodes

### A. Expression et purification de CgMCA-I

#### 1. *Optimisation des gènes*

Toutes les expérimentations de production protéique durant cette thèse ont été réalisées en système hétérologue bactérien chez *E. coli*. Afin d'optimiser le taux d'expression des protéines provenant de l'organisme *C. glabrata* chez *E. coli*, une optimisation des gènes a été réalisée par la société GenScript. Cette dernière, par l'utilisation d'un algorithme OptimumGene™, analyse la séquence d'ADN codant pour la protéine d'intérêt et évalue les paramètres critiques pour le niveau d'expression dans l'organisme d'expression choisi. Parmi ces paramètres sont évalués la fréquence des codons entre organismes, notamment *via* l'indice d'adaptation de codon (« Codon Adaptation Index » ou CAI ; (Sharp and Li, 1987)) pour lequel une valeur supérieure à 0,8 est suffisante pour garantir une bonne expression. Mais également, le nombre de codons rares, le contenu en bases G et C ou la présence d'éléments négatifs pour la transcription et la traduction.

Le gène optimisé codant l'enzyme sauvage (<sup>His</sup>CgMCA-I WT) a été cloné entre les sites de restrictions *NdeI* et *BamH1* du vecteur pET15b (Novagen) possédant une étiquette histidine 6xHis fusionnée à l'extrémité N-terminale (Figure 42).

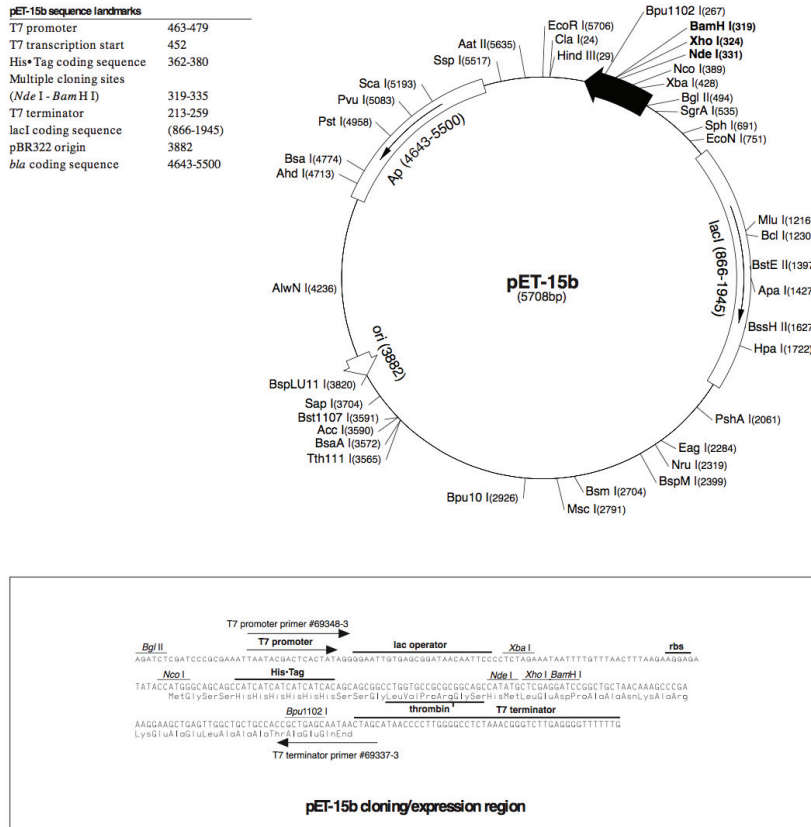


Figure 43. Carte du vecteur pET15b (Novagen).

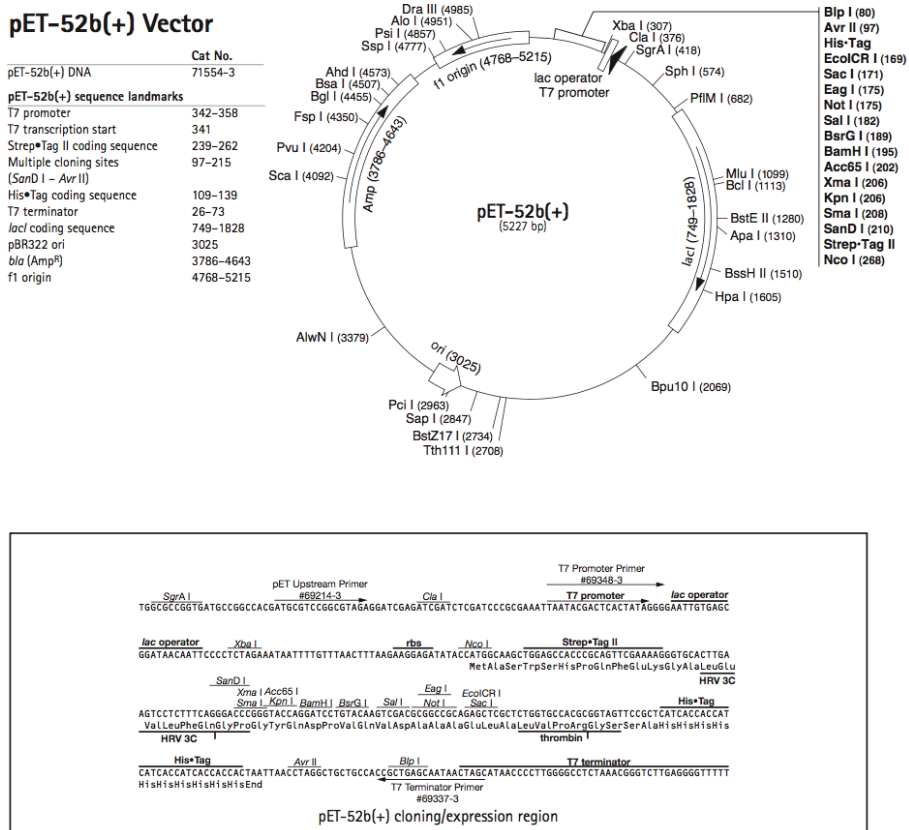


Figure 44. Carte du vecteur pET52b (Novagen).

## 1. Mutagenèse dirigée

Des expériences de clonage et de mutagenèse dirigées ont été réalisées sur CgMCA-I WT afin d'obtenir de nouvelles constructions génétiques et de muter divers résidus en Ala, Asp ou Phe. Pour ce faire, des amorces ADN ont été réalisées et synthétisées par l'entreprise Sigma-Aldrich (Tableau 6).

### *Clonage de CgMCA-I<sup>His</sup>*

Afin de construire CgMCA-I WT avec une étiquette histidine à l'extrémité C-terminale (CgMCA-I<sup>His</sup>), une première étape d'amplification du vecteur pET15b\_CgMCA-I (CgMCA-I<sup>His</sup>) a été réalisée, par réaction de polymérisation en chaîne (PCR), en utilisant des amorces d'ADN contenant à leurs extrémités les sites de restriction *SalI* et *SacI* (Tableau 6). Dans une seconde étape, le fragment amplifié est cloné entre les sites de restriction *SalI* et *SacI* du vecteur pET52 (Novagen) possédant une étiquette 10xHis à l'extrémité C-terminale (Figure 44). La PCR est réalisée à partir de 10 ng de plasmide additionné de 100  $\mu$ M d'amorce sens, 100  $\mu$ M d'amorce anti-sens, de 2 mM de dNTP (désoxyribonucléoside triphosphate), 1  $\mu$ L de polymérase (enzyme Q5, NEB Biolabs) dans un volume final de 50  $\mu$ L. Le cycle de PCR utilisé est le suivant : 1 x 98 °C pendant 5 min, 30 x [98 °C pendant 30 sec ; 54 °C pendant 1 min ; 72 °C pendant 2min] et 72 °C pendant 20 min. Le produit PCR est digéré par l'enzyme DpnI (10 U/ $\mu$ L). L'amplification du produit de PCR est contrôlée en déposant 3  $\mu$ L d'échantillon sur un gel agarose 1 %. Après extraction de la bande sur gel, le produit est purifié à l'aide du kit PCR clean-up (Macherey-Nagel) dans un volume final de 30  $\mu$ L et la concentration est vérifiée à l'aide d'un spectromètre à 260 nm (Nanodrop2000 ; Thermo Fisher Scientific). La digestion du produit d'amplification (3  $\mu$ g) et du vecteur pET52 (1  $\mu$ g) a été réalisée en ajoutant aux échantillons les enzymes de restriction *SalI* et *SacI* (2 U.mL<sup>-1</sup>) en présence du tampon CutSmart® (NEB Biolabs) (10 X) et d'eau en quantité suffisante pour un volume final de 30  $\mu$ L. L'incubation a été réalisée à 37 °C pendant 2h30, puis 1  $\mu$ L de phosphatase alcaline a été ajouté pendant 30 min à 37 °C. Le vecteur et l'insert digérés ont été déposés sur gel d'agarose 1 %, extraits puis purifiés à l'aide du kit PCR clean-up (Macherey-Nagel). Le mix de ligation est réalisé à partir de 50 ng de vecteur et de 150 ng d'insert en présence de la T4 ligase (NEB Biolabs), du tampon de la ligase (10 X) et d'eau en quantité suffisante pour 20  $\mu$ L. L'incubation est réalisée sur la nuit à 14 °C puis l'échantillon est soumis à une température de 70 °C pendant 5 min. Enfin, 100  $\mu$ L de bactéries thermocompétentes NEB 5 $\alpha$  sont transformées avec 10  $\mu$ L du produit de ligation. La construction a été vérifiée par séquençage ADN.

## Mutagenèse dirigée

Le gène CgMCA-I WT a été muté à l'aide des amorces listées dans le tableau 6, selon le principe de la mutagenèse dirigée par PCR (Heckman and Pease, 2007) afin d'obtenir les simples mutants C218A, R54A, K263A et le double mutant C238A/H182A à partir du simple mutant C238A et des amorces permettant de réaliser la mutation H182A. La PCR est réalisée à partir de 200 ng de plasmide additionné de 2 µM d'amorce sens, 2 µM d'amorce anti-sens, de 2 mM de dNTP, 1 µL de polymérase (enzyme Q5, NEB Biolabs) dans un volume final de 50 µL. Le milieu réactionnel est placé dans un thermocycleur 9700 (Hi-Tech Detection Systems) et le programme est le suivant : 1 x 98 °C pendant 5 min, 30 x [98 °C pendant 30 sec; 50 °C pendant 1 min; 72 °C pendant 2 min] et 72 °C pendant 20 min. Le produit PCR est digéré par l'enzyme DpnI (10 U.µL<sup>-1</sup>) et transformé dans des bactéries DH5 α. Toutes les constructions ont été vérifiées par séquençage ADN.

Amorces	Séquences
<b>Amorces externes</b>	
<i>SalI</i>	5'- C GGC <b>GTC GAC</b> ATC CCG GGC AGC GGT AAC-3'
<i>SacI</i>	5'-CG GAT <b>GAG CTC</b> CAT AAT AAA CTG CAG G-3'
<b>Amorces internes</b>	
<b>C238A</b>	
<i>sens</i>	5' - CTGACCGCGCTGTTTGTATAGCGCGCACAGCGGTACCGTTCTGGA - 3'
<i>antisens</i>	5' - TCCAGAACGGTACCGCTGTGCGCGCTATCAAACAGCGCGGTTCAG - 3'
<b>C238A/H182A</b>	
<i>sens</i>	5' - TGTTCTTTCACTACAGCGGTGCGGGCGGTCAAACCAAAGACCTG - 3'
<i>antisens</i>	5' - CAGGTCTTTGGTTTGACCGCCCGCACCGCTGTAGTGAAAGAACA - 3'
<b>H182A</b>	
<i>sens</i>	5' - TGTTCTTTCACTACAGCGGTGCGGGCGGTCAAACCAAAGACCTG - 3'
<i>antisens</i>	5' - CAGGTCTTTGGTTTGACCGCCCGCACCGCTGTAGTGAAAGAACA - 3'
<b>R54A</b>	
<i>sens</i>	5'-GGGCGAGAACGCGGGCCAATACCAGGG-3'
<i>antisens</i>	5'-CCCTGGTATTGGCCCGGTTCTCGCCC-3'
<b>K263A</b>	
<i>sens</i>	5'-GCCGAACATGTGGGCGGATGTTGGTAGCG-3'
<i>antisens</i>	5'-CGCTACCAACATCCGCCACATGTTTCGGC-3'
<b>K263D</b>	
<i>sens</i>	5'-GCCGAACATGTGGGATGATGTTGGTAGCG-3'
<i>antisens</i>	5'-CGCTACCAACATCATCCACATGTTTCGGC -3'
<b>K263F</b>	
<i>sens</i>	5'-GCCGAACATGTGGTTTGTATGTTGGTAGCG -3'
<i>antisens</i>	5'-CGCTACCAACATCAAACCACATGTTTCGGC -3'

**Tableau 6. Amorces sens et anti-sens employées pour la mutagenèse dirigée de C238A, H182A, C238A/H182A, R54A, K263A, K263D et K263F de CgGBE et pour la construction CgMCA-I<sup>His</sup>.**



## 2. *Expression et purification*

Toutes les constructions et mutants de CgMCA-I ont été transformés dans des cellules *E. coli* T7 express (NEB Biolabs). Une colonie unique a été inoculée dans 10 mL de milieu Luria-Bertani (LB) contenant 50  $\mu\text{g.mL}^{-1}$  d'ampicilline puis mis sous agitation et à 37 °C durant la nuit. Le lendemain, la pré-culture a permis d'ensemencer 2 L de milieu LB contenant de l'ampicilline à 50  $\mu\text{g.mL}^{-1}$ . La culture a été réalisée à 37 °C et sous agitation jusqu'à ce que la densité optique (DO) atteigne une valeur d'absorbance à 600 nm de 0,6 environ. L'expression de CgGBE WT et de la protéine mutée a été induite par l'ajout de 0,2 mM d'isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG, Sigma-Aldrich) et s'est poursuivie à 20 °C, pendant la nuit. Les cellules ont été récoltées par centrifugation à 5000 x g pendant 15 min à 4 °C et les culots ont été congelés à -80 °C. Le culot a ensuite été resuspendu dans 60 mL de tampon de lyse composé d'HEPES 10 mM à pH 7,6, NaCl 300 mM et supplémenté par une tablette d'un cocktail d'inhibiteur de protéases sans EDTA (cOmplete ; Roche), de la DNase I et RNase I à 5  $\mu\text{g.mL}^{-1}$  (Sigma-Aldrich) et de lysozyme à 25  $\mu\text{g.mL}^{-1}$  (Sigma-Aldrich). Les bactéries sont lysées à 6 °C en utilisant un Microfluidizer (M110P ; Microfluidics) à 1500 bar. Le lysat est ensuite centrifugé à 11 000 x g à 4 °C pendant 45 minutes. Puis, il est chargé sur une colonne Co-NTA (HiTrap® TALON® crude 5mL ; GE Healthcare, UK) préalablement lavée avec le tampon de lyse contenant 30 mM d'imidazole (Merck) et 1 % (*m/v*) de glycérol (Euromedex) en utilisant un système de chromatographie en phase liquide de type ÄKTA Purifier (GE Healthcare). Après une première étape de lavage ayant pour but d'éluer toutes les protéines aspécifiques, un saut de sel est réalisé avec un tampon HEPES 10 mM à pH 7,6, NaCl 1 M visant à se débarrasser d'éventuels contaminants ADN. L'élution a été réalisée par un gradient linéaire d'une concentration croissante d'imidazole de 30 mM à 500 mM. La protéine éluée est dessalée en utilisant une colonne de dessalage HiPrep© (GE Healthcare, UK) équilibrée avec un tampon composé d'HEPES 10 mM pH 7,6, de NaCl à 150 mM et de 1 % (*m/v*) de glycérol. Après analyse par SDS-PAGE 15 %, les fractions d'élution contenant la protéine ont été mélangées et concentrées à l'aide d'un concentrateur Amicon avec un seuil de coupure à 30 kDa (Millipore™ Corporation). Lors d'une seconde étape de purification la protéine pure et concentrée a été chargée sur une colonne de gel filtration (Superdex-200 ; GE Healthcare) préalablement équilibrée dans le tampon de dessalage. Les aliquots de protéine ont été à nouveau contrôlés par SDS-PAGE 15 %, concentrés, dosés à l'aide d'un spectrophotomètre à 280 nm (Nanodrop2000 ; Thermo Fisher Scientific), utilisés pour des expérimentations ou rapidement congelés à l'azote liquide et stockés à -80 °C jusqu'à leur utilisation.

### 3. Etude du processus de maturation

L'analyse du processus de maturation de CgMCA-I et de ses mutants a été réalisée en purifiant la protéine en absence ou présence de calcium ou d'EGTA en utilisant un BioSprint 96 (Qiagen). Cet appareil permet une purification rapide de protéines recombinantes fusionnées à une étiquette histidine à l'aide de billes d'agarose magnétiques couplées à des ions nickel liés à de l'acide nitrilotriacétique (Ni-NTA). CgMCA-I et ses mutants ont été transformés dans des souches *E. coli* T7 express (NEB Biolabs), cultivées dans 200 mL de Luria-Bertani (LB, Sigma-Aldrich) supplémenté d'ampicilline à 50 µg.mL<sup>-1</sup> puis mis sous agitation à 37 °C. L'expression a été induite par l'ajout de 0,2 mM d'isopropyl-β-D-thiogalactopyranoside (IPTG, Sigma-Aldrich) lorsque la densité optique a atteint une valeur d'absorbance à 600 nm de 0,6 environ, et se poursuit à 20 °C pendant la nuit sous agitation. Les cellules ont été récoltées par centrifugation à 5000 x g pendant 15 min à 4 °C et le culot a été resuspendu dans 5 mL de tampon de lyse avec ou sans additifs (Tableau 7) et supplémenté d'antiprotéase CLAPA 1X (Chymostatine 1 µg.mL<sup>-1</sup>, Leupeptine 1 µg.mL<sup>-1</sup>, Antipaïne 1 µg.mL<sup>-1</sup>, Pepstatine 1 µg.mL<sup>-1</sup> et Aprotinine 1 µg.mL<sup>-1</sup>), de lysozyme à 25 µg.mL<sup>-1</sup> (Sigma-Aldrich), de DNase I et RNase I à 5 µg.mL<sup>-1</sup> (Sigma-Aldrich). Les bactéries sont lysées à 4°C par trois sessions successives de 30 pulsations d'ultrasons à l'aide d'un sonicateur (SONIFIER 250 ; Branson) à une puissance de 50 %. Le lysat est ensuite centrifugé à 11 000 x g à 4 °C pendant 20 minutes. Un volume de 500 µL de lysat est disposé dans les cupules du BioSprint et additionné de 30 µL de billes magnétiques auxquelles la protéine va se lier *via* son étiquette histidine. Les deux cupules de lavages contiennent 1 mL du tampon de lavage adapté et les deux cupules d'élution contiennent 100 µL du tampon d'élution adapté (Tableau 7). L'éluat est prélevé après purification et des aliquots sont prélevés au jour 0, 4 et 8 et supplémenté de tampon Laemmli 1X et stocké à -20 °C avant d'être analysés par SDS-PAGE 15 %.

Tampon	Composition
<i>Tampon de lyse</i>	
∅	HEPES 10 mM pH 7,6, NaCl 150 mM
Ca <sup>2+</sup>	HEPES 10 mM pH 7,6, NaCl 150 mM, CaCl <sub>2</sub> 10 mM
EGTA	HEPES 10 mM pH 7,6, NaCl 150 mM, EGTA 1 mM
<i>Tampon de lavage</i>	
∅	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 30 mM
Ca <sup>2+</sup>	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 30 mM, CaCl <sub>2</sub> 10 mM
EGTA	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 30 mM, EGTA 1 mM
<i>Tampon d'élution</i>	
∅	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 500 mM
Ca <sup>2+</sup>	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 500 mM, CaCl <sub>2</sub> 10 mM
EGTA	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 500 mM, EGTA 1 mM

**Tableau 7. Liste des tampons employés pour l'analyse du processus de maturation de CgMCA-I et de ses mutants, par BioSprint.** Tous les tampons ont été fraîchement préparés pour les expériences. Le symbole ∅ correspond aux tampons de lyse, lavage ou élution ne contenant pas d'additifs.

Cette même étude du processus de maturation a également été réalisée en présence de trois autres ions divalents ( $Mn^{2+}$ ,  $Mg^{2+}$ ,  $Zn^{2+}$ ) (Tableau 8).

Tampon	Composition
<i>Tampon de lyse</i>	
∅	HEPES 10 mM pH 7,6, NaCl 150 mM
Ion	HEPES 10 mM pH 7,6, NaCl 150 mM, $MnCl_2$ ou $MgCl_2$ ou $ZnCl_2$ 10 mM
<i>Tampon de lavage</i>	
∅	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 30 mM
Ion	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 30 mM, $MnCl_2$ ou $MgCl_2$ ou $ZnCl_2$ 10 mM
<i>Tampon d'élution</i>	
∅	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 500 mM
Ion	HEPES 10 mM pH 7,6, NaCl 150 mM, imidazole 500 mM, $MnCl_2$ ou $MgCl_2$ ou $ZnCl_2$ 10 mM

**Tableau 8. Liste des tampons employés pour l'analyse du processus de maturation de CgMCA-I en présence des ions divalents  $Mn^{2+}$ ,  $Mg^{2+}$  et  $Zn^{2+}$ , par BioSprint.** Tous les tampons ont été fraîchement préparés pour les expériences. Le symbole ∅ correspond aux tampons de lyse, lavage ou élution ne contenant pas d'additifs.

## B. Analyse par spectrométrie de masse

### 1. *Identification des sites de coupure de CgMCA-I par LC-MS*

L'identification des sites de coupures de CgMCA-I a été réalisée par spectrométrie de masse LC-MS en collaboration avec la plateforme PSF (« Protein Science Facility ») de UAR 3444/U8 - Lyon.

Quelques nanogrammes de protéines ont été injectés et analysés en utilisant une nano-UHPLC Ultimate 3000 (Thermo Scientific, San Jose California) couplée en ligne avec un spectromètre de masse Q Exactive HF avec une source d'ionisation nano-electrospray (Thermo Scientific, San Jose California).

Les échantillons ont été injectés sur une pré-nano-colonne C4 Acclaim Pepmap 300  $\mu m$  ID x 5 mm, 5  $\mu m$ , 300Å, (Thermo Scientific) pendant 3 minutes à 20  $\mu L \cdot min^{-1}$  avec 2 % d'ACN (Acétonitrile), 0,05 % TFA (Acide trifluoroacétique) dans l'eau. Les fragments protéiques ont ensuite été séparés par une nano-colonne C4 Acclaim Pepmap, 75  $\mu m$  ID 15 cm x 5  $\mu m$ , 300 Å (Thermo Scientific) avec un gradient linéaire de 3,2 % à 80 % de tampon B (A: 0.1 % FA (acide formique) dans l'eau, B: 100 % ACN, 0.1 % FA) pendant 30 min suivi par un gradient de 80 % à 95 % de tampon B pendant 1 min, un maintien à 95 % de tampon B pendant 5 minutes et enfin un retour aux conditions initiales en 1 min pendant 13 min. La durée totale de la manipulation est fixée à 50 minutes et le flux est de 300 nL/min. La température est maintenue constante à 40 °C.

Le spectromètre de masse a fonctionné en mode MS positif, en utilisant une gamme de balayage de 500 à 2000 masse sur charge ( $m/z$ ). La résolution de l'appareil était réglée sur 240 000 FWHM (« full width at half maximum », largeur à mi-hauteur) à 200  $m/z$ , la cible AGC était  $3e6$  et le temps d'injection maximal était de 100 ms.

Les analyses par déconvolution ont été traitées à l'aide du logiciel Biopharma Finder 2.0 (Thermo Scientific) avec la méthode par défaut Xtract-average sur une plage de masse sélectionnée (résolution isotopique) qui utilise un état de charge minimal de 5, avec au moins 3 états de charge différents pour identifier un composé, la tolérance de masse de précision pour la masse déconvoluée a été fixée à 20 ppm. Le spectre de masse déconvolué moyen a été obtenu, ce qui a permis d'obtenir une liste de poids moléculaires correspondant aux fragments détectés présents dans l'échantillon de la métacaspase (CgMCA-I). Les masses déconvoluées dont l'abondance relative est inférieure à 1 % ont été filtrées. La liste des poids moléculaires obtenus est comparée à la masse théorique de la séquence d'acides aminés de CgMCA-I avec la présence potentielle d'oxydations sur la méthionine en utilisant l'outil FindPept sur le site web Expasy (SIB Swiss Institute of Bioinformatics, Expasy).

### C. Analyse enzymatique

L'analyse enzymatique de CgMCA-I a été réalisée en collaboration avec le Dr. Bastien Doumèche, de l'ICBMS UMR 5246 CNRS-Université Lyon 1.

Les paramètres cinétiques de l'enzyme ont été déterminés pour le jour 0 et 8 de maturation de CgMCA-I en présence du substrat Z-GGR-AMC dans une gamme de concentration allant de 0 à 50  $\mu\text{M}$ . Les mesures ont été réalisées à l'aide d'un spectrophotomètre TECAN Infinite M200 (Tecan, Salzburg, Autriche) en utilisant des plaques 96 puits à fond noir. Les réactions dans chaque puits sont réalisées dans un tampon HEPES 10 mM pH 7,6, NaCl 150 mM contenant  $\text{CaCl}_2$  10 mM ainsi que le substrat Z-GGR-AMC à une concentration de 0, 10, 20, 30, 40 ou 50  $\mu\text{M}$  (10  $\mu\text{L}$ ). Le milieu réactionnel est pré-incubé à 25 °C avant de débiter la réaction en ajoutant l'enzyme à une concentration de 4  $\mu\text{M}$  pour un volume final de 100  $\mu\text{L}$ . La libération du produit de la réaction (AMC) est suivie en mesurant sa fluorescence, après excitation à une longueur d'onde d'excitation de 355 nm (9 nm de bande passante), à une longueur d'onde d'émission de 460 nm (20 nm de bande passante). L'expérience a été réalisée en triplicat ( $n=3$ ) et le blanc a été obtenu en remplaçant l'enzyme par du tampon. Les données ont été traitées à l'aide du logiciel DynaFit avec le script ci-dessous, pour l'enzyme

au jour 0 de maturation, incluant l'hydrolyse chimique du substrat. Les paramètres initiaux sont fixés arbitrairement. Un script similaire a été utilisé pour déterminer les constantes de l'enzyme au jour 8 de maturation.

L'hypothèse d'équilibre rapide est utilisée. Les concentrations sont en  $\mu\text{M}$ .

### Script 1 :

```
[task]
  task = fit ;
  data = rate ;
  approximation = rapid-equilibrium ;

[mechanism]
  E + S <==> ES : Km dissociation
  ES ---> E + P : kcat
  S ---> P : kc

[constants]
  kcat = 1 ?
  Km = 25 ?
  kc = 1 ?

[responses]
  P = 0.004309

[concentrations]
  E = 0 ;

[data]
  variable S
  directory ...
  sheet MM.csv
  column 2 | concentration E = 4 | label + Enz
  column 3 | concentration E = 0 | label - Enz

[output]
  directory ...

[settings]
{Output}
  XAxisLabel = [Z-GGR-AMC], microM
  YAxisLabel = rate, Fluorescence units.min-1

[end]
```

Afin de pouvoir étudier le modèle de maturation, les différentes formes de maturation ont été regroupées en deux groupes, les premières formes de maturation sont dites  $E_{\text{low}}$  et la forme de maturation finale est dite  $E_{\text{high}}$ .

La maturation de la population  $E_{\text{low}}$  en  $E_{\text{high}}$  pourrait résulter de trois étapes élémentaires :

- Un mécanisme *cis* où  $E_{\text{low}}$  est convertie en  $E_{\text{high}}$  selon une cinétique d'ordre 1
- Un mécanisme *trans* pour lequel deux formes  $E_{\text{low}}$  sont converties en au moins une forme  $E_{\text{high}}$  (la seconde étant une forme  $E_{\text{low}}$  ou  $E_{\text{high}}$ ) selon une cinétique d'ordre 2

- Un mécanisme *trans* dans lequel une forme  $E_{low}$  est convertie en une forme  $E_{high}$  par une forme  $E_{high}$  préexistante selon une cinétique d'ordre 2. Ce mécanisme suppose que l'un des mécanismes précédents se produise également

Ces mécanismes ne sont pas exclusifs et d'autres pourraient être envisagés par combinaison de ces trois étapes élémentaires.

Les mécanismes *trans* sont généralement associés à des processus de coopérativité, d'autocatalyse ou d'auto-amplification conduisant à un comportement sigmoïdal et semblent plus appropriés pour décrire la maturation des caspases.

Afin d'élucider le mécanisme, les modèles suivants sont mis en œuvre et discriminés à l'aide du logiciel Dynafit (Tableau 9).

Dans tous les modèles, seules deux populations d'enzymes ont pu être considérées car la discrimination du modèle est effectuée à l'aide du seul paramètre d'activité enzymatique. Par conséquent, les premières étapes de la maturation, telles que la conversion de  $E_0$  en  $E_1$ , deux formes à faible activité, n'ont pas pu être différenciées.

Modèle	Réactions	Constantes	Ordre	Unités
<i>Cis</i>	$E_{low} \rightarrow E_{high}$	$k_1$	premier	jour <sup>-1</sup>
<i>Trans A</i>	$E_{low} + E_{low} \rightarrow E_{high} + E_{low}$	$k_1$	second	$\mu M^{-1} \cdot jour^{-1}$
<i>Trans B</i>	$E_{low} + E_{low} \rightarrow E_{high} + E_{high}$	$k_1$	second	$\mu M^{-1} \cdot jour^{-1}$
<i>Trans C</i>	$E_{low} + E_{low} \rightarrow E_{high} + E_{low}$	$k_1$	second	$\mu M^{-1} \cdot jour^{-1}$
	$E_{low} + E_{high} \rightarrow E_{high} + E_{high}$	$k_2$	second	
<i>Trans D</i>	$E_{low} + E_{low} \rightarrow E_{high} + E_{high}$	$k_1$	second	$\mu M^{-1} \cdot jour^{-1}$
	$E_{high} + E_{low} \rightarrow E_{high} + E_{high}$	$k_2$	second	
<i>Cis-Trans</i>	$E_{low} \rightarrow E_{high}$	$k_1$	premier	jour <sup>-1</sup>
	$E_{high} + E_{low} \rightarrow E_{high} + E_{high}$	$k_2$	second	$\mu M^{-1} \cdot jour^{-1}$

**Tableau 9. Modèles proposés décrivant la conversion de la forme  $E_{low}$  à la forme  $E_{high}$ .**

CgMCA-I (16,5  $\mu\text{M}$ ) a été stockée dans le tampon HEPES 10 mM pH 7,6, NaCl 150 mM à 22 °C en absence de calcium pendant 10 jours. L'activité de l'enzyme a été mesurée chaque jour à l'aide d'un spectrophotomètre TECAN Infinite M200 (Tecan, Salzbourg, Autriche) en utilisant des plaques 96 puits à fond noir. La réaction a été réalisée dans un tampon HEPES 10 mM pH 7,6, NaCl 150 mM contenant  $\text{CaCl}_2$  10 mM ainsi que le substrat Z-GGR-AMC à une concentration de 50  $\mu\text{M}$  (10  $\mu\text{L}$ ). Le milieu réactionnel est pré-incubé à 25 °C avant de débiter la réaction en ajoutant l'enzyme à une concentration de 2  $\mu\text{M}$  pour un volume final de 100  $\mu\text{L}$ . La libération du produit de la réaction (AMC) est suivie en mesurant sa fluorescence, après excitation à une longueur d'onde d'excitation de 355 nm (9 nm de bande passante), à une longueur d'onde d'émission de 460 nm (20 nm de bande passante). L'expérience a été réalisée au moins en duplicat ( $n=2$ ) et le blanc a été obtenu en remplaçant l'enzyme par du tampon. Les vitesses sont moyennées et sont exprimées en  $\mu\text{M}\cdot\text{min}^{-1}$ .

La vitesse de réaction est directement liée à la concentration de  $E_{\text{low}}$  et  $E_{\text{high}}$ , chaque forme ayant sa propre réponse spécifique :

$$\text{Vitesse } (\mu\text{M}\cdot\text{min}^{-1}) = r_{\text{low}} \times [E_{\text{low}}] + r_{\text{high}} \times [E_{\text{high}}]$$

où  $r_{\text{low}}$  et  $r_{\text{high}}$  sont respectivement les réponses spécifiques de  $E_{\text{low}}$  et  $E_{\text{high}}$ , initialement fixées à  $5 \times 10^{-6} \mu\text{M}\cdot\text{min}^{-1}\cdot\mu\text{M}^{-1}$  et à  $30 \times 10^{-6} \mu\text{M}\cdot\text{min}^{-1}\cdot\mu\text{M}^{-1}$  et où  $[E_{\text{low}}]$  et  $[E_{\text{high}}]$  sont respectivement les concentrations molaires de  $E_{\text{low}}$  et  $E_{\text{high}}$ .

La concentration initiale de  $E_{\text{low}}$  est fixée à 2  $\mu\text{M}$  et n'est pas ajustée aux conditions expérimentales.

Les paramètres qui peuvent être ajustés sont les réponses spécifiques ( $r_{\text{low}}$  et  $r_{\text{high}}$ ) et les constantes cinétiques ( $k_1$  et  $k_2$ ).

La discrimination des modèles a été effectuée en utilisant trois jeux de paramètres initiaux pour chaque modèle, ce qui a conduit à 18 ajustements principaux différents (Tableau 10).

Jeu de paramètres	1		2		3	
$k_1$	0,001	variable	0,001	variable	0,001-0,1	fixée <sup>a</sup>
$k_2$	1	variable	1	variable	1	variable
$r_{low}$	$5 \times 10^{-6}$	variable	$5 \times 10^{-6}$	fixée	$5 \times 10^{-6}$	variable
$r_{high}$	$30 \times 10^{-6}$	variable	$30 \times 10^{-6}$	fixée	$30 \times 10^{-6}$	variable

**Tableau 10. Jeux de paramètres initiaux pour la discrimination du modèle.** <sup>a</sup>: plusieurs ajustements sont effectués avec des valeurs fixes de  $k_1$ : 0,001 ; 0,01 et 0,1 (voir tableau 5 pour les unités).

Le logiciel Dynafit possède un mode de discrimination permettant la détermination du modèle d'activation de l'enzyme le plus acceptable. Cette discrimination est notamment réalisée en calculant l'AIC (« Akaike Information Criteria »), un paramètre compris entre 0 et 1 étant une estimation de la qualité relative d'un modèle statistique pour un ensemble de données. Une explication détaillée de la méthodologie de discrimination se trouve dans le manuel du logiciel ou dans les références suivantes : (Wei *et al.*, 2016; Policarpo *et al.*, 2019).

Le modèle *Cis-Trans* utilise le jeu de paramètres 3 où la constante  $k_1$  est fixée à 0,001 jour<sup>-1</sup>.

## D. Analyse structurale

### 1. *Analyse par diffraction aux rayons X*

#### a) *Cristallogénèse*

Le criblage des conditions de cristallisation a été réalisé en utilisant la technique de diffusion de vapeur en gouttes assises à l'aide de kits commerciaux de cristallisation : MB Class suite (Qiagen), MB Class II (Qiagen), PEG I (Qiagen), PEG II (Qiagen), PEG ION (Hampton Research), Crystal screen (Hampton Research), MPD (Qiagen), MIDAS+ (Molecular Dimensions) and JCSG+ suite (Qiagen). Des plaques 96 puits MRC (Molecular Dimensions) ont été remplies en utilisant le robot de cristallisation Mosquito® (STP Labtech) en effectuant deux ratios de protéines/agent de cristallisation (200 nL + 200 nL et 200 nL + 100 nL équilibré contre 70  $\mu$ L de solution de cristallisation dans le réservoir). Parmi le criblage des conditions de cristallisation, seules deux préparations de CgMCA-I ont permis d'obtenir des cristaux dans deux conditions différentes. La première, a permis d'obtenir des cristaux en complexe avec du calcium (CgMCA-I<sup>Ca</sup>), à partir de CgMCA-I au jour 0 de maturation et concentrée à 10,3 mg.mL<sup>-1</sup> dans le tampon 10 mM HEPES pH 7,6, 150 mM NaCl, 10 mM



CaCl<sub>2</sub>, dans une condition contenant 0,2 M de sulfate d'ammonium ((NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>), et 30 % (m/v) de PEG 4000. La deuxième préparation, a permis d'obtenir des cristaux en complexe avec du magnésium (CgMCA-I<sup>Mg</sup>), à partir de CgMCA-I au jour 4 de maturation concentrée à 8,2 mg.mL<sup>-1</sup> dans le tampon 10 mM HEPES pH 7,6, 150 mM NaCl, 10 mM CaCl<sub>2</sub>, dans une condition contenant 0,2 M d'acétate de magnésium et 20 % (m/v) PEG 3350. Les cristaux ont été pêchés et cryo-protégés dans la solution de cristallisation à laquelle 15 % d'éthylène glycol a été rajouté.

b) *Enregistrement des jeux de données, résolution et affinement des structures*

Les données de diffraction aux rayons X des cristaux ont été collectées sur la ligne de lumière PXIII (SLS, Suisse) à une longueur d'onde de 0,99999 Å pour le jeu de données CgMCA-I<sup>Ca</sup> et sur la ligne de lumière PROXIMA1 (SOLEIL, Paris) à une longueur d'onde de 0,97857 Å pour le jeu de données CgMCA-I<sup>Mg</sup>. Les données ont été indexées, intégrées et mises à l'échelle en utilisant le logiciel XDS (Kabsch, 2010). Le problème des phases du jeu de données de CgMCA-I<sup>Ca</sup> a été résolu par remplacement moléculaire en utilisant phenix.phaser de la suite PHENIX (Adams *et al.*, 2010) à l'aide de la structure de la métacaspase de *Saccharomyces cerevisiae* (ScMCA-I, code PDB : 4F6O) utilisé comme modèle guide et possédant un pourcentage d'identité de séquence de 66,5 % (Wong, Yan and Shi, 2012). Pour CgMCA-I<sup>Mg</sup>, le modèle guide pour le remplacement moléculaire utilisé a été la structure de CgMCA-I<sup>Ca</sup>. Des cycles d'affinement du modèle ont été réalisés en utilisant phenix.phaser de la suite PHENIX et des corrections manuelles ont été apportées à l'aide du logiciel COOT (Emsley *et al.*, 2010). Les statistiques d'enregistrement, de traitement de données de diffraction et d'affinement de CgMCA-I<sup>Ca</sup> et CgMCA-I<sup>Mg</sup> sont indiquées dans le Tableau 3 (cf. II. Résultats, p 85). Le logiciel PyMol (Schrödinger, LLC) a été utilisé pour la représentation tridimensionnelle des structures.

2. *Analyse par diffusion des rayons X aux petits angles (SAXS)*

Des analyses SEC-SAXS ont été réalisées sur CgMCA-I en utilisant le couplage d'une colonne de gel filtration Superdex 200 5/150 GL (GE Healthcare) équilibrée dans un tampon HEPES 10 mM à pH 7,6 NaCl 150 mM et CaCl<sub>2</sub> 10 mM. Les échantillons de protéine purifiée et concentrée à 9,6 mg.mL<sup>-1</sup> sont injectés dans la colonne et les jeux de données ont été collectés sur la ligne de lumière SWING à Soleil (Paris-Saclay, France). Les données ont été traitées à l'aide de PRIMUS (Konarev *et al.*, 2003) et le calcul de la fonction P(r) a été réalisé en utilisant CRY SOL (Svergun, Barberato and Koch, 1995). Les modèles *ab initio* de CgMCA-I ont été générés en utilisant DAMMIF (Franke and Svergun, 2009). La

superposition du modèle avec la structure tridimensionnelle cristalline de CgMCA-I ont été réalisés avec CHIMERA (University of California, San Francisco).

## VI. Références

- Acosta-Maspons A, Sepúlveda-García E, Sánchez-Baldoquín L, Marrero-Gutiérrez J, Pons T, Rocha-Sosa M, González L. 2014. Two aspartate residues at the putative p10 subunit of a type II metacaspase from *Nicotiana tabacum* L. may contribute to the substrate-binding pocket. *Planta*. 239(1):147–160. doi:10.1007/s00425-013-1975-0.
- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr*. 66(Pt 2):213–221. doi:10.1107/S0907444909052925.
- Aravind L, Koonin EV. 2002. Classification of the caspase-hemoglobinase fold: detection of new families and implications for the origin of the eukaryotic separins. *Proteins*. 46(4):355–367. doi:10.1002/prot.10060.
- Basmaciyan L, Azas N, Casanova M. 2018. Different apoptosis pathways in *Leishmania* parasites. *Cell Death Discov*. 4(1):90, s41420-018-0092-z. doi:10.1038/s41420-018-0092-z.
- Büttner S, Ruli D, Vögtle F-N, Galluzzi L, Moitzi B, Eisenberg T, Kepp O, Habernig L, Carmona-Gutierrez D, Rockenfeller P, et al. 2011. A yeast BH3-only protein mediates the mitochondrial pathway of apoptosis. *EMBO J*. 30(14):2779–2792. doi:10.1038/emboj.2011.197.
- Cao YingYing, Huang S, Dai B, Zhu Z, Lu H, Dong L, Cao YongBing, Wang Y, Gao P, Chai Y, et al. 2009a. *Candida albicans* cells lacking *CaMCA1*-encoded metacaspase show resistance to oxidative stress-induced death and change in energy metabolism. *Fungal Genet Biol*. 46(2):183–189. doi:10.1016/j.fgb.2008.11.001.
- Carafoli E, Lehninger AL. 1971. A survey of the interaction of calcium ions with mitochondria from different tissues and species. *Biochem J*. 122(5):681–690.
- Carmona-Gutierrez D, Bauer MA, Zimmermann A, Aguilera A, Austriaco N, Ayscough K, Balzan R, Bar-Nun S, Barrientos A, Belenky P, et al. 2018. Guidelines and recommendations on yeast cell death nomenclature. *Microb Cell*. 5(1):4–31. doi:10.15698/mic2018.01.607.
- Carraro M, Bernardi P. 2016. Calcium and reactive oxygen species in regulation of the mitochondrial permeability transition and of programmed cell death in yeast. *Cell Calcium*. 60(2):102–107. doi:10.1016/j.ceca.2016.03.005.
- Chai J., Shiozaki E, Srinivasula SM, Wu Q, Datta P, Alnemri ES, Shi Y, Dataa P. 2001. Structural basis of caspase-7 inhibition by XIAP. *Cell*. 104(5):769–780. doi:10.1016/s0092-8674(01)00272-0.
- Chai Jijie, Wu Q, Shiozaki E, Srinivasula SM, Alnemri ES, Shi Y. 2001. Crystal Structure of a Procaspase-7 Zymogen: Mechanisms of Activation and Substrate Binding. *Cell*. 107(3):399–407. doi:10.1016/S0092-8674(01)00544-X.
- Couzinet A, Hérincs Z, Hueber A-O. 2002. Régulation de la mort cellulaire programmée : vers une conception plus dynamique. *médecine/sciences*. 18(8–9):841–852. doi:10.1051/medsci/20021889841.

- Czabotar PE, Lessene G, Strasser A, Adams JM. 2014. Control of apoptosis by the BCL-2 protein family: implications for physiology and therapy. *Nat Rev Mol Cell Biol.* 15(1):49–63. doi:10.1038/nrm3722.
- Du L, Su Y, Sun D, Zhu W, Wang J, Zhuang X, Zhou S, Lu Y. 2008. Formic acid induces Yca1p-independent apoptosis-like cell death in the yeast *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 8(4):531–539. doi:10.1111/j.1567-1364.2008.00375.x.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* 66(Pt 4):486–501. doi:10.1107/S0907444910007493.
- Falcone C, Mazzoni C. 2016. External and internal triggers of cell death in yeast. *Cell Mol Life Sci.* 73:2237–2250. doi:10.1007/s00018-016-2197-y.
- Fortin J, Lam E. 2018. Domain swap between two type-II metacaspases defines key elements for their biochemical properties. *Plant J Cell Mol Biol.* 96(5):921–936. doi:10.1111/tpj.14079.
- Franke D, Svergun DI. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J Appl Crystallogr.* 42(2):342–346. doi:10.1107/S0021889809000338.
- Fuentes-Prior P, Salvesen GS. 2004. The protein structures that shape caspase activity, specificity, activation and inhibition. *Biochem J.* 384(Pt 2):201–232. doi:10.1042/BJ20041142.
- Galluzzi L, Vitale I, Aaronson SA, Abrams JM, Adam D, Agostinis P, Alnemri ES, Altucci L, Amelio I, Andrews DW, et al. 2018. Molecular mechanisms of cell death: recommendations of the Nomenclature Committee on Cell Death 2018. *Cell Death Differ.* 25(3):486–541. doi:10.1038/s41418-017-0012-4.
- Gilio JM, Marcondes MF, Ferrari D, Juliano MA, Juliano L, Oliveira V, Machado MFM. 2017. Processing of metacaspase 2 from *Trypanosoma brucei* (TbMCA2) broadens its substrate specificity. *Biochim Biophys Acta Proteins Proteomics.* 1865(4):388–394. doi:10.1016/j.bbapap.2017.01.002.
- Guaragnella N, Ždralović M, Antonacci L, Passarella S, Marra E, Giannattasio S. 2012. The role of mitochondria in yeast programmed cell death. *Front Oncol.* 2:70. doi:10.3389/fonc.2012.00070.
- Heckman KL, Pease LR. 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc.* 2(4):924–932. doi:10.1038/nprot.2007.132.
- Herker E, Jungwirth H, Lehmann KA, Maldener C, Fröhlich K-U, Wissing S, Büttner S, Fehr M, Sigrist S, Madeo F. 2004. Chronological aging leads to apoptosis in yeast. *J Cell Biol.* 164(4):501–507. doi:10.1083/jcb.200310014.
- Hill SM, Nyström T. 2015. The dual role of a yeast metacaspase: What doesn't kill you makes you stronger. *Bioessays.* 37(5):525–531. doi:10.1002/bies.201400208.
- Horn S, Hughes MA, Schilling R, Sticht C, Tenev T, Ploesser M, Meier P, Sprick MR, MacFarlane M, Leverkus M. 2017. Caspase-10 Negatively Regulates Caspase-8-Mediated Cell Death, Switching the Response to CD95L in Favor of NF-κB Activation and Cell Survival. *Cell Rep.* 19(4):785–797. doi:10.1016/j.celrep.2017.04.010.

- Ivanovska I, Hardwick JM. 2005. Viruses activate a genetically conserved cell death pathway in a unicellular organism. *J Cell Biol.* 170(3):391–399. doi:10.1083/jcb.200503069.
- Julien O, Wells JA. 2017. Caspases and their substrates. *Cell Death Differ.* 24(8):1380–1389. doi:10.1038/cdd.2017.44.
- Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Židek A, Potapenko A, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature.* 596(7873):583–589. doi:10.1038/s41586-021-03819-2.
- Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr.* 66(2):125–132. doi:10.1107/S09074444909047337.
- Kerr JFR, Wyllie AH, Currie AR. 1972. Apoptosis: A Basic Biological Phenomenon with Wide-ranging Implications in Tissue Kinetics. *Br J Cancer.* 26(4):239–257.
- Klemenčič M, Funk C. 2018. Structural and functional diversity of caspase homologues in non-metazoan organisms. *Protoplasma.* 255(1):387–397. doi:10.1007/s00709-017-1145-5.
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr.* 36(5):1277–1282. doi:https://doi.org/10.1107/S0021889803012779.
- Kruman I, Guo Q, Mattson MP. 1998. Calcium and reactive oxygen species mediate staurosporine-induced mitochondrial dysfunction and apoptosis in PC12 cells. *J Neurosci Res.* 51(3):293–308. doi:10.1002/(SICI)1097-4547(19980201)51:3<293::AID-JNR3>3.0.CO;2-B.
- Kumar A, Pandey SC, Samant M. 2018. Slow pace of antileishmanial drug development. *Parasitol Open.* 4:e4. doi:10.1017/pao.2018.1.
- Lamkanfi M, Declercq W, Kalai M, Saelens X, Vandenameele P. 2002. Alice in caspase land. A phylogenetic analysis of caspases from worm to man. *Cell Death Differ.* 9(4):358–361. doi:10.1038/sj.cdd.4400989.
- Lamy L, Ngo VN, Emre NCT, Shaffer AL, Yang Y, Tian E, Nair V, Kruhlak MJ, Zingone A, Landgren O, et al. 2013. Control of Autophagic Cell Death by Caspase-10 in Multiple Myeloma. *Cancer Cell.* 23(4):435–449. doi:10.1016/j.ccr.2013.02.017.
- Lavrik IN, Golks A, Krammer PH. 2005. Caspases: pharmacological manipulation of cell death. *J Clin Invest.* 115(10):2665–2672. doi:10.1172/JCI26252.
- Lee REC, Brunette S, Puente LG, Megeney LA. 2010. Metacaspase Yca1 is required for clearance of insoluble protein aggregates. *Proc Natl Acad Sci.* 107(30):13348–13353. doi:10.1073/pnas.1006610107.
- Lee REC, Puente LG, Kærn M, Megeney LA. 2008. A Non-Death Role of the Yeast Metacaspase: Yca1p Alters Cell Cycle Dynamics. Cookson MR, editor. *PLoS ONE.* 3(8):e2956. doi:10.1371/journal.pone.0002956.
- Léger T, Garcia C, Ounissi M, Lelandais G, Camadro J-M. 2015. The Metacaspase (Mca1p) has a Dual Role in Farnesol-induced Apoptosis in *Candida albicans*. *Mol Cell Proteomics.* 14(1):93–108. doi:10.1074/mcp.M114.041210.

- Li J, Yuan J. 2008. Caspases in apoptosis and beyond. *Oncogene*. 27(48):6194–6206. doi:10.1038/onc.2008.297.
- Li P, Nijhawan D, Budihardjo I, Srinivasula SM, Ahmad M, Alnemri ES, Wang X. 1997. Cytochrome c and dATP-Dependent Formation of Apaf-1/Caspase-9 Complex Initiates an Apoptotic Protease Cascade. *Cell*. 91(4):479–489. doi:10.1016/S0092-8674(00)80434-1.
- Liang Q, Zhou B. 2007. Copper and Manganese Induce Yeast Apoptosis via Different Pathways. Newmeyer D, editor. *Mol Biol Cell*. 18(12):4741–4749. doi:10.1091/mbc.e07-05-0431.
- Ludovico P, Rodrigues F, Almeida A, Silva MT, Barrientos A, Côrte-Real M. 2002. Cytochrome c Release and Mitochondria Involvement in Programmed Cell Death Induced by Acetic Acid in *Saccharomyces cerevisiae*. *Mol Biol Cell*. 13(8):2598–2606. doi:10.1091/mbc.E01-12-0161.
- Ludovico P, Sousa MJ, Silva MT, Leão CL, Côrte-Real M. 2001. *Saccharomyces cerevisiae* commits to a programmed cell death process in response to acetic acid. *Microbiol Read Engl*. 147(Pt 9):2409–2415. doi:10.1099/00221287-147-9-2409.
- Machado MFM, Marcondes MF, Juliano MA, McLuskey K, Mottram JC, Moss CX, Juliano L, Oliveira V. 2013. Substrate specificity and the effect of calcium on *Trypanosoma brucei* metacaspase 2. *FEBS J*. 280(11):2608–2621. doi:10.1111/febs.12248.
- Madeo F, Carmona-Gutierrez D, Ring J, Büttner S, Eisenberg T, Kroemer G. 2009. Caspase-dependent and caspase-independent cell death pathways in yeast. *Biochem Biophys Res Commun*. 382(2):227–231. doi:10.1016/j.bbrc.2009.02.117.
- Madeo F, Fröhlich E, Fröhlich K-U. 1997. A Yeast Mutant Showing Diagnostic Markers of Early and Late Apoptosis. *J Cell Biol*. 139(3):729–734.
- Madeo F, Fröhlich E, Ligr M, Grey M, Sigrist SJ, Wolf DH, Fröhlich K-U. 1999. Oxygen Stress: A Regulator of Apoptosis in Yeast. *J Cell Biol*. 145(4):757–767.
- Madeo F, Herker E, Maldener C, Wissing S, Lächelt S, Herlan M, Fehr M, Lauber K, Sigrist SJ, Wesselborg S, et al. 2002. A Caspase-Related Protease Regulates Apoptosis in Yeast. *Mol Cell*. 9(4):911–917. doi:10.1016/S1097-2765(02)00501-4.
- Mazzoni C, Herker E, Palermo V, Jungwirth H, Eisenberg T, Madeo F, Falcone C. 2005. Yeast caspase 1 links messenger RNA stability to apoptosis in yeast. *EMBO Rep*. 6(11):1076–1081. doi:10.1038/sj.embor.7400514.
- McLuskey K, Mottram JC. 2015. Comparative structural analysis of the caspase family with other clan CD cysteine peptidases. *Biochem J*. 466(Pt 2):219–232. doi:10.1042/BJ20141324.
- McLuskey K, Rudolf J, Proto WR, Isaacs NW, Coombs GH, Moss CX, Mottram JC. 2012. Crystal structure of a *Trypanosoma brucei* metacaspase. *Proc Natl Acad Sci*. 109(19):7469–7474. doi:10.1073/pnas.1200885109.
- Minina EA, Coll NS, Tuominen H, Bozhkov PV. 2017. Metacaspases versus caspases in development and cell fate regulation. *Cell Death Differ*. 24(8):1314–1325. doi:10.1038/cdd.2017.18.



- Minina EA, Staal J, Alvarez VE, Berges JA, Berman-Frank I, Beyaert R, Bidle KD, Bornancin F, Casanova M, Cazzulo JJ, et al. 2020. Classification and nomenclature of metacaspases and paracaspases: no more confusion with caspases. *Mol Cell*. 77(5):927–929. doi:10.1016/j.molcel.2019.12.020.
- Moss CX, Westrop GD, Juliano L, Coombs GH, Mottram JC. 2007. Metacaspase 2 of *Trypanosoma brucei* is a calcium-dependent cysteine peptidase active without processing. *FEBS Lett*. 581(29):5635–5639. doi:10.1016/j.febslet.2007.11.009.
- Nagata S. 2018. Apoptosis and Clearance of Apoptotic Cells. *Annu Rev Immunol*. 36(1):489–517. doi:10.1146/annurev-immunol-042617-053010.
- Orrenius S, Zhivotovsky B, Nicotera P. 2003. Regulation of cell death: the calcium–apoptosis link. *Nat Rev Mol Cell Biol*. 4(7):552–565. doi:10.1038/nrm1150.
- Pereira C, Silva RD, Saraiva L, Johansson B, Sousa MJ, Côrte-Real M. 2008. Mitochondria-dependent apoptosis in yeast. *Biochim Biophys Acta BBA - Mol Cell Res*. 1783(7):1286–1302. doi:10.1016/j.bbamcr.2008.03.010.
- Phillips AJ, Crowe JD, Ramsdale M. 2006. Ras pathway signaling accelerates programmed cell death in the pathogenic fungus *Candida albicans*. *Proc Natl Acad Sci*. 103(3):726–731. doi:10.1073/pnas.0506405103.
- Phillips AJ, Sudbery I, Ramsdale M. 2003. Apoptosis induced by environmental stresses and amphotericin B in *Candida albicans*. *Proc Natl Acad Sci*. 100(24):14327–14332. doi:10.1073/pnas.2332326100.
- Pinton P, Ferrari D, Magalhães P, Schulze-Osthoff K, Di Virgilio F, Pozzan T, Rizzuto R. 2000. Reduced Loading of Intracellular Ca<sup>2+</sup> Stores and Downregulation of Capacitative Ca<sup>2+</sup> Influx in Bcl-2–Overexpressing Cells. *J Cell Biol*. 148(5):857–862.
- Pinton P, Ferrari D, Rapizzi E, Virgilio FD, Pozzan T, Rizzuto R. 2001. The Ca<sup>2+</sup> concentration of the endoplasmic reticulum is a key determinant of ceramide-induced apoptosis: significance for the molecular mechanism of Bcl-2 action. *EMBO J*. 20(11):2690–2701. doi:10.1093/emboj/20.11.2690.
- Pinton P, Giorgi C, Siviero R, Zecchini E, Rizzuto R. 2008. Calcium and apoptosis: ER-mitochondria Ca<sup>2+</sup> transfer in the control of apoptosis. *Oncogene*. 27(50):6407–6418. doi:10.1038/onc.2008.308.
- Policarpo RL, Decultot L, May E, Kuzmič P, Carlson S, Huang D, Chu V, Wright BA, Dhakshinamoorthy S, Kannt A, et al. 2019. High-Affinity Alkynyl Bisubstrate Inhibitors of Nicotinamide N-Methyltransferase (NNMT). *J Med Chem*. 62(21):9837–9873. doi:10.1021/acs.jmedchem.9b01238.
- Pop C, Salvesen GS. 2009. Human Caspases: Activation, Specificity, and Regulation. *J Biol Chem*. 284(33):21777–21781. doi:10.1074/jbc.R800084200.
- Pressman BC. 1976. Biological Applications of Ionophores. *Annu Rev Biochem*. 45(1):501–530. doi:10.1146/annurev.bi.45.070176.002441.
- Ramirez MLG, Salvesen GS. 2018. A primer on caspase mechanisms. *Semin Cell Dev Biol*. 82:79–85. doi:10.1016/j.semcdb.2018.01.002.

- Rawlings ND, Barrett AJ, Thomas PD, Huang X, Bateman A, Finn RD. 2018. The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res.* 46(D1):D624–D632. doi:10.1093/nar/gkx1134.
- Rebeaud F, Hailfinger S, Posevitz-Fejfar A, Tapernoux M, Moser R, Rueda D, Gaide O, Guzzardi M, Iancu EM, Rufer N, et al. 2008. The proteolytic activity of the paracaspase MALT1 is key in T cell activation. *Nat Immunol.* 9(3):272–281. doi:10.1038/ni1568.
- Reiter J, Herker E, Madeo F, Schmitt MJ. 2005. Viral killer toxins induce caspase-mediated apoptosis in yeast. *J Cell Biol.* 168(3):353–358. doi:10.1083/jcb.200408071.
- Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42:W320–324. doi:10.1093/nar/gku316.
- Scariot FJ, Jahn LM, Maianti JP, Delamare APL, Echeverrigaray S. 2016. The fungicide Mancozeb induces metacaspase-dependent apoptotic cell death in *Saccharomyces cerevisiae* BY4741. *Apoptosis Int J Program Cell Death.* 21(7):866–872. doi:10.1007/s10495-016-1251-4.
- Sleich K, Warnken U, Fricker N, Öztürk S, Richter P, Kammerer K, Schnölzer M, Krammer PH, Lavrik IN. 2012. Stoichiometry of the CD95 Death-Inducing Signaling Complex: Experimental and Modeling Evidence for a Death Effector Domain Chain Model. *Mol Cell.* 47(2):306–319. doi:10.1016/j.molcel.2012.05.006.
- Scorrano L, Oakes SA, Opferman JT, Cheng EH, Sorcinelli MD, Pozzan T, Korsmeyer SJ. 2003. BAX and BAK Regulation of Endoplasmic Reticulum Ca<sup>2+</sup>: A Control Point for Apoptosis. *Science.* 300(5616):135–139. doi:10.1126/science.1081208.
- Sharp PM, Li WH. 1987. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 15(3):1281–1295.
- Shi J, Zhao Y, Wang K, Shi X, Wang Y, Huang H, Zhuang Y, Cai T, Wang F, Shao F. 2015. Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature.* 526(7575):660–665. doi:10.1038/nature15514.
- Shi Y. 2002. Mechanisms of Caspase Activation and Inhibition during Apoptosis. *Mol Cell.* 9(3):459–470. doi:10.1016/S1097-2765(02)00482-3.
- Shirazi F, Kontoyiannis D. 2015. Micafungin triggers caspase-dependent apoptosis in *Candida albicans* and *Candida parapsilosis* biofilms, including caspofungin non-susceptible isolates. *Virulence.* 6(4):385–394. doi:10.1080/21505594.2015.1027479.
- Shirliff ME, Krom BP, Meijering RAM, Peters BM, Zhu J, Scheper MA, Harris ML, Jabra-Rizk MA. 2009. Farnesol-induced apoptosis in *Candida albicans*. *Antimicrob Agents Chemother.* 53(6):2392–2401. doi:10.1128/AAC.01551-08.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol.* 7:539. doi:10.1038/msb.2011.75.



- Silva A, Almeida B, Sampaio-Marques B, Reis MIR, Ohlmeier S, Rodrigues F, Vale A do, Ludovico P. 2011. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) is a specific substrate of yeast metacaspase. *Biochim Biophys Acta*. 1813(12):2044–2049. doi:10.1016/j.bbamcr.2011.09.010.
- Strasser A, Jost PJ, Nagata S. 2009. The Many Roles of FAS Receptor Signaling in the Immune System. *Immunity*. 30(2):180–192. doi:10.1016/j.immuni.2009.01.001.
- Svergun D, Barberato C, Koch MHJ. 1995. *CRY SOL* – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J Appl Crystallogr*. 28(6):768–773. doi:10.1107/S0021889895007047.
- Thornberry NA, Bull HG, Calaycay JR, Chapman KT, Howardt AD, Kostura MJ, Millert DK, Molineaux SM, Weidner JR, Elliston KO, et al. 1992. A novel heterodimeric cysteine protease is required for interleukin-1 $\beta$  processing in monocytes. 356:7.
- Thornberry NA, Rano TA, Peterson EP, Rasper DM, Timkey T, Garcia-Calvo M, Houtzager VM, Nordstrom PA, Roy S, Vaillancourt JP, et al. 1997. A Combinatorial Approach Defines Specificities of Members of the Caspase Family and Granzyme B. *J Biol Chem*. 272(29):17907–17911. doi:10.1074/jbc.272.29.17907.
- Tsiatsiani L, Van Breusegem F, Gallois P, Zavialov A, Lam E, Bozhkov PV. 2011. Metacaspases. *Cell Death Differ*. 18(8):1279–1288. doi:10.1038/cdd.2011.66.
- Uren AG, O'Rourke K, Aravind L, Pisabarro MT, Seshagiri S, Koonin EV, Dixit VM. 2000. Identification of Paracaspases and Metacaspases: Two Ancient Families of Caspase-like Proteins, One of which Plays a Key Role in MALT Lymphoma. *Mol Cell*. 6(4):961–967. doi:10.1016/S1097-2765(05)00086-9.
- Van Opdenbosch N, Lamkanfi M. 2019. Caspases in Cell Death, Inflammation, and Disease. *Immunity*. 50(6):1352–1364. doi:10.1016/j.immuni.2019.05.020.
- Vandana, Dixit R, Tiwari R, Katyal A, Pandey KC. 2019. Metacaspases: Potential Drug Target Against Protozoan Parasites. *Front Pharmacol*. 10. doi:10.3389/fphar.2019.00790.
- Vercammen D, van de Cotte B, De Jaeger G, Eeckhout D, Casteels P, Vandepoele K, Vandenberghe I, Van Beeumen J, Inzé D, Van Breusegem F. 2004. Type II Metacaspases Atmc4 and Atmc9 of *Arabidopsis thaliana* Cleave Substrates after Arginine and Lysine\*. *J Biol Chem*. 279(44):45329–45336. doi:10.1074/jbc.M406329200.
- Walker NP, Talanian RV, Brady KD, Dang LC, Bump NJ, Ferez CR, Franklin S, Ghayur T, Hackett MC, Hammill LD. 1994. Crystal structure of the cysteine protease interleukin-1 $\beta$ -converting enzyme: a (p20/p10) $_2$  homodimer. *Cell*. 78(2):343–352. doi:10.1016/0092-8674(94)90303-4.
- Walter D, Matter A, Fahrenkrog B. 2010. Bre1p-mediated histone H2B ubiquitylation regulates apoptosis in *Saccharomyces cerevisiae*. *J Cell Sci*. 123(11):1931–1939. doi:10.1242/jcs.065938.
- Watanabe N, Lam E. 2005. Two *Arabidopsis* metacaspases AtMCP1b and AtMCP2b are arginine/lysine-specific cysteine proteases and activate apoptosis-like cell death in yeast. *J Biol Chem*. 280(15):14691–14699. doi:10.1074/jbc.M413527200.

- Watanabe N, Lam E. 2011. Calcium-dependent activation and autolysis of *Arabidopsis* metacaspase 2d. *J Biol Chem*. 286(12):10027–10040. doi:10.1074/jbc.M110.194340.
- Wei Y, Kuzmič P, Yu R, Modi G, Hedstrom L. 2016. Inhibition of IMPDH from *Bacillus anthracis*: Mechanism revealed by pre-steady state kinetics. *Biochemistry*. 55(37):5279–5288. doi:10.1021/acs.biochem.6b00265.
- Weinberger M. 2005. Apoptosis in budding yeast caused by defects in initiation of DNA replication. *J Cell Sci*. 118(15):3543–3553. doi:10.1242/jcs.02477.
- Wilson KP, Black JA, Thomson JA, Kim EE, Griffith JP, Navia MA, Murcko MA, Chambers SP, Aldape RA, Raybuck SA. 1994. Structure and mechanism of interleukin-1 beta converting enzyme. *Nature*. 370(6487):270–275. doi:10.1038/370270a0.
- Wong AH-H, Yan C, Shi Y. 2012. Crystal Structure of the Yeast Metacaspase Yca1 \*♦. *J Biol Chem*. 287(35):29251–29259. doi:10.1074/jbc.M112.381806.
- Yuan J, Shaham S, Ledoux S, Ellis HM, Horvitz HR. 1993. The *C. elegans* cell death gene *ced-3* encodes a protein similar to mammalian interleukin-1 beta-converting enzyme. *Cell*. 75(4):641–652. doi:10.1016/0092-8674(93)90485-9.
- Yuan S, Akey CW. 2013. Apoptosome Structure, Assembly, and Procaspase Activation. *Structure*. 21(4):501–515. doi:10.1016/j.str.2013.02.024.
- Zhang Y, Lam E. 2011. Sheathing the swords of death. *Plant Signal Behav*. 6(12):2051–2056. doi:10.4161/psb.6.12.18247.
- Zhu P, Yu X-H, Wang C, Zhang Q, Liu W, McSweeney S, Shanklin J, Lam E, Liu Q. 2020. Structural basis for Ca<sup>2+</sup>-dependent activation of a plant metacaspase. *Nat Commun*. 11(1):2249. doi:10.1038/s41467-020-15830-8.



## **CHAPITRE II**

**Caractérisation biochimique et structurale de l'enzyme  
branchante de glucane en  $\alpha$ -(1,4) chez l'organisme pathogène  
*Candida glabrata***



## Liste des figures et des tableaux

### FIGURES

1. **Arbre phylogénétique de la famille des GH13** p 158
2. **Arbre phylogénétique d'enzymes branchantes de la sous-famille GH13\_9** p 159
3. **Représentation schématique simplifiée de la voie de synthèse et de dégradation du glycogène** p 161
4. **Représentation schématique de l'organisation en granules du glycogène et de l'amidon** p 162
5. **Représentation simplifiée du mécanisme à double déplacement avec rétention de la configuration  $\alpha$  du carbone anomérique dans le cas d'une hydrolyse ou d'une transglycolyse** p 165
6. **Organisation des domaines de *EcBE*** p 167
7. **Structure tridimensionnelle de *Ec* $\Delta$ 113BE tronquée des 113 premiers acides aminés** p 167
8. **Structure tridimensionnelle de *Ec* $\Delta$ 113BE et les différents sites de liaisons aux sucres** p 168
9. **Organisation des domaines de *MtBE*** p 169
10. **Structure tridimensionnelle de *MtBE*** p 170
11. **Organisation des domaines de *CsBE*** p 170
12. **Structure tridimensionnelle de *CsBE*** p 171
13. **Structure tridimensionnelle du mutant W610N en complexe avec un maltoheptaose** p 172
14. **Représentation schématique du modèle de fixation des chaînes de glucanes sur *CsBE*** p 173
15. **Organisation des domaines de *RoBE*** p 174
16. **Structure tridimensionnelle de *RoBE*** p 175
17. **Organisation des domaines de *OsBEI*** p 176
18. **Structure tridimensionnelle de *OsE399Q* $\Delta$ .BE en complexe avec du maltopentaose** p 177
19. **Organisation des domaines de *HsBE*** p. 178
20. **Structure tridimensionnelle de *Hs* $\Delta$ 38BE** p 179



## TABLEAUX

1. **Résumé des informations structurales obtenues dans la littérature pour *EcBE*, *MtBE*, *CsBE*, *HsBE*, *OsBE*** p 180





# I. L'enzyme branchante de glucane en $\alpha$ -(1,4) de *Candida glabrata*

## A. Les glycoside hydrolases

Les sucres sont les biomolécules les plus abondantes sur terre et jouent un rôle important contribuant à la diversité des organismes vivants. Ces molécules peuvent participer à la composition des parois fongiques, être impliqués dans la reconnaissance et les interactions cellule-cellule mais aussi avoir un rôle dans la mise en place de la virulence et de l'évasion immunitaire de certains pathogènes humain (Ardèvol and Rovira, 2015). La large variété de sucres présents dans la nature découle de l'action d'un groupe d'enzymes nommées « Carbohydrate-Active Enzymes » ou « CAZymes » (Lombard *et al.*, 2014).

### 1. *Classification*

A la fin des années 1950, la découverte d'un nombre croissant d'enzymes agissant sur les sucres a nécessité la mise en place d'un système de classification et d'une nomenclature unique. Ainsi, en 1956, l'Union Internationale de Biochimie et de Biologie Moléculaire (« International Union of Biochemistry et Molecular Biology »; IUB-MB) crée une Commission d'Enzymes (« Enzyme Commission »; EC). Cette nomenclature divise les enzymes en 7 sous-groupes en fonction de leur activité catalytique et de leur spécificité de substrat : les oxydoréductases (EC. 1), les transférases (EC. 2), les hydrolases (EC. 3), les lyases (EC. 4), les isomérases (EC.5), les ligases (EC. 6) et les translocases (EC. 7). Les glycoside hydrolases, par exemple, sont classées avec le numéro EC 3.2.1.x, le chiffre 2 correspondant aux glycosidases, le chiffre 1 aux glycosidases pouvant hydrolyser des composés O- ou S- glycosylés et x représentant la spécificité de substrat. Cependant, bien que ce système soit toujours utilisé, il ne reflète pas les caractéristiques structurales et mécanistiques des enzymes. Dans les années 1990, un autre système basé sur la similarité de séquences en acides aminés est mis au point et donnera naissance à la base de données des enzymes actives sur les sucres CAZy (« Carbohydrate Active enZyme ») ([www.cazy.org](http://www.cazy.org)) (Lombard *et al.*, 2014). Cette dernière recense tous les modules catalytiques et fonctionnels des enzymes synthétisant, dégradant ou modifiant des liaisons glycosidiques. Cette classification reflète non seulement la similarité structurale et mécanistique mais permet également d'établir des relations évolutives puisque des enzymes ayant des spécificités de substrat différentes peuvent être classées au sein d'une même famille.

Actuellement on peut retrouver sur CAZy la classification des glycoside hydrolases (« Glycoside Hydrolases » ; GHs), glycoside transférases (« GlycosylTransferases » ; GTs), des polysaccharide lyases (« Polysaccharide Lyases » ; PLs), des estérases (« Carbohydrate Esterases » ; CEs), des domaines annexes non catalytiques étant des modules de liaison aux sucres ou CBM (« Carbohydrate Binding Module ») et des enzymes ayant des activités dites auxiliaires (« Auxiliary Activities » ; AAs) car n'agissant pas directement sur des sucres mais en association avec les autres classes d'enzymes.

Les glycosides hydrolases sont des enzymes hydrolysant des liaisons glycosidiques qui constituent actuellement 173 familles dans CAZy (février 2022) et dont certaines familles peuvent présenter jusqu'à 56 sous-familles (Henrissat, 1991; Henrissat and Bairoch, 1993; Bourne and Henrissat, 2001). L'identification de ce grand nombre d'enzymes et la résolution de leur structure tridimensionnelle associée a mené à la création d'un sous classement des familles par clans, de A à R. En effet, le repliement structural pouvant être plus conservé que la séquence, un même clan peut contenir des enzymes de diverses familles. Ce regroupement dans un même clan peut également traduire une origine évolutive commune entre les enzymes (Davies and Henrissat, 1995).

#### a) *La famille des GH13*

La famille des GH13, historiquement appelée famille des  $\alpha$ -amylases, est une famille majeure des GHs contenant à ce jour 124 377 enzymes (décembre 2021). Parmi elles on retrouve un grand nombre d'activités catalytiques :  $\alpha$ -amylase (EC 3.2.1.1), pullulanase (EC 3.2.1.41), cyclomaltodextrine glucanotransférase (EC 2.4.1.19), cyclomaltodextrinase (EC 3.2.1.54), tréhalose-6-phosphate hydrolase (EC 3.2.1.93), oligo- $\alpha$ -glucosidase (EC 3.2.1.10), amylase maltogénique (EC 3.2.1.133), neopullulanase (EC 3.2.1.135),  $\alpha$ -glucosidase (C 3.2.1.20), isoamylase (EC 3.2.1.68), glucodextranase (EC 3.2.1.70), enzyme branchante (EC 3.2.1.18), trehalose synthase (EC 5.4.99.16), 4- $\alpha$ -glucanotransferase (EC 2.4.1.25), amylosucrase (EC 3.2.1.33), sucrose phosphorylase (EC 2.4.1.7), malto-oligosyltrehalose trehalohydrolase (EC 3.2.1.141), isomaltulose/trehalulose synthase (EC 5.4.99.11), malto-oligosyltrehalose synthase (EC 5.4.99.15), amylo- $\alpha$ -1,6-glucosidase (EC 3.2.1.33),  $\alpha$ -1,4-glucane phosphate  $\alpha$ -maltosyltransferase (EC 2.4.99.16), 6(F)-phosphate phosphorylase (EC 2.4.1.329), glucosylglycerol phosphorylase (EC 2.4.1.359), glucosylglycerate phosphorylase (EC 2.4.1.352), sucrose  $\alpha$ -glucosidase (EC 3.2.1.48), oligosaccharide  $\alpha$ -4-glucosyltransferase (EC 2.4.1.161).

La grande diversité en termes d'activité a induit la création de 44 sous-familles (février 2022) qui possèdent cependant un mécanisme catalytique commun à double déplacement et rétention de la configuration du carbone anomérique (Koshland, 1953; Stam *et al.*, 2006).

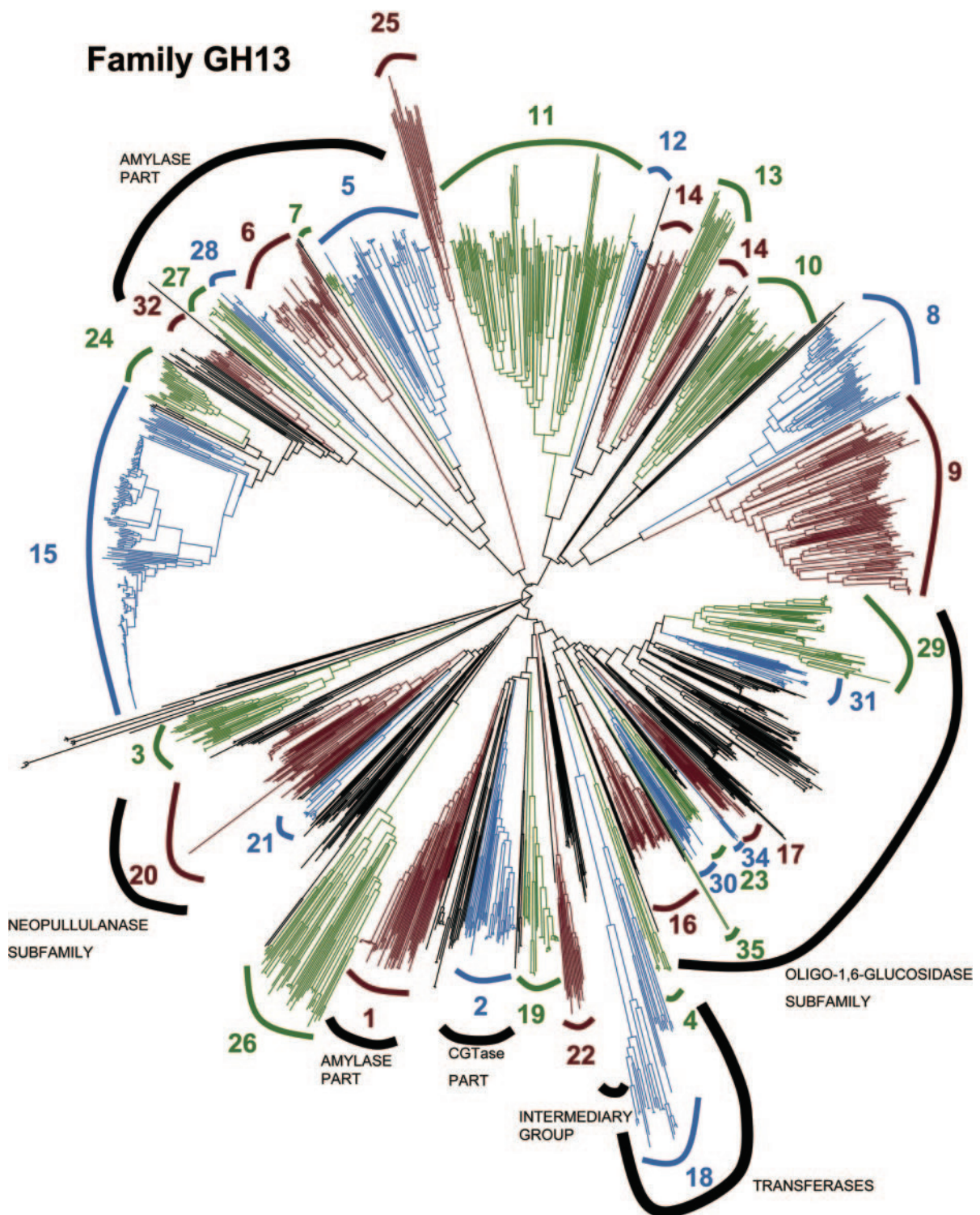
Bien que certaines enzymes de cette famille possèdent peu d'identité de séquence entre elles, les 145 structures tridimensionnelles aujourd'hui référencées dans CAZy ont permis d'observer un repliement structural commun. La famille des GH13, contenant des enzymes agissant sur des liaisons  $\alpha$ -glucosidiques, a donc été classée dans le clan GH-H avec les familles GH70 et GH77 du fait qu'elles partagent un repliement du domaine catalytique de type tonneau ( $\beta/\alpha$ )<sub>8</sub> ainsi qu'un domaine de liaison aux sucres 48 (CBM48). D'un point de vue hiérarchique, le regroupement en clan est la plus haute classification de CAZy puisque les familles d'enzymes d'un même clan sont censées partager un ancêtre commun et une même machinerie catalytique (Stam *et al.*, 2006). Pour le clan GH-H, la conservation du repliement structural est également accompagnée d'une conservation de trois acides aminés catalytiques essentiels : Asp (nucléophile), Glu (acide/base), Asp (stabilisateur de l'état de transition de type ion oxocarbénium).

#### b) *Les sous-familles GH13\_8 et GH13\_9*

Au sein de la grande famille des GH13 se trouvent les sous-familles 8 et 9 regroupant des enzymes classées comme enzymes branchantes de glucane en  $\alpha$ -(1,4) (« branching enzymes » ; BEs) (EC 2.4.1.18). De précédentes analyses de phylogénie ont déterminé que ces deux sous-familles appartenaient à une même branche d'un point de vue phylogénétique avec chaque sous-famille formant un groupe reflétant en grande partie la taxonomie (Stam *et al.*, 2006) (Figure 1). En effet, alors que les GH13\_9 regroupent principalement des enzymes procaryotes, les GH13\_8 regroupent des enzymes eucaryotes et procaryotes (Stam *et al.*, 2006).

La sous-famille 9 présente quelques exceptions avec trois enzymes d'apicomplexes (*Cryptosporidium parvum*, *Neospora caninum* et *Toxoplasma gondii*) et une enzyme de métazoaire (*Cyprideis torosa*). Mis à part ces exceptions, les membres des GH13\_9 sont majoritairement retrouvés dans le phylum des Bacteroidetes, Protéobactéries ou Verrucomicrobes. En revanche, une analyse des séquences en acides aminés a montré que dans ces phyla, la répartition des GH13\_9 ne reflétait pas la taxonomie (Figure 2) (Suzuki and Suzuki, 2016, 2020). En effet, certaines enzymes branchantes d'une même espèce ou d'espèces différentes mais appartenant au même groupe peuvent être sur des branches

distantes de l'arbre phylogénétique suggérant ainsi un transfert latéral de gènes et/ou une disparition de gènes (Suzuki and Suzuki, 2016).



**Figure 1.** Arbre phylogénétique de la famille des GH13 d'après de Stam *et al*, 2006. Les séquences ont été classées en sous-familles de 1 à 35, de couleur *bleu*, *verte* ou *rouge*. Les séquences en *noires* sont des séquences qui n'ont pas été classées à l'époque de la publication. Les arcs *noirs* représentent la précédente classification des sous-familles réalisée (Oslancová and Janecek, 2002 ; Janeček, Svensson and MacGregor, 2003).





**Figure 2. Arbre phylogénique d'enzymes branchantes de la sous-famille GH13\_9 d'après Suzuki et Suzuki, 2020.** Les protéines sont indiquées par le nom du genre de l'organisme et du numéro d'accèsion du gène tiré de GenBank. L'échelle indiquée correspond à 0,2 substitutions par position de séquence.

La sous-famille des GH13\_8 regroupe quant à elle des enzymes eucaryotes et des enzymes procaryotes appartenant aux Firmicutes, Bacteroidetes, Verrucomicrobes et Delta-protéobactéries (Stam *et al.*, 2006). Ceci représente actuellement 606 séquences d'organismes procaryotes (41 %) sur 1477 séquences de GH13\_8 dont 868 d'organismes eucaryotes d'après CAZy.

A part la famille des GH13, des enzymes branchantes ont également été classées parmi la famille des GH57 aussi connue sous le nom de seconde famille d' $\alpha$ -amylases (Murakami *et al.*, 2006 ; Martinovičová and Janeček, 2018).

## B. Les enzymes branchantes

Les enzymes branchantes sont des acteurs majeurs de la synthèse et de la modification d'amidon ou de glycogène chez les organismes procaryotes et eucaryotes. Il est admis que deux types d'enzymes branchantes existent en fonction de leur spécificité en substrat. Les GBEs pour « Glycogen Branching Enzyme » et les SBEs pour « Starch Branching Enzymes ». Cette spécificité de substrat vient du type d'organisme considéré (Suzuki and Suzuki, 2016, 2020). En effet, alors que les levures utilisent le glycogène comme réserve énergétique, les plantes utilisent l'amidon.

### 1. *Substrat*

#### a) *Le glycogène*

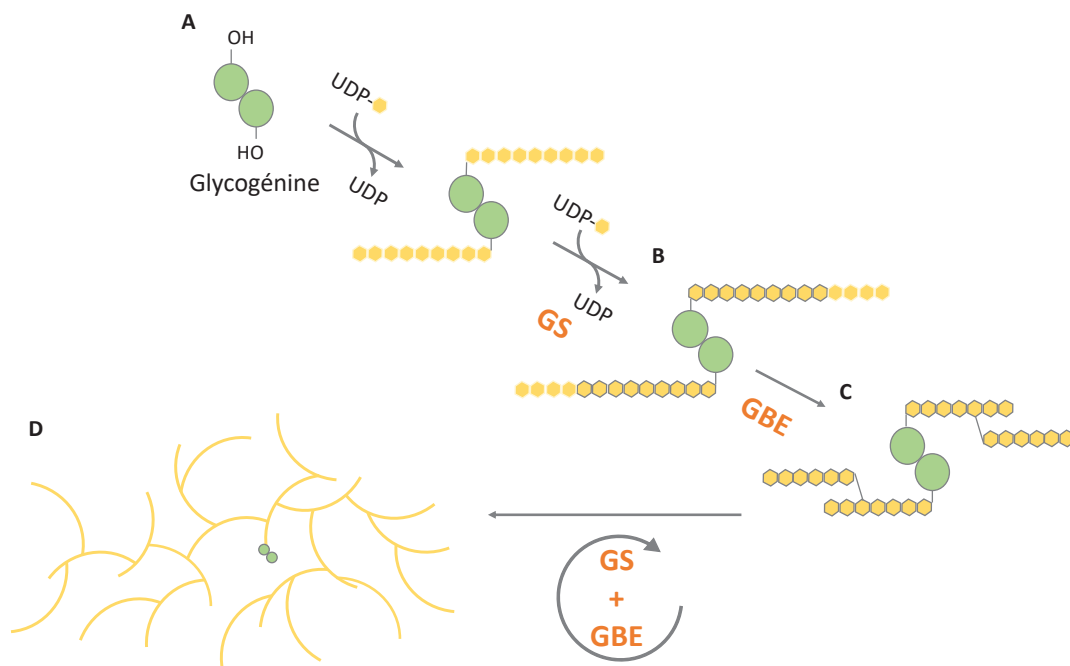
Le glycogène représente une réserve de carbone et d'énergie indispensable au fonctionnement du métabolisme cellulaire et à la survie dans des conditions de carences nutritives (François and Parrou, 2001). Chez *S. cerevisiae*, il a été mis en évidence par Errera en 1885 (Errera, 1885). C'est un polymère de chaînes de glucose liées en  $\alpha$ -(1,4) et ramifié par des liaisons  $\alpha$ -(1,6) avec un point de branchement toutes les 11 à 12 unités glucose (Northcote, 1953 ; Wilson *et al.*, 2010). A l'exception des chaînes extérieures non branchées, chaque chaîne est ramifiée deux fois par des liaisons en  $\alpha$ -(1,6) représentant 7 à 10 % des liaisons totales (Shearer and Graham, 2002; Ball *et al.*, 2011). Ces ramifications sont essentielles car elles permettent un arrangement spatial compact sous la forme de granules hydrosolubles. D'autre part, elles augmentent le nombre d'extrémités non réductrices, lieu de la synthèse et de la dégradation du polymère, permettant une réponse métabolique rapide (Shearer and Graham, 2002). Alors que chez les animaux ces dernières ont un diamètre de 10 à 40 nm (Zmasek and Godzik, 2014), chez *S. cerevisiae* sa taille est d'environ 20 nm (Mundkur, 1960).

La synthèse du glycogène est initiée par une protéine appelée glycogénine (EC 2.4.1.186) (Aon and Curtino, 1984 ; Rodriguez and Whelan, 1985). Les deux isoformes codés par les gènes *GLG1* et *GLG2* chez *S. cerevisiae* (Cheng *et al.*, 1995) sont des glycosyltransférases pouvant s'autoglycosyler en attachant de manière covalente un glucose sur la chaîne latérale d'un résidu tyrosine (Mu, Cheng and Roach, 1996) jusqu'à l'obtention d'une chaîne de 8 à 10

unités glycosidiques liées en  $\alpha$ -(1,4) (Pitcher *et al.*, 1987 ; Smythe and Cohen, 1991 ; Wilson *et al.*, 2010) (Figure 3A).

La suite de la synthèse du glycogène est réalisée par l'action de deux isoformes de la glycogène synthase (« Glycogen Synthase » ; GS) (EC 2.4.1.11) (Farkas *et al.*, 1991) codée par les gènes *GSY1* et *GSY2* chez *S. cerevisiae* (Farkas *et al.*, 1991). Cette enzyme participe à l'allongement des chaînes de glucanes en formant des liaisons  $\alpha$ -(1,4) entre un donneur UDP-glucose et l'extrémité non réductrice de la chaîne (Leloir and Cardini, 1957) (Figure 3B)

Après l'initiation et l'élongation des chaînes de glucose liés en  $\alpha$ -(1,4) intervient une enzyme branchante de glycogène (EC 2.4.1.18) ou GBE codée par le gène *GLC3* chez *S. cerevisiae* (Rowen, Meinke and LaPorte, 1992). Cette dernière intervient dans la ramification du glycogène en transférant des chaînes de 6 à 8 unités glycosidiques de l'extrémité non réductrice d'une chaîne de glucane et en branchant cette chaîne sur l'unité glycosidique d'une chaîne adjacente par une liaison  $\alpha$ -(1,6) (Caudwell and Cohen, 1980) (Figure 3C). La répétition des étapes d'élongations par la GS et de création de ramifications par la GBE permet la synthèse de la molécule de glycogène (Figure 3D).

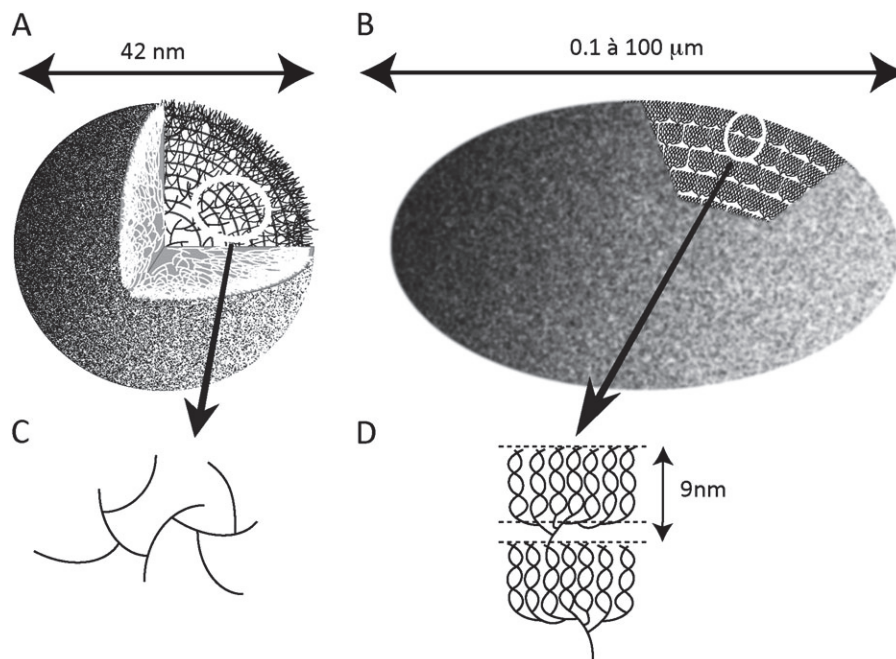


**Figure 3. Représentation schématique simplifiée de la voie de synthèse et de dégradation du glycogène adapté de Roach, 2002.** A. Autoglycosylation de la glycogénine avec une représentation de la glycogénine sous sa forme dimérique par deux ronds *verts*. Les unités glycosidiques liées en  $\alpha$ -(1,4) sont représentées par des hexagones *jaunes*. B. Ajout d'unités glycosidiques par la glycogène synthase (GS) *via* des liaisons  $\alpha$ -(1,4). Les unités glycosidiques précédemment ajoutées lors de l'étape d'auto-glycosylation sont représentées par des hexagones *jaunes* avec un contour *gris*. C. Hydrolyse d'une partie de la chaîne de glucane et branchement de cette dernière par l'action de l'enzyme branchante de glycogène (GBE) *via* une liaison  $\alpha$ -(1,6). D. Molécule de glycogène créée par la répétition des phases (B) et (C).



b) L'amidon

L'amidon, comme le glycogène, est un polymère ramifié intracellulaire nécessaire pour les végétaux et les algues puisqu'il constitue une source d'énergie et de carbone majeure. Il se présente sous la forme de granule insoluble et est composé de deux types de polymères de glucose, l'amylose et l'amylopectine. Ces deux macromolécules sont composées de chaînes de glucanes liées en  $\alpha$ -(1,4) ramifiées par des liaisons en  $\alpha$ -(1,6). Ces dernières représentent 5 à 6 % des liaisons pour l'amylopectine contre moins de 1 % pour l'amylose (Buléon *et al.*, 1998). L'amylose ne constitue que 20 à 30 % de la composition de l'amidon contre 70 à 80 % pour l'amylopectine (Manners, 1989). L'organisation spatiale de l'amidon est bien plus complexe que celle du glycogène du fait de son faible nombre de ramifications et de la répartition asymétrique des points de branchement (Ball *et al.*, 2011). On dit que l'amidon possède une forme semi-cristalline composée de lamelles amorphes correspondant aux points de branchement en  $\alpha$ -(1,6) et de lamelles cristallines correspondant aux chaînes de glucanes liées en  $\alpha$ -(1,4) organisées en doubles hélices parallèles (Figure 4) (Buléon *et al.*, 1998; Ball *et al.*, 2011). Cet arrangement se répète tous les 9 nm et est très conservé dans le règne végétal, suggérant une régulation très fine du processus de synthèse (Cameron and Donald, 1993).



**Figure 4. Représentation schématique de l'organisation en granules du glycogène et de l'amidon d'après Ball *et al.*, 2011.** A et B. Granule de glycogène (A) et d'amidon (B). Les lignes représentent des chaînes de glucane liées en  $\alpha$ -(1,4) et les intersections de ces lignes correspondent à des liaisons  $\alpha$ (1,6). C et D. Vue agrandie de l'arrangement spatial des chaînes de glucanes du glycogène (C) et de l'amidon (D).

A la différence du glycogène, la synthèse de l'amidon ne nécessite pas la présence d'une protéine précurseur capable de s'autoglycosyler. D'après des études, la présence d'ADP-glucose serait suffisante pour induire la synthèse *de novo* du polymère par l'action d'une amidon synthase (EC 2.4.1.2) (Mukerjea, Yu and Robyt, 2002). Cette même enzyme, permet l'allongement des chaînes en formant des liaisons  $\alpha$ -(1,4) entre un donneur ADP-glucose et l'extrémité non réductrice de la chaîne d'amidon. L'enzyme branchante de l'amidon ou SBE (EC 2.4.1.18) hydrolyse des chaînes de glucane liées en  $\alpha$ -(1,4) et introduit des points de branchement par la création de liaison en  $\alpha$ -(1,6).

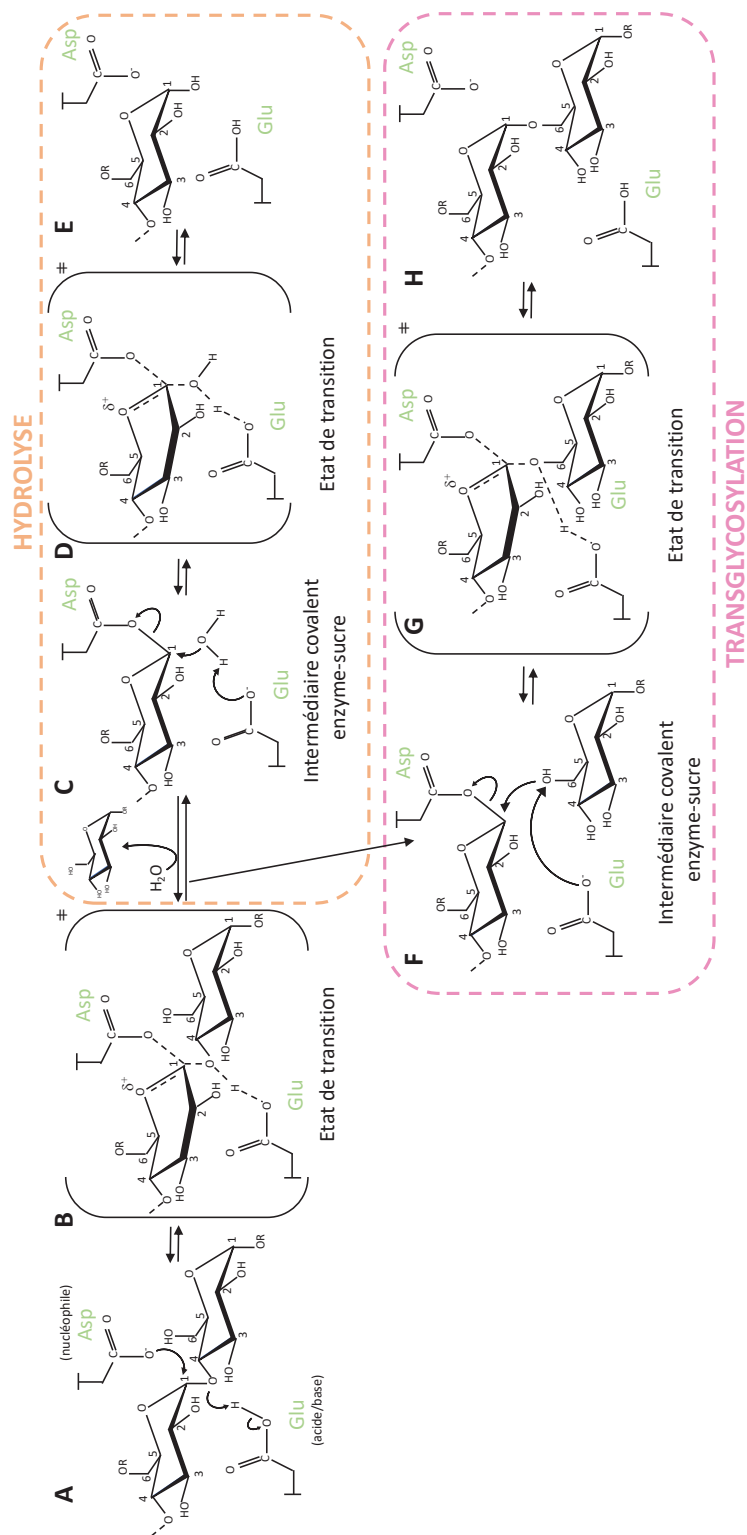
Alors qu'un seul type de GBE a été identifié dans la synthèse du glycogène, il existe plusieurs isoformes de SBE qui sont classés en deux familles, les SBE de type I (SBEI) ou de type II (SBEII). Ces dernières pouvant être codées par deux gènes différents chez certains organismes sont sous-différenciées en SBEIIa ou SBEIIb (Sun *et al.*, 1998). Alors que les deux types de SBE sont retrouvées chez les plantes, seul le type II et ses sous-classes sont retrouvées chez les céréales (Suzuki and Suzuki, 2020). La principale distinction entre les types de SBE provient de leur spécificité de longueur de chaînes branchées. En effet, alors que SBEI préfère brancher des chaînes de 6 à 12 unités glycosidiques, SBEIIa transfère majoritairement des chaînes de 6 à 7 unités et, dans des proportions moindres, des chaînes de 8 à 15 unités. SBEIIb produit, quant à lui, des chaînes branchées de 6 à 7 unités glycosidiques et rarement des chaînes plus longues de 8 à 15 unités (Suzuki and Suzuki, 2020). Plus récemment, un troisième type de SBE (SBEIII) a été identifié chez des espèces de plantes mais est très éloigné des SBEI et SBEII en termes de séquence (Suzuki and Suzuki, 2020).

## 2. Mécanisme catalytique

Toutes les glycosides hydrolases de la famille GH13 et par conséquent toutes les BEs partagent un mécanisme catalytique commun de double déplacement avec rétention de la configuration  $\alpha$  du carbone anomérique initialement décrit par Koshland (Koshland, 1953) (Figure 5). Dans la première étape, l'acide/base (Glu) sous sa forme acide protone l'oxygène interglycosidique d'une chaîne de glucane liée en  $\alpha$ -(1,4) (Figure 5A). Cela engendre un état de transition de type ion oxocarbenium stabilisé par le second Asp. Le nucléophile (Asp), attaque le carbone anomérique (C1) entraînant la formation d'un intermédiaire covalent enzyme-sucre en configuration  $\beta$  (Figure 5 B) et le départ du reste de la chaîne de glucane côté réducteur du site actif (Figure 5 C). Après ce clivage de la liaison interglycosidique, la seconde étape diffère dans le cas d'un mécanisme hydrolytique ou transglycolytique. Dans le premier cas, le résidu Glu (maintenant sous sa forme basique) déprotone une molécule d'eau

formant un ion hydroxyl réalisant une attaque nucléophile sur le C1 de l'intermédiaire covalent enzyme-sucre (Figure 5C). Ceci entraîne un état de transition de type ion oxocarbénium comme précédemment ainsi qu'une seconde inversion de l'anométrie du C1 se retrouvant en configuration  $\alpha$  (Figure 5 D, E). Dans le cas d'une réaction de transglycosylation, le mécanisme d'action est identique mais la molécule d'eau est remplacée par un autre sucre permettant ainsi la formation d'une liaison entre deux chaînes de glucane (Figure 5 F, G, H). Qu'il s'agisse d'un mécanisme hydrolytique ou transglycolytique, le carbone anomérique ciblé par l'action de l'enzyme retrouve son anométrie initiale après deux inversions (Koshland, 1953 ; MacGregor, Janeček and Svensson, 2001). Si la chaîne de glucane est transférée sur la même chaîne, on parle d'un mécanisme intramoléculaire, si elle est transférée sur une chaîne de glucane adjacente, on parle d'un mécanisme intermoléculaire (Tetlow and Emes, 2014).

Les trois résidus catalytiques impliqués dans le mécanisme catalytique des BEs ont été mis en évidence pour la première fois chez l'organisme *Bacillus stearothermophilus*. Pour ce faire, les auteurs ont mesuré l'activité catalytique de l'enzyme mutée au niveau de plusieurs résidus. Ils ont observé une perte d'activité lorsque les résidus Asp<sup>308</sup>, Glu<sup>351</sup> et Asp<sup>419</sup> étaient mutés respectivement en Asn, Gln et Asn (Takata *et al.*, 1994). La même expérience a été réalisée chez l'endosperme du maïs et a permis de mettre en évidence la conservation de ces trois acides aminés essentiels Asp/Glu/Asp chez les BEs (Kuriki *et al.*, 1996).



**Figure 5. Représentation simplifiée du mécanisme à double déplacement avec rétention de la configuration  $\alpha$  du carbone anomérique dans le cas d'une hydrolyse ou d'une transglycolyse adapté de MacGregor, Janeček et Svensson, 2001.** L'acide/base, Glu, protone l'oxygène interglycosidique de la liaison à cliver (A), générant un état de transition de type ion oxocarbenium suivi de l'attaque nucléophile de l'Asp sur le carbone anomérique C1 (B) et conduisant à un intermédiaire covalent enzyme-sucre ainsi qu'au départ du reste de la chaîne (C). Le résidu Glu, sous sa forme basique, active une molécule d'eau (dans le cadre d'une hydrolyse) (C), l'ion hydroxyl résultant fait une attaque nucléophile sur le C1 et rompt l'intermédiaire covalent en passant par un état de transition de type ion oxocarbenium (D), conduisant à l'hydrolyse complète de la chaîne avec rétention de l'anomérie. Dans le cas d'une transglycosylation, le résidu Glu active un autre sucre (F) conduisant à la création d'une nouvelle liaison en  $\alpha$ -(1,6) (H) en passant par un état de transition de type ion oxocarbenium (G).

### 3. Caractérisation structurale des BEs GH13\_9

Les glycosides hydrolases et les enzymes branchantes ont suscité un grand intérêt depuis de nombreuses années. Pourtant, le manque d'information structurale ne permet pas encore une compréhension complète des relations structure-fonction-activité de ces enzymes. Parmi les BEs de la sous-famille 9, seulement 4 structures tridimensionnelles ont été résolues. Le résumé du descriptif structural de chacune des BEs se trouve en fin d'introduction (Tableau 1, p 180).

Les structures résolues ayant été nommées de manière hétérogène dans la littérature, dans ce manuscrit elles seront renommées par les initiales de l'espèce suivi de l'abréviation BE.

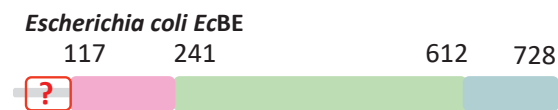
#### a) *Escherichia coli* – EcBE

La structure tridimensionnelle de l'enzyme branchante de *Escherichia coli* (*E. coli*) fut la première résolue en 2002 (PDB : 1M7X) (Abad *et al.*, 2002). Les auteurs ayant rencontré des difficultés à cristalliser la protéine pleine longueur, que l'on renommara EcBE, la structure est celle de l'enzyme tronquée des 113 premiers résidus (EcΔ113BE) par protéolyse limitée. Cette dernière a été déterminée à 2,3 Å de résolution.

#### (1) Caractérisation structurale

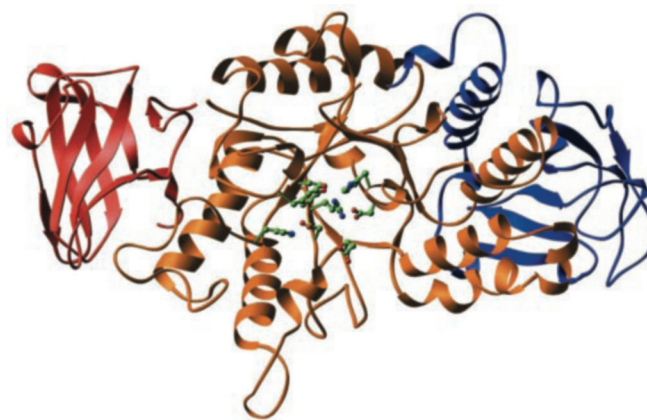
Les auteurs décrivent la structure de EcΔ113BE comme étant organisée en trois domaines, le domaine N-terminale dit NH<sub>2</sub>-terminal, le domaine C-terminale dit COOH-terminal et le domaine catalytique dit tonneau (β/α)<sub>8</sub> (Figure 6). Le domaine N-terminale (117-245) est constitué de 7 brins β formant un repliement de type sandwich β (Figure 7). Le domaine catalytique (241-612), commun à la famille des GH13, est un tonneau (β/α)<sub>8</sub> dont le repliement classique est une alternance d'hélices α et de brins β répétée 8 fois. Pourtant, dans le cas de EcΔ113BE, l'hélice 5 (H5), supposée se trouver entre le brin β5 et β6, est absente. De plus, en comparaison au domaine décrit pour les enzymes de la famille des GH13, de nombreuses insertions d'hélices ont été constatées : les hélices H1a, H3a, H6a, H7a, H7b et H8b. Les boucles, faisant le lien entre les brins et les hélices du domaine catalytique de EcBE, semblent plus courtes que chez la majorité des enzymes de la famille GH13, ayant pour conséquence une cavité catalytique plus ouverte. Cette différence a été attribuée au fait que les BEs prennent en charge des substrats polysaccharidiques plus grands et branchés (Abad *et al.*, 2002). Dans ce domaine les auteurs ont identifié 7 résidus conservés : Asp<sup>335</sup>, His<sup>340</sup>, Arg<sup>403</sup>, Asp<sup>405</sup>, Glu<sup>458</sup>, His<sup>525</sup>, Asp<sup>526</sup> parmi lesquels Asp<sup>405</sup> (nucléophile), Glu<sup>458</sup> (acide/base)

et Asp<sup>526</sup> (stabilisateur de l'état de transition de type ion oxocarbénium) sont les résidus catalytiques essentiels. Le domaine C-terminale (613-728), est séparé du domaine catalytique par 13 résidus, et possède un repliement de type sandwich  $\beta$ .



**Figure 6. Organisation des domaines de EcBE.** La région N-terminale non présente dans la structure tridimensionnelle est indiquée par un rectangle *rouge* avec un point d'interrogation puis le domaine NH<sub>2</sub>-terminal, catalytique ( $\beta/\alpha$ )<sub>8</sub> et le domaine COOH-terminal sont respectivement représentés par des rectangles *rose*, *vert* et *bleu*.

Les 113 premiers résidus de la protéine n'étant pas visibles dans la structure tridimensionnelle déterminée, ils ne sont pas indiqués comme formant un domaine par les auteurs. En revanche, des études réalisées sur la protéine tronquée *Ec* $\Delta$ 113BE ont démontré que l'enzyme perdait 60 % de son activité catalytique et préférait transférer des chaînes de glucanes plus longues par rapport à la protéine native (Binderup, Mikkelsen and Preiss, 2000, 2002). Les auteurs émettent ainsi l'hypothèse que cette région N-terminale serait impliquée dans la spécificité de longueur du transfert des chaînes de glucanes.



**Figure 7. Structure tridimensionnelle de Ec $\Delta$ 113BE tronquée des 113 premiers acides aminés d'après Abad *et al*, 2002.** Les résidus impliqués dans le mécanisme catalytique sont montrés en *vert*. Le domaine NH<sub>2</sub>-terminal est indiqué en *rouge*, le domaine catalytique en *orange* et le domaine COOH-terminal est indiqué en *bleu*.

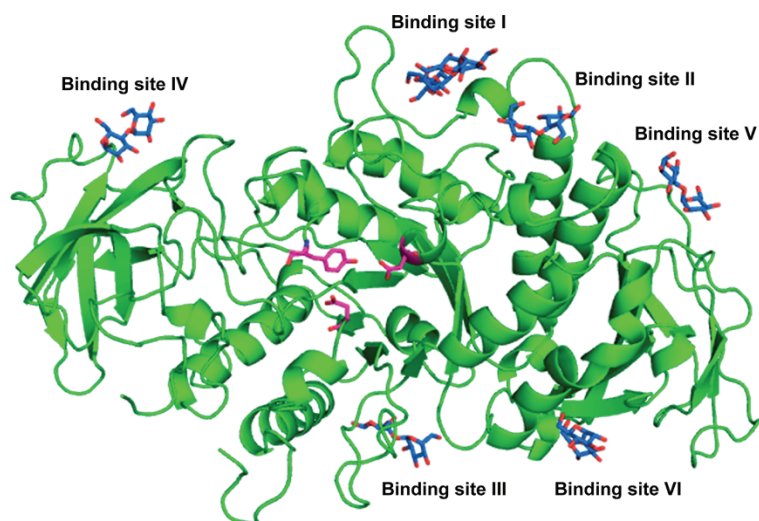
## (2) Sites de liaisons aux sucres

D'autres études ont permis de résoudre la structure de *EcBE* tronquée de sa région N-terminale en complexe avec des oligosaccharides après trempage avec du maltohexaose (G6) (*Ec* $\Delta$ 113BE -G6) (PDB : 4LQ1) et du maltoheptaose (G7) (*Ec* $\Delta$ 113BE -G7) (PDB : 4LPC) (Feng *et al.*, 2015). Les deux structures présentent le même repliement que celui



précédemment observé. Cependant, le domaine appelé « NH2-terminal » a été défini comme étant un domaine de liaison aux sucres (CBM) classé parmi la famille des CBM48 (classification CAZy). Les expérimentations de trempage ont été réalisées avec du maltose, maltotriose, maltotrétraose, maltopentaose, maltohexaose et maltoheptaose pendant des temps relativement longs allant de 3 à 72 heures et avec des concentrations de 50 à 200 mM. La résolution de *Ec*Δ113BE-G6 et *Ec*Δ113BE-G7 suggère une spécificité de l'enzyme pour des longs substrats d'après les auteurs (Feng *et al.*, 2015). En revanche, l'absence de fixation de sucre au site actif de l'enzyme évoque une plus grande affinité pour les polysaccharides que pour des substrats oligosaccharidiques (Figure 8).

Bien que les auteurs indiquent que les structures ont été résolues après trempage avec du G6 et du G7, aucun de ces deux oligosaccharides ne sont entièrement visibles. Ceci est probablement dû à une stabilisation partielle des substrats en interaction. En effet, au niveau du site I, II, III, VI se trouvant au domaine catalytique, ne sont respectivement visibles qu'un triose, maltose, maltose et maltotriose. Le site VI présente une interaction hydrophobe avec l'oligosaccharide *via* trois résidus hydrophobes Leu<sup>512</sup>, Phe<sup>509</sup> et Trp<sup>268</sup>. Le site IV, situé au niveau du domaine CBM48, lie un maltose par une interaction avec trois résidus aromatiques Trp<sup>159</sup>, Lys<sup>189</sup> et Gln<sup>211</sup>. Le site V, se trouvant au niveau du domaine C-terminale, fixe un maltose (Figure 8).



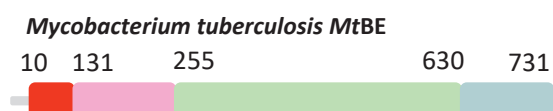
**Figure 8. Structure tridimensionnelle de *Ec*Δ113BE et les différents sites de liaisons aux sucres d'après Feng *et al.*, 2015. Les sucres sont montrés par des bâtonnets bleus et les résidus du site actif sont montrés en rose.**

## b) *Mycobacterium tuberculosis* – MtBE

La seconde structure de GH13\_9 déterminée par cristallographie aux rayons X, à 2,33 Å de résolution, est celle de l'organisme *Mycobacterium tuberculosis* (*M. tuberculosis*), appelée MtbGlgBWT (PDB : 3K1D) que l'on nommera MtBE (Pal *et al.*, 2010).

### (1) Caractérisation structurale

La structure permet d'observer quatre domaines, un domaine N-terminale dit « N-terminal sandwich  $\beta$  N1 » (10-130), un domaine CBM48 « N-terminal sandwich  $\beta$  N2 » (131-254), un domaine catalytique  $(\beta/\alpha)_8$  (255-630) et enfin un domaine C-terminale dit « C-terminal sandwich  $\beta$  domaine » (631-731) (Figure 9, 10) (Pal *et al.*, 2010). Le tunnel du tonneau  $(\beta/\alpha)_8$  possède un diamètre de 13,5 Å et les boucles le composant sont assez courtes pour permettre la fixation d'un large sucre branché tel que le glycogène.

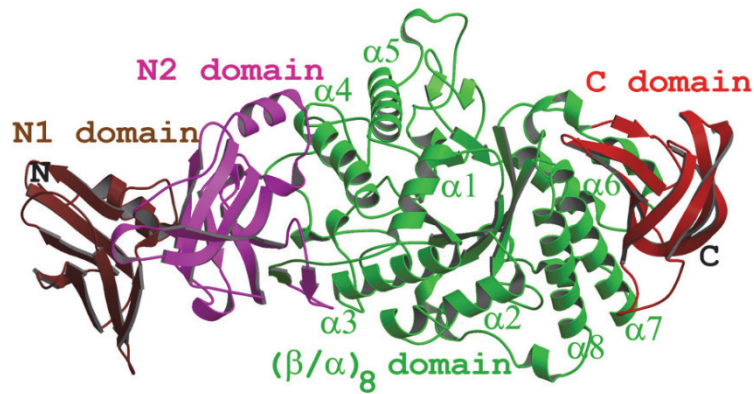


**Figure 9. Organisation des domaines de MtBE.** Le domaine N-terminale N1 sandwich  $\beta$ , N-terminale N2 sandwich  $\beta$ , le domaine catalytique  $(\beta/\alpha)_8$  et le domaine C-terminale sandwich  $\beta$  sont respectivement représentés par des rectangles rouge, rose, vert et bleu.

Comme pour *EcBE*, sept résidus ont été identifiés comme participant au mécanisme catalytique : Asp<sup>341</sup>, His<sup>346</sup>, Arg<sup>409</sup>, Asp<sup>411</sup>, Glu<sup>464</sup>, His<sup>531</sup> et Asp<sup>532</sup> parmi lesquels Asp<sup>411</sup> (nucléophile), Glu<sup>464</sup> (acide/base), Asp<sup>532</sup> (stabilisateur de l'état de transition de type ion oxocarbénium) représenteraient les trois résidus catalytiques essentiels.

Des mesures d'activité ont été réalisées sur la protéine pleine longueur MtBE et la protéine tronquée de son domaine N-terminale appelée Mtb $\Delta$ 108GlgB que l'on nommera Mt $\Delta$ 108BE. Les résultats montrent que Mt $\Delta$ 108BE est 50 % moins active que la protéine pleine longueur, suggérant ainsi l'implication du domaine N-terminale dans la reconnaissance et la liaison au substrat (Pal *et al.*, 2010). En revanche, les activités de branchement ne sont pas altérées par l'absence du domaine N-terminale.





**Figure 10. Structure tridimensionnelle de MtBE d'après Pal *et al.*, 2010.** Représentation des domaines N1 sandwich  $\beta$ , N2 sandwich  $\beta$ , catalytique  $(\beta/\alpha)_8$ , C-terminale sandwich  $\beta$  respectivement en *marron*, *magenta*, *vert*, *rouge*.

### c) *Crocospaera subtropica* – CsBE

La troisième structure de GH13\_9 BE déterminée est celle issue de l'organisme *Crocospaera subtropica* (*C. subtropica*) (Hayashi *et al.*, 2017). Elle a été déterminée sous la forme pleine longueur, dite « LF-WtBE1 » que l'on nommera CsBE et sous une forme mutée CsW610NBE, toutes les deux sans ligand ou en complexe avec un maltohexaose et un maltoheptaose respectivement (CsBE-G6 et CsW610NBE-G7).

#### (1) Caractérisation structurale

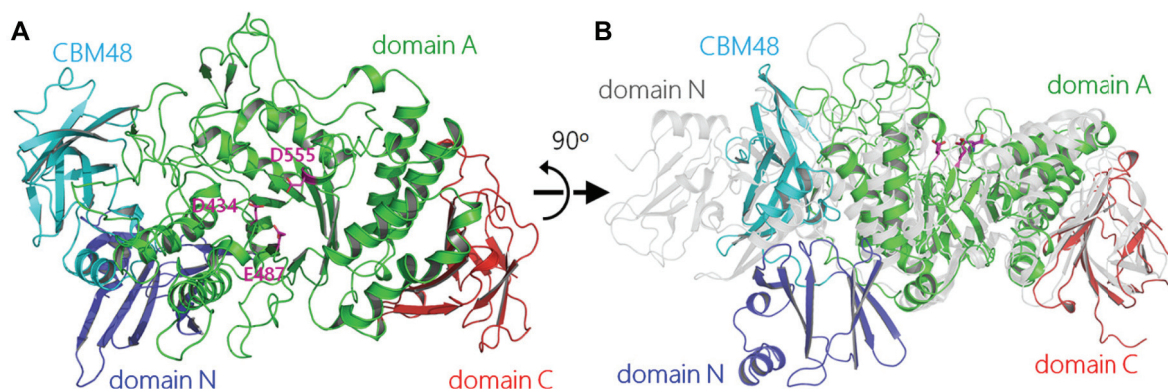
Les structures ont été déterminées à 1,85 Å, 3,0 Å, 1,8 Å et 2,3 Å de résolution respectivement pour CsBE (PDB : 5GQU), CsBE-G6 (PDB : 5GQV), CsW610NBE (PDB : 5GQW), CsW610NBE-G7 (PDB : 5GQX). La structure révèle une organisation multi-domaine, comme les précédentes, avec un domaine N-terminale dit « domaine N » (5-99), un domaine CBM48 (100-206), un domaine catalytique dit « domaine A » (207-651) et un domaine C-terminale dit « domaine C » (652-759) (Figure 11).



**Figure 11. Organisation des domaines de CsBE.** Le domaine N, CBM48, A et C sont respectivement représentés par des rectangles *rouge*, *rose*, *vert* et *bleu*.

Le domaine catalytique présente un repliement de type tonneau  $(\beta/\alpha)_8$  dans lequel on peut identifier les trois résidus catalytiques Asp<sup>434</sup> (nucléophile), Glu<sup>487</sup> (acide/base) et Asp<sup>555</sup> (stabilisateur de l'état de transition de type ion oxocarbénium) (Figure 12 A). Les domaines N-terminale, CBM48 et C-terminale possèdent un repliement de type sandwich  $\beta$ . Une

particularité dans l'arrangement structural par rapport à *MtBE* est que le domaine N-terminale se situe derrière le domaine catalytique au lieu d'être dans le prolongement du CBM48 selon un axe domaine N-terminale - CBM48 - catalytique - C-terminale, (Figure 12B).



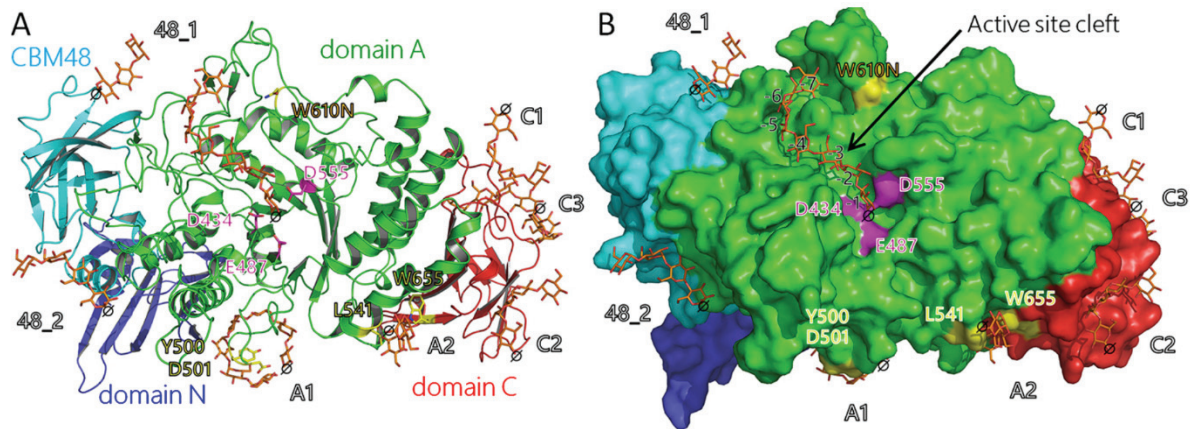
**Figure 12. Structure tridimensionnelle de *CsBE* d'après Hayashi *et al.*, 2017.** A. Les domaines N, CBM48, A et C de *CsBE* sont représentés respectivement en *bleu, cyan, vert, rouge*. Les résidus catalytiques Asp<sup>434</sup>, Glu<sup>487</sup> et Asp<sup>555</sup> sont représentés en bâtonnets *roses*. B. Superposition des structures tridimensionnelles de *CsBE* et de *MtBE* (PDB : 3K1D) respectivement représentées en couleurs mentionnées en (A) et en *gris*. Les structures en (B) sont montrées avec une rotation de 90° sur un axe horizontal par rapport à (A).

## (2) Sites de liaisons aux sucres

Les structures *CsBE*-G6 et *CsW610NBE*-G7 sont les premières résolues en complexe avec un oligosaccharide se trouvant au niveau du site actif. La présence du G7 de *CsW610NBE*-G7 a notamment permis l'identification de sept sous-sites de fixation numérotés de -1 (où est positionné l'extrémité réductrice du sucre) à -7 ainsi que les résidus impliqués : Pro<sup>217</sup>, Glu<sup>284</sup>, Trp<sup>285</sup>, Phe<sup>323</sup>, Asp<sup>324</sup>, Trp<sup>327</sup>, Tyr<sup>329</sup>, Gln<sup>330</sup>, Asp<sup>373</sup>, Trp<sup>399</sup>, His<sup>554</sup>, Asp<sup>555</sup>. L'oligosaccharide adopte une forme de « S » dans le site actif comme cela a déjà pu être observé pour l'enzyme débranchante de *Chlamydomonas reinhardtii* (GH13\_11) (Sim *et al.*, 2014) (Figure 13). Le G6 de *CsBE*-G6 est également positionné au niveau du site actif et des sous sites -2 à -7 dont les résidus impliqués sont essentiellement les mêmes que pour *CsW610NBE*-G7.

En plus de la fixation au site actif, sept autres sites ont été identifiés, les domaines 48\_1 et 48\_2 dans le CBM48, A1 et A2 dans le domaine catalytique, C1, C2 et C3 dans le domaine C-terminale (Figure 13 A) (Hayashi *et al.*, 2017). Ces sites sont appelés SBSs pour « *Surface Binding Site* » ou site de liaisons en surface. Le site 48\_1 fixe un maltotriose et est similaire au site de fixation III chez *EcBE* (Feng *et al.*, 2015) et au site 2 chez *Oriza sativa* (*O. sativa*) (Chaen *et al.*, 2012). Le site 48\_2 fixe un maltotetraose de manière équivalente au site décrit chez *Homo sapiens* (Froese *et al.*, 2015) et au site 1 de *O. sativa* (Chaen *et al.*, 2012). Le site de fixation A1, équivalent au site III de *EcBE*, fixe un G7 en conformation hélicoïdale via l'interaction de deux acides aminés aromatiques Tyr<sup>500</sup> et Asp<sup>501</sup>. Le site A2, équivalent au

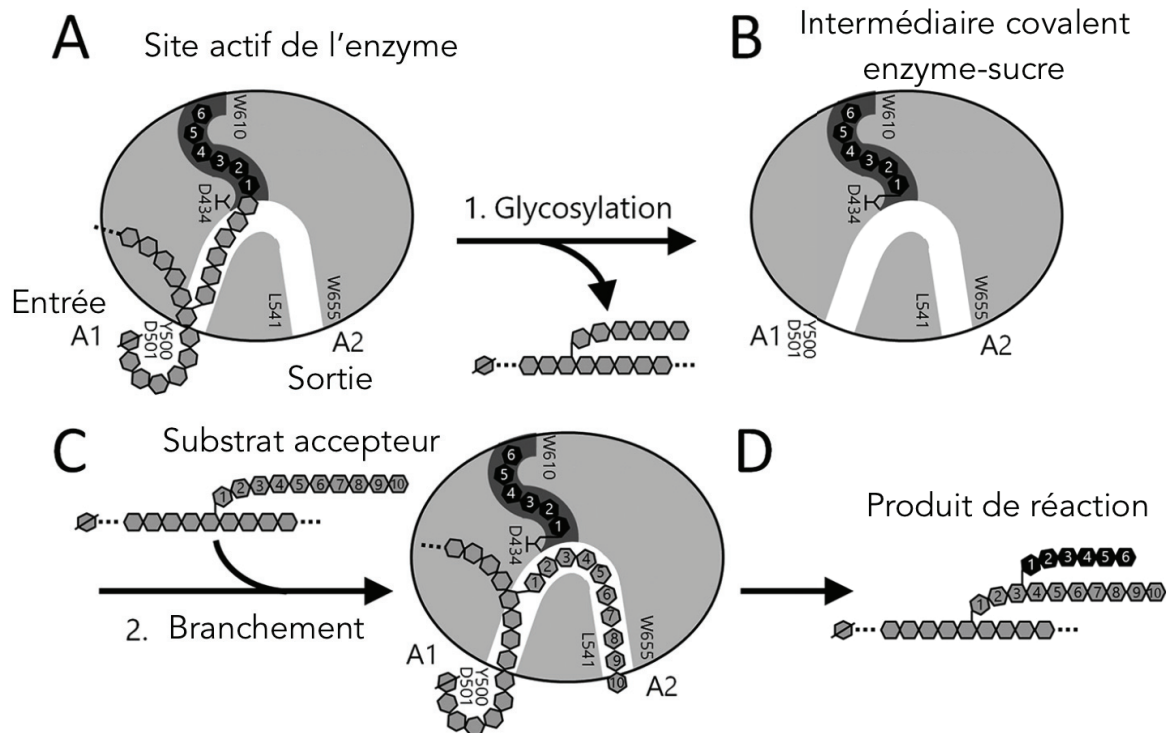
site VI de *EcBE*, fixe un G7 *via* l'interaction avec deux résidus Leu<sup>541</sup> et Trp<sup>655</sup> dont seulement les trois premières unités glycosidiques (G3) ont pu être déterminées dans la structure.



**Figure 13. Structure tridimensionnelle du mutant W610N en complexe avec un maltoheptaose d'après Hayashi *et al.*, 2017.** A. Structure tridimensionnelle avec les différents domaines colorés comme décrit pour la Figure 12 A. Les oligosaccharides sont représentés en bâtonnets *orange*. Les extrémités réductrices des sucres sont représentées par  $\emptyset$ . Les résidus catalytiques (Asp<sup>434</sup>, Glu<sup>487</sup>, Asp<sup>555</sup>) sont représentés par des bâtonnets *roses* et le résidu muté W610N ainsi que les résidus impliqués dans le site de fixation A1 et A2 sont représentés par des bâtonnets *jaunes*. Les différents sites de surface de fixation (A1, A2, C1, C2, C3, 48\_1, 48\_2) sont indiqués sur l'image. B. Représentation de la surface de W610N avec le site actif indiqué par une flèche *noire*. Les sous-sites du site catalytique sont indiqués de -1 à -7.

L'identification de ces deux derniers sites de surface, A1 et A2, a conduit les auteurs à faire la première hypothèse du modèle de fixation des chaînes de glucanes sur les enzymes branchantes (Figure 14) (Hayashi *et al.*, 2017). En effet, la diminution de l'activité spécifique des mutants Y500A/D501A et L541A/W655A pour l'amylose et l'amylopectine a suggéré aux auteurs l'implication du site A1 et A2 dans la fixation des glucanes. L'activité spécifique étant réduite de moitié pour le mutant Y500A/D501A et quasiment nulle pour L541A/W655A. Ainsi, ils ont émis l'hypothèse que le site A1 correspond au site d'entrée d'une chaîne « donneur » de glucane dont l'extrémité réductrice est orientée vers A1 (Figure 14 A). Cette même chaîne pouvant être stabilisée par de multiples sites de fixation à la surface de l'enzyme en plus de la partie de la chaîne stabilisée aux sous sites -1 à -7 du site actif, créant ainsi un complexe enzyme-sucre (Figure 14 B). L'hydrolyse au niveau du sous site -1 induit le départ du reste de la chaîne de glucane. Dans une seconde étape, une chaîne de glucane « accepteur » se positionne entre le site d'entrée A1 et de sortie A2, notamment le long d'une région riche en alanine et glycine (Figure 14 C). L'extrémité non réductrice de la chaîne de glucane se trouve au site de sortie A2. La chaîne de glucane, constituant

l'intermédiaire covalent enzyme-sucre retenue au site actif, est alors branchée par une liaison  $\alpha$ -(1,6) par l'activité transglycolytique de l'enzyme branchante (Figure 14 D).



**Figure 14. Représentation schématique du modèle de fixation des chaînes de glucanes sur CsBE adapté de Hayashi *et al*, 2017.** A. La chaîne de glucane « donneur » se lie du site d'entrée A1 du domaine catalytique, en gris, aux sous sites -1 à -7 du site actif coloré en gris foncé. Les glucanes sont représentés par des hexagones gris et ceux qui vont être transférés sont représentés en noir, l'extrémité réductrice est repérable par un hexagone barré. Les résidus catalytiques participant à ce mécanisme chez CsBE ont été positionnés sur le schéma. B. L'hydrolyse de la chaîne de glucane au niveau du site actif induit sa libération et la rétention de la chaîne prête à être transférée. Cette dernière formant un intermédiaire covalent enzyme-sucre. C. Une chaîne de glucane « accepteur » se positionne du site d'entrée A1 au site de sortie A2. D. Produit de réaction, une chaîne de glucane possédant un nouveau branchement en  $\alpha$ -(1,6) après l'action de transglycosylation de CsBE.

Plus récemment, les mêmes auteurs ont effectué des recherches plus approfondies sur les sites de liaison en surface identifiés chez CsBE. En plus des mutations réalisées précédemment, ils ont également muté les résidus impliqués dans les sites de liaisons 48\_1 (N150A/E410A), 48\_2 (W163A/E214A), C1 (W573A/Q574A), C2 (D657A/Y693A), C3 (N718A/L722A). Ces mutants ont été soumis à une migration électrophorétique sur gels non dénaturants composés d'amylose, amylopectine et glycogène. Un retard de migration étant associé à une interaction avec les polysaccharides, les auteurs ont ainsi pu déduire des constantes de dissociation ( $K_d$ ) associées aux interactions (El Mannai *et al.*, 2021). En comparaison à CsBE WT, ils ont mis en évidence des  $K_d$  significativement plus importantes pour les mutants des sites A2, C1 et C2, suggérant leur interaction avec les polysaccharides. En revanche, aucune différence significative pour les autres sites de liaisons n'a été observée. Les auteurs concluent que ces sites peuvent interagir avec des oligosaccharides comme l'indiquent les structures en



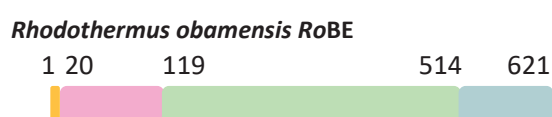
complexe obtenues par co-cristallisation mais interagissent difficilement avec des polysaccharides (El Mannai *et al.*, 2021). Ces résultats remettent en question le mécanisme de fixation des polysaccharides aux enzymes branchantes précédemment décrit ainsi que la fonctionnalité du site A1 comme site d'entrée (Hayashi *et al.*, 2017; El Mannai *et al.*, 20).

#### d) *Rhodothermus obamensis* – RoBE

Récemment, la structure tridimensionnelle de l'organisme *Rhodothermus obamensis* (*R. obamensis* ; RoBE) aussi connue sous le nom de *Rhodothermus marinus*, a été résolue à 2,39Å de résolution (PDB : 6JOY) (Jiang *et al.*, 2021).

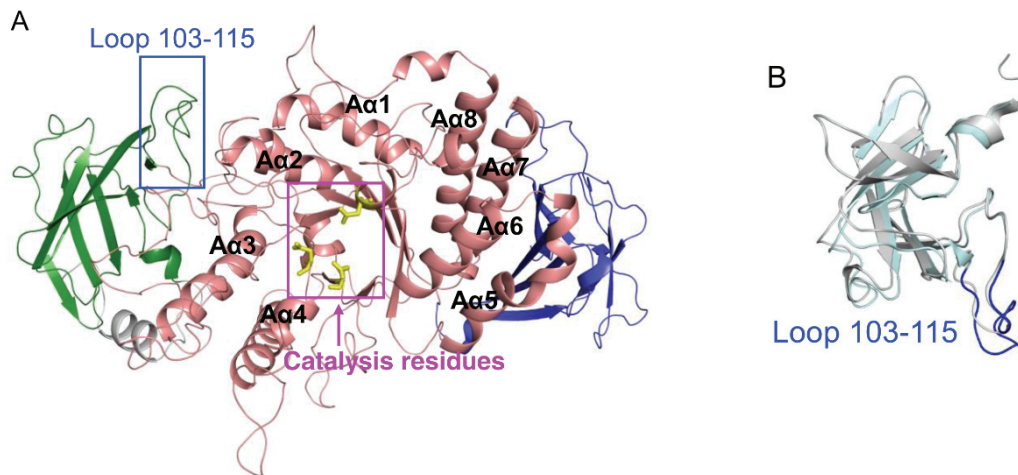
##### (1) Description structurale

RoBE est composée des trois domaines conservés, le domaine CBM48 (20-118), le domaine catalytique dit « domaine A » (119-514) et le domaine C-terminale dit « domaine C » (515-621) (Figure 15). Les auteurs ont identifié à l'extrémité N-terminale de RoBE une courte séquence de 20 résidus ayant un repliement en hélice  $\alpha$  mais n'étant pas considérée comme un domaine N-terminale à part entière. De manière intéressante, le repliement en hélice de cette extrémité N-terminale présente une caractéristique structurale unique chez les GH13\_9 (Figure 15). Le domaine catalytique possède un repliement conservé de tonneau ( $\beta/\alpha$ )<sub>8</sub> où se trouvent les trois résidus catalytiques essentiels Asp<sup>305</sup>, Glu<sup>356</sup> et Asp<sup>424</sup> de la triade catalytique. Le domaine CBM48 et le domaine C-terminale possède un repliement en sandwich  $\beta$  (Figure 16 A)



**Figure 15. Organisation des domaines de RoBE.** L'hélice  $\alpha$  présente à l'extrémité N-terminale, le domaine CBM48, A et C sont respectivement représentés par des rectangles orange, rose, vert et bleu.

Des études enzymatiques sur RoBE tronquée de son domaine CBM48 (*Ro* $\Delta$ CBM48BE) ont montré que l'activité spécifique pour l'amylopectine était réduite de 94,6 % par rapport à l'enzyme sauvage. De plus, la diminution importante du taux d'expression de l'enzyme tronquée et l'augmentation du *K<sub>m</sub>* par rapport à RoBE suggère que le domaine CBM48 est indispensable à l'intégrité structurale et à l'activité catalytique de l'enzyme notamment en ce qui concerne la fixation du substrat (Jiang *et al.*, 2021).



**Figure 16. Structure tridimensionnelle de *RoBE* adaptée de Jiang *et al.*, 2021.** A. L'hélice  $\alpha$  présente à l'extrémité N-terminale est représentée en *gris*, les domaines CBM48, catalytique et C-terminale sont respectivement représentés en *vert*, *rose*, *bleu*. Les hélices du tonneau ( $\beta/\alpha$ )<sub>8</sub> sont indiquées en *noir* et les résidus catalytiques Asp<sup>434</sup>, Glu<sup>487</sup> et Asp<sup>555</sup> sont représentés sous la forme de bâtonnets *jaunes* entourés par un carré *magenta*. La boucle 103-105 est entourée par un rectangle *bleu*. B. Superposition de la boucle flexible 103-115 de *RoBE*, représentée en *bleu*, et de *EcBE* représentée en *gris*.

La superposition du domaine CBM48 de *RoBE* avec *EcBE* a conduit à l'identification d'une boucle flexible (103-115) plus longue chez *RoBE* et interagissant avec les domaines catalytique et CBM48 (Figure 16 B). Un mutant tronqué de ces résidus (*Ro* $\Delta$ 107-114BE) créant ainsi une boucle de même taille que *EcBE* a montré une affinité réduite pour le substrat. Ainsi, les auteurs suggèrent que cette boucle flexible puisse se comporter comme un « couvercle » qui, dans un cas, est ouvert et permet l'accessibilité et la fixation d'un sucre au site actif. Dans l'autre cas, lorsqu'un sucre est fixé au site actif, la boucle bascule afin de refermer le site et de maintenir l'intermédiaire covalent enzyme-sucre (Jiang *et al.*, 2021).

#### 4. Caractérisation structurale des BEs GH13\_8

La sous-famille 8 des GH13 est celle possédant le moins d'information structurale. En effet, seules deux structures ont été résolues entre 2011 et 2021, la structure tridimensionnelle de l'enzyme branchante de la plante *Oryza sativa* (*O. sativa*) et celle d'*Homo sapiens* (*H. sapiens*). Comme pour les GH13\_9, le résumé du descriptif structural de ces enzymes se trouve en fin d'introduction (Tableau 1, p 180).

##### a) *Oryza sativa* – *OsBEI*

Chez *O. sativa*, trois isoformes d'enzymes branchantes ont précédemment été identifiés (Nakamura, 2002) : BEI, BEIIa et BEIIb. Parmi eux, la structure tridimensionnelle de

l'isoforme BEI, que l'on nommera *OsBEI*, a été déterminée à 1,9 Å de résolution (PDB: 3AMK) (Noguchi *et al.*, 2011).

### (1) Description structurale

La résolution de la structure d'*OsBEI* a permis l'identification de quatre domaines, le domaine N-terminale dit « segment N-terminal  $\alpha$ -hélicoïdale » (8-58), CBM48 (59-160), le domaine catalytique dit «  $\alpha$ -amylase » (161-587) et le domaine C-terminale dit «  $\alpha$ -amylase C » (588-702) (Figure 17) (Noguchi *et al.*, 2011). Comme pour les autres structures, le domaine CBM48 et C-terminale ont un repliement en sandwich  $\beta$  et le domaine catalytique a un repliement en tonneau ( $\beta/\alpha$ )<sub>8</sub>. Cependant, pour la première fois, le repliement observé du domaine N-terminale est différent. En effet, il est constitué de trois hélices  $\alpha$  consécutives ( $\alpha$ 1 à  $\alpha$ 3) alors que pour *MtBE* et *CsBE* ce domaine est présent sous la forme d'un sandwich  $\beta$  (Pal *et al.*, 2010; Hayashi *et al.*, 2017). Des études réalisées sur l'enzyme branchante du maïs (Kuriki, Stewart and Preiss, 1997), ont démontré l'importance du domaine N-terminale dans la sélection des chaînes de glucanes transférées en fonction de leur longueur. En connaissance de ces résultats, les auteurs suggèrent que la différence structurale du domaine N-terminale par rapport aux organismes eucaryotes *M. tuberculosis* et *C. subtropica* pourrait être attribuable à une différence de substrat, le glycogène étant plus branché que l'amidon (Noguchi *et al.*, 2011).



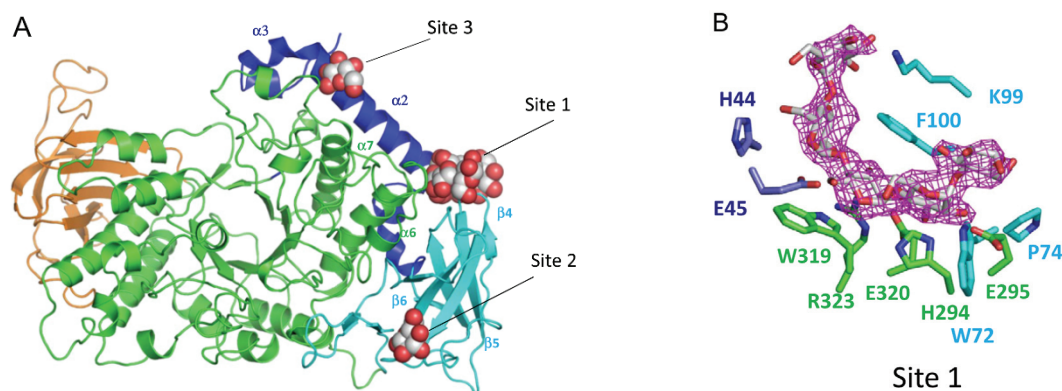
**Figure 17. Organisation des domaines de *OsBEI*.** Le segment N-terminale  $\alpha$ -hélicoïdale et les domaines CBM48,  $\alpha$ -amylase et  $\alpha$ -amylase C sont respectivement représentés par des rectangles orange, rose, vert et bleu.

Le site actif présent dans le domaine catalytique de l'enzyme contient les résidus catalytiques essentiels Asp<sup>344</sup> (nucléophile), Glu<sup>399</sup> (acide/base) et Asp<sup>468</sup> (stabilisateur de l'état de transition de type ion oxocarbenium).

### (2) Les sites de liaisons aux sucres

Afin d'obtenir une structure tridimensionnelle en complexe avec un sucre, des auteurs ont réalisés des expériences de protéolyse limitée et de mutagenèse dirigée. Ainsi, la structure tridimensionnelle déterminée est celle de l'enzyme tronquée des 53 derniers résidus appelée  $\Delta_c$  (Noguchi *et al.*, 2011). De plus, le résidu catalytique acide/base a été muté en glutamine, E399Q. Les essais de co-cristallisation de *OsE399Q* $\Delta_c$ BE avec 300 mM de maltopentaose ont

permis la détermination de la structure tridimensionnelle de cette enzyme à 2,2 Å de résolution (PDB : 3VU2) (Chaen *et al.*, 2012). Trois sites d'interaction en surface ont été identifiés, le site 1, 2 et 3 (Figure 18).



**Figure 18. Structure tridimensionnelle de *OsE399QΔcBE* en complexe avec du maltopentaose d'après Chaen *et al.*, 2012.** **A.** Le segment N-terminale  $\alpha$ -hélicoïdale et les domaines CBM48,  $\alpha$ -amylase,  $\alpha$ -amylase C sont respectivement représentés en *bleu foncé*, *cyan*, *vert* et *orange*. Les trois sites de liaison en surface sont indiqués par des flèches. Les unités oligosaccharidiques se liant aux sites sont représentés en boules *grises* pour les atomes de carbone et *rouges* pour les atomes d'oxygène. **B.** Vue rapprochée du site d'interaction du maltopentaose au site 1. Les résidus participant à la fixation de cet oligosaccharide sont colorés en fonction du domaine dans lequel ils se trouvent comme indiqués dans (A).

Le site 1, équivalent au site 48\_2 de *CsBE*, se situe à l'interface entre les domaines catalytique, CBM48 et l'hélice  $\alpha 2$  du domaine N-terminale. En effet, le G5 est lié *via* une interaction hydrophobe par les chaînes latérales des résidus Trp<sup>72</sup> (CBM48), Phe<sup>100</sup> (CBM48) et Trp<sup>319</sup> (domaine catalytique) (Figure 18 B). De nombreuses liaisons hydrogènes viennent stabiliser cette fixation ainsi que les chaînes latérales des résidus His<sup>44</sup> et Glu<sup>45</sup> (domaine N-terminale) interagissant respectivement avec la seconde et troisième unité glucose du G5. Le résidu Lys<sup>99</sup> (CBM48) interagit avec la dernière unité glucose et His<sup>294</sup>, Glu<sup>295</sup>, Trp<sup>319</sup>, Glu<sup>320</sup>, Arg<sup>323</sup> (domaine catalytique) interagissent avec la troisième et quatrième unité glucose.

Le site 2, équivalent au site 48\_1 de *CsBE*, fixe une unité glucose par une interaction avec un résidu aromatique Trp<sup>88</sup> (CBM48). Enfin, le site 3 est le premier site identifié dans un domaine N-terminale. Il fixe une unité glucose par l'interaction avec le résidu aromatique Tyr<sup>29</sup> (domaine N-terminale) et deux résidus hydrophobes Leu<sup>370</sup> et Val<sup>374</sup> (domaine catalytique).

#### b) *Homo sapiens* – *HsBE*

La structure tridimensionnelle de l'enzyme branchante d'*Homo sapiens* (*H. sapiens*) tronquée de ses 38 premiers résidus dite « hGBEI<sub>trunc</sub> » et que l'on nommera *HsΔ38BE*, a été déterminée à 2,75 Å de résolution (PDB : 4BZY) mais également en complexe avec un



maltoheptaose (*Hs*Δ38BE-G7 ; PDB : 5CLW) et un acarbose (*Hs*Δ38BE-ACR ; PDB : 5CLT) respectivement à 2,8 Å et 2,79Å de résolution (Froese *et al.*, 2015).

### (1) Description structurale

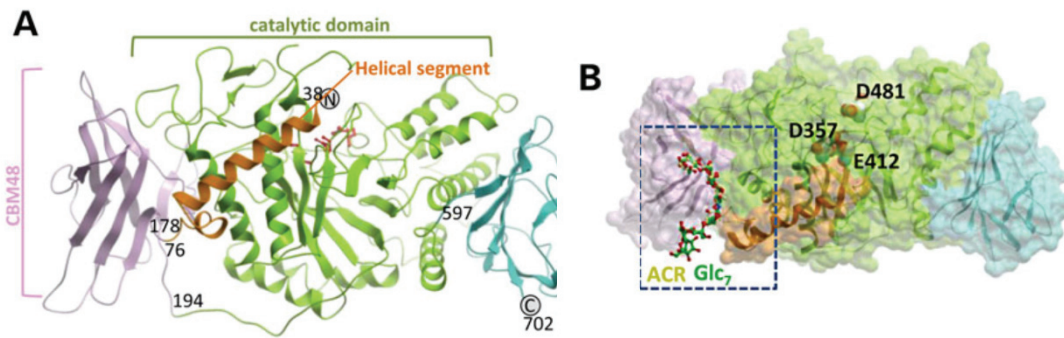
Comme pour *Os*BEI, le domaine N-terminale (38-75), dit « segment N-terminal  $\alpha$ -hélicoïdale » est sous la forme d'hélices  $\alpha$ . Le domaine CBM48 (76-183) est sous la forme d'un sandwich  $\beta$ , le domaine catalytique (184-597) est sous la forme d'un tonneau ( $\beta/\alpha$ )<sub>8</sub> contenant les résidus catalytiques essentiels Asp<sup>357</sup> (nucéophile), Glu<sup>412</sup> (acide/base) et Asp<sup>481</sup> (stabilisateur de l'état de transition de type ion oxocarbénium) et enfin le domaine C-terminale (598-702) dit « domaine C-terminal  $\alpha$ -amylase » est replié en sandwich  $\beta$  (Figure 19).



**Figure 19. Organisation des domaines de *Hs*BE.** Le segment N-terminale  $\alpha$ -hélicoïdale et les domaines CBM48, noyau central catalytique et C-terminale de type tonneau  $\alpha$ -amylase sont respectivement représentés par des rectangles orange, rose, vert et bleu.

### (2) Les sites de liaisons aux sucres

Des expériences de co-cristallisation ont été réalisées en présence de 5 mM de maltoheptaose (G7) et 5 mM d'acarbose, un pseudo tétra-saccharide identifié comme inhibiteur d' $\alpha$ -glucosidases de la famille GH13 (Søgaard *et al.*, 1993). Ces dernières ont permis la résolution de deux structures, *Hs*Δ38BE-ACR et *Hs*Δ38BE-G7 (Figure 20). Alors que les auteurs s'attendaient à une fixation des ligands au site actif, ils furent surpris de les identifier à un site à l'interface du domaine N-terminale, CBM48 et catalytique (Froese *et al.*, 2015). On le nommera site 1, équivalent au site 1 de *Os*BEI (Chaen *et al.*, 2012) et au site 48\_2 de *Cs*BE (Hayashi *et al.*, 2017). De la même manière, le G7 ou ACR est lié par une interaction hydrophobe aux chaînes latérales des résidus Trp<sup>91</sup> et Tyr<sup>119</sup> du domaine CBM48 et du résidu Trp<sup>332</sup> du domaine catalytique. La conservation des sites de fixation en surface, fait suggérer aux auteurs, que ces sites pourraient avoir un rôle d'ancrage des chaînes de glucane et de spécificité de la longueur des chaînes branchées. Les sites de fixation en surface agiraient donc comme des « règles moléculaires » (Froese *et al.*, 2015).



**Figure 20. Structure tridimensionnelle de *HsΔ38BE* d'après Froese *et al.*, 2015.** A. Le segment N-terminale  $\alpha$ -hélicoïdale et les domaines CBM48, noyau central catalytique, domaine C-terminale  $\alpha$ -amylase sont respectivement représentés en *orange*, *rose*, *vert* et *cyan*. B. Le site de fixation du maltoheptaose (Glc<sub>7</sub>) et de l'acarbose (ACR) est entouré d'un rectangle *bleu*. Les résidus catalytiques sont indiqués en *noir*. Les différents domaines sont colorés comme indiqué dans (A).



## II. Caractérisation structurale de CgGBE

Dans le cadre de ce projet de recherche, une caractérisation des relations structure-fonction-activité a été réalisée pour l'enzyme branchante de glucane en  $\alpha$ -(1,4) (CgGBE) afin de mieux comprendre le mode d'action et de reconnaissance des enzymes de la famille des GH13\_8 et leurs différences avec celles de la famille des GH13\_9. Pour ce faire, la structure tridimensionnelle de CgGBE a été résolue sous une forme libre et en complexe avec un maltotriose. Ces résultats ont permis de mettre en évidence un domaine N-terminale sous la forme d'hélices  $\alpha$ , précédemment observé chez *O. sativa* et *H. sapiens*. De plus, une étude phylogénétique a permis de confirmer que ce domaine était une particularité structurale des GH13\_8. Deux sites de liaisons aux sucres ont également pu être mis en évidence et l'analyse phylogénétique de ces régions a démontré que cette organisation structurale était unique et propre au phylum des *Saccharomycotina*.

Ce travail a fait l'objet d'un article accepté dans le journal *Glycobiology* à comité de lecture international en Octobre 2021 (version finale ci-après).

## A. Article

### **The *Candida glabrata* glycogen branching enzyme structure reveals unique features of branching enzymes of the *Saccharomycetaceae* phylum**

**Key words:** *Candida glabrata*, crystal structure, binding site, glycogen branching enzyme, *Saccharomycetaceae*.

**Léa Conchou<sup>1</sup>, Juliette Martin<sup>1</sup>, Isabelle R. Gonçalves<sup>2</sup>, Frédéric Galisson<sup>1</sup>, Sébastien Violot<sup>1</sup>, Florence Guillière<sup>3</sup>, Nushin Aghajari<sup>1\*</sup>, Lionel Ballut<sup>1\*</sup>**

<sup>1</sup> Molecular Microbiology and Structural Biochemistry, UMR 5086, CNRS Université de Lyon, 7 passage du Vercors, 69367 Lyon, France.

<sup>2</sup> Microbiologie Adaptation et Pathogénie, UMR 5240 CNRS Université de Lyon, Villeurbanne 69622, France.

<sup>3</sup> Institut des Sciences Analytiques, UMR 5280 CNRS Université de Lyon, Villeurbanne, France

\*To whom correspondence should be addressed: e-mails: [lionel.ballut@ibcp.fr](mailto:lionel.ballut@ibcp.fr); [nushin.aghajari@ibcp.fr](mailto:nushin.aghajari@ibcp.fr)

**Running head: Structural analysis of a GH13\_8 glycogen branching enzyme**

**Supplementary data:** Supplementary Figure S1 to S12, Supplementary Table SI

## Abstract

Branching enzymes (BE) are responsible for the formation of branching points at the 1,6 position in glycogen and starch, by catalyzing the cleavage of  $\alpha$ -1,4-linkages and the subsequent transfer by introducing  $\alpha$ -1,6-linked glucose branched points. BEs are found in the large GH13 family, eukaryotic BEs being mainly classified in the GH13\_8 subfamily, GH13\_9 grouping almost exclusively prokaryotic enzymes. With the aim of contributing to the understanding of the mode of recognition and action of the enzymes belonging to GH13\_8, and to the understanding of features distinguishing these enzymes from those belonging to subfamily 13\_9 we solved the crystal structure of the glycogen branching enzyme (GBE) from the yeast *Candida glabrata*, CgGBE, in ligand free forms and in complex with a maltotriose. The structures revealed the presence of a domain already observed in *Homo sapiens* and *Oryza sativa* BEs and that we named  $\alpha$ -helical N-terminal domain, in addition to the three conserved domains found in BE. We confirmed by phylogenetic analysis that this  $\alpha$ -helical N-terminal domain is always present in the GH13\_8 enzymes suggesting that it could actually present a signature for this subfamily. We identified two binding sites (BS) in the  $\alpha$ -helical N-terminal domain and in the carbohydrate binding module 48 (CBM48), respectively, which show a unique structural organization only present in the *Saccharomycotina* phylum.

Our structural and phylogenetic investigation provides new insight into the structural characterization of GH13\_8 GBE revealing unique structural features only present in the *Saccharomycotina* phylum thereby conferring original properties to this group of enzymes.

## Introduction

Glycogen is an intracellular highly branched polymer consisting of  $\alpha$ -1,4-linked glucose polymers with  $\alpha$ -1,6-linked glucose polymer branches. This molecule is the main strategy of carbon and energy storage and is used by many organisms such as archaea, bacteria, animals and yeast. Its synthesis is initiated by glycogenin, a self-glycosylating initiator protein, which is then elongated with  $\alpha$ -1,4-linkages by a glycogen synthase (Mu et al. 1996) and branched with  $\alpha$ -1,6-linkages due to the activity of a group of enzymes known as  $\alpha$ -1,4-glucan branching enzymes (abbreviated as BE; EC 2.4.1.18).

Branching enzymes are part of the widespread GH group of enzymes that is composed of 171 families, amongst others comprising amylolytic enzymes such as  $\alpha$ -amylases (EC 3.2.1.1), pullulanases (EC 3.2.1.41), isoamylases (EC 3.2.1.68), cyclomaltodextrin glucanotransferases (EC 2.4.1.19) and cyclomaltodextrinases (EC 3.2.1.54). Family GH13, also known as the  $\alpha$ -amylase family, constitutes the largest group of glycoside hydrolases which gathers 43 subfamilies (CAZy, <http://www.cazy.org/>; Lombard et al. 2014; Janeček 1997; MacGregor et al. 2001; Stam et al. 2006). Amongst these, subfamilies GH13\_8 and GH13\_9 are classified as  $\alpha$ -1,4-glucan branching enzymes. Previous phylogenetic analysis indicated that they both belong to the same subtree, with each subfamily forming a monophyletic group. GH13\_9 gathers almost exclusively prokaryotic enzymes with the exception of three apicomplexan (*Cryptosporidium parvum*, *Neospora caninum* and *Toxoplasma gondii*) and one metazoan, the brackish water ostracods *Cyprideis torosa* as referenced in CAZy (Stam et al. 2006). GH13\_9 members are found in many prokaryotic phyla like Bacteroidetes, Proteobacteria or Verrucomicrobia. A phylogenetic analysis based on amino acid sequences show that the tree branches do not necessarily reflect taxonomy, suggesting lateral gene transfer and/or differential gene loss (Suzuki and Suzuki 2016). GH13\_8 groups both eukaryotic and prokaryotic enzymes (1477 members in the CAZy database) with respectively 868 members



from Eucarya and 606 from Procarya belonging to Firmicutes, Bacteroidetes, Verrucomicrobia and Deltaproteobacteria (Stam et al. 2006; Suzuki and Suzuki, 2016). Apart from GH family 13, BEs have been identified and classified in the GH57 family known as the second  $\alpha$ -amylase family (Murakami et al. 2006; Martinovičová and Janeček 2018). They use either starch (SBE) or glycogen (GBE) as substrate. All BEs share a common double displacement mechanism with net retention of the  $\alpha$ -configuration of the anomeric carbon. After the proton donor (a glutamate) has donated a proton to the glycosidic oxygen to be cleaved, the  $\alpha$ -1,4-glucan chain is cleaved *via* a nucleophilic attack performed by an aspartate and forming a glycosyl-enzyme covalently bound intermediate with this latter (Takata et al. 1994; Kuriki et al. 1997). In the second step, the glutamate now acting as a base deprotonates a water molecule that will hydrolyze the glycosyl-enzyme intermediate and release the glucan chain. If the water molecule is replaced by the C-6 hydroxyl group of the newly formed glucan chain, the enzyme serves as a glycosyl-transferase rather than a hydrolase, resulting in the formation of a new  $\alpha$ -1,6-glucan ramification (Koshland 1953; Davies and Henrissat 1995).

Three-dimensional structures of six different branching enzymes have been solved yet either in their ligand free- or substrate-bound form. The four prokaryotic enzymes belong to GH13\_9 BEs, with *Rhodothermus marinus* (Jiang et al. 2021; PDB: 6JOY), *Crocospaera subtropica* (Hayashi et al. 2017; PDB: 5GQV), *Escherichia coli* (Abad et al. 2002; PDB: 1M7X) and *M. tuberculosis* (Pal et al. 2010; PDB: 3K1D), and two eukaryotic enzymes belong to GH13\_8 BEs either in plant with *Oryza sativa* (Noguchi et al. 2011; PDB: 3AMK) or in mammalian with *Homo sapiens* (Froese et al. 2015; PDB: 4BZY). They all share a common structural architecture with three conserved domains, namely a carbohydrate-binding module (CBM48), a central catalytic ( $\beta/\alpha$ )<sub>8</sub> barrel domain or catalytic core, also referred to as domain A and a carboxy-terminal domain referred to as domain C. The catalytic barrel contains eight parallel  $\beta$ -strands surrounded by eight  $\alpha$ -helices with the conserved catalytic

residues (Asp, Glu, Asp) carried by the loops that link  $\beta$ -strands to the adjacent helices (Abad et al. 2002). In GH13\_9, the CBM48 is either the first domain encountered at the N-terminus (Jiang et al. 2021) or it can be preceded by an additional  $\beta$ -sandwich fold N-terminal domain referred to as N1 (Pal et al. 2010) or N domain (Hayashi et al. 2017). In GH13\_8, an additional N-terminal domain has also been described (Noguchi et al. 2011; Froese et al. 2015). Suzuki and Suzuki (Suzuki and Suzuki 2020) described this N-terminal domain as an  $\alpha$ -helix preceded by a N-terminal extra sequence (N-EXS). The N-terminal domains preceding the central catalytic domain A could play a major role in substrate recognition (Devillers et al. 2003; Pal et al. 2010), binding and in some cases in the specificity of the chain length transferred (Kuriki et al. 1997; Abad et al. 2002)

In this study, we solved the three-dimensional structure of a fungal glycogen branching enzyme, namely that of the opportunistic pathogen *Candida glabrata* (Ciurea et al. 2020), belonging to the eukaryotic GH13\_8 subfamily. Crystal structures of both ligand-free (CgGBE) and maltotriose-bound (CgGBE<sub>A</sub>-G3) enzyme revealed a conserved multi-domain arrangement with the three canonical domains comprising the CBM48, the catalytic domain A and the domain C. A secondary structure prediction made on 137 GH13\_8 sequences shows the systematic presence of two  $\alpha$ -helices. In CgGBE, these two helices are completed by two additional short  $\alpha$ -helices observed in the structure, we name all four helices the  $\alpha$ -helical N-terminal domain, preceded by a short unstructured segment that could correspond to the N-EXS. A phylogenetic study confirms that this  $\alpha$ -helical N-terminal domain exists in all GH13\_8 enzyme and could represent a *bona fide* signature for this subfamily. The maltotriose binding site (BS<sub>B</sub>) identified on the CBM48 and the binding site (BS<sub>A</sub>) we identified by structural homology on the  $\alpha$ -helical N-terminal domain both revealed original structural features that seems to be restricted to the fungi and more specifically to the *Saccharomycotina* phylum.

## Results

### *Candida glabrata Glycogen Branching Enzyme (CgGBE) structure and activity*

CgGBE, 706 amino acids long and encoded by *GLC3*, belongs to the GH13 family and has been classified in subfamily GH13\_8 including eukaryotic glycogen branching enzymes. The full-length form of the protein was produced in *E. coli* cells and the protein purified and eluted as a soluble monomer (Sup Figure 1). Two crystal structures of the ligand free enzyme (CgGBE<sub>A</sub> and CgGBE<sub>B</sub>) and one structure of an enzyme complex obtained by soaking with a maltotriose (CgGBE<sub>A</sub>-G3) have been solved respectively to 2.4 Å, 2.2 Å and 1.93 Å resolution by molecular replacement using the *H. sapiens* GBE (*hGBE1*) structure as search model (PDB: 4BZY) (Table I). The first ligand free CgGBE<sub>A</sub> and the complex CgGBE<sub>A</sub>-G3 both display two monomers per asymmetric unit whereas the second CgGBE<sub>B</sub> displays only one monomer per asymmetric unit. SAXS confirmed that the enzyme is monomeric in solution (Sup. Figure 2).

The protein can be divided in four different parts, an  $\alpha$ -helical N-terminal domain (residues 1-64) with two short and two long helices (helix 1 (Q11-D17), helix 2 (P18, F23), helix 3 (A24-H43) and helix 4 (L53-Y64)), a CBM48 domain (residues 65-195), the central catalytic canonical domain A (residues 196-609) and domain C (610-706) (Figure 1, Figure 2, Figure 5A). All three structures are overall identical with alpha carbon (C $\alpha$ ) RMSD of 1.88 Å between CgGBE<sub>B</sub> and CgGBE<sub>A</sub> or CgGBE<sub>A</sub>-G3 and 0.6 Å between CgGBE<sub>A</sub> and CgGBE<sub>A</sub>-G3. Every residue shows clear electronic density except for loop  $\beta$ 4- $\alpha$ 4 which is partially disordered and CgGBE<sub>B</sub> N-terminal domain lacking the first 30 residues (Figure 1D,6B). In addition, a local rearrangement has been observed in the catalytic domain A segment 504-512 folded either as a one turn helix in CgGBE<sub>A</sub> (507-510) or a two turn helix in CgGBE<sub>B</sub> (504-508) (Sup Figure 3). The catalytic domain A adopts the canonical ( $\beta/\alpha$ )<sub>8</sub> barrel fold common to GH13 family enzymes whereas the CBM48 and domain C both adopt a  $\beta$ -sandwich fold

(Figure 1C). Loops interconnecting the eight  $\beta$ -strands of the barrel delimit the active site nested in a central cleft where the conserved catalytic residues (D358 (nucleophile), E419 (acid/base) and D488 (transition state stabilizer) reside (Figure 1B) (Takata et al. 1994; Kuriki et al. 1997). More generally, residues shown to interact with a maltoheptaose (G7) in the *C. subtropica* active site are well conserved when compared to all other available structures (Sup. Figure 4A). Loop  $\beta$ 3- $\alpha$ 3 which forms one side of the cleft extended from sugar binding subsites -1 to -5 is well conserved from a structural and sequence point of view compared to other eukaryotic and prokaryotic enzymes (Sup. Figure 4B, 4C). The overall surface around the active site pocket is highly negatively charged as observed for all members of  $\alpha$ -amylase family 13 promoting interaction with the substrate (Sup. Figure 4D) (Abad et al. 2002; Noguchi et al. 2011).

When *C. glabrata* act as an opportunistic pathogen for human cells, its infection will cause the activation of the immune system of the host especially macrophages that will engulf and kill the pathogen by exposing it to a combination of low pH and carbon starvation. However, it has been demonstrated that *C. glabrata* yeasts are not only able to resist this confrontation with macrophages but also use them to replicate and hide from other immune cells (Kasper et al. 2015). One of the evasion strategies of the fungi is to prevent the maturation process of the phagosome and its acidification thus avoiding the hostile phagolysosomal stage. Therefore, *in vitro* studies showed that the yeast growth is unaffected by low pH values even at pH 2.0 (Bairwa and Kaur 2011). To investigate the enzymatic activity of CgGBE under several pH and temperature conditions, iodine assays were performed using Amylose V as substrate. Interestingly, the relative activity of the enzyme is only 35 % at 35 °C (Sup. Figure 5), which is lower than what has been observed for *O. sativa* plant BEs, for the bacterial *E. coli* BE and the yeast *S. cerevisiae* BE retaining 50 % of their activity against amylopectin at 35 °C. Hence, the relative activity of CgGBE is 10 % at pH 5.0 suggesting that

the cells reroute their metabolism under drastic conditions as low pH to ensure efficient carbon and energy supply.

### ***Maltotriose binding to the CBM48 domain***

To characterize the oligosaccharide binding type in CgGBE, we performed soaking experiments with maltotriose (G3), maltotetraose (G4) and maltopentaose (G5) using the crystals leading to the crystal structure of CgGBE<sub>A</sub>. Whereas soaking with G4 and G5 resulted in data which were not exploitable, soaking with G3 resulted in X-ray diffraction data giving electron density of a good quality allowing us to fit the trisaccharide in the CBM48 domain. Starting from the reducing end, the first and second sugar moieties of the trisaccharide are stacked by W96 and Y145, respectively, on one side and Y165 on the opposite side forming a “pair of sugar tongs” like organization as seen in barley  $\alpha$ -amylase (Robert et al. 2003) and that we named BS<sub>B</sub> (Figure 3). Y165 is present on an  $\alpha$ -helix at the top of a loop that we named “CBM48 binding loop” (T155-E170) (Figure 3). Interestingly, when we compare the ligand free structure CgGBE<sub>A</sub> with that of the enzyme complex CgGBE<sub>A</sub>-G3, it can be observed that the helix bearing Y165 no longer forms a canonical  $\alpha$ -helix and starts to form a loop. It seems that in the presence of the maltotriose, the ligand could force segment 162-165 to form an additional turn, the stabilization being assured by the reorientation of Y165 which stacks to one side of the tri-saccharide (Figure 3A, 3B). G3 binding is completed by a hydrogen bond between E296, whose side chain is reoriented, and the second sugar moiety (Figure 3A). It can be noted that the ligand free CgGBE<sub>B</sub> adopts the same structured two turn helix as observed for CgGBE<sub>A</sub>-G3, suggesting a high flexibility of this segment that, even in the absence of substrate, could easily switch from a loop to a helix, likely to accommodate a branched substrate (Figure 3C, Sup. Figure 11). The loop/helix bearing Y165 is either lacking or shorter in the branching enzymes of *O. sativa* (OsBEI, PDB: 3VU2) and *H. sapiens* (hGBEI, PDB: 5CLW) limiting a potential structuration (Sup. Figure 6A, 6B). In OsBEI, the two aromatic residues W88 and W133 correspond respectively to

W96 and Y145 in *C. glabrata*. In addition, a phenylalanine (F151) present in the loop above the binding site could replace Y165 to adopt a “pair of sugar tongs” like organization (Figure 3A). However, this loop seems too short to form a helix. In *hGBEI* only a tyrosine is conserved (Y155) suggesting another type of interaction (Sup. Figure 6B). In the structures of GH13\_9 BE from *C. subtropica* (*CsBE*, PDB: 5GR1) or from *E. coli* (*EcBE*, PDB: 4LPC) determined in complex with oligosaccharides, the authors observed the presence of a trisaccharide or a disaccharide, respectively, similarly oriented and positioned at binding sites named respectively 48\_1 or IV (Feng et al. 2016; Hayashi et al. 2017). However, due to the absence of an additional loop in prokaryotes, the binding mode is different as both proteins form a single stacking interaction with W163 (*CsBE*) or W159 (*EcBE*), in addition to respectively three (*CsBE*) and five (*EcBE*) hydrogen bonds in each site (Figure 3D, 3E). We moreover observed a significant displacement of the trisaccharide compared to the ones observed in *CsBE* and *EcBE* (Figure 3F).

#### ***$\alpha$ -Helical N-terminal domain of CgGBE***

Differences in the domain composition and organization of GH13\_8 and GH13\_9 BE have already been observed (Abad et al. 2002; Pal et al. 2010; Noguchi et al. 2011; Froese et al. 2015; Hayashi et al. 2017; Jiang et al. 2021). In the structures of *M. tuberculosis* (*MtBE*; PDB: 3K1D) and *C. subtropica* (*CsBE*, PDB: 5GR1), the N-terminal domain adopts a  $\beta$ -sandwich fold resembling the CBM48. When looking at the structure, this additional domain is either located at one end of the elongated protein or hidden behind the catalytic domain A (Figure 4A, 4B). In *CgGBE* the  $\alpha$ -helical N-terminal domain is located between the CBM48 and the catalytic domain A and interacts both with the CBM48, helix 4, and with the catalytic domain A, helices 1, 3 and 4 (Figure 5A, 5B). It superposes well with the  $\alpha$ -helical N-terminal domains observed in *OsBEI* and *hGBEI* (Figure 5A).

In *Os*BEI and *h*GBEI the  $\alpha$ -helical N-terminal domains have been shown to bind either a pentaose (G5) or a heptaose (G7) (Sup. Figure 7A, 7B); Chaen *et al.* 2012; Froese *et al.* 2015). This BS is referred to as site 1 in *O. sativa* or 48\_2 in *C. subtropica* (Chaen *et al.* 2012; Hayashi *et al.* 2017). When superimposing the structures of *h*GBEI, or *Os*BEI onto that of *Cg*GGE, it seems that the same BS putatively exists in *C. glabrata* and that it could accommodate either a pentaose or a heptaose (Figure 5C). Indeed, the bound oligosaccharides as seen in the *Os*BEI and *h*GBEI crystal structures are stacked by three conserved aromatic residues forming a sandwich-like interaction with residues W91, Y120 and W332 in *h*GBEI, W72, F100 and W319 in *Os*BEI and W80, F109 and F333 in *Cg*GGE (Figure 5C). We named this binding site BS<sub>A</sub>. Interestingly, in *Cg*GGE, a longer loop is present between helices 3 and 4 (Figure 5A, 5C). This loop does not seem to impair the binding to a heptaose even if both residues N62 in *h*GBEI and E45 in *Os*BEI stabilizing the fifth glucose (Glc-5) are not conserved and substituted by A44 in *Cg*GGE. Residues T45, P46, D47 present in the loop, could even play a stabilizing role as they extend the binding site covering two additional sugar moieties (Figure 5C).

### ***Phylogenetic analysis of GH13\_8***

In order to improve our understanding of the structural organization of the GH13\_8 subfamily, the repartition of its members was examined through a phylogenetic analysis. Conservation of the  $\alpha$ -helical N-terminal domain and the repartition of the two structurally original binding sites found at the surface of the enzyme were investigated. For analysis of the distribution of  $\alpha$ -(1,4)-glucan branching enzymes in GH13\_8, we first selected sequences from UniProtKB ([www.uniprot.org](http://www.uniprot.org)) and the CAZy database representatives of Chromalveolata, Hacrobia, Plantae, Excavata, Amoebozoa, Mycota, Filasterea, Choanoflagellida and Metazoa. We then added 17 GH13\_8 sequences from prokaryotic members present in the CAZy database corresponding to a total of 137 sequences. Phylogenetic analysis showed a clear separation between prokaryotes and eukaryotes (Sup.



Figure 8). The repartition of the different phyla respects the taxonomy (Sup. Figure 8). All the prokaryotic groups (Firmicutes, Bacteroidetes, Verrucomicrobia, Deltaproteobacteria) present in GH13\_8 are present in GH13\_9 as well (CAZy). Among the 17 prokaryotic GH13\_8 members added, three firmicutes, *Anaerococcus prevotii*, *Clostridioides difficile* and *Eubacterium callanderi* which possess one copy for both GH13\_8 and GH13\_9 enzyme coding genes are included, suggesting, as proposed by Suzuki and Suzuki (Suzuki and Suzuki 2016) that the appearance of GH13\_8 in prokaryotes could result from a lateral gene transfer followed, in some phyla, by the GH13\_9 gene deletion (Sup. Figure 8).

A sequence alignment and a secondary structure prediction made with the 137 sequences used for the phylogenetic analysis, show that all the studied GH13\_8 enzymes possess the characteristic additional  $\alpha$ -helical N-terminal domain composed of a minimum of two helices in addition to an N-EXS of various length preceding the CBM48 (Figure 6; Sup. Figure 9). The closest GH13\_8 and GH13\_9 subfamilies are GH13\_10, 11, 12, 13 and 14 (Stam et al. 2006). Whereas all possess a CBM48, for subfamilies 12, 13 and 14, CBM48 is not positioned at the N-terminal part of the protein but rather embedded into the sequence and surrounded by other domains (Møller et al. 2016). Regarding subfamilies 10 and 11, the CBM48 is sometimes observed at the N-terminus, however, even in this case it does not seem preceded by an  $\alpha$ -helical domain (Feese et al. 2000; Møller et al. 2016). Altogether, we suggest that the  $\alpha$ -helical N-terminal domain could have appeared in the GH13\_8 ancestor conferring to this family some of its branching properties. Another interesting feature of CgGBE are the two BSs found either at the surface of the CBM48 (BS<sub>B</sub>) or at the surface of the  $\alpha$ -helical N-terminal domain (BS<sub>A</sub>). BS<sub>B</sub> is referred to as 48\_1 in *C. subtropica*, 2 in *O. sativa* or IV in *E. coli* (Chaen et al. 2012; Feng et al. 2016; Hayashi et al. 2017). In GH13\_11-14, the residues forming this canonical BS are not conserved (Janeček et al. 2011; Møller et al. 2016). In GH13\_8 only W96 seems to be well conserved (Sup. Figure 10). Interestingly, Y165, the third aromatic residue forming the “pair of sugar tongs” like organization, is



present in *S. cerevisiae* as well (Y155). Analysis of the *Saccharomycotina* subphylum show that this tyrosine along with the helix forming loop which bears this residue are only present in the *Saccharomycetaceae* and *Saccharomycodaceae* groups (Figure 7). The aromatic Y165 in *C. glabrata* corresponds to either a phenylalanine or tryptophan throughout this group, suggesting that the “pair of sugar tongs” like organization could form in all species of the group. The second binding site (BS<sub>A</sub>) referred to as 48\_2 in *C. subtropica* or 1 in *O. sativa*, is completed in *C. glabrata* by an extended loop (H44-L52) present between helices 3 and 4 (Figure 5A, 5C). Surprisingly, in addition to *C. glabrata*, this longer loop is only present in *S. cerevisiae* (Sup. Figure 10). Again, this feature seems restricted to the *Saccharomycotina* subphylum and more specifically to the *Saccharomycetaceae*, as this characteristic seems not to be present in *Saccharomycodaceae* indicating that it could have either been lost or acquired independently of BS<sub>B</sub> (Figure 7).

## Discussion

Although GH13 enzymes have been widely studied over decades, branching enzymes are still poorly understood due to a lack of functional and structural studies and a subsequent analysis of the relation between these. In this study we present the three-dimensional structure of the third eukaryotic- and the first fungal glycogen branching enzyme. Branching enzymes play a crucial role in glycogen ramification. Our two ligand-free and one maltotriose-bound structures confirmed once again the very strong conservation of the multi-domain arrangement of GBEs with an  $\alpha$ -helical N-terminal domain, a CBM48, a catalytic domain A and a C-terminal domain C (Figure 1).

Our studies suggested that the  $\alpha$ -helical N-terminal domain could be a signature of the GH13\_8. Interestingly, in the three existing structures of GH13\_8, this domain is always positioned between the CBM48 and the catalytic domain A with the helices similarly oriented.

For a cyanobacterial branching enzyme, a reaction scheme was suggested on the basis of a substrate binding model explaining the hydrolytic/transglycolytic activity of the enzyme, which highlighted a so called “entry” and “exit” gate located at the edge of the catalytic domain A (Hayashi et al. 2017). The authors identified two binding sites named A1 and A2 that they respectively identified as the entry and exit gates of the acceptor binding site, the acceptor oligosaccharide extending from A1 to A2 by forming a U shape (Sup. Figure 11). The A1 site is a one turn helix and the authors suggest that two residues, Y500 and D501 could participate in the binding of the oligosaccharide (Sup. Figure 11A). Interestingly, this structural feature is conserved in all eukaryotic and prokaryotic structures available, but the two identified residues are not conserved and are substituted by a threonine and an alanine, respectively, in *CgGBE* (Sup. Figure 11A, 11B). More surprisingly, we observed that the  $\alpha$ -helical N-terminal domain masks the putative entry gate in *CgGBE*, *OsBE* and *hGBEI* and seems to prevent the passage of an extended oligosaccharide from this putative entry site to the catalytic site (Sup. Figure 11C). This amino-terminal organization which differs between subfamilies 8 and 9 was suggested to be due to a difference in substrate since glycogen is more highly branched than amylopectin (Noguchi et al. 2011). In this domain the presence of two highly conserved aromatic residues in helix 4, a phenylalanine (F56) and a tyrosine (Y61), suggest potential roles of these in interactions with a glucan chain in this domain (Figure 6A, 6B). Furthermore, two other residues are very well conserved, an arginine (R31) in helix 3, and an aspartate (D17) at the end of the N-EXS, again, suggesting potential interactions with the substrate. In contrast, a recent study investigated the substrate affinity of all seven BSs identified in *CsBE* (El Mannai et al. 2021). The authors concluded that among all, only A2, C1 and C2 sites showed a relevant affinity for polysaccharides, questioning the binding mechanism previously described and the functionality of A1 as entrance site (Hayashi et al. 2017).

Interestingly, in the CgGBE<sub>B</sub> ligand free structure, no electron density is visible for the first 30 residues (Figure 6B). This lack is due to the crystal packing and does not necessarily reflect what could happen *in vivo*. However, this observation suggests that the  $\alpha$ -helical N-terminal domain shows a certain flexibility which could participate in the reorganization of an entry gate specific to GH13\_8.

Regarding the exit gate, it has been shown that in CsBE, the interaction of a G3 with the A2 binding site depends on residues L541 and W655 carried by a long helix (D511-E546) which is part of loop  $\beta$ 6- $\alpha$ 6. As for the A1 binding site, Leu541 and Trp655 are not conserved in the different available structures of the GH13 branching enzymes family. Moreover, the helix 511-546 carrying these two residues is shorter in the eukaryotic BEs and Leu541 has no counterpart in the corresponding structures (Sup. Figure 11D). In CgGBE, OsBE or hGBEI, W655 correspond respectively to V611, V591 and V604.

Charge surface representation shows a highly negative charged cleft corresponding to the catalytic site as observed for other GH13 family members (Sup. Figure 4D). Our attempts to bind oligosaccharides to this cleft remained unsuccessful and indeed only one three-dimensional structure exists with an oligosaccharide bound in the catalytic cleft compared to the numerous structures showing complexes with oligosaccharides bound at other locations or BSs (Pal et al. 2010; Chaen et al. 2012; Froese et al. 2015; Feng et al. 2016; Hayashi et al. 2017). This is most likely due to the fact that branching enzymes do not transfer short chains as shown by Feng and colleagues (Feng et al. 2016) for the *E. coli* BE which does not transfer chains being shorter than five residues and which has a strong preference for chains consisting of 10–12 moieties. The authors argued that the BEs which display very similar active sites with other GH13 enzymes, in order to avoid the typical reactivity of these latter (hydrolysis or transfer of relatively short glucan chains), distributes the substrate binding sites all over the enzyme space in a manner that only long glucans would be able to occupy simultaneously external binding sites and the active site (Feng et al. 2015). For this reason, the shorter

oligosaccharides, as e.g. maltotriose used in this study, would not have sufficient binding affinity for the enzyme thereby explaining the lack of binding. It was suggested that the BSs may act as an anchor for glycogen donor and acceptor chains and determine the chain lengths specificity as a “molecular ruler” (Chaen et al. 2012; Froese et al. 2015). As for *OsBE* and *CsBE* we could identify two BSs, one experimentally (BS<sub>B</sub>) and one by homology (BS<sub>A</sub>) to earlier determined crystal structures, on the CBM48 domain in a manner similar to what has been observed for CBM20 (Janeček et al. 2019). Both carbohydrate binding domains are suggested to be evolutionary close relatives even if residues of both sites have been shown to be only partially conserved (Janeček et al. 2019). Whereas binding to the homologous BS<sub>A</sub> site of *OsBE* does not correspond to what is seen in CBM20, the second BS of *OsBE* seems to be well conserved between these CBMs with an aromatic residue, Trp88, and the side chain of a Lys99 (Chaen et al. 2012). In contrast, BS<sub>B</sub> presents a different type of binding with more resemblance to that of the “pair of sugar tongs” like as identified in barley  $\alpha$ -amylase (Robert et al., 2003).

Binding of maltotriose to *CgGBE* from soaking experiments allowed a clear identification of a trisaccharide bound to the BS<sub>B</sub> of the CBM48 domain. A small segment at the top of the CBM48 binding loop is able to structure into a helix along with the reorientation of Y165 and E296 residues thus resulting in a “pair of sugar tongs” like motif with both W96 and Y145 residues. This strongly suggests, that BS<sub>B</sub> can probably switch from one fold to another in order to accommodate the substrate. Interestingly, phylogenetic studies performed among all GH13\_8 enzymes allowed us to determine that this feature is not unique to *CgGBE* but also present in all BEs in the *Saccharomycotina* subphylum. The role of the CBM48 binding loop in oligosaccharide binding has been investigated by evaluating the stability of native- and mutated enzymes, Y165A and W69A (Sup. Figure 12). We observed that both mutations had only a slight effect on the overall stability of the enzyme in the presence of G7 or glycogen suggesting that other binding sites could compensate for the binding of oligosaccharides.

However, the Trp69 mutant showed an increase in the overall stability of the enzyme in absence of oligo- or polysaccharides suggesting that the mutation might impair the folding of the CBM48 loop to an alpha-helix thereby lowering the flexibility of the enzyme. These results suggest that BS<sub>B</sub> is not directly involved in the binding of glucan but is more likely play a role in the chain length specificity of glucan chains transferred as it has been previously suggested (Kuriki et al. 1997; Binderup et al. 2000; Binderup et al. 2002; Froese et al. 2015).” BS<sub>A</sub> was identified by structural homology in the  $\alpha$ -helical N-terminal domain and shown to be a hydrophobic sandwich-like binding site further stabilized by an elongated loop between helices 3 and 4. We observed that this characteristic is not present in the *Saccharomycodaceae* indicating that the two BSs could confer properties to the yeast independently from one another.

## **Material and methods**

### **Expression and purification of CgGBE and mutants**

A codon-optimized gene encoding the GBE from *Candida glabrata* was synthesized by GenScript and cloned into *Nde*I and *Bam*HI sites of the pET15b vector (Novagen) with a 6xHis-tag fused at the N-terminal. CgGBE WT was produced in *Escherichia coli* T7 express strains (New England Biolabs), cultivated in 2 L of Luria-Bertani (LB, Sigma-Aldrich) medium supplemented with 50  $\mu\text{g}\cdot\text{mL}^{-1}$  ampicillin at 160 rpm and 37 °C. CgGBE expression was induced by 0.2 mM isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG, Sigma-Aldrich) when the culture reached an OD<sub>600nm</sub> of 0.6; the induction was carried out at 160 rpm and 25 °C for 20 hours. Cells were harvested by centrifugation at 5,000 x g for 15 min at 4 °C and the pellet was resuspended in 60 mL of 10 mM HEPES pH 7.6, 300 mM NaCl and supplemented by one tablet of cOmplete, EDTA free protease inhibitor cocktail (Roche), 25  $\mu\text{g}\cdot\text{mL}^{-1}$  lysozyme (Sigma-Aldrich), 5  $\mu\text{g}\cdot\text{mL}^{-1}$  DNaseI and RNase I (Sigma-Aldrich). Bacteria were lysed by

using the Microfluidizer (M-110P, Microfluidics) at 1,500 bar. The lysate was then centrifugated at 11,000 x g for 45 min at 4 °C and the supernatant was applied to a Co-NTA column (HiTrap® TALON® crude 5 mL, GE Healthcare, UK) beforehand washed with lysis buffer containing 30 mM imidazole (Merck). Elution was performed with an ÄKTA Purifier (GE Healthcare) with a linear gradient from lysis buffer to 100 % of elution buffer containing 500 mM imidazole. The eluted protein was then desalted using a HiPrep© Desalting column (GE Healthcare, UK) equilibrated with a desalting buffer composed of 10 mM HEPES pH 7.6, 150 mM NaCl prior to the experiment. The eluted protein was concentrated using an Amicon Ultra Centrifugal filter with a 50 kDa molecular weight cut-off (Millipore™ Corporation) and loaded on a gel filtration column (Superdex-200, GE Healthcare). The protein concentration was assessed with a NanoDrop (Thermo Fisher Scientific) and the purity and molecular mass of the protein were checked by SDS-PAGE using gel composed of 4 % stacking and 15 % running gel after Coomassie staining.

The CgGBE-WT 1242-bp cDNA fragment was mutated at residues Tyr165 and Trp69 using oligonucleotides 5'-GGAGACCGCGCAAAAAGCGGGTCCGACCTATGAAGGC-3' (sense)/5'-GCCTTCATAGGTTCGGACCCGCTTTTTGCGCGGTCTCC-3' (antisense) and 5'-GGCGAGTTCAACAACGCGAACGAGGAAAGCC-3' (sense)/5'-GGCTTTCCTCGTTCGCGTTGTTGAACTCGCC-3' (antisense) respectively according to the principle of PCR-mediated point mutagenesis. Expression and purification of mutants were performed as described above.

## Enzyme kinetics and characterization

### *Iodine assay*

The enzyme activity was measured with Amylose V as substrate since it has no detectable  $\alpha$ -1,6-linkages by NMR. Reaction of the enzyme was followed using the iodine staining assay by monitoring the decrease of absorbance of the glucan-iodine complex (Krisman 1962). Amylose 0.5 % (w/v) from Sigma-Aldrich was dissolved in sodium 50 mM phosphate buffer pH 7.0 and incubated with 210  $\mu\text{g}\cdot\text{mL}^{-1}$  CgGBE at 25 °C. Aliquots of 10  $\mu\text{L}$  were taken at different times, added to 150  $\mu\text{L}$  of iodine reagent (aqueous solution of I<sub>2</sub> 0.01 % (w/v), KI 0.1 % (w/v), HCl 1 M) and the absorbance was measured at 660 nm.

### *Influence of pH and temperature on activity*

The influence of temperature on CgGBE was measured at pH 7.0 in 50 mM sodium phosphate buffer using the iodine staining assay in the temperature range 5 to 40 °C. The influence of pH on CgGBE activity was measured at 25 °C in 50 mM Tris-HCl pH 5.5, 50 mM MES pH 6.5, 50 mM sodium phosphate pH 8.0 and citrate pH 9.0.

### *Thermostability*

The thermostability of CgGBE was assessed using nanoscale differential scanning fluorimetry (nanoDSF; NanoTemper) by measuring the intrinsic dual-UV fluorescence change in tryptophan and tyrosine residues at emission wavelengths of 330 nm and 350 nm. The melting temperature of CgGBE WT and mutants (CgGBE-Y165A and CgGBE-W69A) was measured in the presence or absence of G7 (10, 20 and 40 mM) or glycogen (1 and 10  $\text{mg}\cdot\text{mL}^{-1}$ ) and screened over a temperature range from 25 °C to 95 °C using a heat gradient of 1 °C/min.

## Crystallization, Data collection and Structure Determination

CgGBE WT was concentrated to 42.7 mg.mL<sup>-1</sup> and 36.9 mg.mL<sup>-1</sup> in 10 mM HEPES pH 7.6, 150 mM NaCl after gel filtration. Screening of crystallization conditions was carried using the sitting-drop vapor-diffusion method employing commercial crystallization kits: MB Class suite (Qiagen), PEG I (Qiagen), PEG II (Qiagen), PEG ION (Hampton Research), AmSO<sub>4</sub> (Qiagen) and JCSG+ suite (Qiagen). MRC crystallization plates (Molecular Dimensions) were filled using a Mosquito<sup>®</sup> crystallization robot (STP Labtech) and two protein/crystallization agent ratios were tested (200 nL + 200 nL and 200 nL + 100 nL drops equilibrated against 70 μL in the reservoir). Crystals were grown in 0.1 M trisodium citrate pH 5.6, 10 % PEG 4000, 10 % isopropanol and 0.1 M MES pH 6.5, 25 % (w/v) PEG 4000. Crystals of CgGBE WT in complex with G3 (CgGBE<sub>A</sub>-G3) were obtained by soaking crystals in a crystallization condition drop containing 5 mM G3 for 15 min at room temperature. Crystals were harvested and further cryo-protected in the same solution to which 15 % ethylene glycol had been added. X-ray diffraction data were collected at ID23-2 – ESRF (Grenoble, France) and PX1 – Soleil (Paris-Saclay, France) at wavelengths of 0.87313 Å and 0.9786 Å, respectively. Data were indexed, integrated and scaled using XDS (Kabsch 2010). The phase problem was solved by molecular replacement using phenix.phaser (PHENIX, (Adams et al. 2010) with the human glycogen branching enzyme (*h*GBE1, PDB entry 4BZY) as search model (59 % identity). The model was then built using Coot (Emsley et al. 2010) and refinement was carried out using PHENIX (Adams et al. 2010). Data collection and refinement statistics for native and maltotriose-bound forms of CgGBE are shown in Table I. Pymol (Schrödinger, LLC) was used to render the molecular structures.



## Small Angle X-ray Scattering

Samples were concentrated to 5.5 mg.mL<sup>-1</sup> in 10 mM HEPES pH 7.6, 150 mM NaCl for SAXS data collection at the SWING beamline, SOLEIL (Paris-Saclay). Data were processed using PRIMUS (Konarev et al. 2003) and calculation of the P(r) was carried out using CRY SOL (Svergun et al. 1995). *Ab initio* envelope models of CgGBE were generated using DAMMIF (Franke and Svergun 2009). Models were superimposed onto the crystal structure using the CHIMERA software (University of California, San Francisco).

## Phylogenetic analysis

Homologs were searched using BLASTP and CgGBE as a query sequence (Altschul et al. 1990). The search was done against the UniProtKB database. In order to have the most complete coverage of the diversity of GH13\_9 among eukaryotes, 38 representative eukaryotic groups were used for the analysis (Sup. Table I). For each of these, a protein BLASTP search was done on the NCBI website against the UniProtKB database or the non-redundant protein sequences database, when no sequences were recovered from UniProtKB. The groups are represented in the phylogenetic tree (Sup. Fig. S5) by 38 different names and colors. Each search set was limited to one of the taxonomic groups represented in Supplementary Table I. All sequences that aligned at least on 80 % of the length of the three canonical domains (CBM48, Catalytic core domain, C domain) of CgGBE sequence were kept for multiple sequence alignment. In parallel, a Pfam analysis was done to confirm the presence of these three domains in the kept sequences. The  $\alpha$ -helical N-terminal domain of CgGBE was not considered during these two steps in order not to introduce bias in the process. Sequences were then clustered with CD-HIT to keep a significant number of sequences for each group. Other prokaryotic GH13\_8 sequences from the CAZy database were then manually added, eight from Bacteroidetes, four from Deltaproteobacteria, four from Firmicutes and one from Verrucomicrobia. Multiple sequence alignment was done with the

MAFFT program (Kato et al. 2002) and conserved blocks were selected by using BMGE 1.12 (Criscuolo and Gribaldo 2010). The BLOSUM30 (Henikoff and Henikoff 1992) matrix was used with both programs. 487 sites were kept for further analysis after character trimming performed by BMGE (Criscuolo and Gribaldo 2010). Phylogenetic analyses were performed, with the LG model and a gamma correction, using a bootstrapped maximum-likelihood approach with PhyML (Guindon et al. 2010). 100 bootstrap replicates presented the default value when using PhyML. The phylogenetic tree was generated and visualized with iTOL software (Letunic and Bork 2019).

## **Bioinformatics**

For the BEs with an available three-dimensional structure, sequence and structure alignments were done using PROMALS3D (Pei et al. 2008) and figures were made with ESPript v3.0 (Robert and Gouet 2014). Regarding the 137 selected sequences used for phylogeny, sequence analyses were done using the Bioinformatics Toolkit resource at <https://toolkit.tuebingen.mpg.de> (Zimmermann et al. 2018; Gabler et al. 2020). The protein sequences were aligned using Clustal Omega version 1.2.1 (Sievers et al. 2011). Secondary structure predictions of aligned sequences was performed by PSIPRED (Jones 1999) employing the Ali2D Tool of the Bioinformatics Toolkit resource. Sequence logos were rendered using the Alignment Viewer tool (Yachdav et al. 2016) of the Bioinformatics Toolkit resource. The pixel representation of the multiple sequence alignments was generated using the Protein multiple sequence alignment viewer (<https://alignmentviewer.org/>). As for the 23 sequences used to study the distribution of the two BSs in the *Saccharomycotina* subphylum, *C. glabrata* CgGBE (Cgl) was aligned with orthologs in *Saccharomycotina* species. Orthologs were identified by searching “EnsemblFungi”.

## Figure rendering

Figures of three-dimensional structures were drawn with PyMol (Schrödinger, <http://pymol.org>)

## Data availability

Coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 7P44 (CgGBE<sub>A</sub>), 7P45 (CgGBE<sub>B</sub>) and 7P43 (CgGBE<sub>A</sub>-G3), respectively. Other data are available from the corresponding authors upon reasonable request.

## Abbreviations

BE, glucan branching enzyme; BS, binding site; CAZy, Carbohydrate-Active enZymes; CBM, carbohydrate binding module; *Cg*, *Candida glabrata*; Co-NTA, cobalt-nitriloacetic acid; GBE, glycogen branching enzyme; GH13, glycoside hydrolase 13; G3, maltotriose; G4, maltotetraose; G5, maltopentaose; G7, maltoheptaose; IPTG, Isopropyl- $\beta$ -D-thiogalactopyranoside; LB, Luria-Bertani; N-EXS, N-terminal extra sequence; PEG, polyethylene glycol; RMSD, root mean square deviation; rpm, revolution per minute; SAXS, Small Angle X-ray Scattering; SBE, starch branching enzyme; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

## Acknowledgments

Technical support from staff on beamlines MX (European Synchrotron Radiation Facility, Grenoble), on beamlines PXIII (Swiss Light Source, Switzerland), as well as on beamlines PX1 (Soleil, France) and SWING (Soleil, France) is gratefully acknowledged. We are also

grateful for assistance from V. Gueguen-Chaignon and F. Delolme of the Protein Science Facility of SFR Biosciences Lyon (UAR3444/US8). Support from the CNRS is acknowledged.

## Competing interests

The authors declare no competing interests.

## References

- Abad MC, Binderup K, Rios-Steiner J, Arni RK, Preiss J, Geiger JH. 2002. The X-ray crystallographic structure of *Escherichia coli* branching enzyme. *J Biol Chem.* 277(44):42164–42170.
- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW, et al. 2010. PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* 66(Pt 2):213–221.
- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. *J Mol Biol.* 215(3):403–410.
- Bairwa G, Kaur R. 2011. A novel role for a glycosylphosphatidylinositol-anchored aspartyl protease, CgYps1, in the regulation of pH homeostasis in *Candida glabrata*. *Mol Microbiol.* 79(4):900–913.
- Binderup K, Mikkelsen R, Preiss J. 2000. Limited proteolysis of branching enzyme from *Escherichia coli*. *Arch Biochem Biophys.* 377(2):366–371.
- Binderup K, Mikkelsen R, Preiss J. 2002. Truncation of the amino terminus of branching enzyme changes its chain transfer pattern. *Arch Biochem Biophys.* 397(2):279–285.
- Chaen K, Noguchi J, Omori T, Kakuta Y, Kimura M. 2012. Crystal structure of the rice branching enzyme I (BEI) in complex with maltopentaose. *Biochem Biophys Res Commun.* 424(3):508–511.
- Ciurea CN, Kosovski I-B, Mare AD, Toma F, Pinteua-Simon IA, Man A. 2020. *Candida* and candidiasis-opportunism versus pathogenicity: a review of the virulence traits. *Microorganisms.* 8(6):857.
- Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. *BMC Evol Biol.* 10:210.

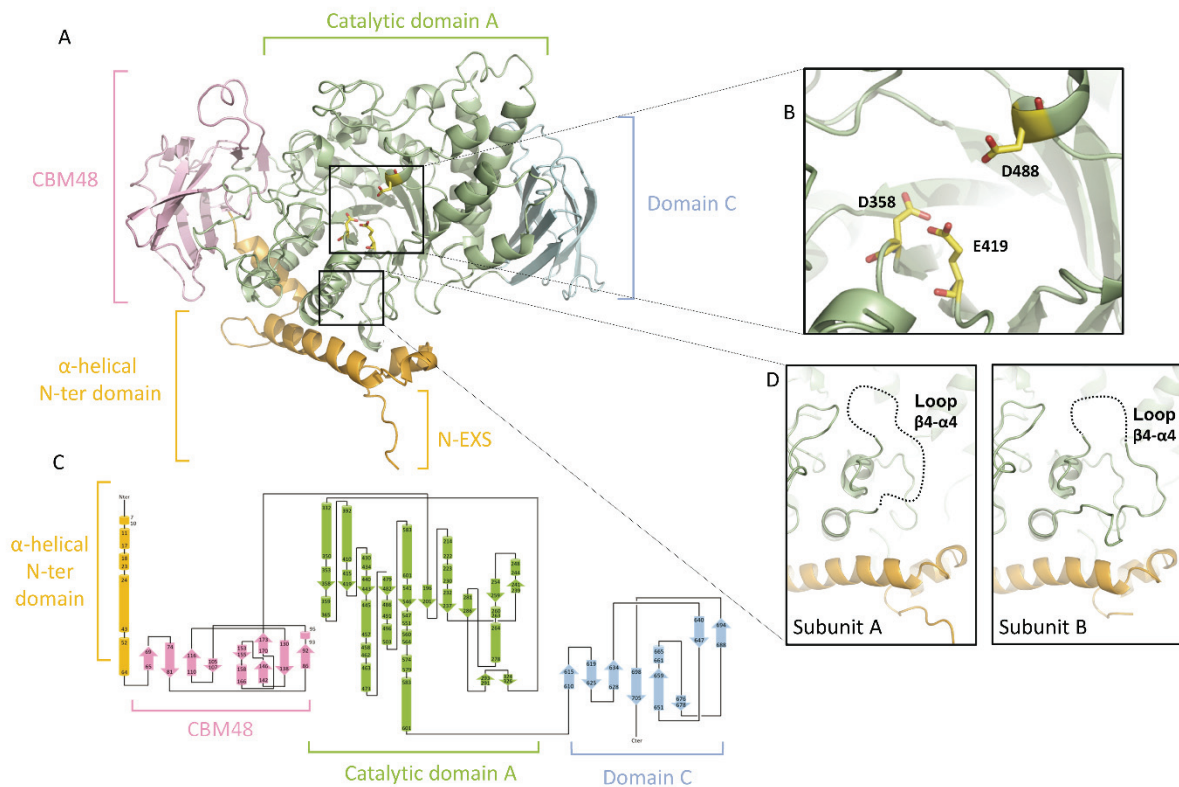
- Davies G, Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure*. 3(9):853–859.
- Devillers CH, Piper ME, Ballicora MA, Preiss J. 2003. Characterization of the branching patterns of glycogen branching enzyme truncated on the N-terminus. *Arch Biochem Biophys*. 418(1):34–38.
- El Mannai Y, Deto R, Kuroki M, Suzuki R, Suzuki E. 2021. Cyanobacterial branching enzymes bind to  $\alpha$ -glucan via surface binding sites. *Arch Biochem Biophys*. 702:108821–108821.
- Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*. 66(Pt 4):486–501.
- Feese MD, Kato Y, Tamada T, Kato M, Komeda T, Miura Y, Hirose M, Hondo K, Kobayashi K, Kuroki R. 2000. Crystal structure of glycosyltrehalose trehalohydrolase from the hyperthermophilic archaeum *Sulfolobus solfataricus* 11 Edited by D. Rees. *J Mol Biol*. 301(2):451–464.
- Feng L, Fawaz R, Hovde S, Sheng F, Nosrati M, Geiger JH. 2016. Crystal structures of *Escherichia coli* branching enzyme in complex with cyclodextrins. *Acta Crystallogr Sect Struct Biol*. 72(Pt 5):641–647.
- Franke D, Svergun DI. 2009. DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering. *J Appl Crystallogr*. 42(2):342–346.
- Froese DS, Michaeli A, McCorvie TJ, Krojer T, Sasi M, Melaev E, Goldblum A, Zatsepin M, Lossos A, Álvarez R, *et al.* 2015. Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. *Hum Mol Genet*. 24(20):5667–5676.
- Gabler F, Nam S-Z, Till S, Mirdita M, Steinegger M, Söding J, Lupas AN, Alva V. 2020. Protein sequence analysis using the MPI bioinformatics toolkit. *Curr Protoc Bioinforma*. 72(1):e108.
- Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol*. 59(3):307–321.
- Hayashi M, Suzuki R, Colleoni C, Ball S, Fujita N, Suzuki E. 2017. Bound substrate in the structure of cyanobacterial branching enzyme supports a new mechanistic model. *J Biol Chem*. 292(13):5465–5475.
- Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. *Proc Natl Acad Sci U S A*. 89(22):10915–10919.
- Janeček S. 1997.  $\alpha$ -amylase family: molecular biology and evolution. *Prog Biophys Mol Biol*. 67(1):67–97.
- Janeček Š, Svensson B, MacGregor EA. 2011. Structural and evolutionary aspects of two families of non-catalytic domains present in starch and glycogen binding proteins from microbes, plants and animals. *Enzyme Microb Technol*. 49(5):429–440.

- Janeček Š, Mareček F, MacGregor EA, Svensson B. 2019. Starch-binding domains as CBM families—history, occurrence, structure, function and evolution. *Biotechnol Adv.* 37(8):107451.
- Jiang H, Xie X, Ban X, Gu Z, Cheng L, Hong Y, Li C, Li Z. 2021. Flexible loop in carbohydrate-binding module 48 allosterically modulates substrate binding of the 1,4- $\alpha$ -glucan branching enzyme. *J Agric Food Chem.* 69(20):5755–5763.
- Jones DT. 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J Mol Biol.* 292(2):195–202.
- Kabsch W. 2010. XDS. *Acta Crystallogr D Biol Crystallogr.* 66(2):125–132.
- Kasper L, Seider K, Hube B. 2015. Intracellular survival of *Candida glabrata* in macrophages: immune evasion and persistence. *FEMS Yeast Res.* 15(5):fov042.
- Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30(14):3059–3066.
- Konarev PV, Volkov VV, Sokolova AV, Koch MHJ, Svergun DI. 2003. PRIMUS: a windows PC-based system for small-angle scattering data analysis. *J Appl Crystallogr.* 36(5):1277–1282.
- Koshland DE. 1953. Stereochemistry and the mechanism of enzymatic reactions. *Biol Rev.* 28(4):416–436.
- Krisman CR. 1962. A method for the colorimetric estimation of glycogen with iodine. *Anal Biochem.* 4:17–23.
- Kuriki T, Stewart DC, Preiss J. 1997. Construction of chimeric enzymes out of maize endosperm branching enzymes I and II: ACTIVITY AND PROPERTIES. *J Biol Chem.* 272(46):28999–29004.
- Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47(W1):W256–W259.
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42:D490–D495.
- MacGregor EA, Janeček Š, Svensson B. 2001. Relationship of sequence and structure to specificity in the  $\alpha$ -amylase family of enzymes. *Biochim Biophys Acta BBA - Protein Struct Mol Enzymol.* 1546(1):1–20.
- Martinovičová M, Janeček Š. 2018. *In silico* analysis of the  $\alpha$ -amylase family GH57: eventual subfamilies reflecting enzyme specificities. *3 Biotech.* 8(7):307.
- Møller MS, Henriksen A, Svensson B. 2016. Structure and function of  $\alpha$ -glucan debranching enzymes. *Cell Mol Life Sci CMLS.* 73(14):2619–2641.
- Mu J, Cheng C, Roach PJ. 1996. Initiation of glycogen synthesis in yeast. Requirement of multiple tyrosine residues for function of the self-glucosylating Glg proteins *in vivo*. *J Biol Chem.* 271(43):26554–26560.
- Murakami T, Kanai T, Takata H, Kuriki T, Imanaka T. 2006. A novel branching enzyme of the GH-57 family in the hyperthermophilic Archaeon *Thermococcus kodakaraensis* KOD1. *J Bacteriol.* 188(16):5915–5924.

- Noguchi J, Chaen K, Vu NT, Akasaka T, Shimada H, Nakashima T, Nishi A, Satoh H, Omori T, Kakuta Y, Kimura M. 2011. Crystal structure of the branching enzyme I (BEI) from *Oryza sativa* L with implications for catalysis and substrate binding. *Glycobiology*. 21(8):1108–1116.
- Pal K, Kumar S, Sharma S, Garg SK, Alam MS, Xu HE, Agrawal P, Swaminathan K. 2010. Crystal structure of full-length *Mycobacterium tuberculosis* H37Rv glycogen branching enzyme: insights of N-terminal beta-sandwich in substrate specificity and enzymatic activity. *J Biol Chem*. 285(27):20897–20903.
- Pei J, Kim B-H, Grishin NV. 2008. PROMALS3D: a tool for multiple protein sequence and structure alignments. *Nucleic Acids Res*. 36(7):2295–2300.
- Robert X, Gouet P. 2014. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res*. 42(W1):W320-324.
- Robert X, Haser R, Gottschalk TE, Ratajczak F, Driguez H, Svensson B, Aghajari N. 2003. The structure of barley  $\alpha$ -amylase isozyme 1 reveals a novel role of domain C in substrate recognition and binding: a pair of sugar tongs. *Structure*. 11(8):973–984.
- Shen X-X, Zhou X, Kominek J, Kurtzman CP, Hittinger CT, Rokas A. 2016. Reconstructing the backbone of the *Saccharomycotina* yeast phylogeny using genome-scale data. *G3 GenesGenomesGenetics*. 6(12):3927–3939.
- Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol*. 7:539.
- Stam MR, Danchin EGJ, Rancurel C, Coutinho PM, Henrissat B. 2006. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of alpha-amylase-related proteins. *Protein Eng Des Sel PEDS*. 19(12):555–562.
- Suzuki E, Suzuki R. 2016. Distribution of glucan-branching enzymes among prokaryotes. *Cell Mol Life Sci CMLS*. 73(14):2643–2660.
- Suzuki R, Suzuki E. 2020. Structure and function of branching enzymes in eukaryotes. *Trends Glycosci Glycotechnol*. 32(185):E21–E30.
- Svergun D, Barberato C, Koch MHJ. 1995. CRY SOL – a program to evaluate X-ray solution scattering of biological Macromolecules from Atomic Coordinates. *J Appl Crystallogr*. 28(6):768–773.
- Takata H, Takaha T, Kuriki T, Okada S, Takagi M, Imanaka T. 1994. Properties and active center of the thermostable branching enzyme from *Bacillus stearothermophilus*. *Appl Environ Microbiol*. 60(9):3096–3104.
- Yachdav G, Wilzbach S, Rauscher B, Sheridan R, Sillitoe I, Procter J, Lewis SE, Rost B, Goldberg T. 2016. MSASViewer: interactive JavaScript visualization of multiple sequence alignments. *Bioinforma Oxf Engl*. 32(22):3501–3503.
- Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J Mol Biol*. 430(15):2237–2243.

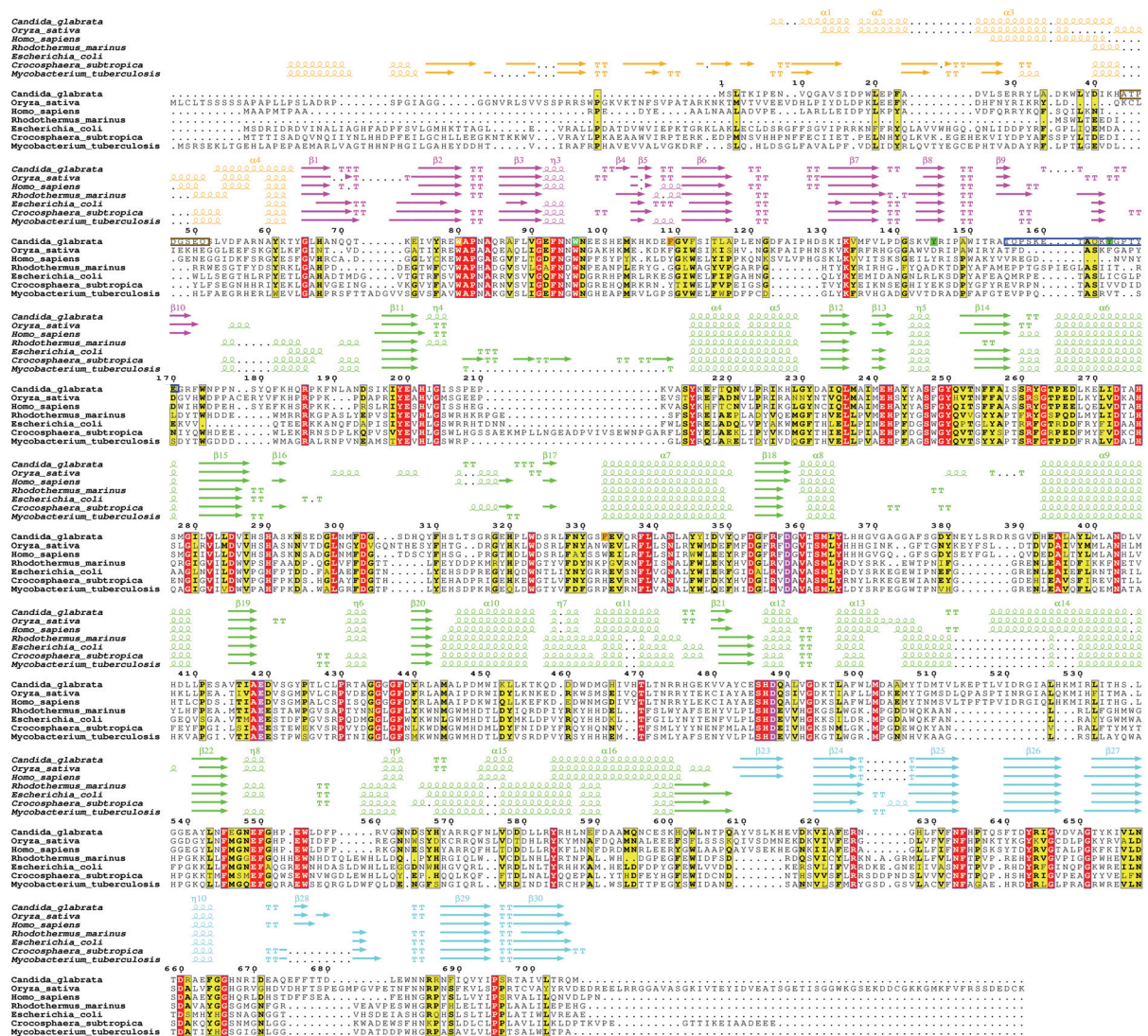


## B. Figures et tableaux

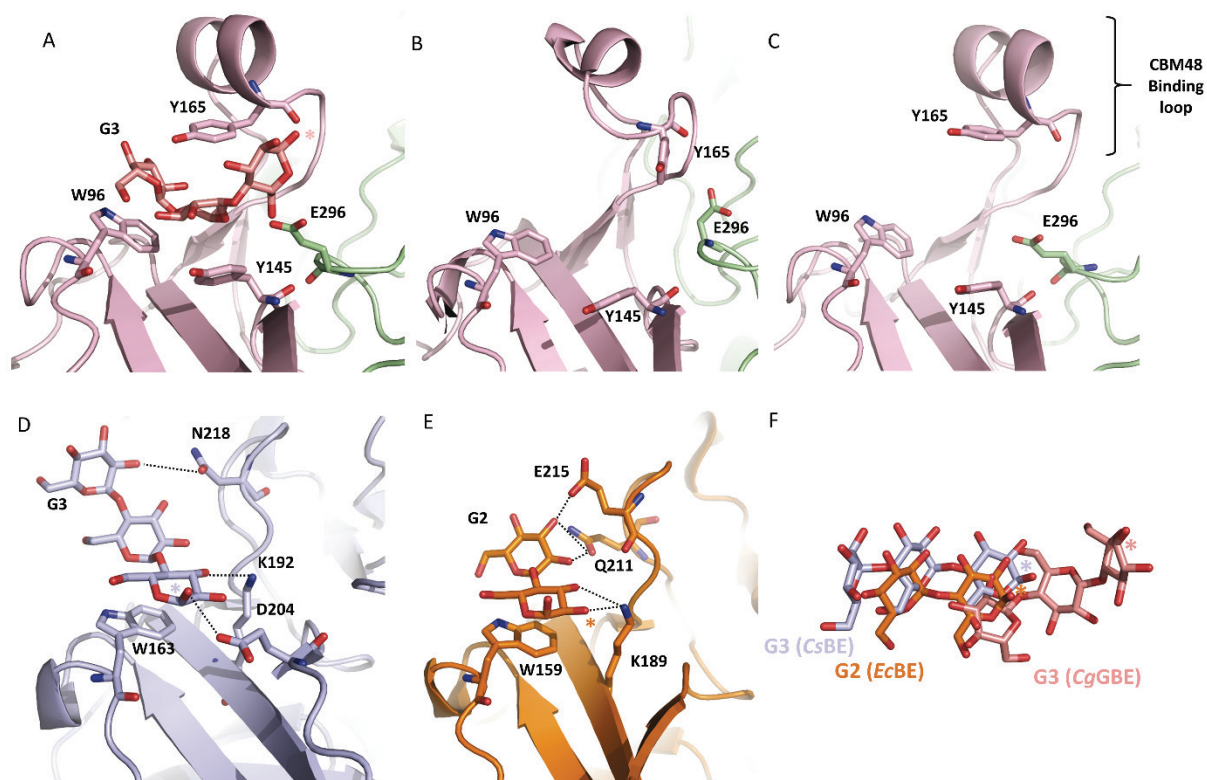


**Figure 1. Structural organization of CgGBE.** **A.** Three-dimensional structure of ligand free CgGBE<sub>A</sub> with the  $\alpha$ -helical N-terminal domain preceded by the N-terminal extra sequence (N-EXS) in *light orange*, the CBM48 in *pink*, the catalytic domain A in *green* and the C-terminal domain in *blue*. **B.** Close-up view of the active site with the three catalytic residues, D358 (nucleophile), E419 (catalytic acid/base) and D488 (transition state stabilizer) in *yellow*. **C.** Topological representation of the secondary structures of CgGBE. **D.** Close-up view of the partially disordered loop  $\beta 4$ - $\alpha 4$  for subunits A and B found in the asymmetric unit (CgGBE<sub>A</sub>).

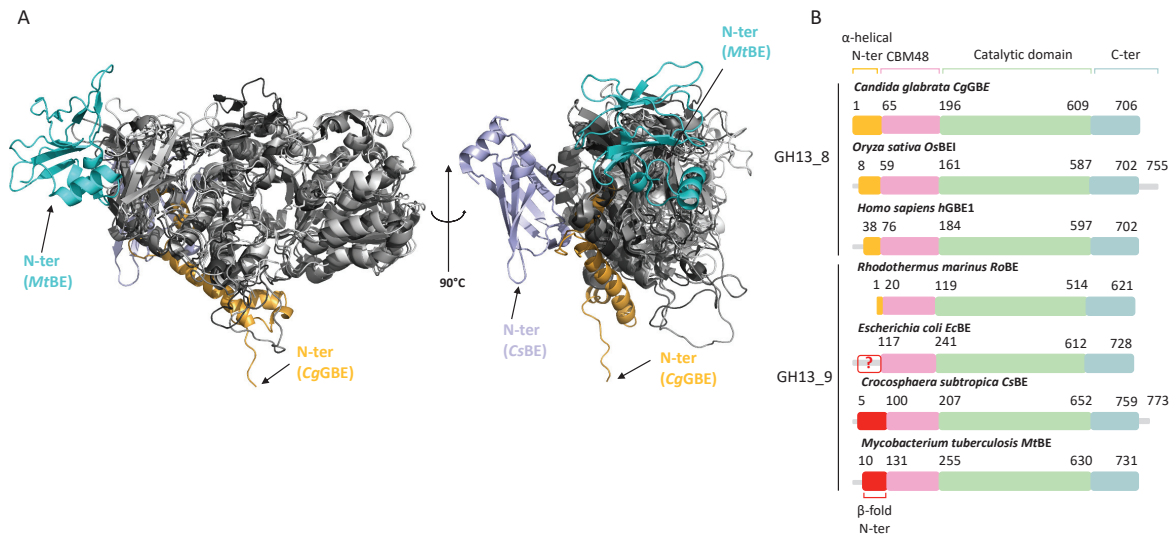




**Figure 2. Structure-based protein sequence alignment of *C. glabrata*, *O. sativa*, *H. sapiens*, *R. marinus*, *C. subtropica*, *M. tuberculosis*.** Primary- and secondary structure alignments were performed using PROMALS3D. The figure was rendered with ESPrnt 3.0 (Robert and Gouet 2014), red boxes indicate residues with strict identities and yellow boxes indicate similar residues. Secondary structure colors refers to each BE domain, *light orange* for the  $\alpha$ -helical N-terminal domain, *pink* for the CBM48 domain, *green* for the catalytic domain A and *cyan* for the C-terminal domain. Catalytic residues for each organism are highlighted in *pink*, residues from CgGBE implicated in BS<sub>A</sub> binding are highlighted in *green* and those from CgGBE BS<sub>B</sub> are highlighted in *orange*. The CBM48 binding loop (T155-E170) of CgGBE is indicated by a *dark blue* rectangle and the  $\alpha$ -helical N-terminal domain elongated loop (H44-L52) is indicated by a *brown* rectangle.

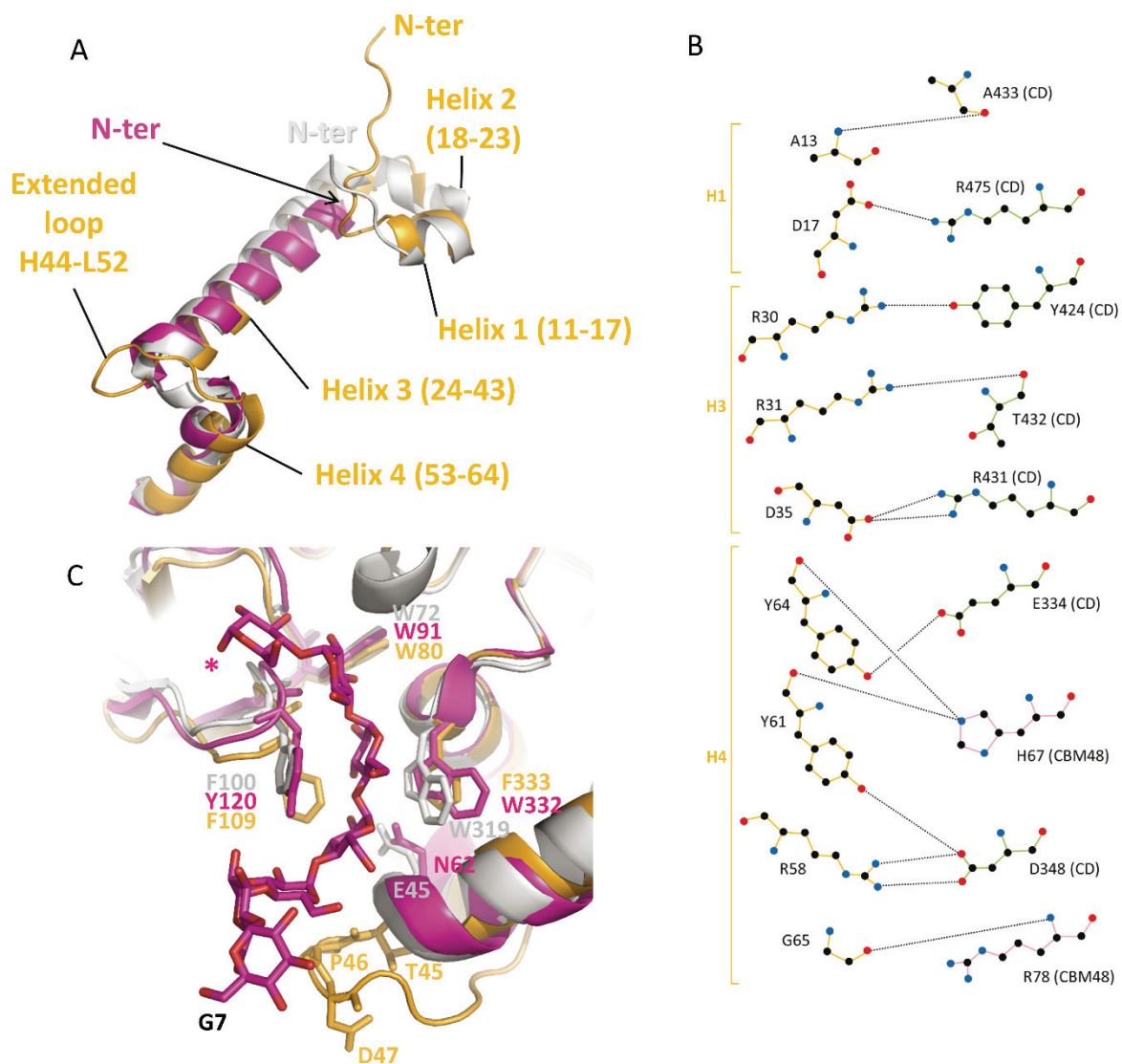


**Figure 3. Structuring of the CBM48 binding site of CgGBE.** A, B, C. CBM48 binding site (BS<sub>B</sub>) of CgGBE<sub>A</sub> bound to a maltotriose (CgGBE<sub>A</sub>-G3) (the reducing end of the G3 is indicated by \*) (A), or ligand free CgGBE<sub>A</sub> (B) and CgGBE<sub>B</sub> (C). CBM48 is in *light pink* and the catalytic domain A is in *green*. D. Binding site 48\_1 in the CBM48 domain of GH13\_9 *C. subtropica* BE (CsBE, PDB: 5GR1) colored in *light blue* in complex with G3 shown as sticks. E. Binding site IV in CBM48 domain of GH13\_9 *E. coli* BE (EcBE, PDB: 4LPC) colored in *orange* in complex with G2 shown as sticks. F. Superimposition of G3 or G2 bound to CBM48 respectively in CsBE (G3) in *light blue*, EcBE (G2) in *orange* and CgGBE (G3) in *light pink*. Reducing ends are identified by \*.

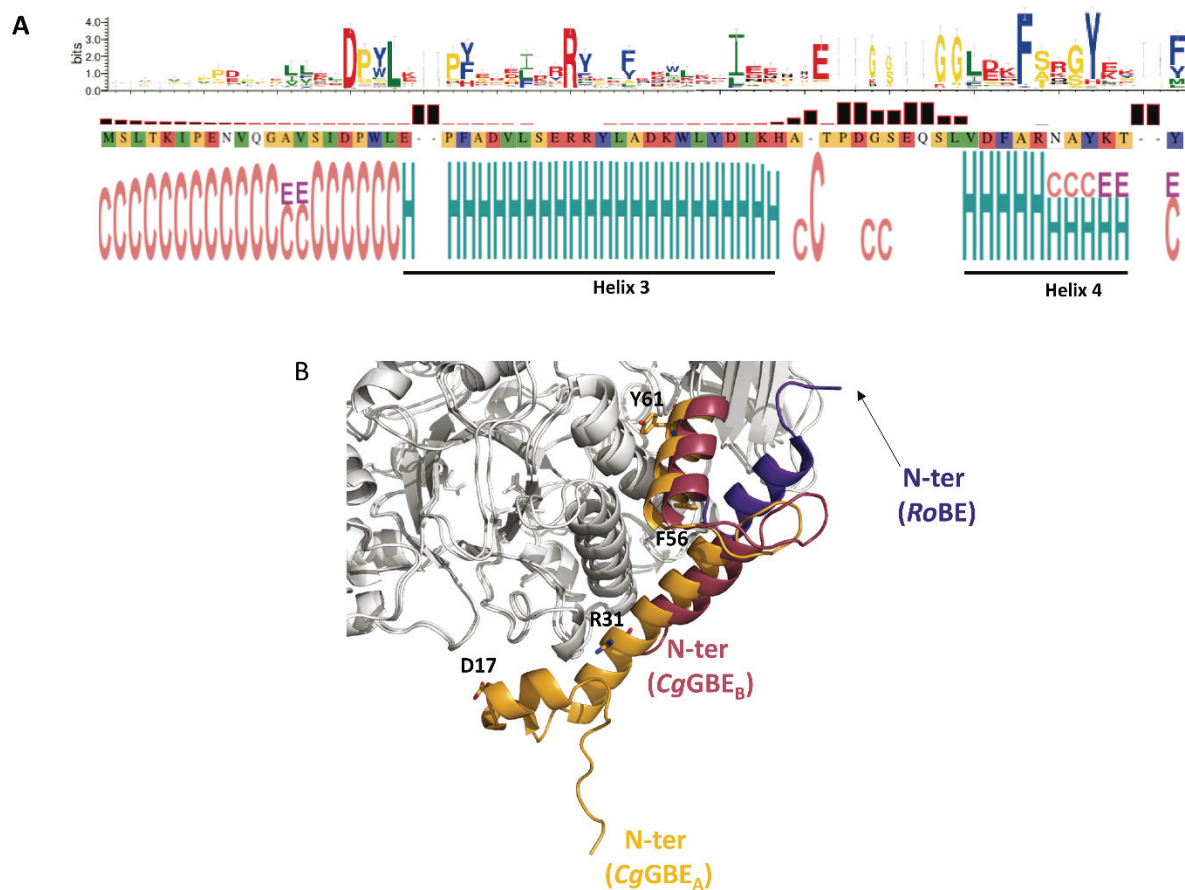


**Figure 4. Superposition of GH13\_8 and GH13\_9 branching enzymes.** **A.** Superposition of CgGBE, MtBE (PDB: 3K1D) and CsBE (PDB: 5GR1). The N-terminal domains represented as ribbon drawings are shown respectively in *light orange* (CgGBE), *cyan* (MtBE) and *light blue* (CsBE). **B.** Domain organization of *C. glabrata*, *O. sativa*, *H. sapiens*, *R. marinus*, *E. coli*, *M. tuberculosis* and *C. subtropica* enzymes, with the N-terminal domain shown in *light orange* when  $\alpha$ -helical fold and *red* when  $\beta$ -sandwich fold, CBM48 in *pink*, central catalytic domain A in *green* and C-terminal domain in *blue*. The empty *red* square represents the yet unknown fold of the N-terminal domain of *E. coli* (PDB: 4LPC).





**Figure 5. Structural properties of the  $\alpha$ -helical N-terminal domain.** **A.** Superposition of the  $\alpha$ -helical N-terminal domain of *Cg*GBE (*light orange*) onto *h*GBEI (*pink*) and onto *Os*BEI (*grey*). An elongated loop H43-L53 in *Cg*GBE, is not present in other structures. **B.** Representation of the contacts between the  $\alpha$ -helical N-terminal domain residues from helix 2 (H2), 4 (H4) and 5 (H5) with catalytic domain A (CD) residues in *green* and CBM48 residues in *pink*, respectively **C.** Close-up view of the predicted (by homology)  $\alpha$ -helical N-terminal domain binding site (BS<sub>A</sub>) of *Cg*GBE (*light orange*), *h*GBEI (*pink*) and *Os*BEI (*grey*) with the superimposition of a G7 observed in *h*GBEI. The reducing end of the G7 polysaccharide is indicated by \*.



**Figure 6. Secondary structure conservation of the  $\alpha$ -helical N-terminal domain in the GH13\_8 subfamily.** **A.** The secondary structure predictions by PSIPRED (Jones 1999) for the 137 aligned sequences are depicted as a sequence logo on the bottom line: H=helix, C=coil, E=beta-strand. Black columns indicate the lack of conservation of the corresponding residues. For clarity, the alignment is cut after position 64 of the CgGBE sequence. **B.** Superposition of CgGBE<sub>A</sub>, CgGBE<sub>B</sub> and RoBE (PDB: 6JOY) structures respectively in *light orange*, *raspberry* and *purple*. Conserved residues D17, R31, F56 and Y61 (CgGBE numbering) of the  $\alpha$ -helical N-terminal domain shown in (A) are highlighted as *light orange* sticks.

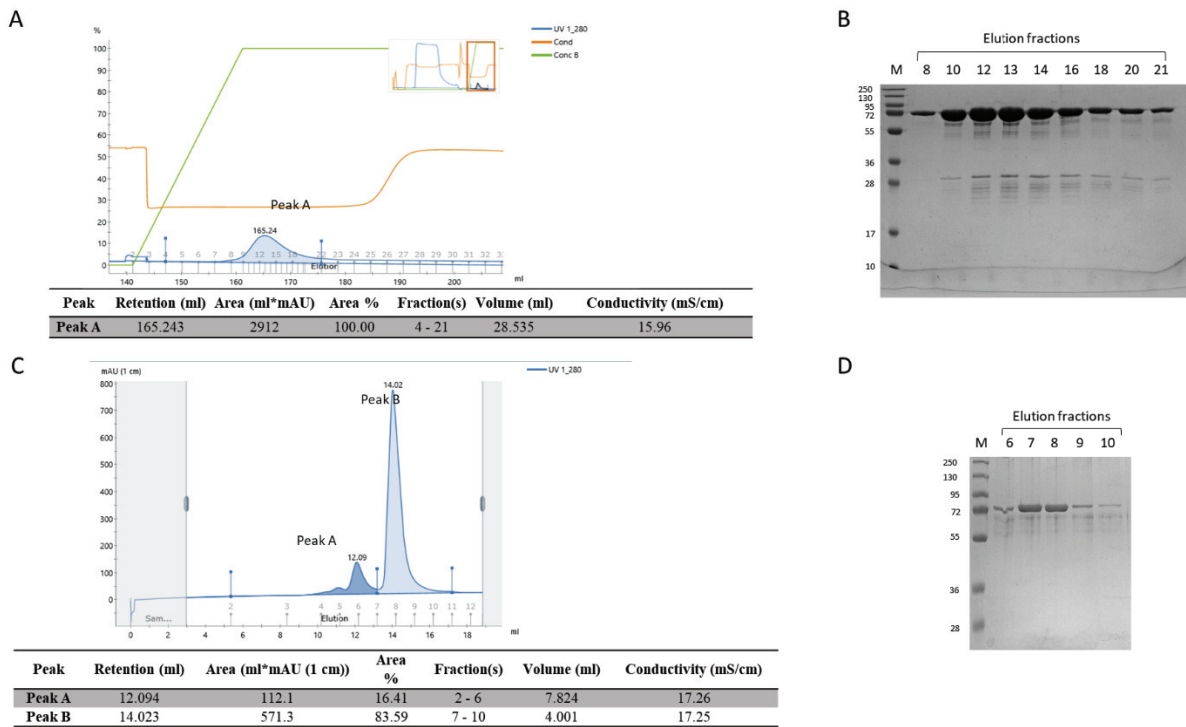


**Table I. Data collection and refinement statistics**

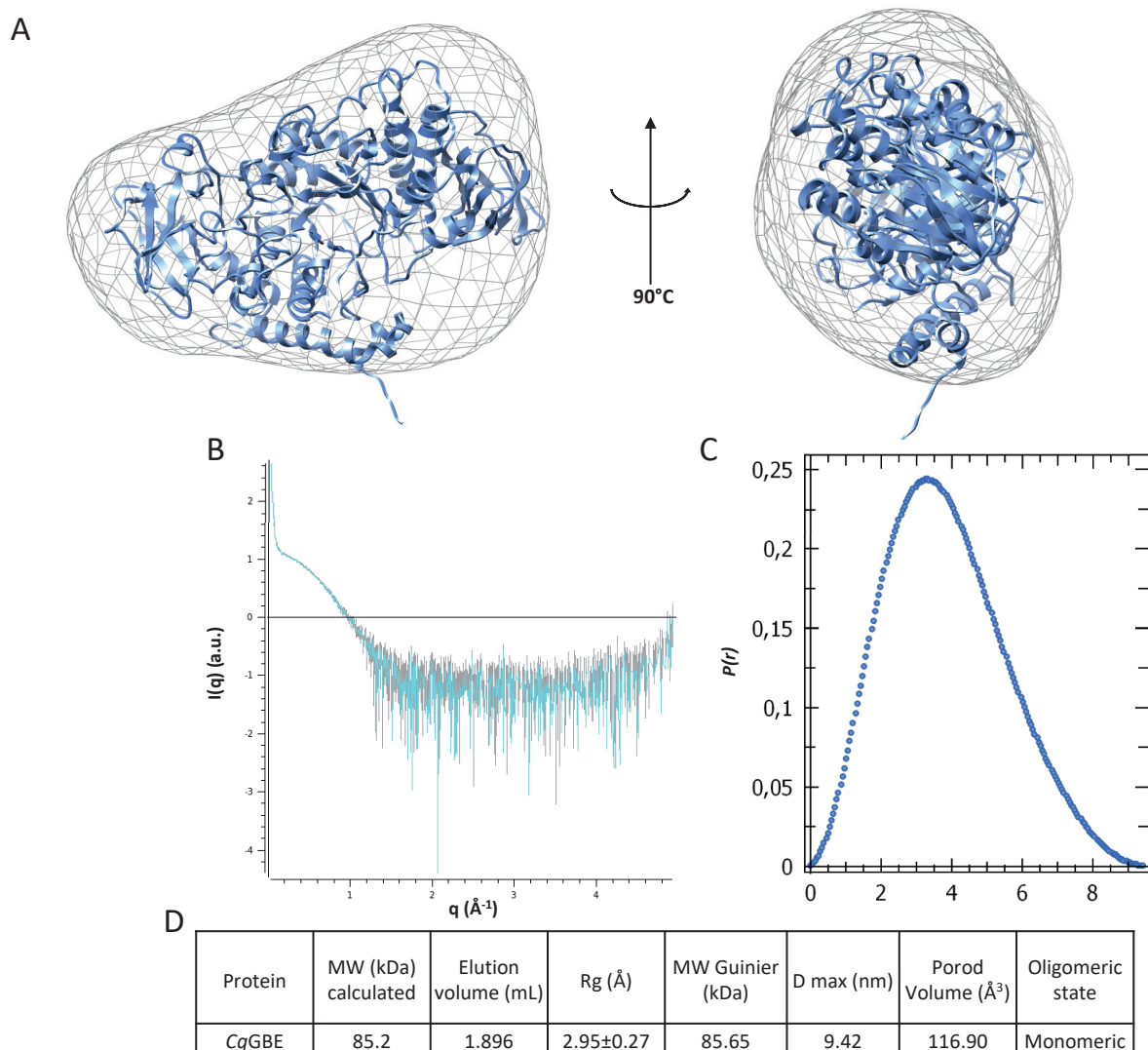
Structure-ID	<i>Cg</i> GBE <sub>A</sub>	<i>Cg</i> GBE <sub>B</sub>	<i>Cg</i> GBE-G3
<b>Data collection</b>			
Beamline	ID23-2	PX1	PX1
Wavelength (Å)	0.87313	0.97857	0.97857
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2
Cell dimensions			
a, b, c (Å)	210.03 88.61 89.72	79.58 89.23 112.06	88.72 210.70 90.52
α, β, γ (°)	90.0 90.0 90.0	90.0 90.0 90.0	90.0 90.0 90.0
Resolution range (Å)	46.85-2.4	47.45-2.09	47.05-1.93
Total reflections	323401	589906	834910
Unique reflections	65906	43538	128265
R <sub>meas</sub> (%)	17.4 (85.4)	11.7 (126.8)	9.1 (118.7)
CC <sub>1/2</sub> (%)	99.4 (77.7)	99.9 (79.3)	99.9 (86.0)
I/σ(I)	6.62 (1.58)	14.7 (1.7)	13.38 (1.76)
Multiplicity	4.9 (5.1)	13.5 (13.9)	6.5 (6.4)
Completeness (%)	99.4 (99.8)	99.9 (99.4)	99.6 (97.6)
No. mol. /asymm. unit	2	1	2
<b>Refinement</b>			
R <sub>work</sub> /R <sub>free</sub> (%)	21.07/24.75	19.05/23.33	19.32/21.75
No. atoms			
Protein	11089	5358	10881
Ligand/ion	16	40	68
Water	406	372	692
Average B-factor (Å <sup>2</sup> )			
Protein	55.1	44.7	57.3
Ligand/ion	40;4	52.1	93.8
Water	43.8	47.2	48.2
r.m.s.d.			
Bond lengths (Å)	0.006	0.011	0.010
Angles (°)	0.768	1.094	0.985
Ramachandran			
Favored (%)	96.3	96.5	96.6
Allowed (%)	3.5	3.4	3.4
Outliers (%)	0.2	0.2	0.0



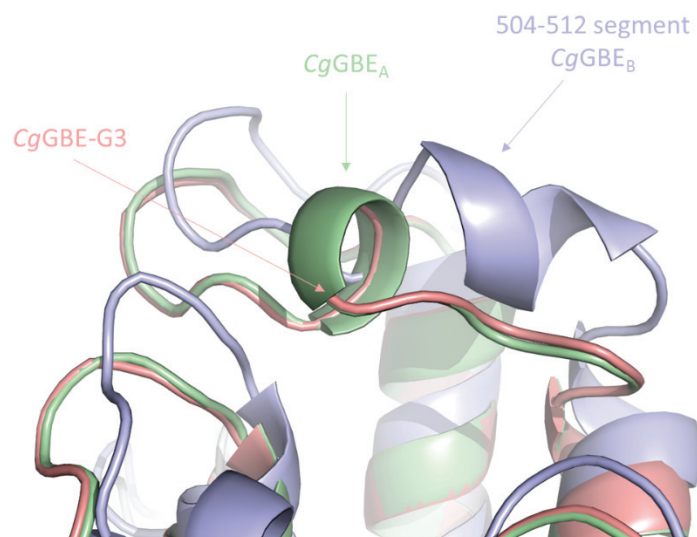
## C. Figures supplémentaires



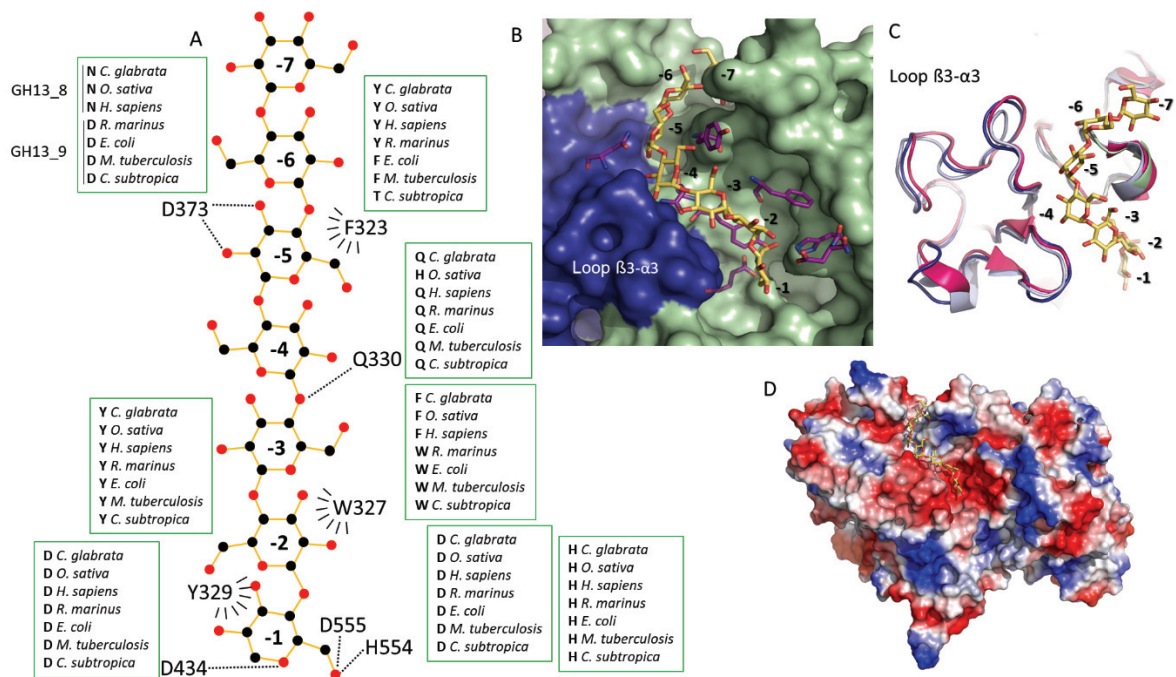
**Supplementary Figure 1. Purification of CgGBE.** **A.** Elution chromatogram of CgGBE affinity chromatography using TALON cobalt crude column. Elution was performed by a linear gradient of elution buffer (30 mM HEPES pH 7.6, 300 mM NaCl and 500 mM imidazole) following by a step at 100% of elution buffer. **B.** Elution peak A fractions 8 to 21 were loaded on a 15 % SDS-PAGE gel, CgGBE (85.2 kDa) is present in each fractions. **C.** Elution chromatogram of CgGBE size exclusion chromatography using S200 increase column. Elution was performed by an isocratic elution using elution buffer (30 mM HEPES pH 7.6, 300 mM NaCl). **D.** Elution fractions from peak A and B were loaded on a 15 % SDS-PAGE gel and fractions 7 to 9 were concentrated using an Amicon Ultra Centrifugal filter with a 50 kDa molecular weight cut-off for a final 20 $\mu$ L volume of CgGBE at 42.7 mg.mL<sup>-1</sup>.



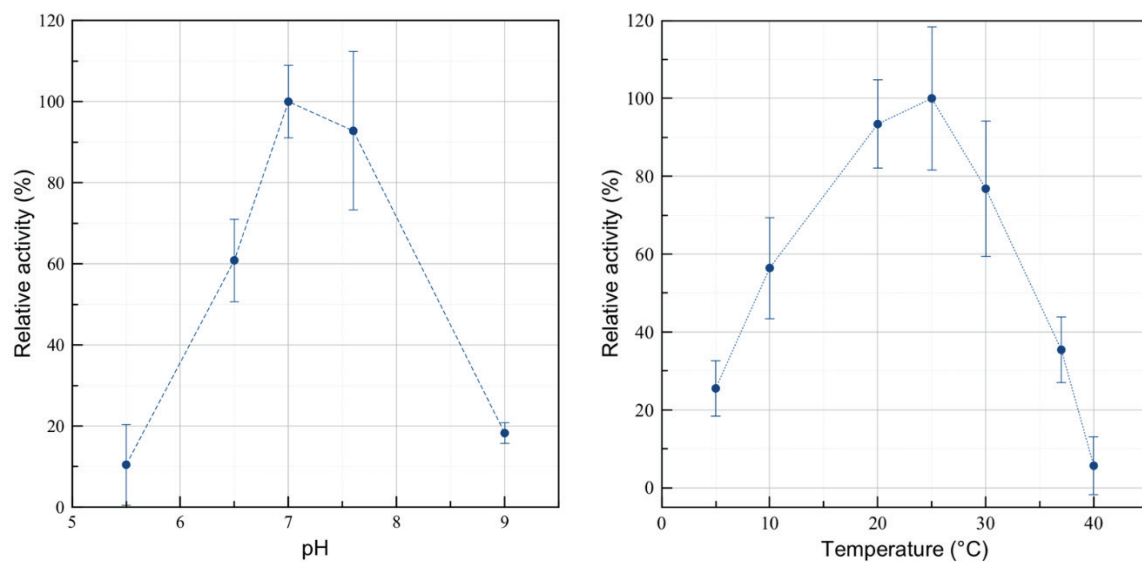
**Supplementary Figure 2. Small angle X-ray scattering studies of CgGBE.** **A.** *Ab initio* model of CgGBE generated from the SEC-SAXS experimental data using the program DAMMIF (Franke and Svergun 2009). The crystal structure of CgGBE<sub>A</sub> was superimposed onto the calculated (17 Å) SAXS envelope and confirmed a monomeric structure of the enzyme in solution using the software CHIMERA (University of California, San Francisco). **B.** Log  $I(q)$  versus  $q$  was generated using PRIMUS (Konarev, P *et al.*, 2003) **C.**  $P(r)$  versus  $r$  profile from the data in (B) was carried out using CRY SOL (Svergun, Barberato and Koch, 1995). **D.** SAXS data for CgGBE analysis.



**Supplementary Figure 3. Structural local rearrangement of the 504-512 segment between CgGBE<sub>A</sub>, CgGBE<sub>B</sub>, CgGBE-G3.** Segment 504-512 is structurally different between all three structures being folded either in a loop for CgGBE-G3, colored in *light red*, in a one turn helix for CgGBE<sub>A</sub>, colored in *light green*, or in a two turn helix for CgGBE<sub>B</sub>, colored in *light blue*.

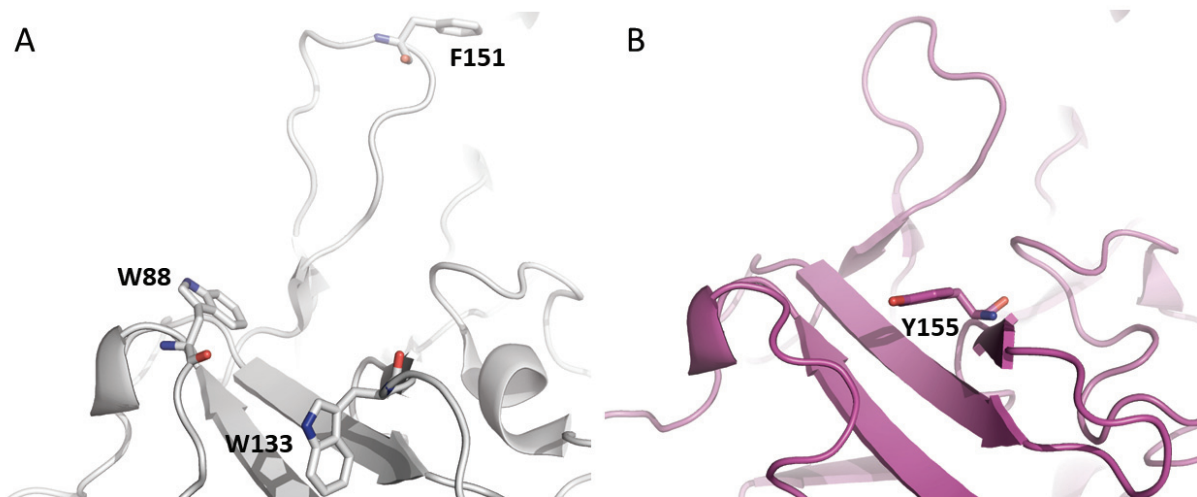


**Supplementary Figure 4. Conservation of the CgGBE catalytic site.** **A.** Schematic representation of the residues interacting with a G7 in the active site of *C. subtropica* CsBE and the counterparts in *C. glabrata*, *O. sativa*, *H. sapiens*, *R. marinus*, *E. coli* and *M. tuberculosis* BEs. **B.** Surface representation of CgGBE with superposition of a G7 (yellow sticks) identified in the CsBE structure. The conserved loop  $\beta 3$ - $\alpha 3$  is shown in dark blue, residues thought to participate in the interaction are highlighted as purple sticks. **C.** Superimposition of the conserved loop  $\beta 3$ - $\alpha 3$  in CgGBE (dark blue), hGBEI (pink) and CsBE (light blue) in complex with G7 shown in yellow sticks. **D.** Electrostatic surface potential of CgGBE with positive charged areas shown in blue and negative charged area in red. G7 from CsBE is superposed in the structure and shown in yellow sticks.



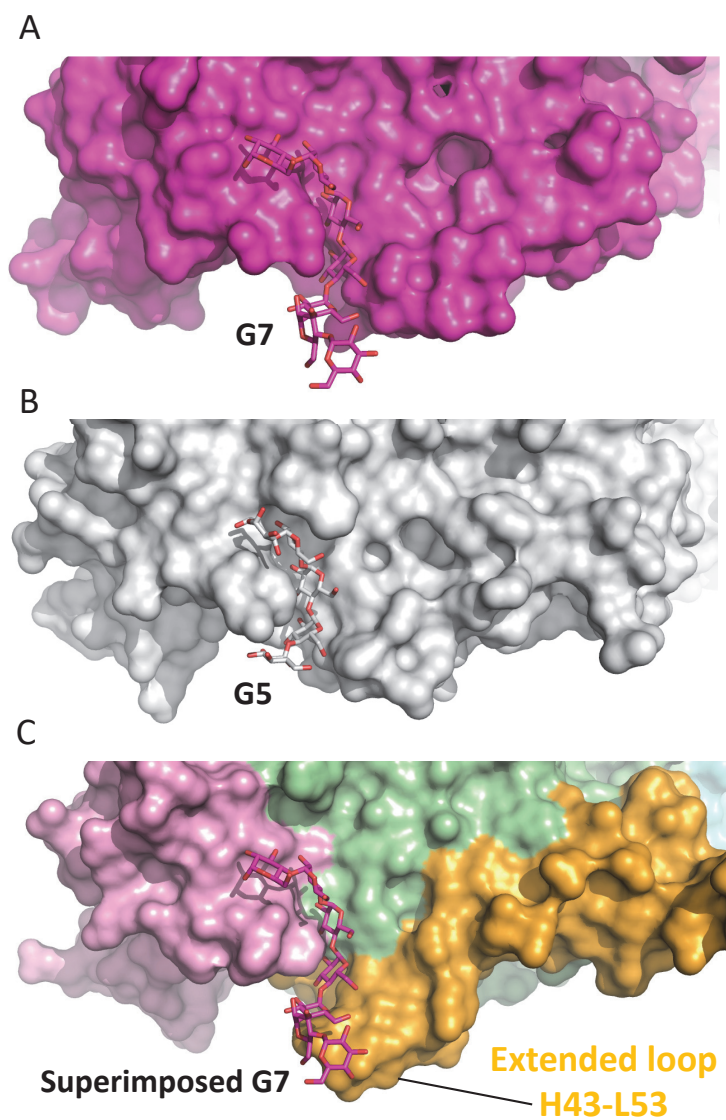
**Supplementary Figure 5. Temperature and pH effect on CgGBE hydrolytic activity. A.**

Hydrolytic activity of CgGBE was measured by iodine staining assay using Amylose V as substrate. The influence pH and temperature was measured with pH ranging from 5.5 to 9.0 and temperature ranging from 5 to 40°C.



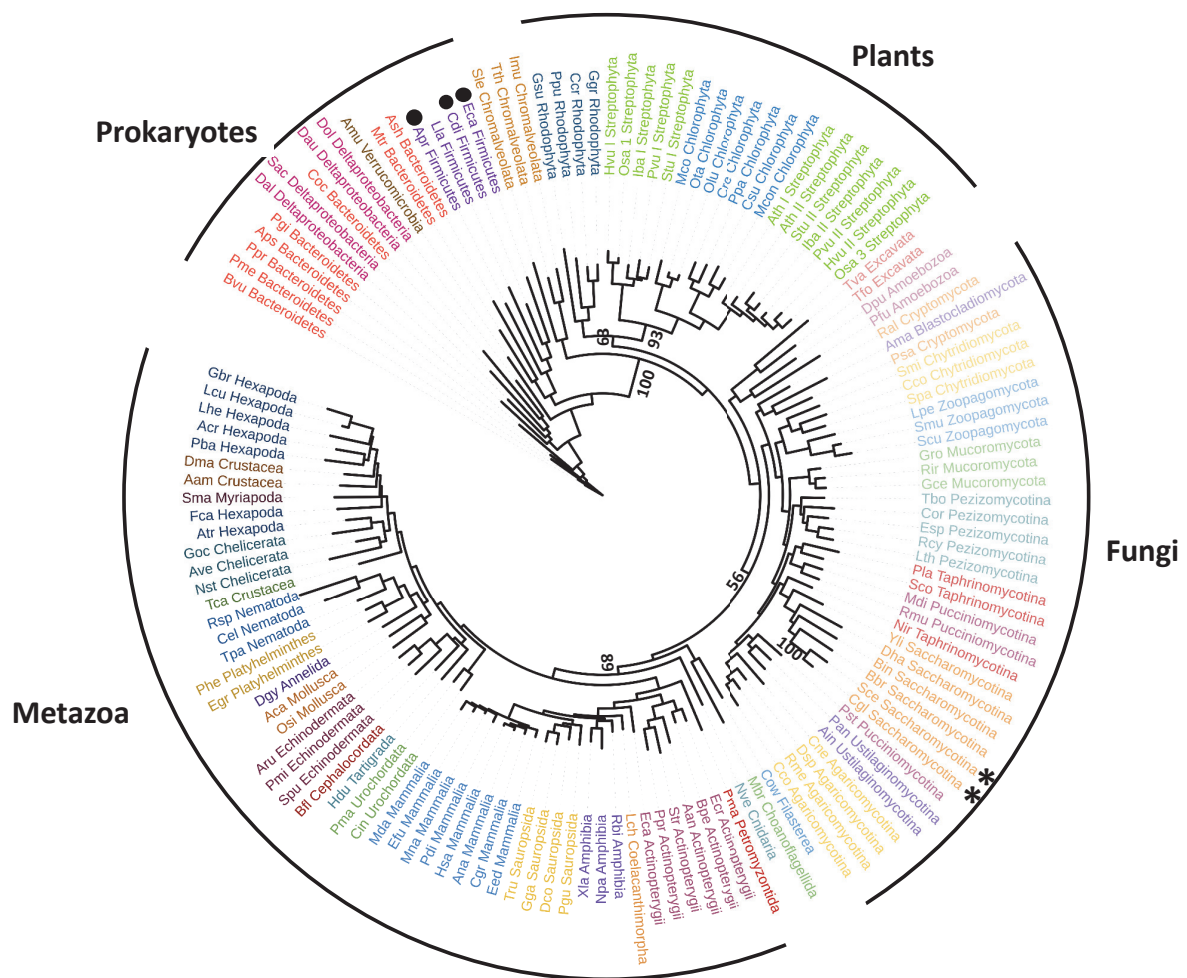
**Supplementary Figure 6. Comparisons of GH13\_8 CBM48 binding modes. A.** Surface binding site 2 observed in the *O. sativa* glycogen branching enzyme (*Os*GBE1, PDB: 3VU2). **B.** Surface binding site predicted in human glycogen branching enzyme (*h*GBE, PDB: 5CLW).





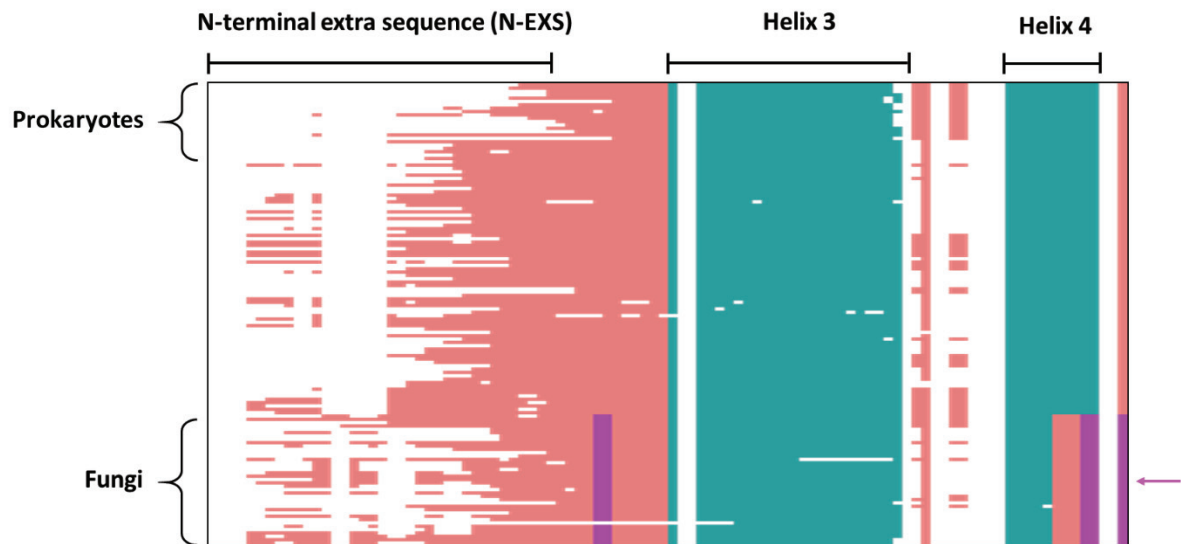
**Supplementary Figure 7. Surface binding site in the  $\alpha$ -helical domain.** **A.** Surface representation of the *h*GBEI structure with a G7 bound in the CBM48 domain. **B.** Surface representation of *Os*BEI structure with a bound (CBM48 domain, site 1). **C.** Surface representation of the *Cg*GBE  $\alpha$ -helical N-terminal domain (*light orange*), CBM48 (*light pink*), catalytic domain (*light green*) and superimposition of a G7 identified from *h*GBEI in complex (PDB: 4CLW).



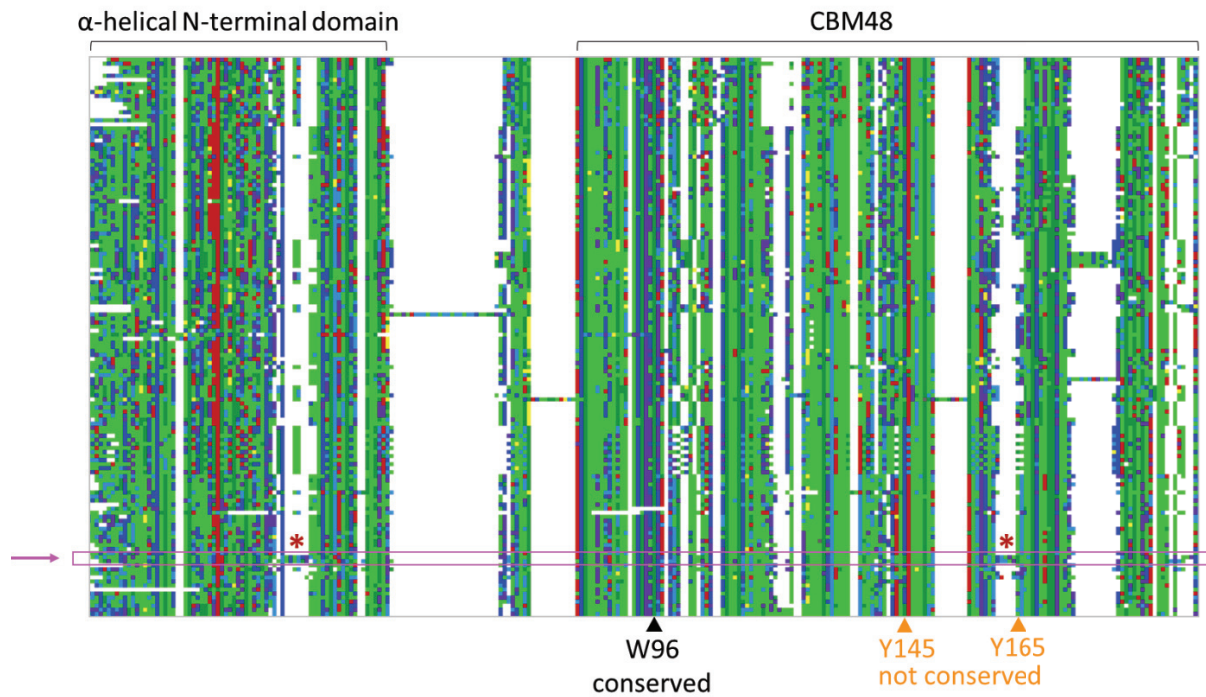


**Supplementary Figure 8. Phylogenetic tree of GH13\_8 enzymes.** A ML phylogeny based on 487 amino acid alignment positions of 137 sequences was constructed with phyML using GH13\_8 from eukaryotes and bacteria. Eukaryotic species phylogeny was adapted from Gonçalves *et al*, 2016 with modification for *mycota*. Bootstraps of interest  $\geq 55$  are shown.

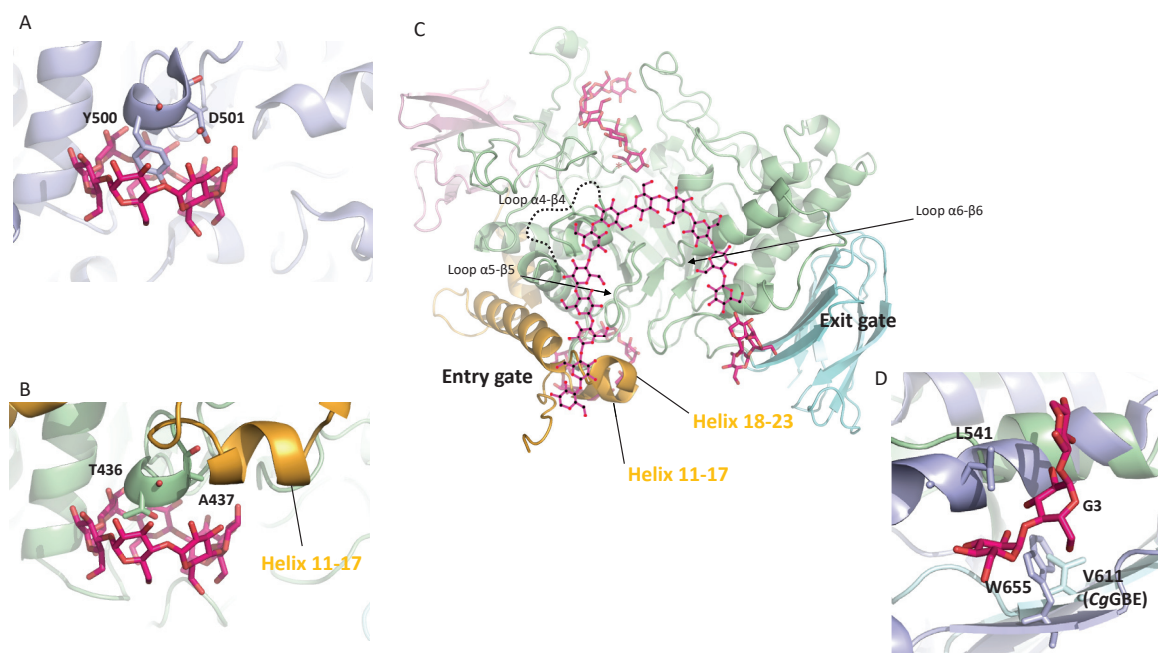
Black circles indicate the presence of both GH13\_8 and GH13\_9 in the corresponding species. Saccharomycetaceae are indicated by a star (\*).



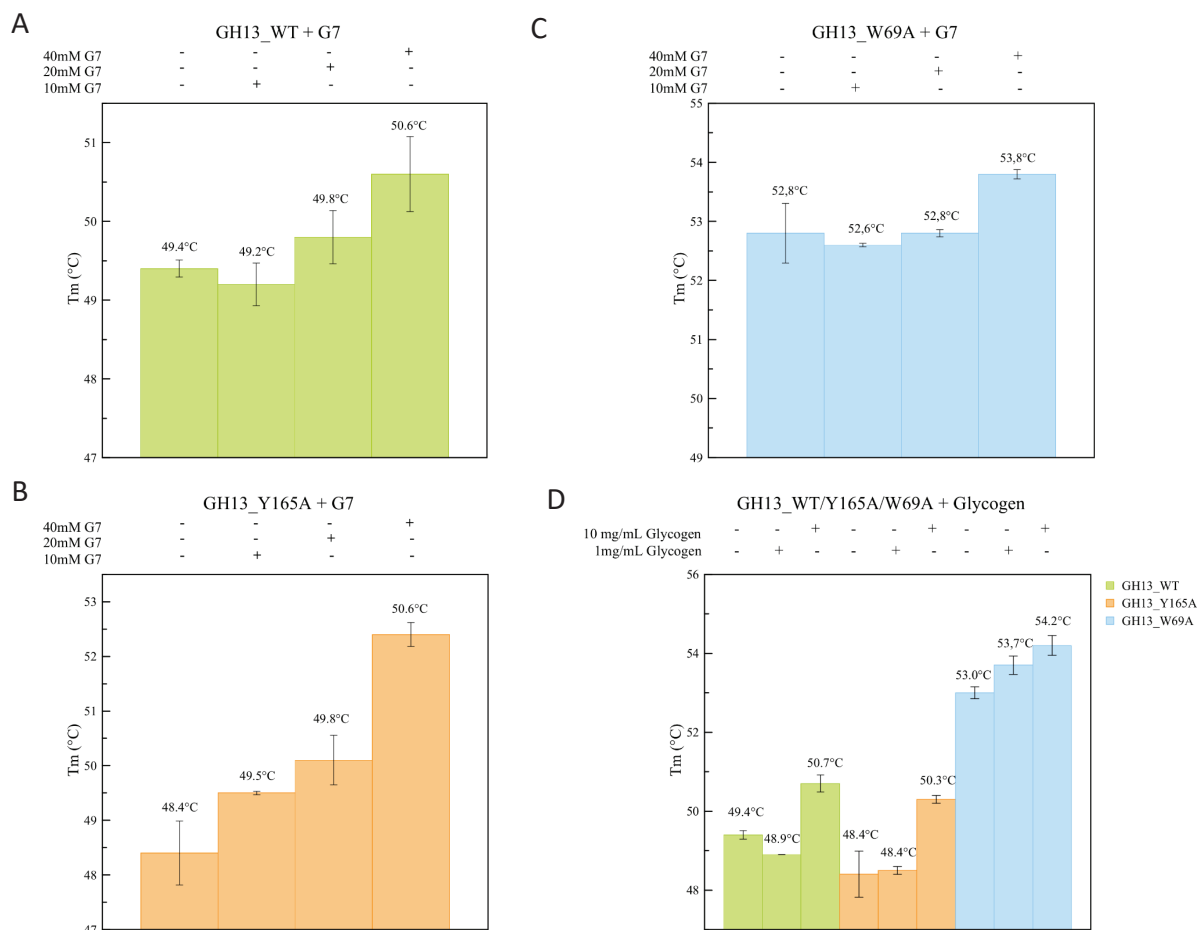
**Supplementary Figure 9. Secondary structure conservation of the  $\alpha$ -helical N-terminal domain in the GH13\_8 subfamily.** The secondary structure predictions by PSIPRED (Jones DT, 1999) for the 137 aligned sequences are depicted as pixels for the 137 aligned sequences with coil in *pink*, helix in *blue* and beta-strand in *magenta*. Secondary structure predictions are run with the Ali2D tool, and the sequence logo is drawn with the Alignment Viewer tool of the Bioinformatics Toolkit resource (<https://toolkit.tuebingen.mpg.de>) (Gabler *et al.*, 2020). The CgGBE and ScGBE sequences are indicated by *pink* arrow. For clarity, the alignment is cut after position 64 of the CgGBE sequence.



**Supplementary Figure 10. Additional motifs presence in the  $\alpha$ -helical N-terminal domain and CBM48 domain of *Cg*GBE.** Image view of the 137 GH13\_8 aligned sequences (<https://alignmentviewer.org/>). Each pixel represents an amino-acid (*light green*: KLMVPAIG, *dark green*: HFWY, *red*: R, *light blue*: ST, *dark blue*: DE, *purple*: NQ, *yellow*: C). *Candida glabrata* and *Saccharomyces cerevisiae* sequences are indicated by *pink* arrows and framed in a *pink* box. For clarity the alignment is limited to region 1-180 of the *Candida glabrata* sequence. The two insertions are indicated by stars (\*). Conservation of the residues forming the CBM48 pair of sugar tongs are indicated by arrow heads.



**Supplementary Figure 11. Predicted position of the entry and exit gates in CgGBE. A.** A1 Surface-binding site in the catalytic domain of the CsBE structure (PDB: 5GR1) in complex with G7. The bound G7 is shown in *magenta* sticks. **B.** Superposition of G7 from CsBE onto CgGBE structure with the catalytic domain in *green*,  $\alpha$ -helical N-terminal domain in *light orange* and G7 in *hot pink*. **C.** CgGBE structure to which two G7 and one G3 as identified in the CsBE structure have been superposed. The first G7 is bound to the active site and the second G7 is bound to the “entry gate” (A1 SBS). G3 is bound to the “exit gate” (A2 SBS). An acceptor glucan chain in the CgGBE active site is represented in *magenta* sticks. **D.** The surface binding site A2 in CsBE shown in *light blue* in complex with G3 shown in *hot pink* sticks superposed onto CgGBE with the catalytic domain shown in *light green* and the C-terminal domain in *cyan*. Residues participating in the recognition of the oligosaccharide are shown in *light blue* sticks in CsBE and the residue equivalent to W655 position in CsBE, V611 in CgGBE is shown in *cyan*.



**Supplementary Figure 12. T<sub>m</sub> (°C) determination by nDSF of CgGBE-WT and its mutants CgGBE-Y165A and CgGBE-W69A in presence or absence of G7 oligosaccharide or glycogen.** The melting temperature of CgGBE-WT (A) CgGBE-Y165A (B) and CgGBE-W69A (C) were measured by using nanoscale differential fluorescence (nano-DSF) in the presence or absence of G7 oligosaccharide at 10, 20, 40mM (A, B, C) or glycogen at 1,10 mg.mL<sup>-1</sup> (D). All measurement were made in duplicate and the T<sub>m</sub> (°C) indicated is the average T<sub>m</sub> with associated errors.

### III. Conclusion et perspectives

Au cours de ce projet de recherche nous avons pu caractériser d'un point de vue biochimique et structural la première enzyme branchante de glucane en  $\alpha$ -(1,4) de levure issue du pathogène *Candida glabrata*. L'étude structurale de cette dernière nous a non seulement permis de confirmer la forte conservation de l'arrangement structural globale de l'enzyme mais également d'identifier des particularités structurales à la sous-famille des GH13\_8. En effet, en comparaison aux enzymes de la sous famille des GH13\_9, ces dernières possèdent un repliement du domaine amino terminal en hélices  $\alpha$ . Les enzymes de branchement, primordiales au métabolisme énergétique, exercent leur rôle sur des polysaccharides complexes, l'amidon et le glycogène. L'existence d'un repliement unique du domaine amino-terminal chez les GH13\_8 peut poser plusieurs interrogations, est-ce lié à un degré de ramifications spécifique du glycogène ? ou est-ce lié à un environnement spécifique en termes de pH et de température dans lequel évolue le pathogène ?

Bien que la structure tridimensionnelle des BEs ait été décrite chez plusieurs organismes, le manque d'informations fonctionnelles et enzymatiques ne permet pas encore de comprendre complètement leur mode de fonctionnement et notamment la mise en place de leur activité hydrolytique et transglycolytique. En complément à notre caractérisation de l'activité hydrolytique de CgGBE WT, il serait intéressant d'étudier l'activité de branchement de CgGBE et d'identifier sa spécificité concernant la longueur de chaînes de glucanes transférées. Ainsi, il pourrait être envisagé de tronquer la protéine de son domaine amino terminal ou de son domaine de liaison aux sucres (CBM48) afin d'observer l'effet sur l'activité et la spécificité de l'enzyme. Il serait également envisageable de réaliser une protéine chimérique, où le domaine amino terminal de CgGBE serait remplacé par le domaine amino terminal d'une BE de la sous famille GH13\_9 et d'observer les conséquences sur l'activité de l'enzyme et sa spécificité de substrat.

Comme cela a été démontré par la détermination de structures de BEs en complexe avec des oligosaccharides, ces enzymes possèdent de nombreux sites de liaison aux sucres en surface (Tableau 1, p 180). Il a été suggéré que ces sites de liaison seraient responsables de la fixation des chaînes de glucane donneurs et receveurs intervenant dans le mécanisme de branchement. Ils sont également décrits comme des « règles moléculaires » impliquées dans la spécificité de longueurs de chaînes transférées. Nos expériences ont permis de mettre en évidence deux sites de liaisons aux sucres (BS<sub>A</sub> et BS<sub>B</sub>) ainsi que les résidus impliqués dans cette fixation. Des



expériences de mutagenèse dirigée ayant été réalisées pour le site de liaison BS<sub>B</sub>, la caractérisation enzymatique de ces mutants permettrait de valider l'implication du site dans la fixation et/ou dans la spécificité de longueurs de chaînes transférées.

Concernant le site BS<sub>A</sub>, ce dernier ayant été identifié par homologie aux structures de HsBE et OsBEI, il conviendrait de réaliser de nouvelles expériences de cristallogénèse et de trempage afin de déterminer la structure cristalline de CgGBE en complexe avec un sucre positionné sur ce site. Ceci, couplé à des expériences de mutagenèse dirigée et à une caractérisation de l'activité enzymatique de ces mutants, permettrait d'explorer son rôle dans le mécanisme catalytique de l'enzyme. De plus, il serait possible d'observer si la présence de la boucle (H43-L53) permet la fixation de sucres de taille supérieure à ce qui a été précédemment observé chez HsBE et OsBEI.

L'étude de CsBE a permis aux auteurs de faire l'hypothèse d'un modèle de fixation des chaînes de glucanes et du mécanisme de transfert de ces chaînes. Or, comme nous l'avons démontré, les sites « d'entrée » et de « sortie » ne sont pas conservés chez CgGBE. Cette hypothèse ayant été réalisée pour une enzyme de la sous famille GH13\_9, il semble que le mode de fixation des chaînes de glucanes soit différent pour les enzymes de la sous famille GH13\_8. Les résidus conservés du domaine amino terminal pourraient faire l'objet d'expérience de mutagenèse dirigée afin d'observer les conséquences sur l'activité de l'enzyme et plus généralement sur leur rôle dans la fixation des sucres. Alors qu'il apparaît difficile de résoudre la structure tridimensionnelle d'une BE en complexe avec un oligosaccharide présent au site actif, il pourrait être envisagé de réaliser des expériences de modélisation informatique. En effet, ayant connaissance des sous sites de fixation du site actif et des sites de liaison aux sucres en surface de l'enzyme, il serait intéressant de modéliser la fixation d'une longue chaîne de glucane, permettant de mimer au mieux la fixation de polysaccharides complexes comme cela se produit *in vivo*.

La proximité structurale entre l'enzyme branchante de *Candida glabrata* et de l'Homme ne permet pas clairement d'affirmer que CgGBE pourrait être une cible thérapeutique. En effet, ces deux enzymes possédant une identité de séquence élevée, il est probable qu'une molécule thérapeutique pourrait ne pas être spécifique de l'enzyme du pathogène, nécessitant des études comparatives poussées. En revanche, comme nous l'avons indiqué, une partie du fonctionnement de ces enzymes demeure inconnue. Notre travail de caractérisation structurale et enzymatique participe donc à une meilleure compréhension du mécanisme des BEs et plus largement des glycoside hydrolases. Ces connaissances pourront notamment être étendues à l'étude d'autres glycoside hydrolases étudiées au sein du laboratoire intervenant au niveau de



la paroi fongique du pathogène et pouvant représenter des cibles thérapeutiques plus intéressantes.

## IV. Références

- Abad MC, Binderup K, Rios-Steiner J, Arni RK, Preiss J, Geiger JH. 2002. The X-ray crystallographic structure of *Escherichia coli* branching enzyme. *J Biol Chem.* 277(44):42164–42170. doi:10.1074/jbc.M205746200.
- Aon MA, Curtino JA. 1984. Evidence for the glycoprotein nature of retina glycogen. *Eur J Biochem.* 140(3):557–566. doi:10.1111/j.1432-1033.1984.tb08138.x.
- Ardevol A, Rovira C. 2015. Reaction Mechanisms in Carbohydrate-Active Enzymes: Glycoside Hydrolases and Glycosyltransferases. Insights from *ab initio* Quantum Mechanics/Molecular Mechanics Dynamic Simulations. *J Am Chem Soc.* 137(24):7528–7547. doi:10.1021/jacs.5b01156.
- Ball S, Colleoni C, Cenci U, Raj JN, Tirtiaux C. 2011. The evolution of glycogen and starch metabolism in eukaryotes gives molecular clues to understand the establishment of plastid endosymbiosis. *Journal of Experimental Botany.* 62(6):1775–1801. doi:10.1093/jxb/erq411.
- Binderup K, Mikkelsen R, Preiss J. 2000. Limited proteolysis of branching enzyme from *Escherichia coli*. *Arch Biochem Biophys.* 377(2):366–371. doi:10.1006/abbi.2000.1815.
- Binderup K, Mikkelsen R, Preiss J. 2002. Truncation of the amino terminus of branching enzyme changes its chain transfer pattern. *Arch Biochem Biophys.* 397(2):279–285. doi:10.1006/abbi.2001.2544.
- Bourne Y, Henrissat B. 2001. Glycoside hydrolases and glycosyltransferases: families and functional modules. *Curr Opin Struct Biol.* 11(5):593–600. doi:10.1016/s0959-440x(00)00253-0.
- Buléon A, Colonna P, Planchot V, Ball S. 1998. Starch granules: structure and biosynthesis. *International Journal of Biological Macromolecules.* 23(2):85–112. doi:10.1016/S0141-8130(98)00040-3.
- Cameron RE, Donald AM. 1993. A small-angle X-ray scattering study of the absorption of water into the starch granule. *Carbohydrate Research.* 244(2):225–236. doi:10.1016/0008-6215(83)85003-4.
- Caudwell FB, Cohen P. 1980. Purification and subunit structure of glycogen-branching enzyme from rabbit skeletal muscle. *Eur J Biochem.* 109(2):391–394. doi:10.1111/j.1432-1033.1980.tb04806.x.
- Chaen K, Noguchi J, Omori T, Kakuta Y, Kimura M. 2012. Crystal structure of the rice branching enzyme I (BEI) in complex with maltopentaose. *Biochemical and Biophysical Research Communications.* 424(3):508–511. doi:10.1016/j.bbrc.2012.06.145.
- Cheng C, Mu J, Farkas I, Huang D, Goebel MG, Roach PJ. 1995. Requirement of the self-glucosylating initiator proteins Glg1p and Glg2p for glycogen accumulation in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 15(12):6632–6640. doi:10.1128/MCB.15.12.6632.
- Davies G, Henrissat B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure.* 3(9):853–859. doi:10.1016/S0969-2126(01)00220-9.

- El Mannai Y, Deto R, Kuroki M, Suzuki R, Suzuki E. 2021. Cyanobacterial branching enzymes bind to  $\alpha$ -glucan via surface binding sites. *Archives of Biochemistry and Biophysics*. 702:108821. doi:10.1016/j.abb.2021.108821.
- Errera L. 1885. Sur l'existence du glycogène dans la levure de bière. *Comptes rendus hebdomadaires des séances de l'Académie des sciences*. 101:1–3.
- Farkas I, Hardy TA, Goebel MG, Roach PJ. 1991. Two glycogen synthase isoforms in *Saccharomyces cerevisiae* are coded by distinct genes that are differentially controlled. *J Biol Chem*. 266(24):15602–15607.
- Feng L, Fawaz R, Hovde S, Gilbert L, Chiou J, Geiger JH. 2015. Crystal Structures of *Escherichia coli* Branching Enzyme in Complex with Linear Oligosaccharides. *Biochemistry*. 54(40):6207–6218. doi:10.1021/acs.biochem.5b00228.
- François J, Parrou JL. 2001. Reserve carbohydrates metabolism in the yeast *Saccharomyces cerevisiae*. *FEMS Microbiol Rev*. 25(1):125–145. doi:10.1111/j.1574-6976.2001.tb00574.x.
- Froese DS, Michaeli A, McCorvie TJ, Krojer T, Sasi M, Melaev E, Goldblum A, Zatsepin M, Lossos A, Álvarez R, et al. 2015. Structural basis of glycogen branching enzyme deficiency and pharmacologic rescue by rational peptide design. *Hum Mol Genet*. 24(20):5667–5676. doi:10.1093/hmg/ddv280.
- Hayashi M, Suzuki R, Colleoni C, Ball S, Fujita N, Suzuki E. 2017. Bound Substrate in the Structure of Cyanobacterial Branching Enzyme Supports a New Mechanistic Model. *Journal of Biological Chemistry*. 292(13):5465–5475. doi:10.1074/jbc.M116.755629.
- Henrissat B. 1991. A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 280 ( Pt 2):309–316. doi:10.1042/bj2800309.
- Henrissat B, Bairoch A. 1993. New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem J*. 293 ( Pt 3):781–788. doi:10.1042/bj2930781.
- Janeček Š, Svensson B, MacGregor EA. 2003. Relation between domain evolution, specificity, and taxonomy of the  $\alpha$ -amylase family members containing a C-terminal starch-binding domain. *European Journal of Biochemistry*. 270(4):635–645. doi:10.1046/j.1432-1033.2003.03404.x.
- Jiang H, Xie X, Ban X, Gu Z, Cheng L, Hong Y, Li C, Li Z. 2021. Flexible Loop in Carbohydrate-Binding Module 48 Allosterically Modulates Substrate Binding of the 1,4- $\alpha$ -Glucan Branching Enzyme. *J Agric Food Chem*. 69(20):5755–5763. doi:10.1021/acs.jafc.1c00293.
- Koshland DE. 1953. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biological Reviews*. 28(4):416–436. doi:https://doi.org/10.1111/j.1469-185X.1953.tb01386.x.
- Kuriki T, Guan H, Sivak M, Preiss J. 1996. Analysis of the active center of branching enzyme II from maize endosperm. *J Protein Chem*. 15(3):305–313. doi:10.1007/BF01887119.
- Kuriki T, Stewart DC, Preiss J. 1997. Construction of Chimeric Enzymes out of Maize Endosperm Branching Enzymes I and II:: ACTIVITY AND PROPERTIES\*. *Journal of Biological Chemistry*. 272(46):28999–29004. doi:10.1074/jbc.272.46.28999.

Leloir LF, Cardini CE. 1957. BIOSYNTHESIS OF GLYCOGEN FROM URIDINE DIPHOSPHATE GLUCOSE1. J Am Chem Soc. 79(23):6340–6341. doi:10.1021/ja01580a061.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 42(Database issue):D490–D495. doi:10.1093/nar/gkt1178.

MacGregor EA, Janeček Š, Svensson B. 2001. Relationship of sequence and structure to specificity in the  $\alpha$ -amylase family of enzymes. Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology. 1546(1):1–20. doi:10.1016/S0167-4838(00)00302-2.

Manners DJ. 1989. Recent developments in our understanding of amylopectin structure. Carbohydrate Polymers. 11(2):87–112. doi:10.1016/0144-8617(89)90018-0.

Martinovičová M, Janeček Š. 2018. In silico analysis of the  $\alpha$ -amylase family GH57: eventual subfamilies reflecting enzyme specificities. 3 Biotech. 8(7):307. doi:10.1007/s13205-018-1325-9.

Mu J, Cheng C, Roach PJ. 1996. Initiation of glycogen synthesis in yeast. Requirement of multiple tyrosine residues for function of the self-glucosylating Glg proteins in vivo. J Biol Chem. 271(43):26554–26560. doi:10.1074/jbc.271.43.26554.

Mukerjea R, Yu L, Robyt JF. 2002. Starch biosynthesis: mechanism for the elongation of starch chains. Carbohydrate Research. 337(11):1015–1022. doi:10.1016/S0008-6215(02)00067-8.

Mundkur B. 1960. Electron microscopical studies of frozen-dried yeast. I. Localization of polysaccharides. Exp Cell Res. 20:28–42. doi:10.1016/0014-4827(60)90219-6.

Murakami T, Kanai T, Takata H, Kuriki T, Imanaka T. 2006. A Novel Branching Enzyme of the GH-57 Family in the Hyperthermophilic Archaeon *Thermococcus kodakaraensis* KOD1. J Bacteriol. 188(16):5915–5924. doi:10.1128/JB.00390-06.

Nakamura Y. 2002. Towards a better understanding of the metabolic system for amylopectin biosynthesis in plants: rice endosperm as a model tissue. Plant Cell Physiol. 43(7):718–725. doi:10.1093/pcp/pcf091.

Noguchi J, Chaen K, Vu NT, Akasaka T, Shimada H, Nakashima T, Nishi A, Satoh H, Omori T, Kakuta Y, et al. 2011. Crystal structure of the branching enzyme I (BEI) from *Oryza sativa* L with implications for catalysis and substrate binding. Glycobiology. 21(8):1108–1116. doi:10.1093/glycob/cwr049.

Northcote DH. 1953. The molecular structure and shape of yeast glycogen. Biochemical Journal. 53(3):348–352. doi:10.1042/bj0530348.

Oslancová A, Janecek S. 2002. Oligo-1,6-glucosidase and neopullulanase enzyme subfamilies from the alpha-amylase family defined by the fifth conserved sequence region. Cell Mol Life Sci. 59(11):1945–1959. doi:10.1007/pl00012517.

Pal K, Kumar S, Sharma S, Garg SK, Alam MS, Xu HE, Agrawal P, Swaminathan K. 2010. Crystal Structure of Full-length *Mycobacterium tuberculosis* H37Rv Glycogen Branching Enzyme. J Biol Chem. 285(27):20897–20903. doi:10.1074/jbc.M110.121707.

- Pitcher J, Smythe C, Campbell DG, Cohen P. 1987. Identification of the 38-kDa subunit of rabbit skeletal muscle glycogen synthase as glycogenin. *Eur J Biochem.* 169(3):497–502. doi:10.1111/j.1432-1033.1987.tb13637.x.
- Roach PJ. 2002. Glycogen and its Metabolism. *Current Molecular Medicine.* 2(2):101–120. doi:10.2174/1566524024605761.
- Rodriguez IR, Whelan WJ. 1985. A novel glycosyl-amino acid linkage: Rabbit-muscle glycogen is covalently linked to a protein via tyrosine. *Biochemical and Biophysical Research Communications.* 132(2):829–836. doi:10.1016/0006-291X(85)91206-9.
- Rowen DW, Meinke M, LaPorte DC. 1992. GLC3 and GHA1 of *Saccharomyces cerevisiae* are allelic and encode the glycogen branching enzyme. *Mol Cell Biol.* 12(1):22–29.
- Shearer J, Graham T. 2002. New Perspectives on the Storage and Organization of Muscle Glycogen. *Canadian journal of applied physiology = Revue canadienne de physiologie appliquée.* 27:179–203. doi:10.1139/h02-012.
- Sim L, Beeren SR, Findinier J, Dauvillée D, Ball SG, Henriksen A, Palcic MM. 2014. Crystal Structure of the Chlamydomonas Starch Debranching Enzyme Isoamylase ISA1 Reveals Insights into the Mechanism of Branch Trimming and Complex Assembly \*. *Journal of Biological Chemistry.* 289(33):22991–23003. doi:10.1074/jbc.M114.565044.
- Smythe C, Cohen P. 1991. The discovery of glycogenin and the priming mechanism for glycogen biogenesis. *Eur J Biochem.* 200(3):625–631. doi:10.1111/j.1432-1033.1991.tb16225.x.
- Søgaard M, Kadziola A, Haser R, Svensson B. 1993. Site-directed mutagenesis of histidine 93, aspartic acid 180, glutamic acid 205, histidine 290, and aspartic acid 291 at the active site and tryptophan 279 at the raw starch binding site in barley alpha-amylase 1. *Journal of Biological Chemistry.* 268(30):22480–22484. doi:10.1016/S0021-9258(18)41554-2.
- Stam MR, Danchin EGJ, Rancurel C, Coutinho PM, Henrissat B. 2006. Dividing the large glycoside hydrolase family 13 into subfamilies: towards improved functional annotations of -amylase-related proteins. *Protein Engineering Design and Selection.* 19(12):555–562. doi:10.1093/protein/gzl044.
- Sun C, Sathish P, Ahlandsberg S, Jansson C. 1998. The Two Genes Encoding Starch-Branching Enzymes IIa and IIb Are Differentially Expressed in Barley. *Plant Physiol.* 118(1):37–49.
- Suzuki E, Suzuki R. 2016. Distribution of glucan-branching enzymes among prokaryotes. *Cell Mol Life Sci.* 73(14):2643–2660. doi:10.1007/s00018-016-2243-9.
- Suzuki R, Suzuki E. 2020. Structure and Function of Branching Enzymes in Eukaryotes. *Trends in Glycoscience and Glycotechnology.* 32(185):E21–E30. doi:10.4052/tigg.1974.1E.
- Takata H, Takaha T, Kuriki T, Okada S, Takagi M, Imanaka T. 1994. Properties and active center of the thermostable branching enzyme from *Bacillus stearothermophilus*. *Appl Environ Microbiol.* 60(9):3096–3104. doi:10.1128/aem.60.9.3096-3104.1994.
- Tetlow IJ, Emes MJ. 2014. A review of starch-branching enzymes and their role in amylopectin biosynthesis. *IUBMB Life.* 66(8):546–558. doi:10.1002/iub.1297.

Wilson WA, Roach PJ, Montero M, Baroja-Fernández E, Muñoz FJ, Eydallin G, Viale AM, Pozueta-Romero J. 2010. Regulation of glycogen metabolism in yeast and bacteria. *FEMS Microbiology Reviews*. 34(6):952–985. doi:10.1111/j.1574-6976.2010.00220.x.

Zmasek CM, Godzik A. 2014. Phylogenomic analysis of glycogen branching and debranching enzymatic duo. *BMC Evolutionary Biology*. 14(1):183. doi:10.1186/s12862-014-0183-2.







## Caractérisations biochimiques et structurales d'une métacaspase et d'une enzyme branchante de glucane en $\alpha$ -(1,4) chez *Candida glabrata*

### RÉSUMÉ

Les infections causées par diverses espèces *Candida* ont considérablement augmenté ces dernières années notamment dans un cadre hospitalier où elles entraînent une forte mortalité. Parallèlement à *Candida albicans*, *Candida glabrata* est l'une des levures pathogènes ayant la plus forte prévalence chez l'humain. À ce jour, les composés antifongiques utilisés en traitement présentent de nombreux phénomènes de résistance nécessitant l'identification de nouvelles cibles thérapeutiques. Afin d'explorer de nouvelles voies, nous nous sommes intéressés à la métacaspase de *C. glabrata*. Ces protéases à cystéines sont des homologues structuraux des caspases et des effecteurs de la mort cellulaire régulée, un processus physiologique visant à maintenir l'homéostasie cellulaire chez les métazoaires et les levures. L'étude structurale, fonctionnelle et enzymatique a permis d'élucider le déroulement de la maturation de l'enzyme dans un mécanisme *trans* avec ses changements structuraux associés, ainsi que son activation, les deux processus étant dépendants de la présence d'ions calcium. Nous avons également étudié une enzyme branchante de glucane en  $\alpha$ -(1,4) de la sous famille des GH13\_8. Ces dernières sont impliquées dans le métabolisme énergétique du glycogène et plus particulièrement dans l'introduction de ramifications *via* la création de liaisons  $\alpha$ -(1,6). L'étude biochimique et structurale a permis la résolution de la structure tridimensionnelle de l'enzyme sous une forme libre et en complexe avec un maltotriose. Nous avons également mis en évidence des particularités structurales de l'enzyme propres aux levures et plus particulièrement au phylum des *Saccharomycotina*.

Mots clés : *Candida glabrata*, mort cellulaire régulée, métacaspase, glycoside hydrolase, enzyme branchante de glucane, cristallographie aux rayons X