



HAL
open science

Développement d'une méthode de déconvolution pharmacophorique pour la découverte accélérée d'antipaludiques chez les Rhodophytes

Laure Margueritte

► **To cite this version:**

Laure Margueritte. Développement d'une méthode de déconvolution pharmacophorique pour la découverte accélérée d'antipaludiques chez les Rhodophytes. Chimie analytique. Université de Strasbourg, 2018. Français. NNT : 2018STRAF048 . tel-02138926

HAL Id: tel-02138926

<https://theses.hal.science/tel-02138926>

Submitted on 24 May 2019

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.

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES (ED 222)

Laboratoire d'Innovation Thérapeutique

(UMR 7200 CNRS/Unistra)

THÈSE

présentée par :

Laure MARGUERITTE

soutenue le : **21 Septembre 2018**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : **Pharmacognosie – Chimie Analytique**

Développement d'une méthode de déconvolution pharmacophorique pour la découverte accélérée d'antipaludiques chez les Rhodophytes

THÈSE dirigée par :

M. HIBERT Marcel

Mme VONTHRON-SÉNÉCHEAU Catherine

Professeur, Université de Strasbourg

Maître de Conférences (HDR), Université de Strasbourg

RAPPORTEURS :

M. RENAULT Jean-Hugues

M. TOUBOUL David

Professeur, Université de Reims

Chargé de recherche, ICSN, Paris-Saclay

AUTRES MEMBRES DU JURY :

M. DELSUC Marc-André

M. WOLFENDER Jean-Luc

Mme RUSIG Anne-Marie

Directeur de recherche CNRS, IGBMC, Strasbourg

Professeur, Université de Genève

Maître de Conférences, Université de CAEN-Normandie

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES (ED 222)

Laboratoire d'Innovation Thérapeutique

(UMR 7200 CNRS/Unistra)

THÈSE

présentée par :

Laure MARGUERITTE

soutenue le : **21 Septembre 2018**

pour obtenir le grade de : **Docteur de l'université de Strasbourg**

Discipline/ Spécialité : **Pharmacognosie – Chimie Analytique**

Développement d'une méthode de déconvolution pharmacophorique pour la découverte accélérée d'antipaludiques chez les Rhodophytes

THÈSE dirigée par :

M. HIBERT Marcel

Mme VONTHRON-SÉNÉCHEAU Catherine

Professeur, Université de Strasbourg

Maître de Conférences (HDR), Université de Strasbourg

RAPPORTEURS :

M. RENAULT Jean-Hugues

M. TOUBOUL David

Professeur, Université de Reims

Chargé de recherche, ICSN, Paris-Saclay

AUTRES MEMBRES DU JURY :

M. DELSUC Marc-André

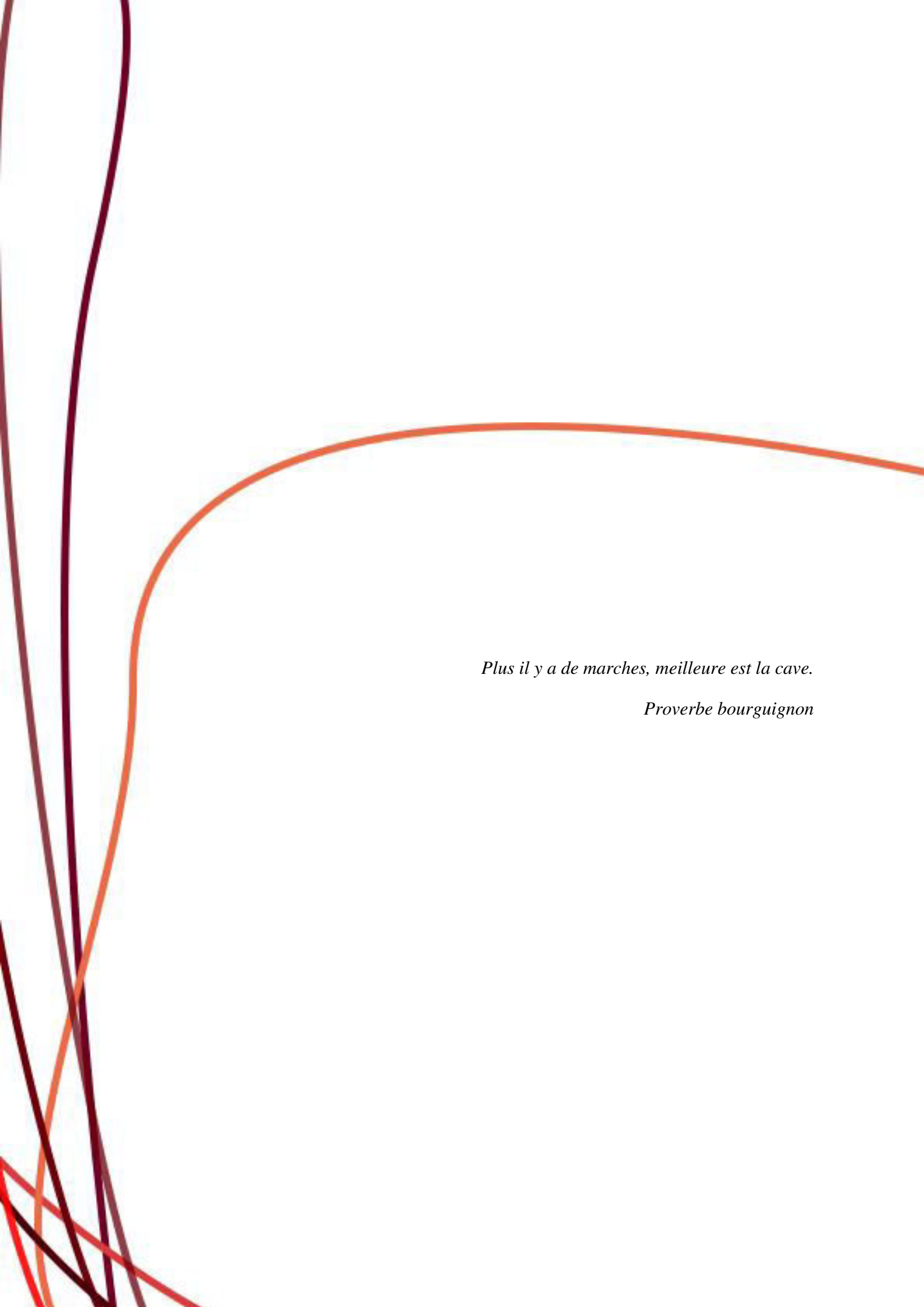
M. WOLFENDER Jean-Luc

Mme RUSIG Anne-Marie

Directeur de recherche CNRS, IGBMC, Strasbourg

Professeur, Université de Genève

Maître de Conférences, Université de CAEN-Normandie



Plus il y a de marches, meilleure est la cave.

Proverbe bourguignon

Remerciements

Je tiens à exprimer toute ma reconnaissance aux membres du jury qui ont accepté d'évaluer ce travail. Je remercie le Pr Jean-Hugues Renault, de l'Université de Reims et le Dr David Touboul, de l'Institut de Chimie des Substances Naturelles (ICSN) d'avoir accepté la charge de rapporteurs. Je remercie le Pr Jean-Luc Wolfender pour l'intérêt et le temps qu'il a consacrés à la lecture de ce travail.

J'ai eu la chance et le plaisir d'effectuer cette thèse au sein du Laboratoire d'Innovation Thérapeutique (LIT) d'Octobre 2015 à Octobre 2018, sous la direction du Pr Marcel Hibert et du Dr Catherine Vonthron-Sénécheau.

Je remercie tout d'abord les directeurs du laboratoire, Pr Marcel Hibert et Dr Didier Rognan, de m'avoir donné l'opportunité de réaliser ma thèse au sein de cet établissement.

Je remercie une nouvelle fois le Pr Marcel Hibert en tant que directeur de thèse, pour son soutien et son implication continus dans ce projet malgré sa fonction de directeur de laboratoire. Ta créativité et ton originalité représentent cette *Marcel'touch* si caractéristique dont il faut s'inspirer.

Je remercie chaleureusement ma directrice de thèse Catherine Vonthron-Sénécheau, pour sa direction et l'aiguillage pertinent de ma thèse. Malgré ton emploi du temps chargé, tu as toujours su être là pour faire le point sur les avancées et l'orientation à privilégier. Ta bienveillance et ton écoute tout au long de ces trois ans m'ont permis de réaliser cette thèse dans d'agréables conditions. Ton sens de la pédagogie m'a beaucoup inspiré et j'espère le transmettre dans de futures situations. J'espère que tu trouveras encore quelque chose à grignoter pour tes pauses de 17h.

Je remercie doublement le Dr Marc-André Delsuc, de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) en tant que membre invité de ce jury et d'encadrant. Officieusement troisième directeur de ma thèse, son avancée n'aurait certainement pas été la même sans son investissement. Je te remercie de m'avoir accueillie dans ton « labo », de m'avoir formée sur le super spectromètre RMN, et d'avoir été aussi

patient lorsque nous ne parlions pas la même langue (de serpent). Ta vision de la recherche et ton encadrement me resteront en exemple.

Je remercie le Dr Mélanie Bourjot pour son encadrement pendant la première moitié de ma thèse, ainsi que pour son dynamisme et sa persévérance dans ce projet. Tes conseils techniques de phytochimiste m'ont été essentiels. Je te souhaite tout le meilleur au niveau professionnel, mais surtout personnel.

Je remercie le Pr Ermanno Candolfi de l'Institut de Parasitologie et de Pathologie Tropicale (IPPTS), pour m'avoir accueillie dans son établissement pendant ces trois années pour la réalisation des tests biologiques. Je remercie bien sûr Flore Nardella et Cécile Lang pour leur transmission de savoir quant au dressage de ce parasite et de leur expertise dans ce domaine (ainsi que la bonne humeur de Lydia et Esterina).

Je remercie toutes les personnes qui ont participé à ce projet :

- Dr. Anne-Marie Rusig de l'Université de Caen, pour la récolte, l'identification et le deal de poudre... d'algues.
- Dr. Andréa Hemmerlin, de l'Institut de Biologie Moléculaire des Plantes (IBMP) et Quentin Chevalier pour les évaluations des algues sur la voie des isoprénoïdes.

Je souhaite remercier une seconde fois le Dr David Touboul pour m'avoir accueillie au sein de l'équipe de spectrométrie de masse de l'ICSN. J'espère que la quantité de spectres RMN dans cette thèse ne rendra pas sa lecture trop pénible...

Je remercie Delphine Garnier, responsable de la plateforme d'Analyse Chimique de Strasbourg-Illkirch (PACSI) pour son investissement et sa persévérance quant au fonctionnement de la méthode couplée. Je remercie également tout le personnel qui est passé par là !

Je tiens à remercier chaleureusement toutes les personnes qui composent cette UMR 7200, qui m'ont aidée et m'ont toujours apporté toutes les informations techniques et autres pendant ces trois années. Je remercie bien entendu Patrick alias « Deus », merci de m'avoir accompagnée techniquement et personnellement pendant ces trois ans. Je remercie également Séverine pour ses petits conseils, surtout vis-à-vis des « marchés » et autres histoires de commande. Merci, à David, le meilleur secrétaire que je connaisse jusqu'à présent. Je remercie l'équipe SOMHet, pour leur expertise de chimistes durs. Je remercie la bande de l'aquarium partie trop tôt : Yann, Malik et Ludo.

Mes pensées vont bien sûr à notre équipe de football ! Fab, Seb, Guilhem, Mickael, Mickael (Momo), Nicolas G, et les filles : Lucie et Laurie !

Merci à Alex (arrivé trop tard cette fois), pour l'observation des Grands Ducs, la grimpe et sa bonne humeur !

Je remercie les Hibert, Steph pour m'avoir laissée de la place sur l'HPLC et à côté d'elle au repas et Sri pour sa bonne humeur, sa vitalité et son éléphant.

Je remercie Lorène pour avoir été là pendant les deux premières années, pour son soutien moral et scientifique (et pour les petites soirées mezzés et autres debriefings d'émissions télévisées dont je ne citerai pas les noms).

Merci au couple strasbourgeois de la promo, Inès et Victor.

Je remercie la bande d'Annecy pour leur bonne humeur, JD, Sophish, AL, Pierre, BonNass, Flo, Charly, Coralie, Victor (et Patrick bien sûr).

Un grand merci à mes deux meilleures amies bourguignonnes, Jess et Zoé, d'avoir toujours été là (Zoé quand est-ce que l'on collabore ? Jess merci pour tout, ton soutien et la fête du vin d'Eguisheim 2016).

Je remercie les meufs, Lucie pour ta joie de vivre et ton rayonnement. Toujours disponible pour boire un verre, papoter. N'oublie pas de penser à toi malgré tout.

Laura, tu es un peu comme le Saint-Agur, fort et fondant même si tu préfères la raclette au petit-déj. J'espère que nous continuerons à nous voir pour boire de la Triple, je demanderai à Thibaut qu'il m'invite.

Enfin, je remercie Nico, mon colocataire de bureau et de thèse depuis trois ans, bref mon bro'. J'aurais pu difficilement trouver mieux comme compagnon de bureau (excepté les jours de boudage intensif). J'écris ces lignes en te tournant le dos pendant que tu stresses pour ta rédaction (rire espiègle). Je résumerai tous les moments de bonheur que nous avons partagés par quelques mots-clés : chaises roulantes, sushis, randos, escalade, café, bières, et reggae !

Je remercie mes parents de m'avoir toujours soutenue et réconfortée, notamment ces derniers mois en prenant leur mal en patience quant à mon absence de nouvelles. Je vous remercie de l'éducation et de l'ouverture d'esprit que vous m'avez apportées. Je vous suis

infiniment reconnaissante de n'avoir jamais influencé un seul choix d'orientation et de m'avoir toujours fait confiance. Je pense à mon frère Fabien, Dilek et le petit Ilyes, préparez-vous à me voir débarquer à New-York prochainement !

Enfin, je termine par mon Geo aux yeux couleur lac d'Annecy. Je te remercie pour ta patience lors de ces plats derniers mois, pour ton soutien et tes encouragements sans faille (... « tu es le seul qui m'aïlle, je te le dis sans faille, reste cool bébé » etc). Merci pour la vie pleine d'aventures, d'insouciance, de bonheur et d'amour que tu m'apportes. Nous partons aujourd'hui, ensemble cette fois, vers de nouvelles aventures !

A ma famille.

Productions Scientifiques

Publications

- Margueritte L, Markov P, Chiron L, Starck J-P, Vonthron-Sénécheau C, Bourjot M, Delsuc M-A. Automatic differential analysis of NMR experiments in complex samples. *Magnetic Resonance in Chemistry*. 2018;56:469–479.
<https://doi.org/10.1002/mrc.4683>
- Nardella F, Margueritte L, Lamure B, Viéville J M P, Bourjot M. Targeted discovery of bioactive natural products using a pharmacophoric deconvolution strategy: Proof of principle with eleganolone from *Bifurcaria bifurcata* R. Ross. *Phytochemistry Letters*. 2018;56:469–479.
<https://doi.org/10.1016/j.phytol.2018.05.034>

Communications orales

- **The 5th AFERP International Conference, 2017 (Juillet 2017, Angers)**
Development of a chemical deconvolution method to accelerate the discovery of antiplasmodial molecules from Rhodophyta. L. Margueritte, M. Bourjot, P. Markov, G. Bret, D. Rognan, M.A. Delsuc, C. Vonthron-Sénécheau
- **Phytoday 2017 (Juin 2017, Reims)**
Development of a pharmacophoric deconvolution method for the targeted discovery of bioactive natural products: application to the antimalarial compounds discovery from Rhodophyta (Red algae). L. Margueritte, M.A. Delsuc, G. Bret, D. Rognan, M. Bourjot, C. Vonthron-Sénécheau
- **Journées du Campus d'Illkirch 2017 (Mars 2017, Illkirch)**
Development of a pharmacophoric deconvolution method to accelerate the discovery of antiplasmodial molecules from Rhodophyta. L. Margueritte, M. Bourjot, P. Markov, G. Bret, M.A. Delsuc, D. Rognan, C. Vonthron-Sénécheau

Communications affichées

- **European Magnetic Resonance Meeting (EUROMAR) (Juin 2018, Nantes)**
Recent improvement in the SPIKE processing program. Lionel Chiron, Laura Duciel, Afef Cherni, Laure Margueritte, Fabrice Bray, Christian Rolando, Marc-André Delsuc
- **7^{ème} Journée du consortium antiparasitaire et fongique (CAPF) (Mai 2017, Strasbourg)**
Development of a pharmacophoric deconvolution method to accelerate the discovery of antiplasmodial molecules from Rhodophyta. L. Margueritte, M. Bourjot, P. Markov, G. Bret, M.A. Delsuc, D. Rognan, C. Vonthron-Sénécheau
- **9th Joint Natural Products Conference 2016 (Juillet 2016, Copenhagen)**
Development of a pharmacophoric deconvolution method for the targeted discovery of bioactive natural products. M. Bourjot, L. Margueritte, P. Markov, G. Bret, M.A. Delsuc, D. Rognan, C. Vonthron-Sénécheau
Planta Medica, 2016, 82(S01): S1-S381
- **Journée du Campus d'Illkirch 2016 (Mars 2016, Illkirch)**
Development of a pharmacophoric deconvolution method to accelerate the discovery of antiplasmodial molecules from Rhodophyta. L. Margueritte, M. Bourjot, P. Markov, G. Bret, M.A. Delsuc, D. Rognan, C. Vonthron-Sénécheau
- **Journée du Labex Medalis (Décembre 2015, Strasbourg)**
Development of a pharmacophoric deconvolution method for the targeted discovery of bioactive natural products. M. Bourjot, L. Margueritte, F. Nardella, J.-B. Gallé, B. Schaeffer, J.M.P. Viéville, C. Vonthron-Sénécheau

Table des matières

Remerciements	i
Productions Scientifiques	v
Table des matières	vii
Liste des Abréviations	xii
Liste des Figures	xiv
Liste des Tableaux.....	xx
Contexte et objectifs de la thèse	3
Chapitre I : Traitement du paludisme et stratégie de recherche.....	9
I. Le paludisme.....	10
I. 1. Cycle de développement de <i>Plasmodium</i>	11
I. 2. Traitements antipaludiques actuels et en développement	13
I. 2. a. Alcaloïdes du quinquina	14
I. 2. b. Amino-alcools	14
I. 2. c. 4-aminoquinoléines	15
I. 2. d. 8-aminoquinoléines	16
I. 2. e. Antifolates	17
I. 2. f. Endoperoxydes	18
I. 2. g. Naphtoquinones	19
I. 2. h. Nouvelles classes d'antipaludiques	19
I. 3. Molécules naturelles à activité antipaludique	20
I. 4. Conclusion.....	24
II. L'Apicoplaste : Une cible thérapeutique d'intérêt.....	24
II. 1. Une origine longtemps discutée	25
II. 2. Une cible thérapeutique.....	25
II. 2. a. Inhibiteurs de la traduction ou transcription dans l'apicoplaste	27
II. 2. b. Inhibiteurs de la voie de biosynthèse de l'hème et des complexes Fer-Soufre	28
II. 2. c. Inhibiteurs de la voie de biosynthèse des acides gras.....	29
II. 2. d. Inhibiteurs de la voie de biosynthèse des isoprénoides	29
II. 3. Conclusion.....	31
III. Les algues rouges (Rhodophytes).....	32
III. 1. Généralités ³⁹	32
III. 2. Les algues rouges et l'activité antipaludique	34
III. 3. Phytochimie des algues rouges	36
III. 3. a. Métabolites des algues rouges.....	36
III. 3. b. Métabolites antipaludiques des algues rouges	39
IV. Conclusion et objectifs du travail	40
Chapitre II : Etude de l'activité antipaludique d'extraits d'algues rouges	43
I. Présentation du matériel algal	44
II. Phytochimie et activités biologiques des espèces étudiées	49

III. Extraits obtenus.....	50
IV. Evaluation biologique des algues	51
IV. 1. Activité antipaludique <i>in vitro</i>	51
IV. 2. Inhibition de la voie des isoprénoïdes et de la prénylation des protéines	54
V. Conclusion.....	58
Chapitre III : Automatisation de l'analyse différentielle par RMN.....	63
I. Méthodes d'identification de molécules naturelles bioactives dans les extraits végétaux.....	64
I. 1. Techniques analytiques	64
I. 2. Le fractionnement bio-guidé	69
I. 3. Les stratégies de déréplication	70
I. 4. Les méthodes couplées	73
I. 5. Pourquoi développer une méthode alternative ?	74
II. Mise en place de la stratégie analytique.....	75
II. 1. Travaux du Dr Frank Schroeder	75
II. 2. Essai de l'analyse différentielle de F. Schroeder sur <i>Bifurcaria bifurcata</i>	77
II. 3. Présentation générale de la méthode de déconvolution pharmacophorique...	81
III. Automatisation de l'analyse différentielle par RMN¹⁶⁰.....	83
III. 1. Le langage Python et son environnement	84
III. 2. SPIKE (<i>Spectrometry Processing Innovative Kernel</i>).....	85
III. 3. Plasmodesma et le <i>bucketing</i>	85
III. 3. a. Principe des opérations.....	86
III. 3. b. Le <i>bucketing</i>	87
III. 4. <i>Bucket Utilities</i>	89
III. 5. Pré-traitement des spectres : Nettoyage et symétrisation.....	89
III. 5. a. Nettoyage des spectres	89
III. 5. b. Symétrisation des spectres homonucléaires	92
III. 6. Conclusion.....	93
IV. Exemplification de l'analyse différentielle sur des échantillons reconstitués.....	93
IV. 1. Présentation de la série CHLORO.....	95
IV. 2. Présentation de la série ARTE	98
IV. 3. Analyse différentielle de spectres deux à deux	99
IV. 3. a. Choix de la variable et du calcul	99
IV. 3. b. Analyse différentielle deux à deux de spectres ¹ H- ¹ H TOCSY	105
IV. 3. c. Analyse différentielle deux à deux de spectres ¹ H- ¹³ C HSQC	108
IV. 3. d. Analyse différentielle deux à deux de spectres ¹ H- ¹³ C HMBC.....	111
IV. 3. e. Conclusion sur l'analyse différentielle de spectres de fractions actives versus inactives	114
IV. 4. Analyse différentielle de l'ensemble des échantillons	114
IV. 4. a. Analyse par régression linéaire : Utilisation de la fonction <code>linear_model.LinearRegression</code>	115
IV. 4. b. Analyse par régression linéaire : Utilisation de l'outil RFE	119
IV. 4. c. Analyse différentielle par régression linéaire sur les autres expériences RMN ..	121

IV. 5.	Conclusion sur l'analyse différentielle de l'ensemble des échantillons.....	127
IV. 6.	Les expériences RMN ^1H - δ DOSY	127
V.	Conclusion.....	130
Chapitre IV : Identification de molécules bio-actives dans des extraits végétaux 135		
I.	Analyse différentielle « en se rapprochant du réel » - Série SMARTE.....	136
I. 1.	Introduction	136
I. 2.	Analyse différentielle : Fractions actives <i>versus</i> inactives.....	137
I. 3.	Analyse différentielle par régression linéaire.....	139
I. 4.	Filtration par les expériences ^1H - δ DOSY	140
I. 5.	Conclusion.....	142
II.	Analyse différentielle de <i>Bifurcaria bifurcata</i> – Preuve de concept.....	143
II. 1.	Analyse différentielle automatique de <i>B. bifurcata</i>	144
II. 1. a.	Analyse différentielle des spectres ^1H - ^1H TOCSY.....	144
II. 1. b.	Analyse différentielle des spectres ^1H - ^{13}C HSQC.....	145
II. 2.	Comparaison des empreintes spectrales et du spectre RMN issu de la méthode couplée	146
II. 2. a.	Comparaison de l'empreinte ^1H - ^1H TOCSY et du spectre ^1H - ^1H COSY de la molécule stockée sur la cartouche SPE	146
II. 2. b.	Comparaison de l'empreinte ^1H - ^{13}C HSQC et du spectre ^1H - ^{13}C HSQC de la molécule stockée sur la cartouche SPE	147
II. 3.	Conclusion.....	147
III.	Analyse différentielle de <i>Cinchona pubescens</i> – Cas complexe.....	149
III. 1.	Analyse différentielle par régression linéaire.....	151
III. 1. a.	Analyse différentielle des expériences ^1H - ^1H TOCSY	151
III. 1. b.	Analyse différentielle des expériences ^1H - ^{13}C HSQC	152
III. 2.	Conclusion.....	153
IV.	Analyse différentielle de <i>Dilsea carnos</i>a.....	154
IV. 1.	Introduction	154
IV. 2.	Analyse différentielle de <i>D. carnos</i> a.....	155
IV. 2. a.	Analyse différentielle des spectres ^1H - ^1H TOCSY	155
IV. 2. b.	Analyse différentielle des spectres ^1H - ^{13}C HSQC.....	156
IV. 3.	Identification des composés par la méthode couplée HPLC-SPE-RMN	157
IV. 4.	Identification des composés par fractionnement bioguidé.....	157
IV. 5.	Comparaison de F2-3-3 à l'empreinte pharmacophorique.....	159
IV. 6.	Comparaison avec l'empreinte de <i>Gracilaria gracilis</i>	160
IV. 6. a.	Analyse différentielle des spectres RMN	161
IV. 6. b.	Comparaison des empreintes de <i>D. carnos</i> a et de <i>G. gracilis</i>	163
IV. 7.	Conclusion.....	167
V.	Conclusion.....	168
Discussion..... 173		
Partie expérimentale 187		
I.	Procédures générales	187

I. 1.	Solvants	187
I. 2.	Fractionnements, purifications et suivi	187
I. 2. a.	Chromatographie sur couche mince (CCM).....	187
I. 2. b.	Chromatographie sur colonne (CC).....	187
I. 2. c.	Chromatographie liquide haute performance (CLHP)	187
I. 3.	Analyses physiques et spectrales des composés	188
I. 3. a.	Spectrométrie de masse (SM).....	188
I. 3. b.	Résonance magnétique nucléaire (RMN).....	188
II.	Matériel végétal	189
II. 1.	Matériel algal.....	189
II. 1. a.	Collection et identification	189
II. 1. b.	Préparation des extraits bruts.....	189
II. 2.	Matériel végétal non algal	190
III.	Méthode de déconvolution pharmacophorique.....	190
III. 1.	Fractionnement par MPLC.....	190
III. 1. a.	Extraits algaux.....	190
III. 1. b.	Extrait non algal : <i>Cinchona pubescens</i> L.....	191
III. 2.	Analyses RMN	191
III. 3.	Bonnes pratiques d'acquisition RMN	192
III. 4.	Traitement des données	193
III. 4. a.	Principe des opérations.....	193
III. 4. b.	Auto-phasage des analyses RMN-1D (fonction FT1D).....	194
III. 4. c.	Auto-zéro.....	195
III. 4. d.	<i>Processing</i>	195
III. 4. e.	<i>Peak-picking</i>	195
III. 4. f.	<i>Bucketing</i>	196
III. 4. g.	Rapport.....	196
III. 5.	<i>Bucket Utilities</i>	196
III. 5. a.	Nettoyage des spectres bucketés	196
III. 5. b.	Symétrisation des spectres bucketés homonucléaires	197
III. 5. c.	Importation et affichage des spectres bucketés	197
III. 6.	Analyse différentielle par régression linéaire.....	198
III. 7.	Filtration des empreintes par les empreintes ¹ H-δ DOSY.....	199
III. 8.	Méthode couplée HPLC-DAD-MS-SPE-RMN	202
IV.	Essais biologiques <i>anti-Plasmodium</i>	202
IV. 1.	Culture de <i>P. falciparum</i>	202
IV. 1. a.	Souche plasmodiale.....	202
IV. 1. b.	Préparation des hématies.....	202
IV. 1. c.	« Milieu de culture Malaria » MCM	203
IV. 1. d.	Décongélation de la souche et mise en culture	203
IV. 1. e.	Frottis sanguin.....	204
IV. 1. f.	Division	204
IV. 1. g.	Congélation	204
IV. 2.	Essais biologiques	204
IV. 2. a.	Screening des extraits.....	204

IV. 2. b. Réalisation des gammes de concentration des composés à tester	205
IV. 2. c. Réalisation de la suspension parasitaire	205
IV. 2. d. Quantification de la Plasmodium Lactate DeHydrogenase (pLDH).....	206
IV. 2. e. Détermination de la CI_{50}	208
IV. 2. f. Test d'inhibition de l'incorporation de l'hypoxanthine tritiée (collab. M. Kaiser, Swiss TPH)	208
IV. 2. g. Intérêt de combiner les deux techniques	208
V. Inhibition de la voie de synthèse des isoprénoides (MEP) ou de la prénylation des protéines.....	209
Bibliographie.....	213
Annexes	229
Annexe 1 : Spectres 1H , 1H - δ DOSY, 1H - 1H COSY, 1H - 1H TOCSY, 1H - ^{13}C HSQC et 1H - ^{13}C HMBC de la chloroquine.....	229
Annexe 2 : Fichiers « Report » et « Analysis » de la série CHLORO	232
Annexe 3 : Spectres 1H , 1H - δ DOSY, 1H - 1H COSY, 1H - 1H TOCSY, 1H - ^{13}C HSQC et 1H - ^{13}C HMBC de l'artémisinine	234
Annexe 4 : Fichiers « Report » et « Analysis » de la série ARTE	237
Annexe 5 : Spectres RMN de F2-3-3.....	240
Annexe 6 : BucketUtilities_LM.....	244
Annexe 7 : Publication	249

Liste des Abréviations

^{13}C : Carbone 13

^1H : Proton

AcOEt : Acétate d'éthyle

APCI : Ionisation chimique à pression atmosphérique

BRMS : Spectrométrie de masse basse résolution

CCM : Chromatographie sur couche mince

CI₅₀ : Concentration inhibitrice 50

COSY : *CORrelation SpectroscopY*

CTA : Combinaison thérapeutique à base d'artémisinine

DAD : Détecteur à barrette de diodes (*Diode-Array Detector*)

DCM : Dichlorométhane

DMAPP : Dimethylallyl pyrophosphate

DMSO : Diméthyl sulfoxyde

DO : Densité optique

DOSY : *Diffusion Ordered SpectroscopY*

ELISA : *Enzyme-Linked Immunosorbent Assay*

ESI : Ionisation par électronébulisation (*Electrospray*)

FAS II : *Fatty Acid Synthase II*

FDA : *Food and Drug Administration*

GC : Chromatographie en phase gazeuse

GFP : *Green Fluorescent Protein*

GNPS : *Global Natural Products Social Molecular Networking*

HMBC : *Heteronuclear multiple-bond correlation spectroscopy*

HPLC : Chromatographie liquide à haute performance (pression)

HPTLC : Chromatographie sur couche mince à haute performance

HRMS : Spectrométrie de masse haute résolution

HSQC : *Heteronuclear Single-Quantum Correlation spectroscopy*

IBMP : Institut de Biologie Moléculaire des Plantes

IGBMC : Institut de Génétique et de Biologie Moléculaire et Cellulaire

IPP : Isopentényl pyrophosphate
IPPTS : Institut de Parasitologie et de Pathologie Tropicale de Strasbourg
J : Constante de couplage
LC : Chromatographie Liquide
LIT : Laboratoire d'Innovation Thérapeutique
MCM : Milieu de Culture Malaria
MeOD : Méthanol deutéré
MeOH : Méthanol
MEP : méthylérythritol phosphate
MPLC : Chromatographie liquide à moyenne pression
MS : Spectrométrie de masse
MVA : Mévalonate
NOESY : *Nuclear Overhauser Effect Spectroscopy*
OS : Système d'exploitation
RFE : *Recursive Feature Elimination*
RMN : Résonance magnétique nucléaire
SI : Indice de sélectivité
SPE : Extraction sur phase solide
SPIKE : *Spectrometry Processing Innovative Kernel*
TLC : *Thin Layer Chromatography (CCM)*
TMS : Tetraméthylsilane
TOCSY : *Total Correlation Spectroscopy*
TOF : Temps de vol (*Time of flight*)
UHPLC : Chromatographie liquide ultra haute performance (pression)
UV : Ultraviolet

Liste des Figures

Figure 1 : Présentation de l'approche phylogénétique employée.....	4
Figure 2 : Pays où le paludisme est endémique en 2000 et 2016.....	10
Figure 3 : Cycle de développement de <i>P. falciparum</i> chez l'homme et le moustique anophèle.	12
Figure 4 : Les alcaloïdes du quinquina.....	14
Figure 5 : Famille des amino-alcools représentée par la méfloquine et la luméfántrine.	15
Figure 6 : Famille des 4-aminoquinoléines.	16
Figure 7 : Famille des 8-aminoquinoléines.	17
Figure 8 : Famille des antifolates.	17
Figure 9 : Famille des endopéroxydes.....	18
Figure 10 : Nouvelles classes d'antipaludiques.	19
Figure 11 : Nombre de composés publiés entre 2005 et 2008 montrant une forte activité ($CI_{50} \leq 2 \mu M$) ou une activité modérée ($CI_{50} \leq 11 \mu M$) <i>in vitro</i> contre différentes souches de <i>P. falciparum</i> , d'après J. Bero ¹³	22
Figure 12 : Sesquiterpènes, diterpènes et triterpènes actifs sur <i>P. falciparum in vitro</i>	23
Figure 13 : Organisation cellulaire d'un mérozoïte.	24
Figure 14 : Voies métaboliques de l'apicoplaste et cibles thérapeutiques.	26
Figure 15 : Antibiotiques évalués pour leur activité vis-à-vis de <i>Plasmodium</i>	27
Figure 16 : Inhibiteurs de la voie de biosynthèse de l'hème et des complexes Fer-Soufre.	28
Figure 17 : Inhibiteurs de la voie de biosynthèse des acides gras.....	29
Figure 18 : Voie de biosynthèse des isoprénoïdes et cibles thérapeutiques chez <i>P. falciparum</i> d'après A. M. Guggisberg <i>et al.</i>	30
Figure 19 : Inhibiteurs de la voie des isoprénoïdes.	31
Figure 20 : Arbre global des eucaryotes à partir d'un consensus des preuves phylogénétiques, de rares signatures génomiques et des caractéristiques morphologiques.....	32
Figure 211 : Classification taxonomique actuelle des algues rouges sur critère génétique (Classe des Florideophyceae).....	45
Figure 222 : Localisation de la zone de récolte des algues normandes.....	49
Figure 233 : Principe du test phénotypique permettant de mettre en évidence des inhibiteurs de la voie du MEP ou de la prénylation.	56

Figure 244 : Fractionnement bio-guidé.	69
Figure 255 : Stratégie d'analyse différentielle d'après F. Schroeder.	75
Figure 266 : Méthode d'analyse différentielle de F. Schroeder appliquée à <i>B. bifurcata</i>	77
Figure 277 : Analyse différentielle par comparaison visuelle des spectres ^1H - ^1H COSY des fractions 3 et 4 actives	78
Figure 288 : Chromatogramme HPLC-UV et MS de la fraction 3 de <i>B. bifurcata</i>	79
Figure 299 : Comparaison de l'empreinte pharmacophorique de <i>B. bifurcata</i> avec le spectre ^1H - ^1H COSY d'une molécule stockée sur SPE.	80
Figure 3030 : Méthode de déconvolution pharmacophorique complète.	81
Figure 311 : Principe des opérations de Plasmodesma.	86
Figure 322 : Représentation de la division de spectres RMN 1D et 2D en <i>buckets</i>	87
Figure 333 : Effet de la taille des <i>buckets</i> sur un spectre RMN ^1H . Spectre non <i>bucketé</i> , spectre <i>bucketé</i> à 0.01 ppm et spectre <i>bucketé</i> à 0.05 ppm.	88
Figure 344 : Représentation d'un spectre RMN ^1H entre 3.5 et 3.9 ppm et le spectre <i>bucketé</i> correspondant et un spectre ^1H - ^1H -TOCSY et le spectre <i>bucketé</i> correspondant.	88
Figure 355 : Histogrammes des données RMN COSY <i>bucketées</i> d'ARTE 3 sans nettoyage, avec fonction « nettoie », avec fonction « nettoie_mieux » et fonction « nettoie_encore_mieux »	90
Figure 366 : Effet des fonctions de nettoyage sur les spectres <i>bucketés</i>	91
Figure 377 : Spectre ^1H - ^1H -COSY <i>bucketé</i>	92
Figure 388 : Représentation schématique de la symétrisation des données RMN.	92
Figure 399 : Spectre ^1H - ^1H -COSY <i>bucketé</i> , nettoyé non symétrisé et symétrisé.	93
Figure 4040 : Echantillons de la série CHLORO et ARTE.	94
Figure 411 : Structure chimique de la chloroquine	95
Figure 422 : Spectre ^1H de la chloroquine	96
Figure 433 : Comparaison du spectre ^1H - ^1H COSY <i>bucketé</i> de l'extrait de <i>S. muticum</i> et de la chloroquine.	96
Figure 444 : Spectres ^1H - ^1H COSY <i>bucketés</i> des échantillons CHLORO 1 à 5 et zoom sur la partie aromatique du spectre	97
Figure 455 : Structure chimique de l'artémisinine	98
Figure 466 : Zoom du spectre ^1H de l'artémisinine entre 0 et 2.5 ppm.	98
Figure 477 : Soustractions des différentes variables des <i>buckets</i> des expériences ^1H - ^1H COSY	101
Figure 488 : Rapports des différentes variables des <i>buckets</i> des expériences ^1H - ^1H COSY.	103

Figure 499 : Production automatique d’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^1H COSY des échantillons ARTE 5 et ARTE 2 <i>versus</i> ARTE 1 (inactif).....	104
Figure 5050 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^1H TOCSY des échantillons CHLORO 5 et CHLORO 2 <i>versus</i> CHLORO 1 (inactif).....	106
Figure 511 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^1H TOCSY des échantillons ARTE 5 et ARTE 2 <i>versus</i> ARTE 1 (inactif).....	107
Figure 522 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^{13}C HSQC des échantillons CHLORO 5 et CHLORO 2 <i>versus</i> CHLORO 1 (inactif).	109
Figure 533 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^{13}C HSQC des échantillons ARTE 5 et ARTE 2 <i>versus</i> ARTE 1 (inactif).....	110
Figure 544 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^{13}C HMBC des échantillons CHLORO 5 et CHLORO 2 <i>versus</i> CHLORO 1 (inactif).	112
Figure 555 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des <i>buckets</i> des expériences ^1H - ^{13}C HMBC des échantillons ARTE 5 et ARTE 2 <i>versus</i> ARTE 1 (inactif).....	113
Figure 566 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des <i>buckets</i> des cinq échantillons CHLORO et les valeurs de concentration en chloroquine dans les fractions.	116
Figure 577 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des <i>buckets</i> des cinq échantillons CHLORO et les pourcentages d’inhibition.	116
Figure 588 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des <i>buckets</i> des échantillons CHLORO 1 , 2 , 3 et 4 et les concentrations en chloroquine.....	117
Figure 599 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des <i>buckets</i> des échantillons CHLORO 1 , 2 , 3 et 4 et les pourcentages d’inhibition.	117

Figure 6060 : Empreintes spectrales ^1H - ^1H COSY obtenues par régression linéaire entre les intensités des <i>buckets</i> des cinq échantillons ARTE et les concentrations en artémisinine ou les pourcentages d'inhibition de croissance.	118
Figure 611 : Empreintes spectrales ^1H - ^1H COSY de la chloroquine obtenues par RFE entre les intensités des <i>buckets</i> des cinq échantillons CHLORO et les concentrations en chloroquine ou les pourcentages d'inhibition de croissance	119
Figure 622 : Empreintes spectrales ^1H - ^1H COSY de la chloroquine obtenues par RFE entre les intensités des <i>buckets</i> des échantillons CHLORO 1 à 4 et les concentrations en chloroquine ou les pourcentages d'inhibition de croissance	120
Figure 633 : Empreintes spectrales ^1H - ^1H TOCSY de la chloroquine obtenues par régression linéaire entre les intensités des <i>buckets</i> des échantillons CHLORO et les concentrations en chloroquine ou les pourcentages d'inhibition de croissance	121
Figure 644 : Empreintes spectrales ^1H - ^1H TOCSY de l'artémisinine obtenues par régression linéaire entre les intensités des <i>buckets</i> des échantillons ARTE et les concentrations en artémisinine ou les pourcentages d'inhibition de croissance	122
Figure 655 : Empreintes spectrales ^1H - ^{13}C HSQC de la chloroquine obtenues par régression linéaire entre les intensités des <i>buckets</i> des échantillons CHLORO et les concentrations en chloroquine ou les pourcentages d'inhibition de croissance	123
Figure 666 : Empreintes spectrales ^1H - ^{13}C HSQC de l'artémisinine obtenues par régression linéaire entre les intensités des <i>buckets</i> des échantillons ARTE et les concentrations en artémisinine ou les pourcentages d'inhibition de croissance	124
Figure 677 : Empreintes spectrales ^1H - ^{13}C HMBC de la chloroquine obtenue par régression linéaire entre les intensités des <i>buckets</i> des échantillons CHLORO et les concentrations en chloroquine ou les pourcentages d'inhibition de croissance	125
Figure 688 : Empreintes spectrales ^1H - ^{13}C HMBC obtenues par régression linéaire entre les intensités des <i>buckets</i> des échantillons ARTE et les concentrations en artémisinine ou les pourcentages d'inhibition de croissance	126
Figure 699 : Spectre ^1H - δ DOSY de la chloroquine (MeOD, 700MHz).	127
Figure 7070 : Empreinte spectrale ^1H - δ DOSY obtenue par régression linéaire entre les intensités des <i>buckets</i> des échantillons CHLORO et les concentrations en chloroquine.....	128
Figure 711 : Empreinte spectrale ^1H - ^{13}C HSQC produite par filtration de l'empreinte HSQC obtenue par régression linéaire par l'empreinte ^1H - δ DOSY.....	129
Figure 722 : Schéma des deux premières étapes de la méthode de déconvolution pharmacophorique appliquées à l'extrait SMARTE..	136

Figure 733 : Empreintes spectrales ^1H - ^1H TOCSY obtenues par rapport des écarts-types des <i>buckets</i> entre la fraction active F2 <i>versus</i> la fraction inactive F1 et entre la fraction active F3 <i>versus</i> la fraction inactive F4 de la série SMARTE	137
Figure 744 : Empreintes spectrales ^1H - ^{13}C HSQC obtenues par le rapport des écarts-types des <i>buckets</i> entre la fraction active F2 <i>versus</i> la fraction inactive F1 et entre la fraction active F3 <i>versus</i> la fraction inactive F4 de la série SMARTE	138
Figure 755 : Empreinte spectrale ^1H - ^1H TOCSY obtenue par régression linéaire entre l'intensité des <i>buckets</i> et le pourcentage d'inhibition des fractions à 10 $\mu\text{g}/\text{mL}$	139
Figure 766 : Empreinte spectrale ^1H - ^{13}C HSQC obtenue par régression linéaire en fonction de l'intensité des <i>buckets</i> et le pourcentage d'inhibition des fractions à 10 $\mu\text{g}/\text{mL}$	140
Figure 777 : Spectre ^1H - δ DOSY de l'artémisinine (MeOD, 700 MHz).	140
Figure 788 : Empreinte spectrale ^1H - δ DOSY obtenue par régression linéaire ou RFE entre les intensités des <i>buckets</i> des échantillons SMARTE et les pourcentage d'inhibition des fractions.	141
Figure 799 : Filtration de l'empreinte ^1H - ^{13}C HSQC obtenue par régression linéaire, par l'empreinte ^1H - δ DOSY obtenue par RFE.....	142
Figure 8080 : Schéma des deux premières étapes de la méthode de déconvolution appliquées à l'extrait AcOEt de <i>B. bifurcata</i>	143
Figure 811 : Comparaison des différentes empreintes ^1H - ^1H TOCSY produites par le rapport des écarts-types des <i>buckets</i> des fractions 2 sur 1 et des fractions 4 sur 5 et par régression linéaire	144
Figure 822 : Comparaison des différentes empreintes ^1H - ^{13}C HSQC produites par le rapport des écarts-types des <i>buckets</i> des fraction 2 sur 1 et des fraction 4 sur 5 et par régression linéaire.	145
Figure 833 : Comparaison du spectre ^1H - ^1H COSY de la molécule stockée sur SPE et de l'empreinte spectrale ^1H - ^1H TOCSY obtenue par régression linéaire entre les intensités des <i>buckets</i> et le pourcentage d'inhibition de fractions	146
Figure 844 : Spectre ^1H - ^{13}C HSQC de la molécule stockée sur SPE et de l'empreinte spectrale HSQC obtenue par régression linéaire entre les intensités des <i>buckets</i> et le pourcentage d'inhibition de fractions	147
Figure 855 : Structures chimiques de la quinine, de la quinidine, de la cinchonine et de la cinchonidine.	149

Figure 866 : Empreinte spectrale ^1H - ^1H TOCSY obtenue par régression linéaire des intensités des <i>buckets</i> des données ^1H - ^1H TOCSY des 6 fractions de quinquina et de leurs activités antipaludiques.....	151
Figure 877 : Comparaison de l’empreinte ^1H - ^1H TOCSY et des spectres des 4 alcaloïdes quinoléiques majoritaires.	152
Figure 888 : Empreinte spectrale ^1H - ^{13}C HSQC obtenue par régression linéaire des données RMN HSQC des 6 fractions de quinquina en fonction de leurs activités antipaludiques.....	153
Figure 899 : Fractionnement par MPLC et évaluation biologique de l'extrait de <i>D. carnos</i> a.....	154
Figure 9090 : Empreintes spectrales obtenues par le rapport des écarts-types des <i>buckets</i> des analyses ^1H - ^1H TOCSY des fractions 2 sur 1 et des fractions 4 sur 5, et par régression linéaire	155
Figure 911 : Empreintes spectrales obtenues par le rapport des écarts-types des <i>buckets</i> des analyses HSQC des fractions 2 sur 1 et des fractions 4 sur 5, et par régression linéaire	156
Figure 922 : Fractionnement de l’extrait AcOEt de <i>Dilsea carnos</i> a.....	158
Figure 933 : Comparaison des spectres ^1H - ^1H COSY et ^1H - ^{13}C HSQC de F2-3-3 à l’empreinte pharmacophorique produite par régression linéaire.	159
Figure 944 : Fractionnement et évaluation biologique de l'extrait de <i>G. gracilis</i>	160
Figure 955 : Comparaison des empreintes spectrales ^1H - ^1H TOCSY de <i>G. gracilis</i>	161
Figure 966 : Comparaison des empreintes spectrales ^1H - ^{13}C HSQC de <i>G. gracilis</i>	162
Figure 977 : Comparaison des empreintes ^1H - ^{13}C HSQC et ^1H - ^1H TOCSY obtenues par régression linéaire de <i>D. carnos</i> a et de <i>G. gracilis</i>	163
Figure 98 : Comparaison des empreintes ^1H - ^{13}C HSQC et ^1H - ^1H TOCSY obtenues par régression linéaire de <i>D. carnos</i> a, <i>G. gracilis</i> et <i>B. bifurcata</i>	165
Figure 999 : Gradient des solvants de la méthode MPLC pour le fractionnement des extraits algaux.	191
Figure 100100 : Principe des opérations de Plasmodesma.	194
Figure 1011 : Test d’activité anti- <i>Plasmodium</i> par la mesure de la sécrétion de pLDH via le kit ELISA Antigen ApDia®.....	207

Liste des Tableaux

Tableau 1 : Algues rouges marines évaluées <i>in vitro</i> contre différentes souches de <i>P. falciparum</i>	35
Tableau 2 : Identité botanique des échantillons algaux étudiés.	47
Tableau 3 : Données bibliographiques recueillies sur les espèces d'algues rouges fournies... ..	50
Tableau 4 : Rendements d'extraction obtenus.	51
Tableau 5 : Activités antipaludiques et cytotoxiques <i>in vitro</i> des différents extraits obtenus. ..	53
Tableau 6 : Activités antipaludiques <i>in vitro</i> des fractions obtenues par MPLC à partir des extraits actifs.	54
Tableau 7 : Inhibition de la voie du MEP dans un modèle végétal (cellules BY-2 de tabac transformées) par les extraits et fractions les plus actives vis-à-vis de <i>P. falciparum</i>	57
Tableau 8 : Expériences communes de spectroscopie RMN 2D.	68
Tableau 9 : Déplacements chimiques ^1H et ^{13}C de la chloroquine (MeOD, 700MHz)	95
Tableau 10 : Déplacements chimiques ^1H et ^{13}C de l'artémisinine (700 MHz, MeOD)	99
Tableau 11 : Déplacements chimiques des quatre alcaloïdes majoritaires, la quinine, la quinidine, la cinchonine et la cinchonidine (700MHz, MeOD).	150
Tableau 12 : Activités antipaludiques des fractions issues de l'extrait hydroalcoolique de quinquina.	151
Tableau 13 : Avantages et inconvénients des méthodes actuelles les plus utilisées pour l'identification de produits naturels dans des mélanges complexes.	180
Tableau 14 : Tableau des solvants deutérés de référence.....	189
Tableau 15 : Méthode MPLC de l'étape de fractionnement des extraits algaux.	190
Tableau 16 : Méthode MPLC de l'étape de fractionnement de l'extrait de quinquina.....	191



Contexte et Objectifs de la Thèse

Contexte et objectifs de la thèse

Le laboratoire d'Innovation Thérapeutique (LIT) hébergé à la Faculté de Strasbourg se donne pour missions principales **la conception, la synthèse et la caractérisation de molécules bioactives afin de mieux comprendre le vivant**. L'équipe de « Chimie-Biologie Intégrative et Pharmacognosie » dirigée par Dr. Dominique Bonnet a pour but de découvrir rapidement des petites molécules permettant de comprendre le Vivant et de le « manipuler » dans une perspective cognitive ou/et thérapeutique. Dans cette thématique, s'inscrit le groupe spécialisé dans les molécules naturelles anti-protazoaires, dirigé par le Dr C. Vonthron-Sénécheau.

Le groupe s'appuie sur différentes approches rationnelles pour puiser dans la biodiversité végétale afin de découvrir de nouveaux châssis moléculaires doués d'activités antiprotazoaires (*Plasmodium*, *Leishmania*) :

- deux approches « traditionnelles », usuelles en pharmacognosie, à savoir : i) l'approche ethnopharmacologique, qui s'appuie sur la connaissance des usages traditionnels des plantes médicinales ; ii) l'approche chimiotaxonomique, qui s'appuie sur le rapport entre la composition chimique des végétaux et leur classification systématique (taxonomie) ;
- et plus récemment, une approche inédite, **l'approche phylogénétique**, qui s'appuie sur un lien de phylogénie entre le phylum des Apicomplexes, auquel appartient *Plasmodium*, et les algues rouges (Rhodophytées).

C'est sur cette approche qu'a porté mon travail de thèse.

En effet, des travaux préliminaires du Dr. Catherine Vonthron réalisés à l'Université de Caen-Normandie suggèrent que les algues rouges puissent contenir des métabolites susceptibles de perturber le développement de *Plasmodium falciparum*, parasite responsable des formes les plus dangereuses de paludisme.

Le paludisme est l'un des axes principaux du groupe car malgré des progrès importants ces dix dernières années, le paludisme reste la première endémie mondiale, en particulier dans les zones les plus pauvres du globe. L'efficacité des traitements disponibles

s'amenuise à cause de l'apparition de résistances, y compris contre les associations à base d'artémisinine (CTA), médicaments les plus efficaces actuellement. Il est donc urgent de développer de nouvelles associations intégrant des molécules inédites présentant une activité conservée vis-à-vis de ces souches et dans l'idéal, présentant un mécanisme d'action original, afin de retarder l'apparition de résistances.

C'est précisément l'apicoplaste qui est visé, un organe relique de *Plasmodium* hérité d'une algue rouge ancestrale. En effet, l'apicoplaste exprime notamment des voies de biosynthèse essentielles à la survie du parasite. Notre hypothèse de travail est que les algues rouges actuelles puissent contenir des métabolites susceptibles d'interférer avec les voies métaboliques apicoplastiques de *Plasmodium* par proximité structurale (Figure 1).

Dans ce contexte, les objectifs de ce projet de thèse sont :

- la confirmation des activités anti-*Plasmodium* des extraits d'algues rouges à partir d'un grand nombre d'espèces ;
- le développement d'une nouvelle méthode analytique adaptée et basée sur la RMN permettant d'accélérer le processus de découverte de molécules antipaludiques à partir de ces espèces d'algues rouges ;
- l'identification de molécules algales responsables de l'activité antipaludique.

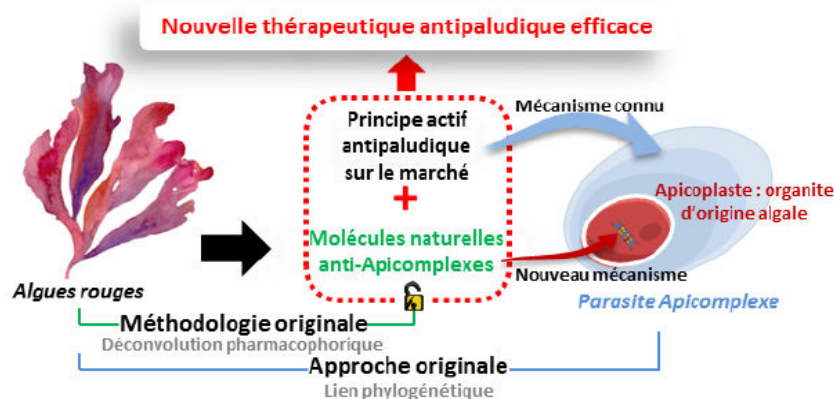


Figure 1 : Présentation de l'approche phylogénétique employée. Cette approche s'appuie sur un lien de phylogénie entre le phylum des Apicomplexes, auquel appartient *Plasmodium*, et les algues rouges. Celles-ci seraient susceptibles de contenir des métabolites capables d'interférer avec les voies métaboliques apicoplastiques de *Plasmodium* par proximité structurale. Pour accélérer le processus d'identification des composés antipaludiques, une méthode de déconvolution pharmacophorique a été développée.

Le premier chapitre se propose de définir le contexte de cette étude par une présentation générale du paludisme et de ses traitements actuels et en développement. Nous y découvrirons également l'apicoplaste et ses voies métaboliques qui représentent une cible d'intérêt pour le développement de nouveaux traitements. Puis, des données bibliographiques relatives aux métabolites des algues rouges et de leurs composés anti-*Plasmodium* seront apportées.

Le second chapitre est consacré à l'évaluation antipaludique de nouveaux extraits d'algues rouges ainsi que de leur évaluation à inhiber la voie de biosynthèse des isoprénoïdes dans un modèle végétal, voie métabolique également exprimée dans l'apicoplaste.

Le troisième chapitre porte sur le développement de la méthode de déconvolution pharmacophorique. Il précise les méthodes actuelles d'identification de composés naturels au sein de mélanges complexes puis il développe l'automatisation de l'analyse différentielle de spectres RMN et son application à des échantillons reconstitués.

Le quatrième chapitre présente l'application de la méthode de déconvolution pharmacophorique à des échantillons réels et plus complexes. Elle est notamment appliquée aux extraits algaux afin d'identifier les métabolites porteurs de l'activité antipaludique et de répondre à l'hypothèse de recherche.

Enfin, une discussion conduira aux conclusions quant aux réponses apportées par ce travail à l'hypothèse de départ.



CHAPITRE I

Chapitre I : Traitement du paludisme et stratégie de recherche

Ce chapitre porte sur la présentation du paludisme, ses enjeux, les traitements actuels et en développement, ainsi que les molécules naturelles antipaludiques les plus importantes.

L'intérêt de rechercher de nouvelles molécules actives aux cibles originales sera mis en avant, notamment contre un organe relictuel, l'apicoplaste. Ses voies métaboliques ainsi que les molécules connues les ciblant seront présentées.

Enfin, la relation évolutive entre l'apicoplaste et les algues rouges formulera l'hypothèse de recherche. Une présentation générale des algues rouges conclura ce chapitre.

I. Le paludisme

Le paludisme est une maladie parasitaire causée par un protozoaire du genre *Plasmodium*. Ce genre compte plus de 100 espèces capables d'infecter les reptiles, les oiseaux et divers mammifères. Cinq espèces sont connues pour infecter l'humain : *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* et *P. knowlesi*. Ces parasites sont transmis par la piqûre d'un moustique du genre *Anopheles*. *P. falciparum* et *P. vivax*, les plus répandus et dangereux, sont présents dans l'ensemble des zones tropicales et intertropicales. *P. falciparum* représente à lui seul plus de 90% des cas de paludisme en Afrique. La population des enfants de moins de 5 ans est la plus exposée, représentant près de 80 % des décès liés au paludisme¹ (Figure 2).

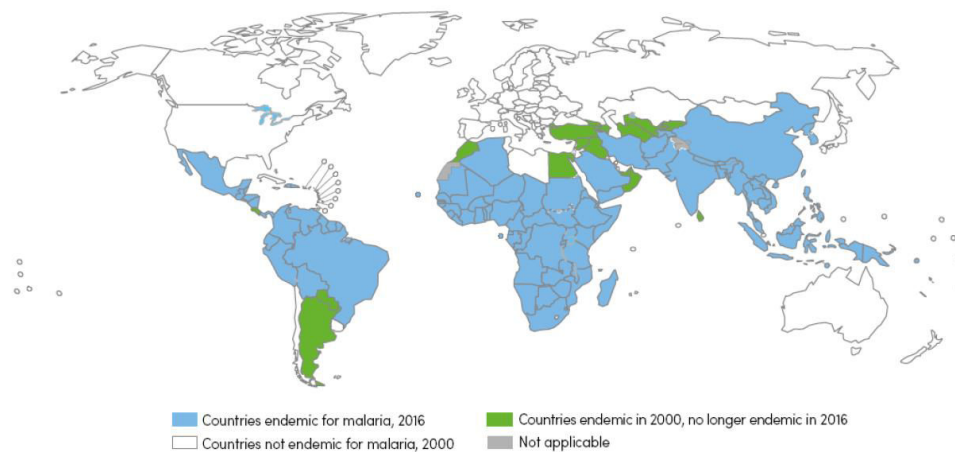


Figure 2 : Pays où le paludisme est endémique en 2000 et 2016. Les pays ne présentant pas de cas autochtone pendant trois années consécutives sont considérés comme ayant éliminé le paludisme. Source : World Malaria Report 2016, WHO.

Les dernières estimations font état de 216 millions de cas de paludisme pour l'année 2016, responsables du décès de près de 445 000 personnes à travers le monde. Malgré ces chiffres toujours dramatiques, le nombre de cas et les taux de mortalité ont diminué respectivement de 30 % et 47 % depuis l'an 2000 grâce au programme global de lutte contre le paludisme mené par l'OMS avec l'aide des ministères de la santé des zones d'endémie et un nombre important d'autres partenaires (Fonds Mondial, ONG, Fondations, ...). De nombreux efforts restent toutefois à effectuer afin de diminuer de façon importante les chiffres toujours alarmants liés à cette pathologie¹.

I. 1. Cycle de développement de *Plasmodium*

Le parasite présente un cycle hétéroxène car il est hébergé par deux hôtes différents : un hôte définitif (chez qui se déroule la reproduction sexuée ou gamogonie) représenté par un moustique femelle du genre *Anopheles* et un hôte intermédiaire vertébré, chez qui se déroule la reproduction asexuée (schizogonie) (Figure 3).

L'infestation se fait par inoculation du parasite sous sa forme **sporozoïte** par l'anophèle femelle lors d'un repas sanguin. Après un premier séjour très court dans la circulation générale, les sporozoïtes infestent les hépatocytes où ils se multiplient de manière asexuée, par schizogonie, donnant naissance à un **schizonte** plurinucléé dans le cytoplasme de l'hépatocyte. Cette phase hépatocytaire dure 1 à 3 semaines. Puis, les schizontes provoquent la lyse des hépatocytes, libérant de nombreux **mérozoïtes** dans la circulation générale qui envahissent les hématies grâce à leur complexe apical. Ils sont alors internalisés dans des vacuoles parasitophores, on parle de **trophozoïtes**. Ceux-ci se multiplient de manière asexuée pour donner des schizontes qui libèrent des mérozoïtes, qui vont infester d'autres hématies, après maturation, et permettre à un autre cycle érythrocytaire de redémarrer. La lyse massive de nombreuses hématies entraîne une anémie chez la personne parasitée. Dans les hématies, le parasite se nourrit d'hémoglobine et rejette l'hème sous forme d'hémozoïne, un pigment pyrogène responsable des accès de fièvre caractéristiques de la maladie

Certains trophozoïtes deviennent des cellules sexuées, les gamétocytes mâles (microgamétocytes) ou femelles (macrogamétocytes), qui peuvent uniquement se développer chez l'anophèle. Ce sont les gamétocytes qui constituent la forme infestante pour le moustique. Au cours de son repas sanguin, l'anophèle femelle ingère des gamétocytes qui vont maturer dans son estomac. Ceux-ci donnent naissance à des gamètes mâles qui vont féconder les gamètes femelles engendrant un œuf mobile, l'ookinète. Après avoir traversé la paroi du mésentéron, celui-ci s'enkyste pour former un oocyste. L'oocyste subit une sporogonie et libère des sporozoïtes qui vont migrer vers les glandes salivaires et pourront contaminer un autre hôte lors du repas sanguin suivant².

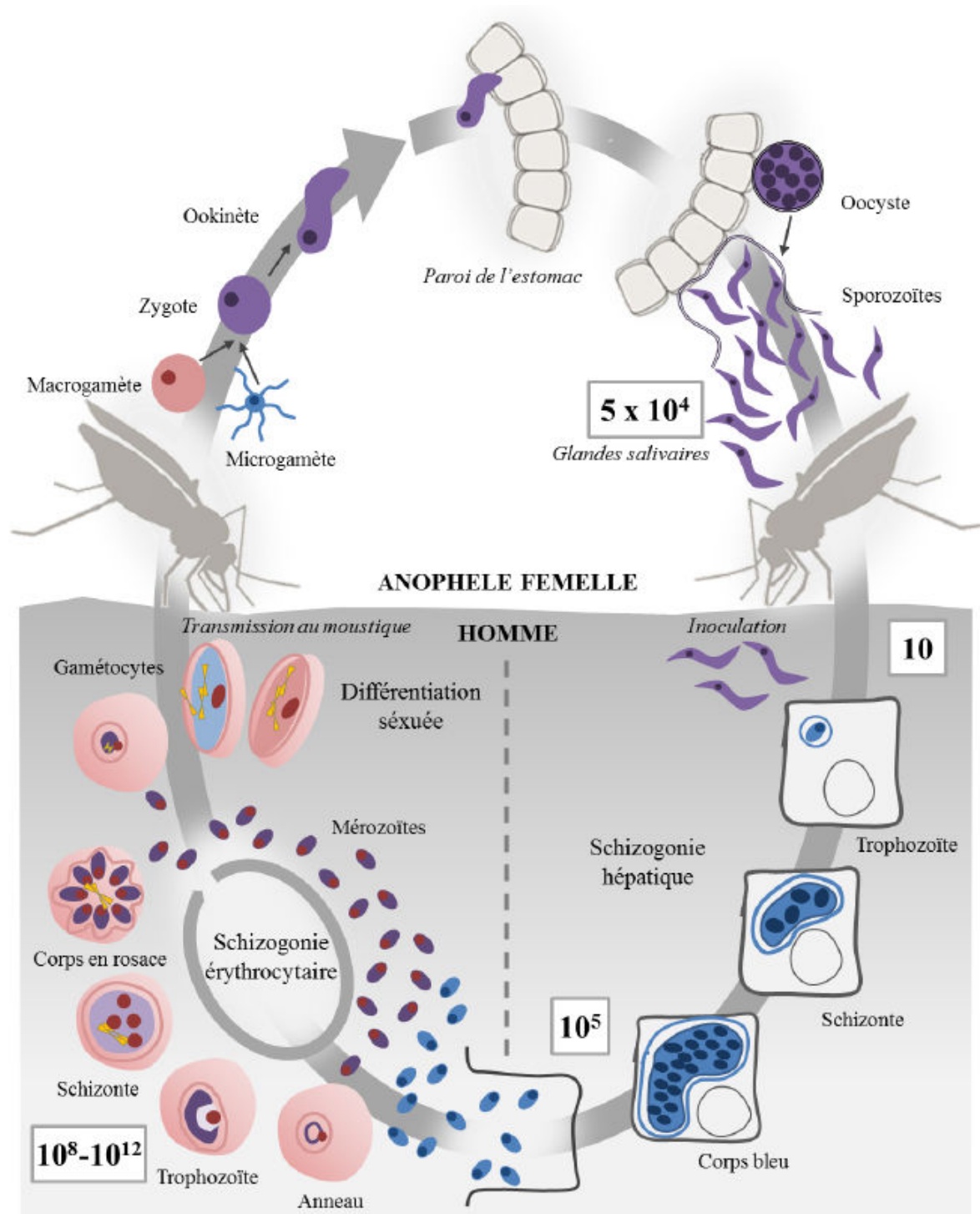


Figure 3 : Cycle de développement de *P. falciparum* chez l'homme et le moustique anophèle. Le cycle débute par l'inoculation de sporozoïtes mobiles dans le derme qui vont ensuite rejoindre le foie. Chaque sporozoïte envahit un hépatocyte et s'y multiplie. Après environ une semaine passée dans le foie, les schizontes (ou corps bleus) éclatent, libérant ainsi dans la circulation sanguine des milliers de mérozoïtes qui vont envahir les hématies et entamer ainsi le cycle érythrocytaire. Les symptômes apparaissent quand le nombre total de parasites en circulation atteint environ 100 millions. Certains parasites vont se différencier en formes sexuées (gamétocytes) qui formeront des gamètes dans l'estomac du moustique après un nouveau repas sanguin. La fusion des gamètes va donner lieu à un ookinète mobile qui va traverser la paroi de l'estomac et se différencier en oocyste, sujet à une intense division. L'éclatement de l'oocyste va libérer des sporozoïtes qui vont migrer vers les glandes salivaires du moustique et attendre le prochain repas sanguin du moustique femelle. Le cycle complet dure environ un mois. Le nombre estimé de parasites impliqués à chaque étape du cycle est indiqué dans les rectangles blancs. Une charge parasitaire totale d'environ 10^{12} parasites correspond approximativement à une parasitémie de 2% chez l'adulte. Copyright Flore Nardella, inspiré d'après le cycle ANOFEL (<http://campus.cerimes.fr/parasitologie/enseignement/paludisme/site/html/1.html>)

I. 2. Traitements antipaludiques actuels et en développement

Les alcaloïdes de quinquina (*Cinchona pubescens* ou *C. succirubra*, Rubiacées) représentent le traitement historique du paludisme en Europe. L'écorce de quinquina rouge était utilisée traditionnellement dans les Andes pour soigner les fièvres et a été importée en Europe par les missionnaires jésuites à la fin du 16^{ème} siècle. La quinine et la quinidine, ont été isolées pour la première fois en 1817 par Pelletier et Caventou. En 1940, l'Allemagne nazie détruit les stocks de quinine lors du bombardement d'Amsterdam, ce qui laisse les troupes américaines et australiennes évoluant en Asie sans traitement antipaludique.

L'effort de guerre mènera à la synthèse de la chloroquine par les américains, s'inspirant de la quinine et du bleu de méthylène dont l'activité avait été démontrée par les chercheurs allemands Guttman et Ehrlich. La chloroquine supplantera rapidement la quinine, son coût de fabrication étant très faible par rapport à l'extraction du produit naturel. La chloroquine a laissé entrevoir la possibilité d'éradiquer le paludisme et a mené au premier programme mondial d'éradication du paludisme de l'OMS en 1955. Cependant, l'émergence de résistance en Asie du Sud-Est à la fin des années 1950 et sa diffusion en Afrique à la fin des années 1970 ont mis fin aux efforts en place. La guerre froide, et notamment la guerre du Vietnam, a été un nouvel élan dans la recherche de nouveaux médicaments antipaludiques, tant sur le front américain (le *Walter Reed Institute* ayant criblé plus de 250 000 composés pour n'en sélectionner qu'un : la méfloquine - composé n°142490), que sur le front asiatique, le criblage d'extraits utilisés en médecine traditionnelle chinoise par l'équipe du Dr Tu Youyou (prix Nobel de médecine en 2015) ayant mené à l'isolement de l'artémisinine.

L'élimination du paludisme dans la plupart des pays du nord et la fin des conflits en Asie ont mené au désintéressement des industries pharmaceutiques pour la lutte contre le paludisme. Un nouvel essor a vu le jour avec la mise en place du second plan de lutte de l'OMS en 2000. Le réinvestissement financier dû à la création de l'Initiative Présidentielle américaine contre le Paludisme (PMI) correspondant à une expansion historique des ressources du gouvernement américain pour la lutte contre le paludisme, à l'implication de la banque mondiale et de fondations comme la *Bill and Melinda Gates Foundation*, ainsi qu'à la mise en place de partenariats public-privé solides ont permis le développement rapide de nouveaux antipaludiques ces dernières années.

1.2.a. Alcaloïdes du quinquina

La quinine a longtemps été utilisée en monothérapie et est encore utilisée dans le traitement du paludisme grave, même si l’OMS recommande préférentiellement l’artésunate. Les autres alcaloïdes du quinquina, à savoir la quinidine, la cinchonine et la cinchonidine possèdent une activité comparable à la quinine mais ont été moins utilisés car l’écorce de quinquina cultivé (*C. ledgeriana*) contient la quinine de façon majoritaire (Figure 4).

La structure de ces alcaloïdes quinoléiques contient une fonction amino-alcool aromatique dont est inspirée la méfloquine. Ces composés sont très basiques et sont toujours administrés sous forme de sels dont certains possèdent des propriétés nécosantes.

Le mécanisme d’action n’est pas exactement connu, bien qu’il soit considéré comme le même que celui de la chloroquine (Cf. Partie I. 2. c)

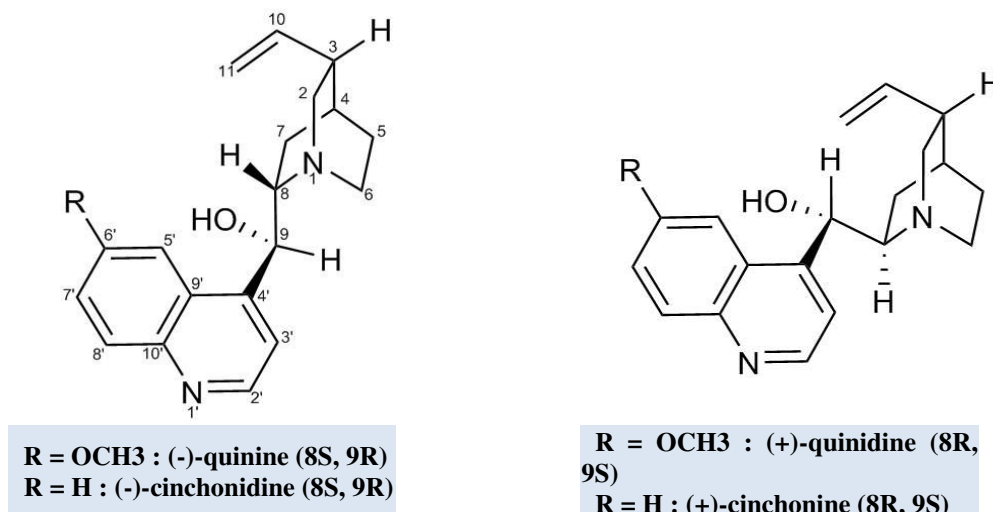


Figure 4 : Les alcaloïdes du quinquina. Ces quatre molécules sont les alcaloïdes quinoléiques majoritaires du quinquina. La quinine a été le premier médicament utilisé pour le traitement du paludisme.

1.2.b. Amino-alcools

La famille des amino-alcools comprend des composés présentant une structure inspirée de celle de la quinine. Cette classe comprend la méfloquine et la luméfantine. Par analogie structurale, leur mécanisme d’action est considéré comme impliquant la polymérisation de l’hème. Cependant, une étude indique que la méfloquine inhiberait également l’endocytose (transport des molécules vers l’intérieur de la cellule) de l’hémoglobine et d’autres substances du cytoplasme de l’érythrocyte vers la vacuole digestive (accumulation d’hémoglobine indigeste dans le parasite). Le mécanisme d’action, de la luméfantine, quant à lui, n’a pas été évalué³ (Figure 5).

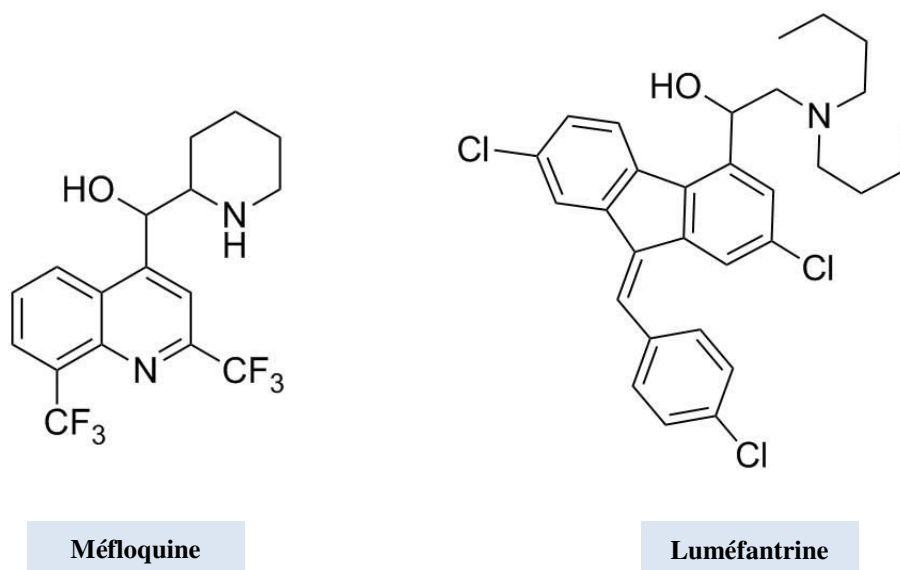


Figure 5 : Famille des amino-alcools représentée par la méfloquine et la luméfantrine. Ces molécules sont inspirées des alcaloïdes quinoléiques du quinquina.

I. 2. c. 4-aminoquinoléines

La chloroquine est le médicament chef-de-file de la classe des 4-aminoquinoléines. Son efficacité est sérieusement impactée depuis l'émergence de résistances dans la quasi-totalité des régions où *P. falciparum* est prévalent : son utilisation pour le traitement de *P. falciparum* est aujourd'hui limitée à l'Amérique Centrale. La synthèse de dérivés a permis de contourner, en partie, les résistances à la chloroquine. L'amodiaquine, la pipéraquline et depuis peu la pyronaridine, sont associées aux dérivés de l'artémisinine dans les CTAs (associations thérapeutiques à base d'artémisinine). La ferroquine et l'AQ-13 sont actuellement en développement clinique en phase II (associés à l'artefenomel) et la naphtoquine en phase III (associée avec l'arthéméther/luméfantrine). La chloroquine inhibe la détoxification de l'hème en hémotoïne lors de la digestion de l'hémoglobine par le parasite. Les autres dérivés semblent avoir le même mécanisme d'action. La ferroquine est le premier dérivé organométallique en essai clinique contre le paludisme. Son mode d'action paraît multiple : il semblerait cibler les lipides, inhiber la formation de l'hémotoïne et générer des espèces réactives de l'oxygène⁴ (Figure 6).

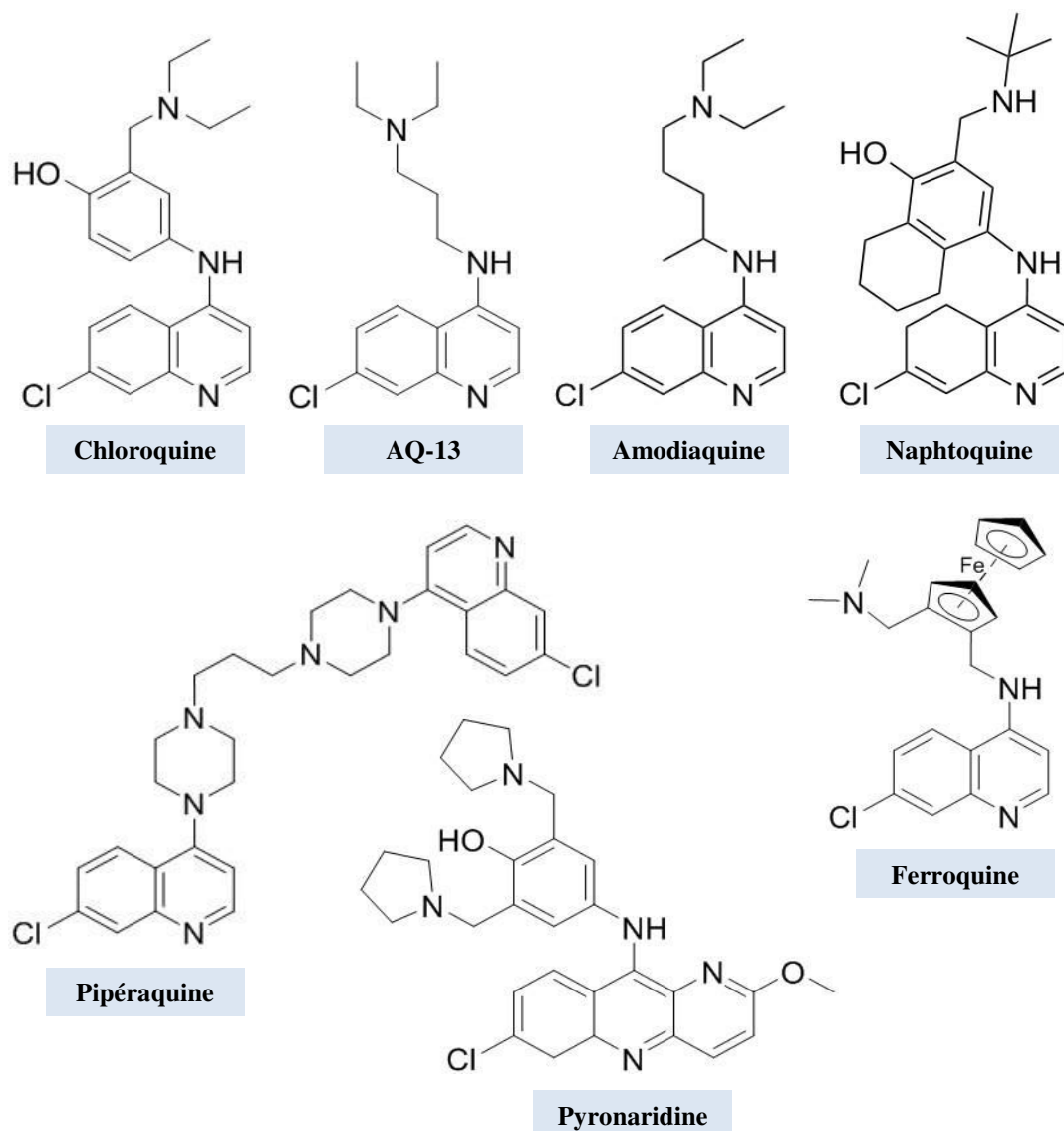


Figure 6 : Famille des 4-aminoquinoléines. Ces molécules sont dérivées des alcaloïdes quinoléiques. Elles sont actuellement utilisées ou en évaluation clinique pour le traitement du paludisme.

1.2.d. 8-aminoquinoléines

Les 8-aminoquinoléines sont une classe importante dans l'arsenal antipaludique car elles sont efficaces contre *P. vivax* et *P. ovale*. Le mécanisme d'action de la primaquine n'est pas connu mais il pourrait impliquer la mitochondrie. La tafénoquine a été approuvée récemment par la *US Foods and Drug Administration* (FDA)⁵ (Figure 7).

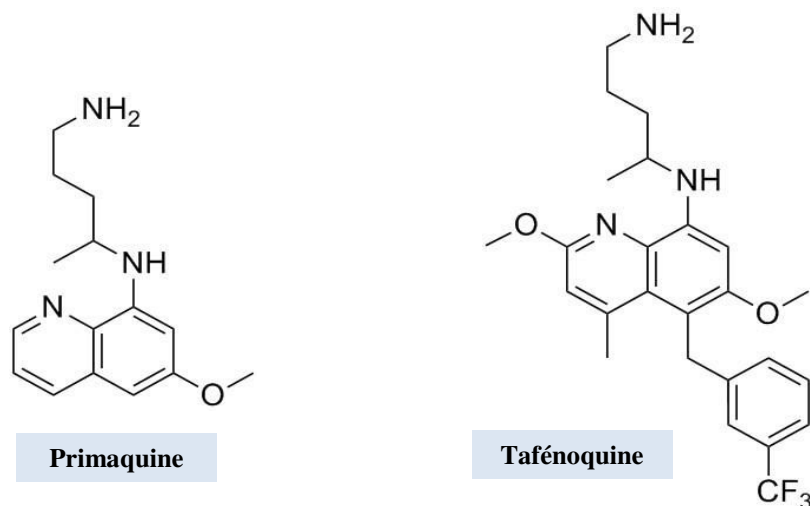


Figure 7 : Famille des 8-aminoquinolines. La primaquine et la tafénoquine sont actuellement utilisées pour le traitement du paludisme.

I. 2. e. Antifolates

Les antifolates inhibent des enzymes de la voie de synthèse de l'acide folique. Cette voie métabolique est essentielle, jouant un rôle dans la synthèse de l'ADN du parasite.

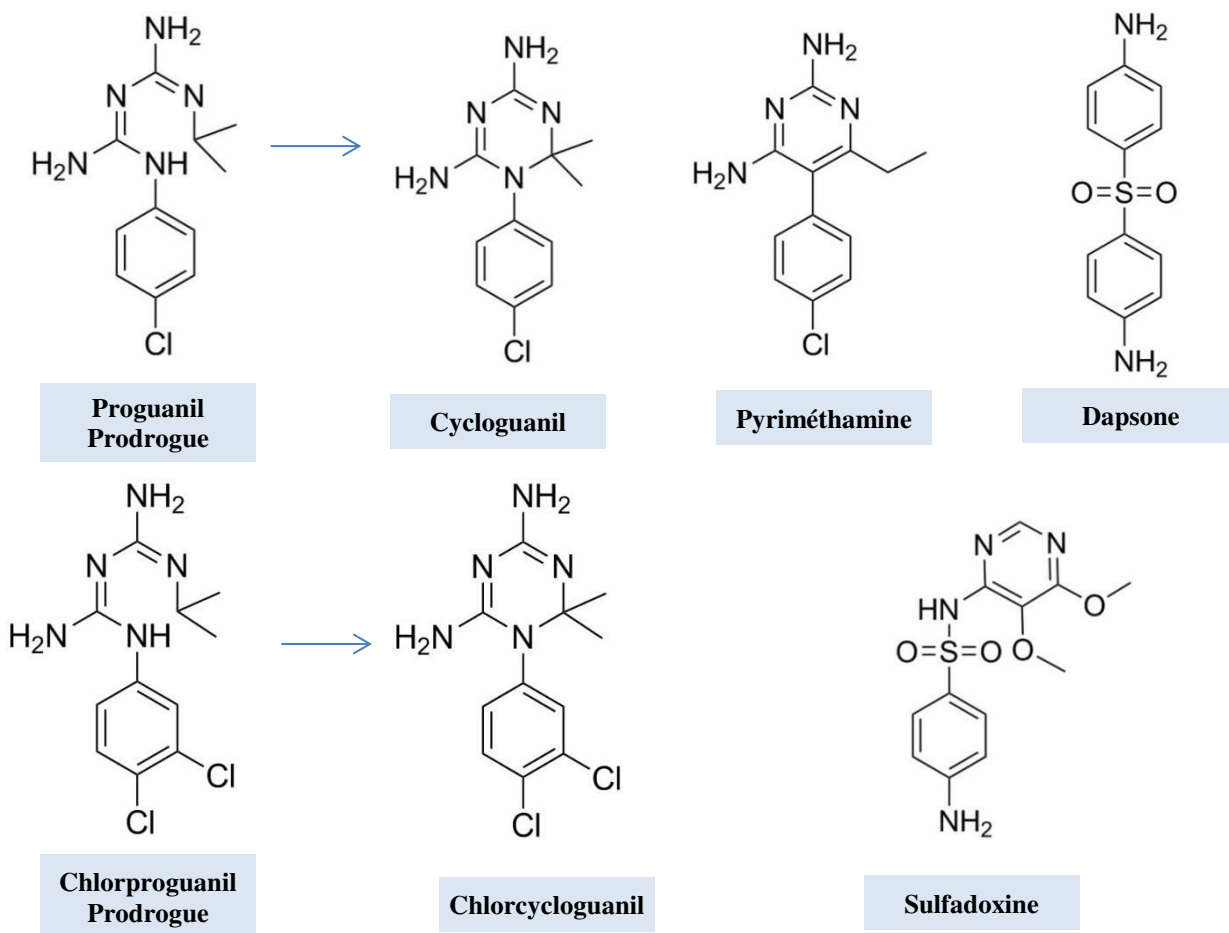


Figure 8 : Famille des antifolates. Ces composés sont utilisés en association avec l'atovaquone, la sulfadoxine ou l'artésunate pour le traitement du paludisme.

Ces inhibiteurs sont utilisés en association avec l'atovaquone, la sulfadoxine ou l'artésunate. Des résistances sont connues pour ces composés (Figure 8).

1.2.f. Endopéroxydes

L'artémisinine est une molécule d'origine naturelle isolée d'*Artemisia annua* (Astéracées), plante utilisée traditionnellement contre les fièvres en Chine. Depuis, des dérivés hémisynthétiques ont été réalisés afin d'optimiser leur absorption, mais le métabolite actif, la dihydroartémisinine, est commun à tous ces dérivés. C'est aujourd'hui la classe thérapeutique la plus utilisée et ces dérivés jouent le rôle de médicament d'action rapide dans les associations à base d'artémisinine (CTAs) car ils permettent de réduire la charge parasitaire d'un facteur 10 000 toutes les 48 heures. Les dérivés de l'artémisinine sont des lactones sesquiterpéniques qui comportent un pont endopéroxyde, déterminant majeur de leur activité (Figure 9).

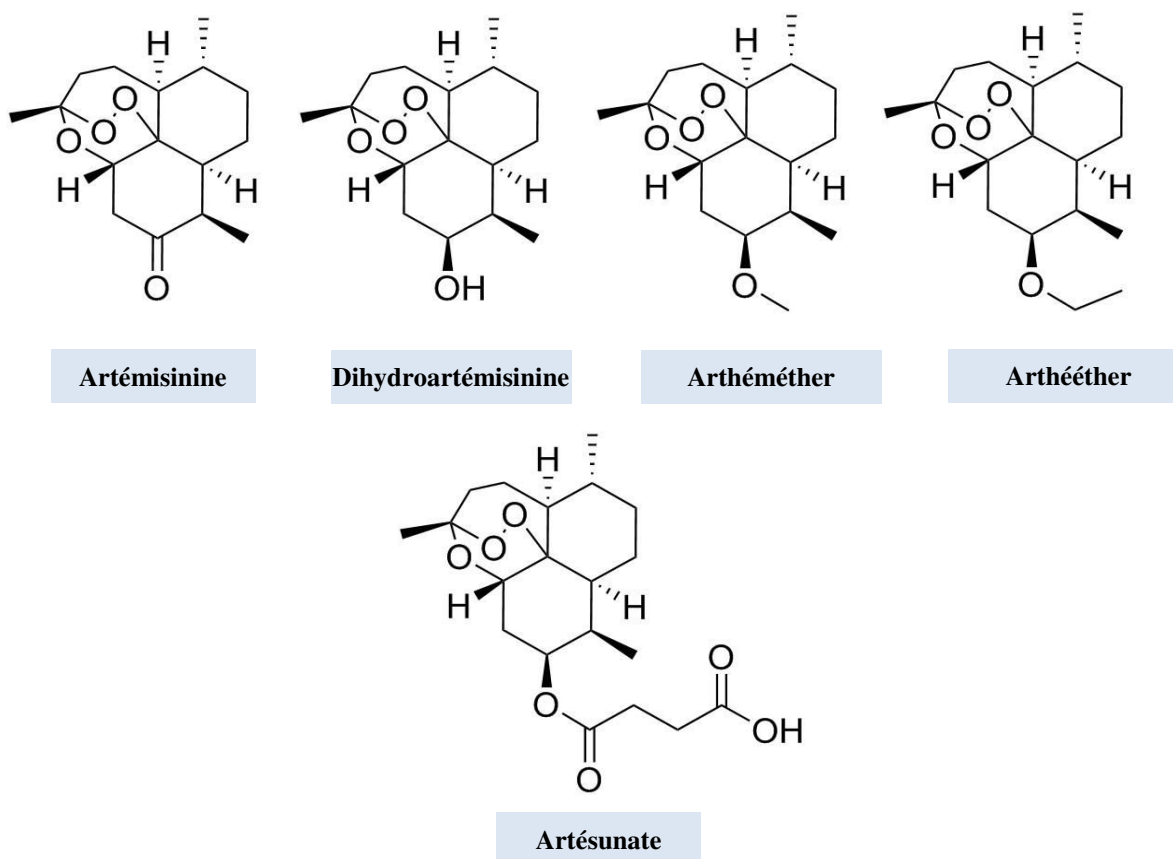


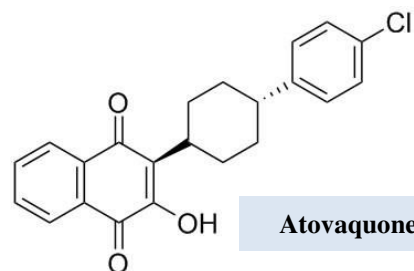
Figure 9 : Famille des endopéroxydes. C'est la classe thérapeutique la plus utilisée pour le traitement du paludisme. Ces composés ont un mode d'action rapide et sont utilisés en association avec d'autres familles d'antipaludiques.

Leur mécanisme d'action n'est pas encore compris mais ferait intervenir l'hème dont le fer catalyse l'ouverture du pont endopéroxyde qui mène à la génération d'espèces radicalaires et induit un stress alkylant pour le parasite.

Des résistances ont été mises en évidence au Cambodge en 2008 et ont été décrites régulièrement depuis en Asie du sud-est, avec notamment de manière plus inquiétante des cas de multirésistances aux CTAs⁶.

1.2.g. Naphtoquinones

L'atovaquone est le seul représentant de cette classe. Il est utilisé en association avec le proguanil. Sa cible est la chaîne respiratoire de la mitochondrie car il inhibe le transport d'électrons et donc la synthèse d'ATP.



1.2.h. Nouvelles classes d'antipaludiques

Quatre autres nouvelles classes d'antipaludiques sont en phase II d'études cliniques.

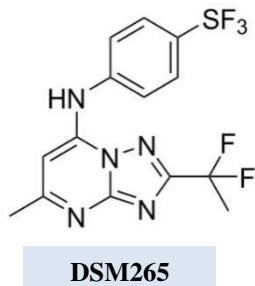
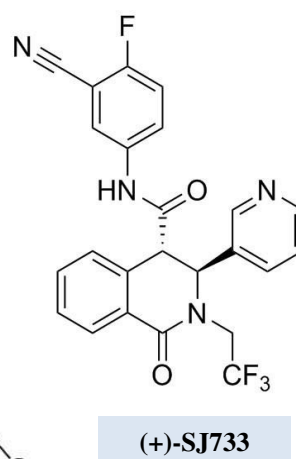
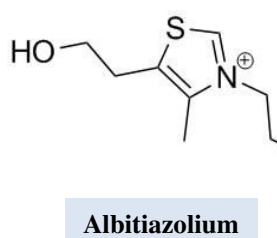
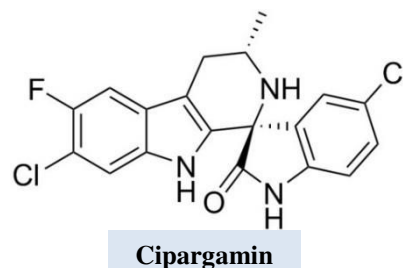
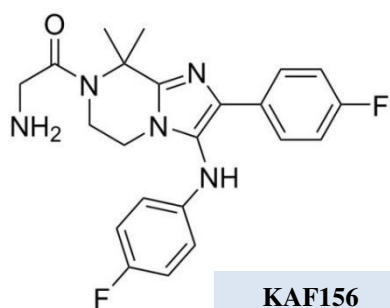


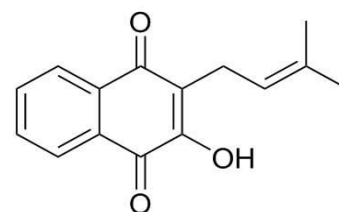
Figure 10 : Nouvelles classes d'antipaludiques. Ces composés sont actuellement en développement clinique pour le traitement du paludisme.

KF156 (Novartis) est le premier représentant de la classe des imidazolopipérazines. Il agit sur tous les stades du parasite. Son mode d'action n'est pas connu. Cipargamin et SJ173 sont des inhibiteurs d'un transporteur présent sur la membrane plasmique du parasite. Il permet de maintenir une concentration cytosolique faible en sodium. DSM265 est un inhibiteur de la dihydroorotate déshydrogénase (DHODH), enzyme impliquée dans la voie de biosynthèse de la pyrimidine, donc de l'ADN. L'albitiazolium est un inhibiteur de l'absorption de choline nécessaire au parasite pour la synthèse de phosphatidylcholine, le lipide principal de ses membranes (Figure 10).

I. 3. Molécules naturelles à activité antipaludique

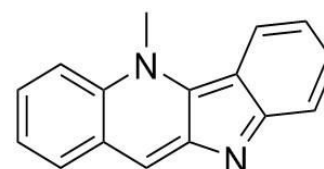
Les alcaloïdes quinoléiques et l'artémisinine démontrent l'importance des molécules naturelles dans le traitement du paludisme. Quelques autres produits naturels emblématiques sont à citer⁷.

Les naphotoquinones isolées de Bognoniaceae tel que le lapachol, connu depuis le début du 20^{ème} siècle, et dont l'optimisation du squelette a conduit à l'atovaquone⁸.



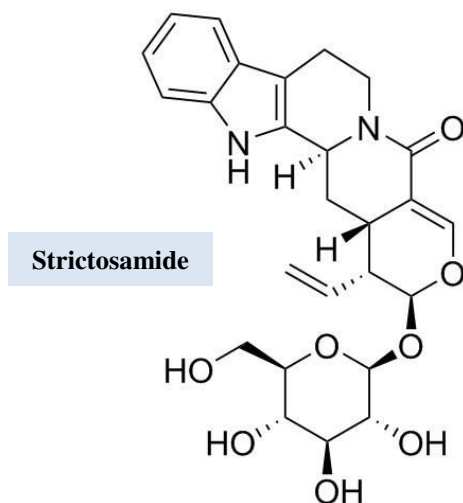
Lapachol

La cryptolépine, un alcaloïde indolique isolé de *Cryptolepis sanguinolenta* (Apocynaceae). Ce composé a montré une activité modérée chez la souris infectée par *P. berghei* (parasitémie réduite de 80% à 50 mg/kg/jour, pendant 4 jours). Une décoction standardisée est commercialisée sous le nom de PHYTO-LARIA® par l'Université de Science et de Technologie Kwame Nkrumah au Ghana⁹.

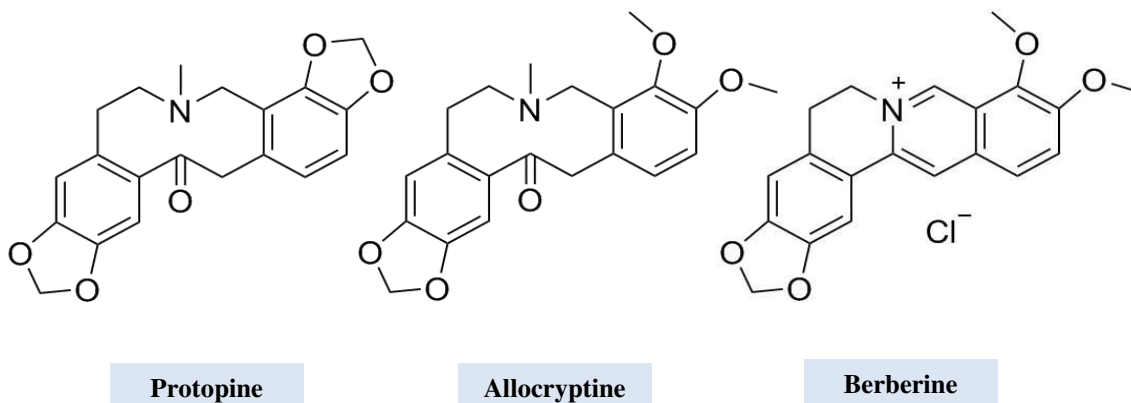


Cryptolépine

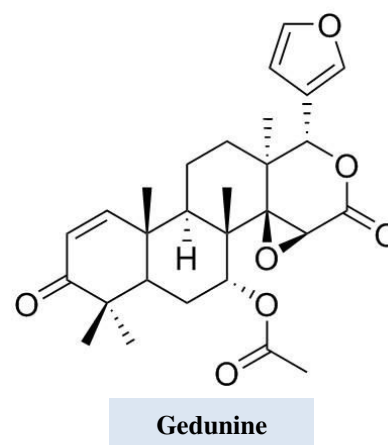
L'extrait hydroalcoolique 80% standardisé de *Nauclea pobeguinii* (Rubiaceae) (PR 259 CT1), plante utilisée en médecine traditionnelle au Congo contre le paludisme, est en cours d'essai clinique phase II. Cet extrait contient majoritairement un alcaloïde, le strictosamide¹⁰.



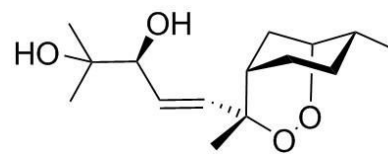
Des extraits et décoctions d'*Argemone mexicana* (Papaveraceae) sont également utilisés en médecine traditionnelle africaine pour le traitement du paludisme. Un essai clinique de la décoction a été réalisé sur 301 patients. Aujourd'hui, la décoction est approuvée par le gouvernement malien comme complément aux CTAs. Ces composants principaux sont des alcaloïdes isoquinoléiques ; la protopine, la berbérine et l'allocryptopine¹¹.



L'usage d'*Azadirachta indica* (Meliaceae) en médecine traditionnelle indienne a été observé au début du 20^{ème} siècle. Aujourd'hui, des capsules de son extrait acétone/eau, IRACARP® ont été enregistrées au Nigeria. Son principal constituant est la gédunine, un triterpène pentacyclique¹².



Le yingzhaosu A isolé de l'*Artabotrys uncinatus* (Ying Zhao, Annonaceae), plante utilisée en médecine traditionnelle chinoise, est un sesquiterpène peroxydé. Son dérivé de synthèse l'arteflene a été évalué en phase clinique II. Les essais ont été abandonnés car il n'apportait pas d'amélioration par rapport à la méfloquine ou l'artémisinine et montrait plus d'effets secondaires.



Yingzhaosu A

Les flavonoïdes représentent une famille chimique qui a été explorée pour ses propriétés antipaludiques. Cependant, aucun antipaludique de cette famille n'est sur le marché. La lanaroflavone, un biflavonoïde isolé de *Camposperma panamense* (Anacardiaceae) présentait *in vitro* une CI_{50} de $0.48 \mu\text{M}$ vis-à-vis de *P. falciparum* (souche multirésistante K1). Aujourd'hui, des dérivés de cette molécule font l'objet d'un brevet déposé par notre laboratoire en collaboration avec le CERM de l'Université de Caen-Normandie¹³ (*Compound for use in the treatment of parasites mediated diseases European patent EP18305121.8*).

De nombreuses autres molécules naturelles sont évaluées pour leurs activités antipaludiques. Pour avoir une vue d'ensemble de ces composés, nous nous appuyons sur la revue de J. Bero¹⁴. Entre 2005 et 2008, 300 composés ont été décrits pour leurs propriétés antipaludiques mais finalement très peu ont une $CI_{50} < 2 \mu\text{M}$ (considérés comme très actifs).

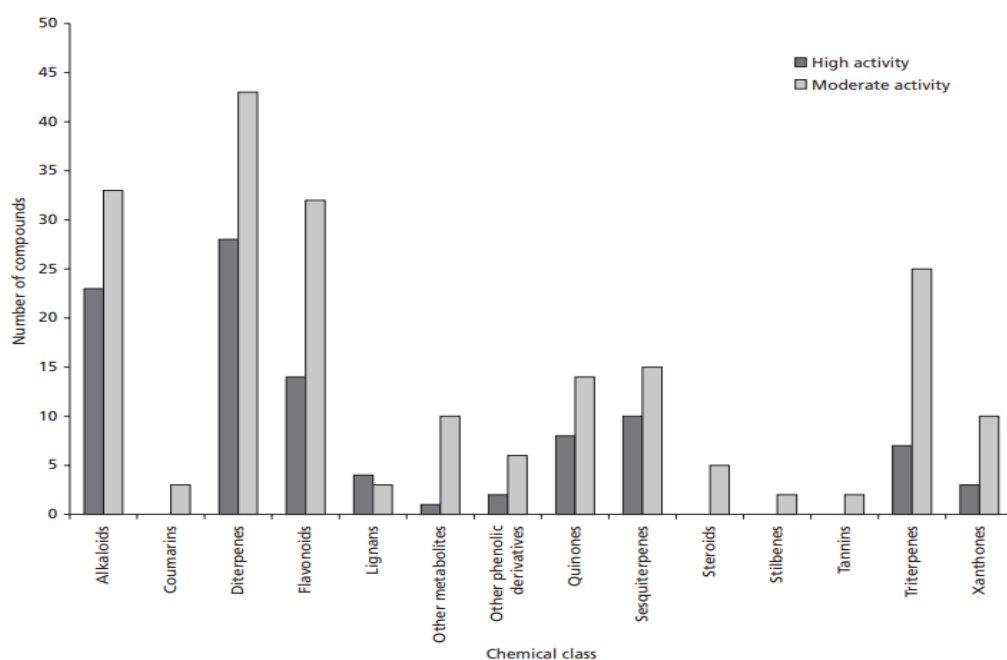


Figure 11 : Nombre de composés publiés entre 2005 et 2008 montrant une forte activité ($CI_{50} \leq 2 \mu\text{M}$) ou une activité modérée ($CI_{50} \leq 11 \mu\text{M}$) *in vitro* contre différentes souches de *P. falciparum*, d'après J. Bero¹³. Les alcaloïdes, les flavonoïdes et les terpènes constituent les principales classes de composés antipaludiques naturels.

Les classes les plus représentées sont les terpénoïdes (diterpènes, sesquiterpènes et triterpènes), les alcaloïdes et flavonoïdes (Figure 11).

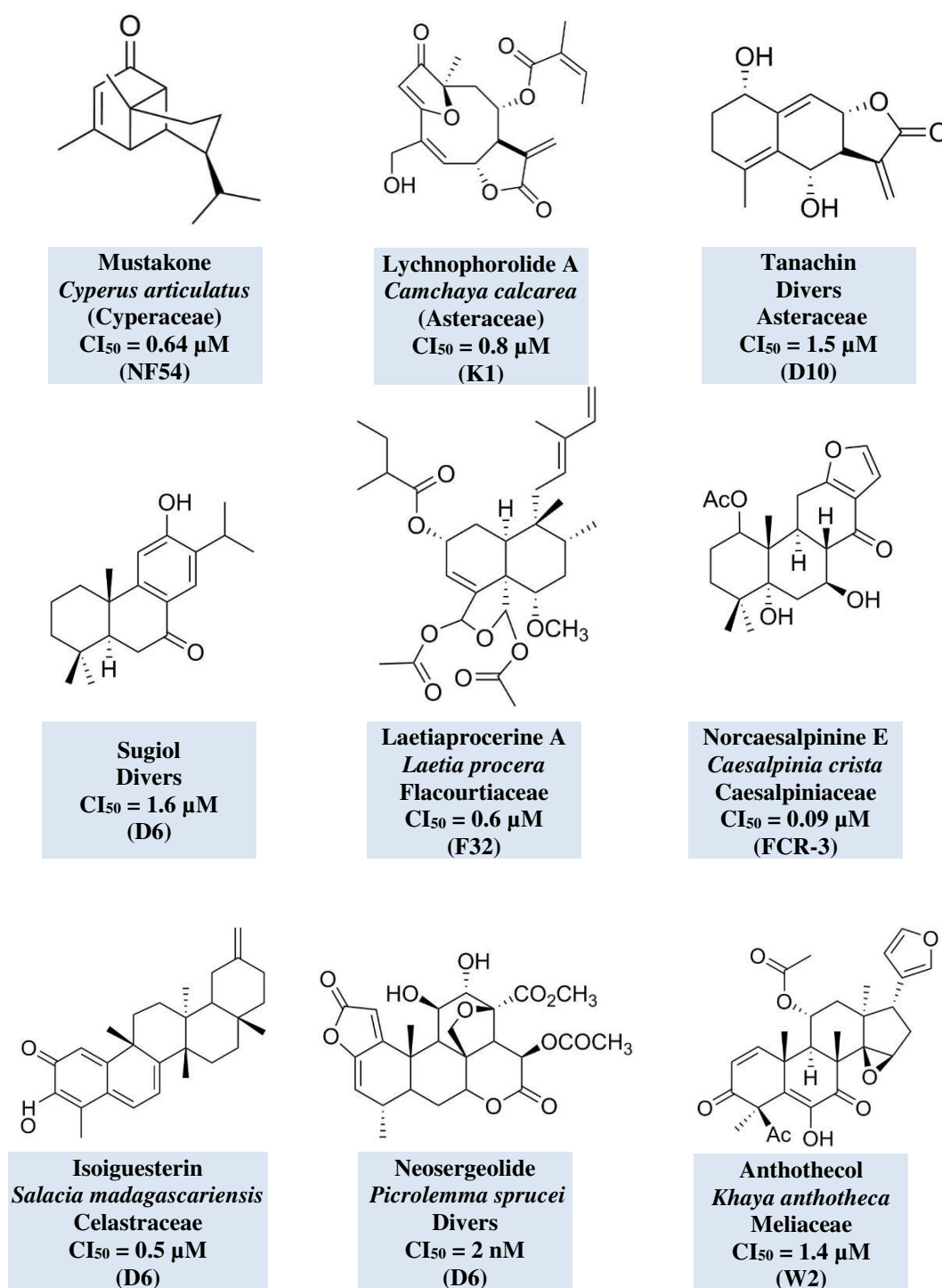


Figure 12 : Sesquiterpènes, diterpènes et triterpènes actifs sur *P. falciparum* *in vitro*. Souches sensibles de *P. falciparum* à la chloroquine : NF54, D6, D10, F32 et souches résistantes à la chloroquine : K1, FCR-3, W2.

Des exemples d'alcaloïdes et de flavonoïdes ont déjà été cités, pour illustrer la vue d'ensemble de composés naturels antipaludiques, quelques exemples de terpènes sont donnés dans la **Figure 12**. A noter que ces composés sont parfois cytotoxiques¹⁵⁻²³.

I. 4. Conclusion

Aujourd'hui, une dizaine de nouvelles molécules sont en phase clinique pour lutter contre les parasites notamment multi-résistants. *Plasmodium* est un organisme qui développe rapidement des résistances aux différentes classes chimiques d'antipaludiques. Des cas de multirésistances de *P. falciparum* aux CTAs ont été détectés en Asie du Sud-Est.

Ainsi, il est essentiel de continuer le développement de nouvelles classes de composés antipaludiques car ces composés ne sont pas encore sur le marché et des associations de médicaments au mécanisme d'action nouveau pourraient limiter le développement de résistances à ces nouvelles classes.

II. L'Apicoplaste : Une cible thérapeutique d'intérêt

L'apicoplaste est un organelle relique commun au phylum des Apicomplexes (hors *Cryptosporidium*) (Figure 13). On le retrouve donc dans plusieurs organismes parasites de l'Homme tels que *Plasmodium* ou *Toxoplasma*. Il est impliqué dans plusieurs voies métaboliques essentielles à la survie du parasite. Cet organelle original unique à ce phylum représente une cible d'intérêt en raison de son absence des cellules eucaryotes humaines.

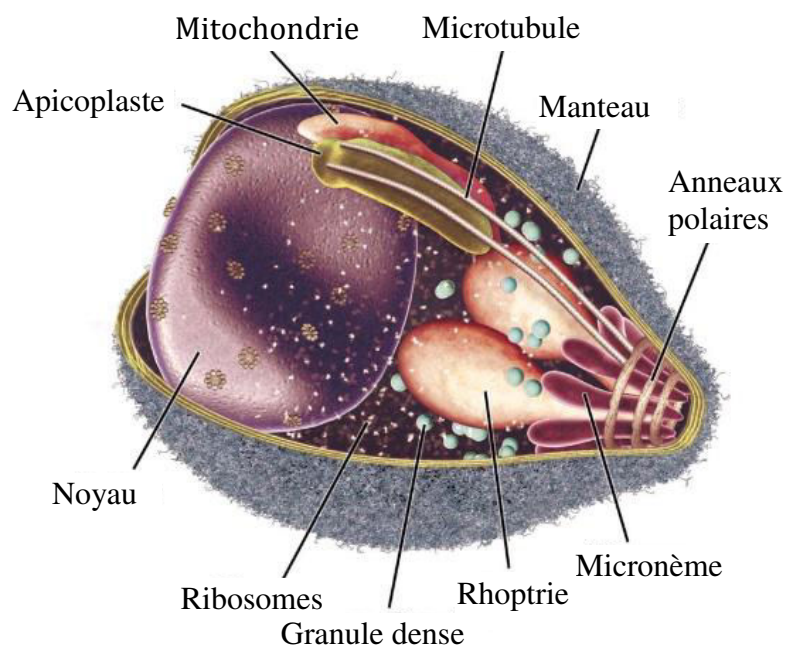


Figure 13 : Organisation cellulaire d'un mérozoïte. L'apicoplaste est accolée à la mitochondrie avec qui il partage des voies métaboliques, schéma modifié d'A. F. Cowman et B. S. Crabb.

II. 1. Une origine longtemps discutée

L'apicoplaste a été clairement identifié comme provenant d'une endosymbiose secondaire. Une cellule ancestrale eucaryote a intégré une seconde cellule eucaryote issue d'une endosymbiose avec un procaryote de type cyanobactérie. Les organites issus d'endosymbioses secondaires sont caractéristiques car ils présentent 3 à 4 membranes, quatre dans le cas de l'apicoplaste. La nature du second endosymbiote a longtemps été discutée. Sur l'étude génétique de l'ADN apicoplastique (circulaire, 35 Kb), deux possibilités ont été présentées, soit cet organite provenait d'une algue verte, soit d'une algue rouge. La compréhension de l'histoire évolutive de ce parasite est restée non résolue jusqu'à la découverte du chaînon manquant : *Chromera velia*. Cet organisme unicellulaire possède le même organite à 4 membranes à la différence qu'il a conservé son appareil photosynthétique. Sa proximité phylogénétique aux Apicomplexes et l'organisation de son appareil photosynthétique ont montré qu'il était bien issu d'une algue rouge²⁴⁻²⁶. Ainsi, l'origine de l'apicoplaste a été résolue.

II. 2. Une cible thérapeutique

En raison de son passé évolutif, l'apicoplaste possède plusieurs voies métaboliques proches de celles rencontrées chez les végétaux. Certaines de ces voies métaboliques sont indispensables à la survie de *Plasmodium*. Elles représentent des cibles privilégiées pour le développement de nouveaux médicaments car elles sont absentes des cellules animales et cela permettrait de s'assurer d'une sélectivité²⁷. L'apicoplaste est impliqué dans les voies de biosynthèse des acides gras, des isoprénoïdes, de l'hème et des centres Fer-Soufre (Figure 14)²⁸.

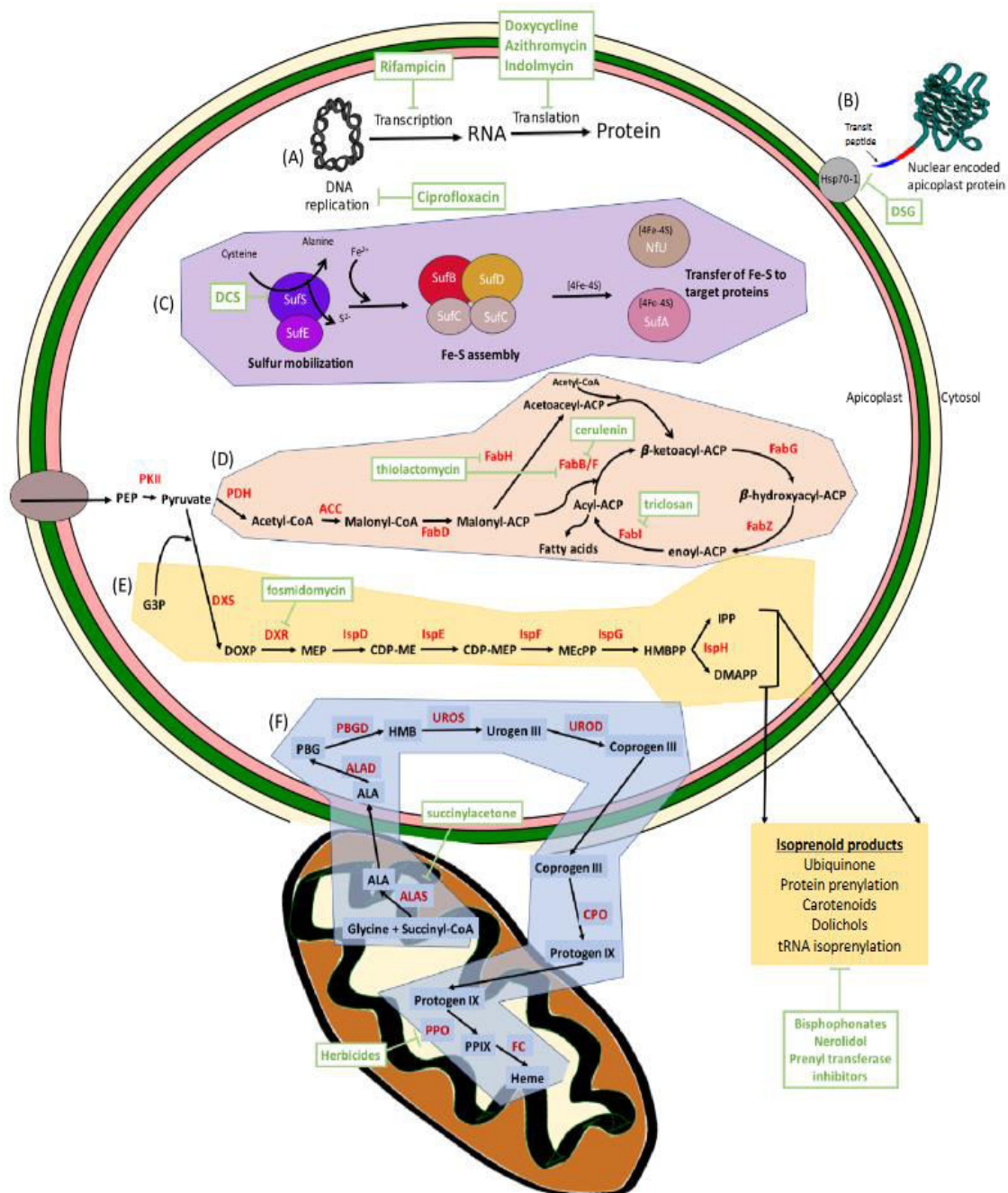


Figure 14 : Voies métaboliques de l'apicoplaste et cibles thérapeutiques. Synthèse de protéines (A), import de protéines nucléaires (B), voie de biosynthèse des complexes fer-Soufre (C), voie de biosynthèse des acides gras (D), voie de biosynthèse des isoprénoïdes (E), voie de biosynthèse de l'hème (F). En vert, les inhibiteurs des enzymes de ces voies, d'après L. M. Low et al (2017).

La doxycycline appartient aux tétracyclines. Elle inhibe la synthèse protéique en se liant à la sous unité 30S du ribosome bactérien. Tout comme la clindamycine, son action est lente et implique l'apicoplaste et un phénomène de mort retardée.

L'azithromycine agit de la même manière que la clindamycine. Tout comme les deux autres antibiotiques, elle agit lentement, son effet étant retardé.

II. 2. b. Inhibiteurs de la voie de biosynthèse de l'hème et des complexes Fer-Soufre

La voie de biosynthèse de l'hème est partagée entre trois compartiments, la mitochondrie, l'apicoplaste et le cytoplasme. Celle-ci est seulement essentielle au parasite au stade exoérythrocyte³¹. Elle peut être inhibée par la succinylacétone.

Des herbicides ont également montré qu'ils étaient capables d'interférer avec cette voie par l'inhibition de la protoporphyrinogène IX oxidase dans la mitochondrie (atovaquone, antimycine A)³².

La voie de biosynthèse des complexes Fer-Soufre est présente dans divers compartiments chez les eucaryotes mais chez *Plasmodium*, cette voie serait uniquement présente dans l'apicoplaste. C'est une voie essentielle à la survie de l'apicoplaste. Elle jouerait un rôle avec la voie de biosynthèse des isoprénoïdes de l'apicoplaste³³. Elle peut être inhibée par la D-cyclosérine *in vitro*³⁴ (Figure 16).

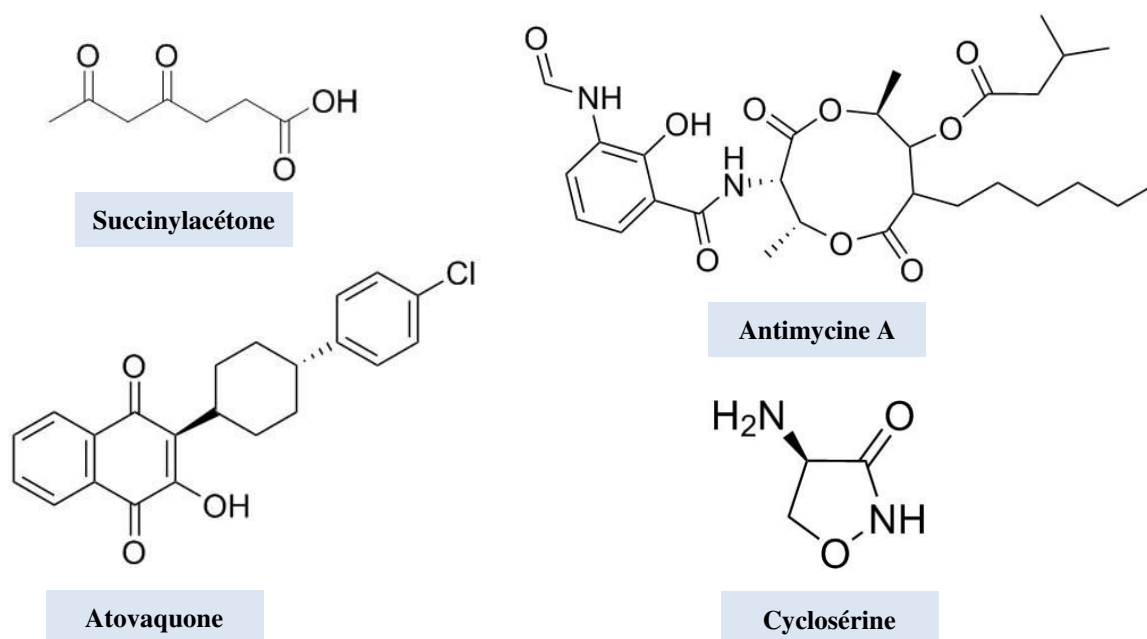


Figure 16 : Inhibiteurs de la voie de biosynthèse de l'hème et des complexes Fer-Soufre.

II. 2. c. Inhibiteurs de la voie de biosynthèse des acides gras

L'apicoplaste possède une voie de biosynthèse des acides gras de Type II (FAS II) qui implique plusieurs enzymes. Cette voie diffère de la voie de biosynthèse des acides gras Type I des cellules animales et représente une cible particulièrement intéressante pour sa spécificité. La thiolactomycine, le triclosan et le cérulenine inhibent des enzymes de cette voie et provoquent la mort du parasite (Figure 17).

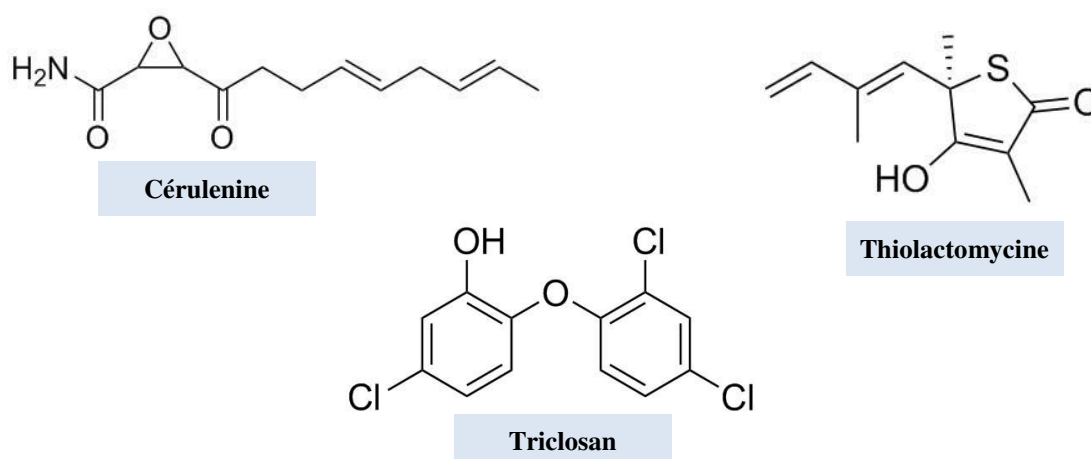


Figure 17 : Inhibiteurs de la voie de biosynthèse des acides gras. Cette voie métabolique exprimée dans l'apicoplaste diffère de celle exprimée dans les cellules animales.

II. 2. d. Inhibiteurs de la voie de biosynthèse des isoprénoides

La voie de biosynthèse des isoprénoides chez *Plasmodium* est une voie clé qui conduit à la formation de l'isopentenyl phosphate (IPP) et le diméthylallyl pyrophosphate (DMAPP) (Figure 18). Elle conduit ensuite à la formation de monoterpènes, sesquiterpènes, diterpènes, de stérols, de caroténoïdes, d'ubiquinone et de dolichols³⁵. Il s'agit de la voie MEP (2-C-méthyl-D-erythritol-4-phosphate) ou DOXP (1-déoxy-d-xylulose 5-phosphate).

C'est une voie alternative de celle retrouvée chez les cellules animales et qui représente donc une cible spécifique potentielle. Cette voie est également présente chez les plantes, les algues et les bactéries.

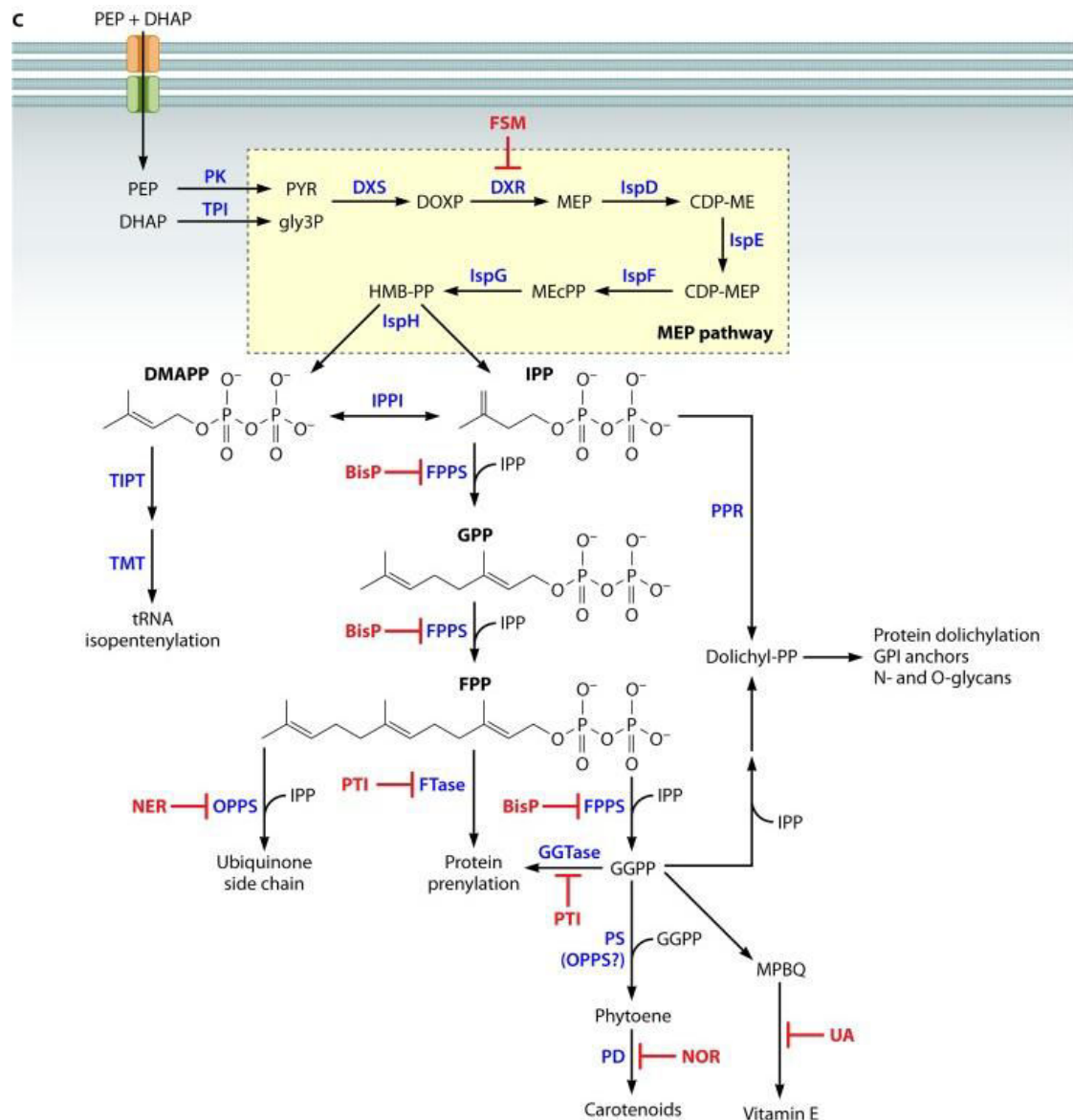


Figure 18 : Voie de biosynthèse des isoprénoides et cibles thérapeutiques chez *P. falciparum* d'après A. M. Guggisberg *et al.* Fosmidomycine (FSM), bisphosphonates (BisP), nérolidol (NER), inhibiteurs de la prényltransférase (PTI), norflurazon (NOR), acide usnique (UA).

La fosmidomycine inhibe la DOXP reductoisomérase et provoque la mort du parasite. Elle est actuellement en essai clinique en combinaison avec la pipéraquine (phase II, NCT02198807). Son dérivé FR-900098 inhibe également la même enzyme et provoque la mort du parasite *in vivo* sans cytotoxicité. Également, le nérolidol, et des terpènes tels que le limonène, le farnesol, le linalol et l'acide S-farnesylthiosalicylique inhibent la voie de synthèse du dolichol et provoquent une diminution de croissance de *Plasmodium*³⁵⁻³⁸.

Plusieurs enzymes de cette voie sont inhibées par des bisphosphonates, notamment le risedronate. Ce composé présente une activité antipaludique *in vitro* et *in vivo*. Le norflurazon, un herbicide capable d'inhiber la phytoène désaturase, est finalement peu actif sur la croissance du parasite (Figure 19).

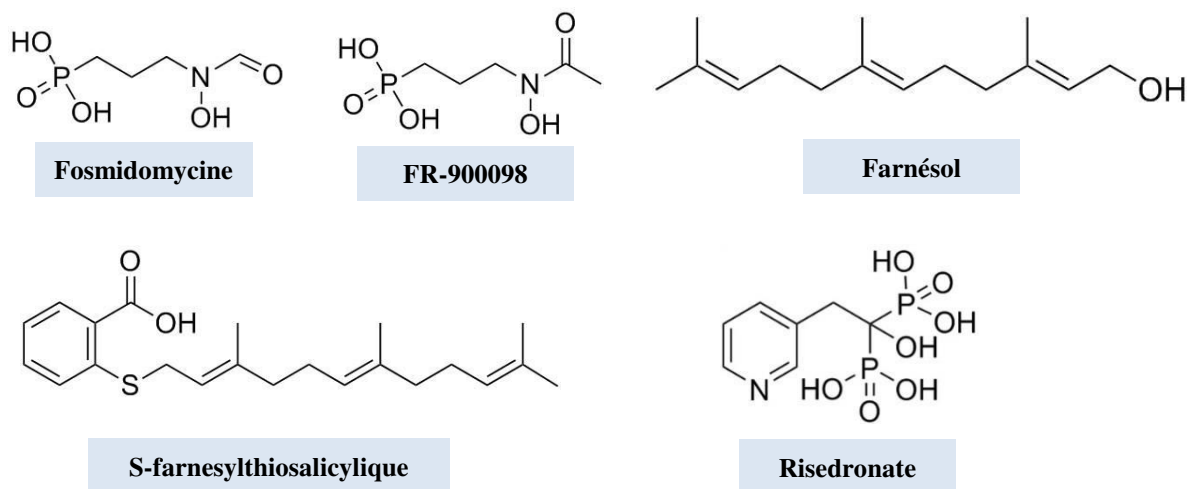


Figure 19 : Inhibiteurs de la voie des isoprénoïdes. Ces molécules sont capables d'inhiber la voie de biosynthèse des isoprénoïdes, voie essentielle et spécifique de *Plasmodium*.

II. 3. Conclusion

L'apicoplaste est spécifique des Apicomplexes, ce qui en fait une cible d'intérêt pour le traitement contre le paludisme. De plus, c'est un organite au passé évolutif original qui possède des voies métaboliques conservées de son origine algale.

Cette proximité évolutive pourrait augmenter la probabilité de rencontrer des métabolites secondaires chez les algues rouges pouvant interférer avec les voies métaboliques de l'apicoplaste par analogie structurale. Nous avons donc émis et exploré l'hypothèse que les algues rouges pourraient représenter une source privilégiée de nouveaux antipaludiques.

III. Les algues rouges (Rhodophytes)

III. 1. Généralités³⁹

La définition des différentes classes d'algues est complexe et implique de nombreux critères tels que la nature et la localisation des réserves des glucanes, la nature de leurs parois cellulaires, le nombre de leurs membranes périplastidiales, la forme de leur mitochondrie, la présence et la structure des flagelles, le type de mitose, le mode de nutrition, de reproduction et leurs habitats. Le mot « algue » ou Phycophyte (du grec *phukos* : algue et *phuton* : plantes) est utilisé pour désigner un groupe large, varié et hétérogène d'organismes dont le statut taxonomique n'est pas clairement défini et qui est régulièrement soumis à des révisions. Les macroalgues sont des organismes physiologiquement proches des plantes supérieures mais se distinguent par leur présence en milieu marin, leur organisation simple et, pour les organismes pluricellulaires, par la présence d'un macrothalle et d'un crampon (Figure 20). Elles se différencient également des plantes par leurs structures reproductives qui possèdent à la fois des cellules fertiles et stériles.

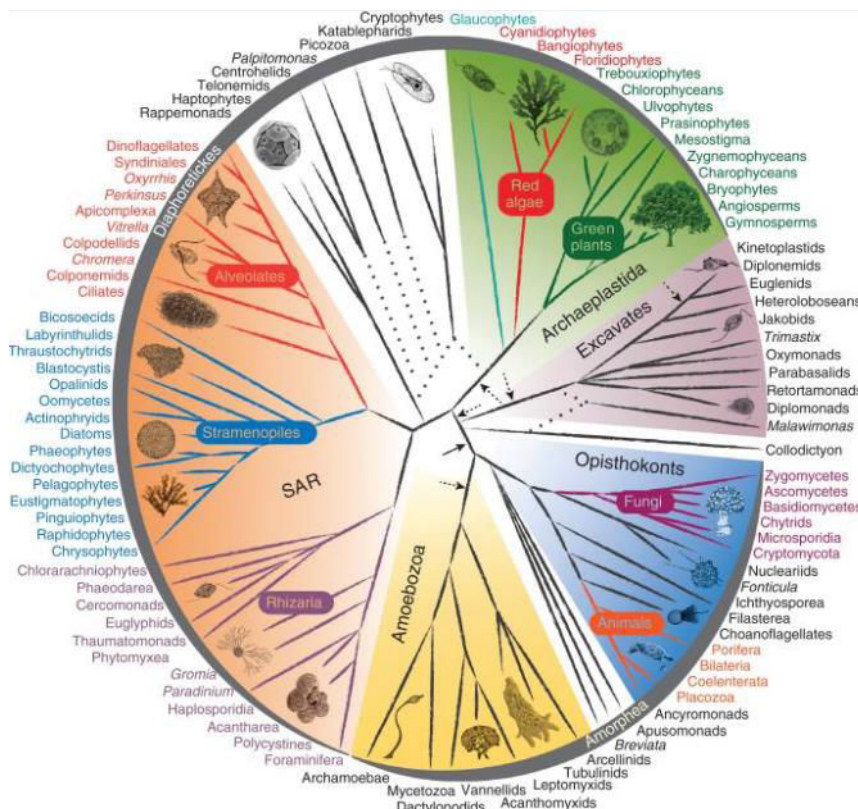


Figure 20 : Arbre global des eucaryotes à partir d'un consensus des preuves phylogénétiques, de rares signatures génomiques et des caractéristiques morphologiques. SAR : straménopiles, alvéolés, et Rhizaires. Les algues rouges sont phylogénétiquement proches des plantes supérieures⁴⁰.

La distribution des algues dépend de nombreux facteurs : physiques (substrat, température, qualité et intensité de la lumière, activité des marées, des vents), chimiques (salinité, pH, nutriments, gaz, pollution), et biologiques (herbivores, pathogènes, épiphytes, endophytes, symbiotes).

Approximativement 9600 espèces de macroalgues ont été identifiées et leur classification est toujours soumise à discussion. Elles sont usuellement classées sur la base de leurs pigments en trois larges groupes : les algues rouges (Rhodophyta), vertes (Chlorophyta) et brunes (Phaeophyta). Les algues rouges sont les plus diversifiées et comptent plus de 7000 espèces, suivies des algues brunes et vertes avec respectivement, 2030 et 600 espèces marines⁴¹. Les algues vertes contiennent de la chlorophylle a et b. Les algues rouges et brunes contiennent les chlorophylles a et c (également d pour les algues rouges) associées à d'autres pigments surnuméraires qui leur confèrent leur couleur. Les algues brunes sont riches en caroténoïdes (carotène, xanthophylle). Les algues rouges contiennent des phycobilines, la phycoérythrine et la phycocyanine, qui sont des protéines colorées.

La plupart des algues rouges sont filamenteuses, pseudoparenchymateuses ou parenchymateuses. Leurs habitats sont diversifiés, elles sont présentes à la fois dans les eaux tropicales, tempérées et arctiques (peu). Grâce à leurs pigments accessoires, elles peuvent se répartir sur de larges niveaux de lumières et en grande profondeur.

Les algues rouges se distinguent principalement par leurs cellules eucaryotes, une absence complète de structure flagellée et de réticulum endoplasmique externe, la présence d'amidon floridéen, de phycobilines et de chloroplastes sans thylakoïdes empilés.

Les algues rouges sont notamment connues pour leurs propriétés nutritionnelles car elles possèdent un taux de protéines élevé, à hauteur de 30 à 50 % de leur poids sec. Ce sont des valeurs comparables aux aliments réputés pour leur teneur en protéines comme le soja. Les algues suscitent donc l'intérêt des industriels des compléments alimentaires. Les espèces de *Gracilaria*, *Chondrus*, *Gelidiella* sont également source de phycocolloïdes qui sont utilisés en formulation de produits cosmétiques et comme gélifiant et stabilisant pour l'industrie agro-alimentaire.

III. 2. Les algues rouges et l'activité antipaludique

D'un point de vue pharmacologique, des extraits d'algues rouges ont montré des activités diverses : antioxydante, antivirale, anticoagulante, antimicrobienne, cytotoxique, antiprotozoaire.

Un criblage d'algues françaises a été réalisé par le Dr. C. Vonthron à l'Université de Caen. Les extraits hydroalcooliques et acétate d'éthyle de 20 espèces d'algues rouges, vertes et brunes ont été évalués *in vitro* pour leurs activités antiprotozoaires sur *P. falciparum*, *Trypanosoma cruzi* et *Leishmania donovani*⁴². Huit des neuf espèces d'algues rouges testées se sont révélées actives vis-à-vis de la souche multirésistante K1 de *P. falciparum* (CI₅₀ < 5µg/mL, test hypoxanthine tritiée), avec une sélectivité satisfaisante. De plus, ces extraits se sont montrés capables d'inhiber la croissance d'un autre parasite Apicomplexe, *T. gondii*, suggérant une activité Apicomplexe-spécifique. Certains extraits d'algues brunes se sont révélés actifs mais présentaient également une cytotoxicité élevée. Enfin, les algues vertes se sont avérées totalement inactives (CI₅₀ > 20 µg/mL).

A ces résultats s'ajoutent les résultats du Dr. Tasdemir. Les espèces d'algues rouges *Corallina granifera*, *Ceramium rubrum*, *Dasya pedicellata*, *Gelidium crinale*, *Osmundea pinnatifida*, *Ceramium virgatum*, *Corallina officinalis* et *Halopitys incurvus* ont montré une activité antipaludique *in vitro* variant de 8.6 à 14.5 µg/mL (*P. falciparum* K1). De manière intéressante, l'extrait de *C. virgatum* et de *H. incurvus* (non actif sur le stade érythrocytaire) ont pu inhiber plusieurs enzymes de la voie de biosynthèse des lipides FAS-II, voie spécifique de l'apicoplaste mais cependant non essentielle à la survie du parasite en conditions normales⁴³.

Des algues rouges originaires d'Afrique du Sud, *Plocamium corallorhiza*, *Plocamium cornutum*, *Amphiroa ephedraea* et *Trematocarpus flabellatus* ont également montré une activité *in vitro* contre *P. falciparum* D10 (test hypoxanthine tritiée) variant de 3.5 à 8.3 µg/mL⁴⁴. Des algues de la côte Est de l'Inde ont été évaluées *in vitro*, *Gracilaria verrucosa*, *Hypnea espera* étaient actives contre *P. falciparum* (test microscopique : comptage des globules rouges infectés après coloration au Giemsa)⁴⁵.

Les algues rouges à activité antipaludique sont présentées dans le tableau 1⁴³⁻⁴⁹.

Tableau 1 : Algues rouges marines évaluées *in vitro* contre différentes souches de *P. falciparum*.

Espèces	Ordre	Produits testés	CI ₅₀	IS
<i>Amphiroa crassa</i> ⁴⁵	Corallinales	δ-Tocopherylquinone	10µM	/
<i>Amphiroa ephedraea</i> ⁴³	Corallinales	Extrait DCM	8.3 µg/mL	>12
<i>Calliblepharis jubata</i> ⁴²	Gigartinales	Extrait AcOEt	5 µg/mL	14
<i>Callophycus serratus</i> ⁴⁵	Gigartinales	Callophycol A	35.7 µM	/
		Bromophycoic acid C	8.7 µM	>4
		Bromophycolide A	0.5 µM	/
		Acetylbromophycolide A	0.24 µM	/
		Bromophycolide D	0.3 µM	/
		Bromophycolide E	0.8 µM	/
Bromophycolide M	0.5 µM	/		
<i>Ceramium rubrum</i> ⁴³	Ceramiales	Extrait MeOH	16.3 µg/mL	>6
<i>Ceramium virgatum</i> ⁴³	Ceramiales	Extrait CHCl ₃ :MeOH (3 :1)	13.6 µg/mL	>7
<i>Chondrus crispus</i> ⁴²	Gigartinales	Extrait AcOEt	2.9 µg/mL	29
<i>Corallina granifera</i> ⁴²	Corallinales	Extrait MeOH	16.3 µg/mL	>6
<i>Corallina officinalis</i> ⁴³	Corallinales	Extrait CHCl ₃ :MeOH (3 :1)	8.6 µg/mL	>10
<i>Dasya pedicellata</i> ⁴³	Ceramiales	Extrait MeOH	0.38 µg/mL	39
<i>Dilsea carnosa</i> ⁴²	Gigartinales	Extrait AcOEt	3.9 µg/mL	19
<i>Gelidium crinale</i> ⁴³	Gelidiales	Extrait MeOH	19.5 µg/mL	>5
<i>Gelidium latifolium</i> ⁴²	Gelidiales	Extrait AcOEt	3.4 µg/mL	18
<i>Gracilaria gracilis</i> ⁴²	Gracilariales	Extrait AcOEt	3.3 µg/mL	21
<i>Gracilaria verrucosa</i> ⁴⁴	Gracilariales	Extrait EtOH:H ₂ O (3:1)	5.6 µg/mL	/
<i>Grateloupia turuturu</i> ⁴²	Halymeniales	Extrait AcOEt	3.1 µg/mL	23
<i>Halurus flosculosus</i> ⁴²	Ceramiales	Extrait AcOEt	4.6 µg/mL	12
<i>Hypnea espora</i> ⁴⁴	Gigartinales	Extrait EtOH:H ₂ O (3:1)	8.9 µg/mL	/
<i>Laurencia corallopsis</i> ⁴⁷	Ceramiales	Extrait EtOH	14.8 µg/mL	/
<i>Laurencia papillosa</i> ⁴⁶	Ceramiales	p-Hydroxybenzaldehyde	7.3 µg/mL	>2.7
<i>Laurencia obtusa</i> ⁴⁷	Ceramiales	(8R*)-8-bromo-10-epi-β-snyderol	2.7µg/mL	/
<i>Mastocarpus stellatus</i> ⁴²	Gigartinales	Extrait AcOEt	2.8 µg/mL	>30
<i>Plocamium corallorhiza</i> ⁴³	Plocamiaceae	Extrait hexane	3.5 µg/mL	10
<i>Plocamium cornutum</i> ⁴³	Plocamiales	Extrait DCM	4.6 µg/mL	13
<i>Osmundea pinnatifida</i> ⁴⁴	Ceramiales	Extrait CHCl ₃ :MeOH (3 :1)	14.5 µg/mL	>6
<i>Pterosiphonia cloiophylla</i>	Ceramiales	Fraction	4.7 µg/mL	16
<i>Trematocarpus flabellatus</i> ⁴³	Plocamiales	Extrait DCM	7.8 µg/mL	6

Vingt-cinq espèces d'algues rouges ont montré une activité antipaludique. CI : Concentration inhibitrice 50, IS : Indice de sélectivité.

De nombreux extraits d'algues rouges montrent une activité antipaludique, cependant, **les molécules responsables de cette activité restent, pour la plupart, à être isolées et identifiées.**

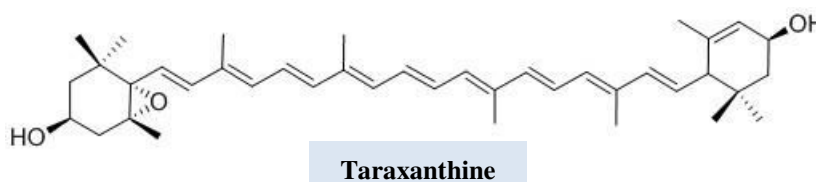
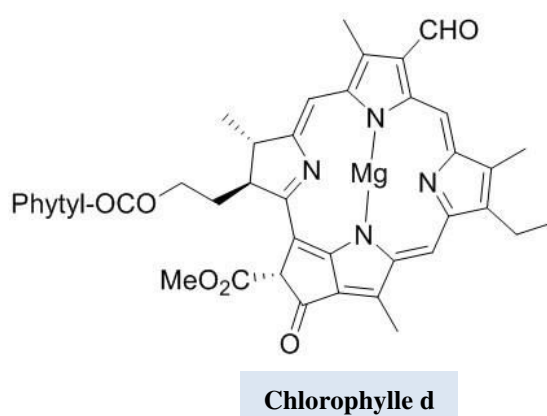
III. 3. Phytochimie des algues rouges

La phytochimie des algues rouges est relativement peu connue, nous présentons quelques métabolites primaires et secondaires.

III. 3. a. Métabolites des algues rouges

- **Pigments photosynthétiques**

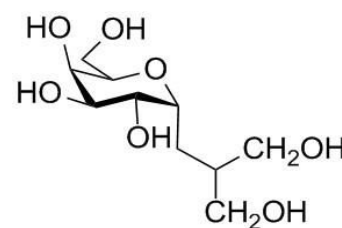
Les algues rouges possèdent un appareil photosynthétique comparable à celui des cyanobactéries, composé de chlorophylles, phycobiliprotéines (influe sur la couleur de l'algue), caroténoïdes et xanthophylles^{50,51}. Les algues rouges ont la particularité de posséder de la chlorophylle d (diffère de la chlorophylle a par un groupement aldéhyde)⁵². Elles possèdent également une xanthophylle spécifique, la taraxanthine.



Ces pigments, en particulier, les caroténoïdes et les xanthophylles, sont déjà bien connus pour leurs activités anti-oxydantes et anti-inflammatoires, et pour d'autres activités reliées telles qu'anti-angiogéniques, cardio-protectrices, anti-cancéreuses, anti-diabétiques, anti-obésité...⁵³.

- **Mono- et Polysaccharides**

Les algues rouges possèdent des polysaccharides de réserve qui forment l'amidon floridéen. Il s'agit de polymères



Floridoside

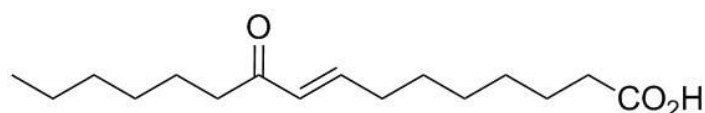
de α -1,4-glucane ramifiés qui utilisent deux isomères de galactosylglycérol, le floridoside et l'isofloridoside⁵⁴.

Les polysaccharides sont très abondants chez les algues et peuvent représenter 30 à 70% de la masse sèche d'un thalle. Les deux principaux types sont les carraghénanes et l'agar. Ce sont des polymères soufrés de galactose et/ou de 3,6- α -anhydrogalactose. Ils diffèrent par leurs proportions de soufre et leurs conformations. Les galactanes et fucanes soufrés sont connus pour leurs propriétés anticoagulantes et antithrombotiques⁵⁵ mais certains ont également montré des activités antivirales⁵⁶, anti-métastatiques⁵⁷, anti-angiogéniques⁵⁸, anti-inflammatoires⁵⁸ et antiadhésives^{56,58}.

- **Phospholipides et acides gras**

Les algues rouges possèdent des phospholipides tels que la phosphatidylcholine présente en forte concentration, et dans une moindre mesure, du phosphatidylglycérol^{59,60}. Elles sont riches en acides gras polyinsaturés à 20 carbones et plus particulièrement en acide arachidonique (20:4n-6) et eicosapentaénoïque (20:5n-3) mais contiennent également de l'acide palmitique en forte quantité (16:0)⁶¹. L'acide arachidonique est essentiel : il est le précurseur de prostaglandines, des eicosanoïdes, des thromboxanes et des oxylipines.

Des acides céto-palmitiques ont été isolés chez le genre *Gracilaria* et *Hypnea* (Gigartinales) et ont montré une activité anti-inflammatoire⁶².

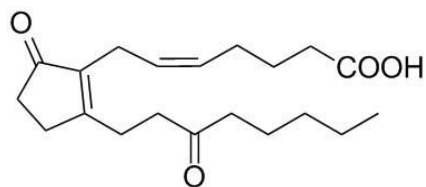


Acide 10-céto-hexadec-8-énoïque
Gracilaria verrucosa

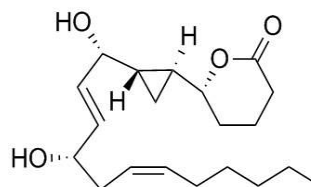
- **Oxylipines et prostaglandines**

Les oxylipines sont des composés à longue chaîne polyinsaturée et oxygénée, dérivés des acides gras (souvent 18 à 20 carbones). Ces molécules sont impliquées dans des mécanismes de défenses des algues⁶³. Certaines présentent des fonctions lactones originales (constanolactone A).

Des oxylipines et des prostaglandines ont été retrouvées dans de nombreuses algues rouges, notamment chez plusieurs espèces de *Gracilaria*^{64,65}. Certains de ces composés possèdent une activité anti-inflammatoire⁶⁶.



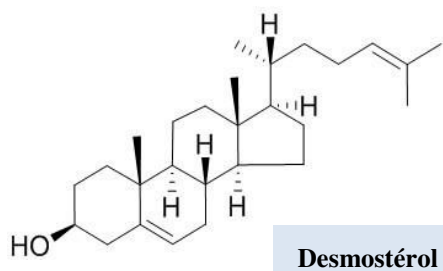
Acide (5Z)-9,15-Dioxoprostanoïque
Prostaglandine
Gracilaria verrucosa



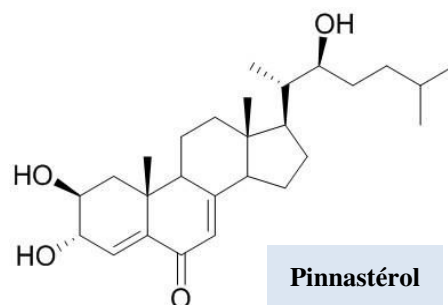
Constanolactone A
Oxylipine
Constantinea simplex

• Stérols et ecdystéroïdes

Le cholestérol est le principal stérol des algues rouges. D'autres ont été identifiés, tels que le desmostérol présent chez plusieurs Gigartinales ou le formate de cholest-5-en-3-ol chez *Grateloupia turuturu*⁶⁷. Des polyhydroxystérols avec des groupements cétones ont aussi été isolés, comme le pinnastérol chez *Laurencia pinnata*⁶⁸. Récemment, le fucostérol a semblé montrer une activité antipaludique⁶⁹. Des stérols glycolsylés ont également été isolés de *Peyssonnelia sp.* et montrent une cytotoxicité contre les lignées de cellules cancéreuses humaines du sein et du poumon (MDA-MB-468 et A549)⁷⁰.



Desmostérol



Pinnastérol

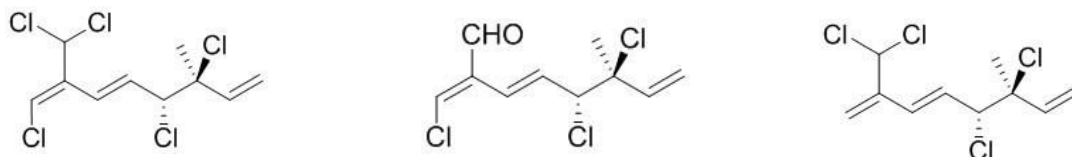
• Autres métabolites

Les rhodophytes ont la caractéristique de montrer une abondance de composés halogénés (bromés et chlorés), notamment de nature terpénique ou dérivés d'acétogénine. Les Cérémiales, Gigartinales et Némaliales sont particulièrement riches en ces structures. Ainsi, de nombreux monoterpènes, sesquiterpènes et quelques diterpènes et triterpènes halogénés ont déjà été isolés. Ces molécules ont montré des activités cytotoxiques⁷¹⁻⁷⁴, répulsives⁷⁵, antioxydantes⁷⁶, antivirales⁷⁷, antibactériennes⁷⁸⁻⁸⁰, anti-inflammatoires⁷⁹...

Des molécules azotées ont été identifiées, à savoir des acides aminés, des bétaines et des sulfoniotéamines et quelques rares alcaloïdes.

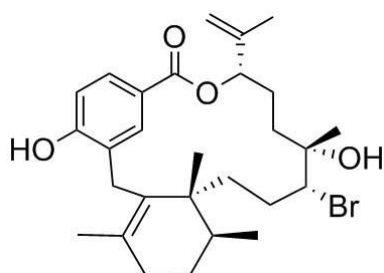
III. 3. b. Métabolites antipaludiques des algues rouges

Des monoterpènes halogénés au squelette de type 3,7-diméthyl-3,4-dichloro-octa-1,5,7-triene ont été isolés de *Plocamium cornutum* et ont présenté une activité antipaludique modérée contre la souche chloroquino-sensible de *P. falciparum* D10 (CI_{50} = 16 à 27 μ M, test pLDH)⁸¹.

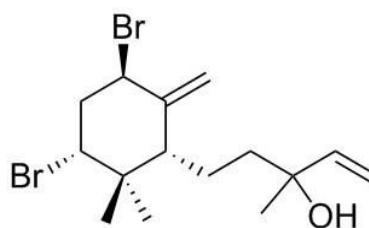


Monoterpènes halogénés
P. cornutum

Dix-huit bromophycolides ont été isolés chez *Callophycus serratus*. Certains de ces composés, les bromophycolides A, D, E, M sont actifs *in vitro* contre *P. falciparum* 3D7 (CI_{50} inférieure à 1 μ M). Certains de ces composés ont été évalués contre des lignées de cellules tumorales et sembleraient également cytotoxiques⁸².

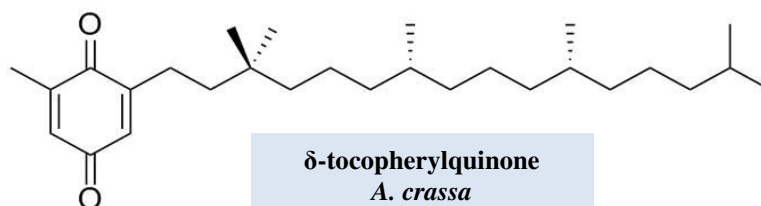


Bromophycolide M
 CI_{50} = 0.5 μ M
C. serratus



(8R*)-8-bromo-10-epi- β -snyderol
L. obtusa

Le δ -tocopherylquinone isolé d'*Amphiroa crassa* a montré une activité *in vitro* intéressante contre *P. falciparum* 3D7 (CI_{50} = 10 μ M). De même, des sesquiterpènes brominés isolés de *Laurencia obtusa* et *papillosa* ont présenté une activité antipaludique contre *P. falciparum* (CI_{50} = 2,7 μ g/mL)⁴⁵.



δ -tocopherylquinone
A. crassa

IV. Conclusion et objectifs du travail

Le traitement du paludisme reste un enjeu actuel majeur. En effet, les résistances répétitives de *Plasmodium* à toutes les classes de molécules mises sur le marché incitent à continuer de découvrir de nouvelles classes d'antipaludiques aux cibles originales. L'apicoplaste et ses voies métaboliques pourraient représenter une de ces cibles car il est spécifique des Apicomplexes. Cela limiterait les interactions avec les voies métaboliques humaines. Cet organe a un passé évolutif particulier car il est issu d'une endosymbiose secondaire avec une algue rouge dont il a conservé certaines voies métaboliques. Ce passé évolutif commun pourrait favoriser la probabilité de trouver des métabolites secondaires chez les algues rouges interférant dans les voies de l'apicoplaste. Cela est appuyé par les travaux préliminaires du Dr. Vonthron qui a montré que la quasi-totalité des algues rouges évaluées jusqu'alors possédaient une activité antipaludique. Ces résultats sont complétés par les données de la bibliographie. Alors que de nombreux extraits algaux ont présenté une activité antipaludique, peu de molécules actives ont été isolées.

Ce partage d'activité biologique inter-espèces laisse supposer qu'une classe de molécules communes aux algues rouges serait capable d'interagir avec *Plasmodium*, et notamment d'interférer avec les voies métaboliques de l'apicoplaste avec un fort potentiel thérapeutique. C'est sur cette hypothèse que reposent notre projet et la stratégie adoptée.

Mener une étude phytochimique classique sur chacune des espèces d'algues rouges actives serait un travail long et fastidieux. Il nous est donc apparu important de :

- **confirmer les activités antipaludiques en élargissant à d'autres espèces d'algues rouges ;**
- **développer, valider et utiliser une méthode analytique rapide permettant de comparer plusieurs échantillons actifs ;**
- **identifier les molécules actives au potentiel mécanisme d'action originale.**



CHAPITRE II

Chapitre II : Etude de l'activité antipaludique d'extraits d'algues rouges

Ce chapitre porte sur la présentation du matériel algal qui fait l'objet de cette étude, ainsi que de son évaluation biologique.

Dans un premier temps, nous présenterons les espèces étudiées et les extraits obtenus à partir de leur thalle.

Nous aborderons ensuite l'activité anti-*Plasmodium* des extraits et fractions obtenus, ainsi que sur leur potentiel d'inhibition d'une voie de biosynthèse apicoplastique essentielle à la survie de *Plasmodium* : la voie de biosynthèse des isoprénoïdes, exprimée dans un modèle végétal.

I. Présentation du matériel algal

Les algues ont premièrement été classifiées par leurs couleurs (Lamouroux), par leurs morphologies (Linnée) puis par leurs données moléculaires. Le phylum Rhodophyte intègre toutes les espèces d'algues rouges. Il est composé d'environ 700 genres et plus de 6 000 espèces⁸³. La classification des algues rouges selon les données moléculaires est présentée dans la **Figure 21**. Elles arborent énormément de morphologies différentes. Elles ont pour caractéristiques communes : la présence de la phycocyanine et la phycoérythrine, un cytoplasme présentant des grains d'amidon floridéen, des cellules adjacentes liées par des synapses, l'absence de flagelles des cellules reproductives⁸⁴.

A la suite des résultats d'activités antipaludiques des espèces d'algues rouges du Dr. Vonthron, six espèces de macro-algues rouges toutes originaires de Normandie ont été réceptionnées au laboratoire⁴⁶. Deux d'entre elles, *Callithamnion tetricum* (Ceramiaceae) et *Plocamium cartilagineum* (Plocamiaceae) n'avaient jamais été évaluées pour leur activité antipaludique. Les récoltes et la détermination taxonomique ont été réalisées par le Dr. A.-M. Rusig (UMR BOREA, Université de Caen-Normandie). Un échantillon de référence de chaque espèce a été déposé à l'Herbier de l'Université de Caen. Six espèces ont été récoltées, il s'agit de *Mastocarpus stellatus* (Phyllophoraceae), *Callithamnion tetricum* (Ceramiaceae), *Dilsea carnosa* (Dumontiaceae), *Gracilaria gracilis* (Gracilariaceae), *Plocamium cartilagineum* (Plocamiaceae) et *Palmaria palmata* (Palmariaceae). Différents lots ont parfois pu être récoltés, en fonction de la disponibilité de la matière première. (Tableau 2).

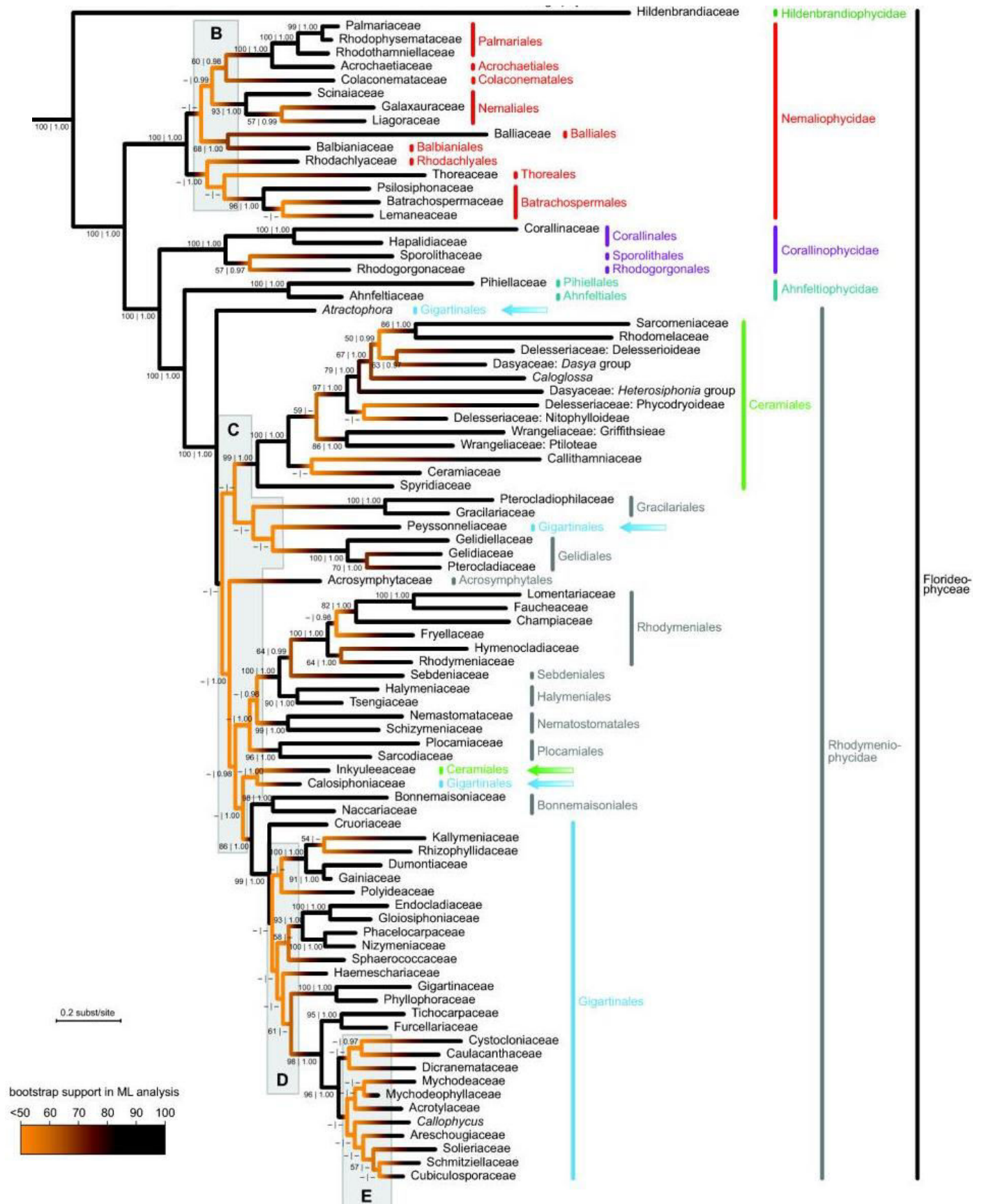






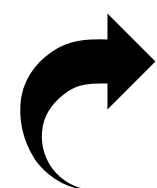


Figure 21 : Classification taxonomique actuelle des algues rouges sur critère génétique (Classe des Florideophyceae). L'arbre a été construit sur les inférences phylogénétiques Bayésiennes des données ADN de GenBank. Les couleurs des branches indiquent la soutenance statistique des clades. Plus une branche est foncée plus elle est statistiquement soutenue, plus elle est claire plus le clade est incertain, d'après H. Berburggen *et al.*

Tableau 2 : Identité botanique des échantillons algaux étudiés.

Phylum	Famille	Identité botanique	Date de collecte	Localité	
Rhodophyta (algues rouges)	Phylloporaceae	<i>Mastocarpus stellatus</i> (Stackhouse) Guiry	16/06/14	Cap Lévi, Fermanville (50)	
	Ceramiales	<i>Callithamnion tetricum</i> (Dillwyn) S. F. Gray	17/04/15	Grandcamp-Maisy (14)	
	Dumontiaceae	<i>Dilsea carnosa</i> (Schmidel) Kuntze (lot 1)	27/10/15	Pointe de la Loge, Vicq-sur-Mer (50)	
	Dumontiaceae	<i>Dilsea carnosa</i> (Schmidel) Kuntze (lot 2)	01/10/16	Pointe de la Loge, Vicq-sur-Mer (50)	
	Gracilariaceae	<i>Gracilaria gracilis</i> (Stackhouse) M. Steentoft, L.M. Irvine & W.F.Farnham (lot 1)	25/11/15	Bernières-sur-Mer (14)	
	Gracilariaceae	<i>Gracilaria gracilis</i> (Stackhouse) M. Steentoft, L.M. Irvine & W.F. Farnham (lot 2)	25/11/15	Courseulles-sur-Mer (14)	
	Plocamiaceae	<i>Plocamium cartilagineum</i> (Linnaeus) P.S. Dixon	25/11/15	Bernières-sur-Mer (14)	
	Palmariaceae	<i>Palmaria palmata</i> (Linnaeus) F. Weber & D. Mohr	14/05/18	Luc-sur-Mer (14)	

Toutes les espèces étudiées appartiennent au phylum des Rhodophytées (algues rouges). Le thalle de chaque espèce a été récolté entre 2014 et 2018 dans les différents départements de Normandie indiqués, lavé à l'eau de mer filtrée puis à l'eau distillée, avant d'être congelé puis lyophilisé, broyé et conservé à 20°C, à l'obscurité. Différents lots ont pu être récoltés pour une même espèce, en fonction de la disponibilité de la matière première. La détermination taxonomique et la récolte ont été réalisées par le Dr. A.-M. Rusig (UMR BOREA, Université de Caen). Un échantillon de référence de chaque espèce a été déposé à l'Herbier de l'Université de Caen. Crédit photo : Algaebase.



Ouvrez la page SVP

Toutes les algues sont originaires de Normandie, les lieux de récoltes sont distribués sur la côte entre Caen et Cherbourg (Figure 22).



Figure 22 : Localisation de la zone de récolte des algues normandes. 1 : Cap Lévi, 2 : Pointe de la Loge, 3 : Courseulles-sur-Mer, 4 : Grandcamp-Maisy, 5 : Bernières-sur-Mer, 6 : Luc-sur-Mer

II. Phytochimie et activités biologiques des espèces étudiées

La phytochimie des macro-algues rouges a été assez peu étudiée et se restreint essentiellement aux molécules qui font l'objet d'une exploitation industrielle pour leurs propriétés texturantes, à savoir des polysaccharides (carraghénanes). Très peu d'études phytochimiques ont été réalisées sur les espèces étudiées. Cependant, les genres *Plocamium* et *Garcilaria* sont les mieux connus d'un point de vue phytochimique. L'originalité du genre *Plocamium* repose notamment sur son abondance en monoterpènes halogénés (Tableau 3).

Tableau 3 : Données bibliographiques recueillies sur les espèces d'algues rouges fournies

Espèces	Informations référencées	
	Composition phytochimique	Activités thérapeutiques
<i>M. stellatus</i> (Stackhouse) Guiry	Floridoside ⁸⁵ , alcaloïde (Horde-nine) ⁸⁶ , carraghénanes ^{87,88} , bétaïne lipidique, phospholipides ⁸⁹ , caroténoïdes ⁹⁰	Antibactérien ⁹¹ , antiprotozoaire ^{42,92} , immunomodulation ⁸⁵ , anticoagulant et antioxydant (carraghénane) ^{88,93}
<i>D. carnosa</i> (Schmidel) Kuntze	Floridoside ⁹⁴	Antifongique modéré ⁹⁵ , antibactérien ⁹⁶ , antiprotozoaire ⁴²
<i>C. tetricum</i> (Dillwyn) S. F. Gray	Phospholipides ⁸⁹	Antibactérien ⁹⁶ , antiprotozoaire ⁴²
<i>G. gracilis</i> (Stackhouse) M. Steentoft	Agar (production industrielle) ⁹⁷	Antiprotozoaire ⁴²
<i>P. cartilagineum</i> (Linnaeus) P. S. Dixon	Acides organiques ⁹⁸ , monoterpènes cycliques ou acycliques halogénés ⁹⁹⁻¹⁰⁴ , acides sesquiterpéniques polyhalogénés ¹⁰⁵	Insecticide, acaricide ¹⁰⁰ , antioxydant ⁹⁸ , antiviral modéré ¹⁰⁶ , antiprotozoaire ⁴²
<i>P. palmata</i> (Linnaeus) F. Weber, D. Mohr	Lipides (phospholipide, diglycéride, galactolipide) ¹⁰⁷ , stérols ¹⁰⁸ , monoterpènes halogénés ^{71,72,109-112} , sesquiterpènes halogénés ¹⁰⁵	Anti-inflammatoire ^{107,113} , antioxydant ^{114,115} , antiprolifératif ¹¹⁵ , cytotoxicité ^{71,72} , anti-viral ¹⁰⁶ , répulsif ⁷⁵

Les mots-clés utilisés pour la recherche bibliographique sont : « genre botanique » puis « espèce + genre » lorsque les données disponibles le permettaient. Les bases de données interrogées sont ScienceDirect, Pubmed et Google Scholar.

III. Extraits obtenus

Le thalle de chaque algue a été extrait à l'AcOEt dans les mêmes conditions que les extraits obtenus lors des travaux antérieurs mentionnés ci-avant⁴² (Cf. Partie Expérimentale : Préparation des extraits bruts). Les rendements d'extraction sont faibles, comme attendu avec le solvant employé, et varient entre 0.39 et 2.44 % (Tableau 4). Ils sont également variables entre les lots de la même espèce. Les deux lots de *Gracilaria gracilis* ont été récoltés le même jour mais dans deux localités différentes, en revanche les lots de *Dilsea carnosa* ont été récoltés dans deux localités différentes et à un an d'intervalle. Plusieurs publications montrent la variabilité de compositions des algues en fonction de la saisonnalité ou de leur localité, ce qui pourrait expliquer ces variations^{116,117}.

L'activité anti-*Plasmodium in vitro* de chaque nouveau lot a été systématiquement évaluée.

Tableau 4 : Rendements d'extraction obtenus.

Espèces	Code	Masse sèche (g)	Masse extraite (g)	Rendement (%)
<i>Mastocarpus stellatus</i>	MB1-30-1	214.96	1.664	0.77
<i>Callithamnion tetricum</i>	LM1-19-1	229.30	1.692	0.77
<i>Dilsea carnosa (lot 1)</i>	LM1-33-1	126.43	1.611	1.27
<i>Dilsea carnosa (lot 2)</i>	LM1-115-1	357.31	2.622	0.73
<i>Gracilaria gracilis (lot 1)</i>	LM1-49-1	152.58	0.596	0.39
<i>Gracilaria gracilis (lot 2)</i>	LM1-88-1	157.46	2.531	1.61
<i>Plocamium cartilagineum</i>	LM1-57-1	113.51	3.264	2.44
<i>Palmaria palmata</i>	LM1-197-1	25.908	0.354	1.36

Le thalle pulvérisé de chaque espèce a été extrait à l'acétate d'éthyle (10% m/v) puis évaporé à l'évaporateur rotatif (35°C). Les extraits obtenus ont été séchés au Genevac® pour éliminer les traces de solvant et conservés à -20°C jusqu'à leur évaluation dans les tests biologiques.

IV. Evaluation biologique des algues

IV. 1. Activité antipaludique *in vitro*

L'activité antipaludique des extraits d'algues est évaluée à l'aide d'un test immunoenzymatique. Il consiste en l'immunodétection de la Lactate Déshydrogénase de *Plasmodium* (pLDH) à l'aide d'un kit ELISA. La pLDH est une protéine sécrétée par le parasite. Sa sécrétion reflète la **viabilité** de *Plasmodium*. Nos tests d'activités sont réalisés sur la souche multi-résistante *P. falciparum* 7G8 (chloroquine, pyriméthamine, et cycloguanine) à deux concentrations 10 et 100 µg/mL. L'activité est exprimée sous la forme d'un pourcentage d'inhibition de l'activité de la pLDH. Ce test est réalisé au sein de l'Institut de Parasitologie et Pathologie Tropicale à la faculté de médecine de l'Université de Strasbourg. Cet établissement est habilité pour la culture de *Plasmodium* (Collab. E. Candolfi), à laquelle j'ai été formée afin de réaliser ces tests. Sa proximité avec notre laboratoire facilite les allers-retours entre le fractionnement bio-guidé des extraits actifs et les tests antipaludiques.

Cette évaluation est complétée par un second test antipaludique *in vitro*, le test de référence radioisotopique¹¹⁸ qui consiste en la mesure de l'incorporation d'hypoxanthine tritiée par le parasite. En effet, l'hypoxanthine est un précurseur de base purique nécessaire à la synthèse nucléique du parasite. Une mesure de CI₅₀ est réalisée et ce sont les souches

multirésistantes de *P. falciparum* K1 (chloroquine, pyriméthamine, cycloguanine et sulfadoxine) et sensibles NF54 qui sont employées dans ce test. Celui-ci est réalisé sur la plateforme de criblage du *Swiss Tropical and Public Health Institute* (Swiss TPHI) de Bâle, en Suisse (Collab. M Kaiser).

Ces deux tests sont complémentaires car ils permettent d'évaluer deux paramètres différents : le premier la **viabilité** du parasite et le second, sa **division cellulaire**. De plus, en fonction du mécanisme d'action mis en jeu, ces tests peuvent répondre différemment¹¹⁹. La stratégie du laboratoire est de mener ces deux tests en parallèle afin de confirmer les touches mais également pour n'en manquer aucune.

La cytotoxicité est systématiquement mesurée vis-à-vis de cellules de mammifères L6 (myoblastes de muscle squelettique de rat¹²⁰) ce qui permet de vérifier que l'activité antipaludique n'est pas due à une cytotoxicité générale. En effet, le rapport des deux CI_{50} mesurées dans le test antipaludique d'une part, et dans le test de cytotoxicité d'autre part, permet d'obtenir l'indice de sélectivité (IS). Il est généralement admis que lorsque l'IS est supérieur à 5, l'activité antipaludique est sélective de *Plasmodium*.

Le test radioisotopique indique que l'ensemble des extraits algaux montre une bonne activité antipaludique (CI_{50} de l'ordre de 10 $\mu\text{g/mL}$) sans cytotoxicité ($43 \mu\text{g/mL} < CI_{50} < 50 \mu\text{g/mL}$, $4.7 < IS < 10$), sauf pour l'extrait AcOEt de *Plocamium cartilagineum* qui montre une cytotoxicité non négligeable ($CI_{50} = 10.4 \mu\text{g/mL}$, $IS = 1.06$). L'activité antipaludique des extraits est équivalente vis-à-vis des deux souches chloroquino-sensibles NF54 et chloroquino-résistantes K (Tableau 5).

Le test immunoenzymatique, bien que réalisé à deux concentrations uniquement, confirme cette tendance, ainsi qu'une bonne reproductibilité entre deux lots de la même espèce.

Par conséquent, comme les espèces d'algues rouges précédemment étudiées, ces nouveaux extraits viennent appuyer notre hypothèse.

Tableau 5 : Activités antipaludiques et cytotoxiques *in vitro* des différents extraits obtenus.

Espèce	Screening pLDH (7G8)		CI ₅₀ (µg/mL) 3 ^H -Hypoxanthine		Cytotoxicité	Indice de sélectivité	
	100 µg/mL	10 µg/mL	NF54	K1	Cellules L6		
<i>M. stellatus</i>	84 %	3 %	11.4	4.76	47.6	10	
<i>D. carnosa</i>	Lot 1	75 %	26 %	9.75	8.07	43.8	5.4
	Lot 2	98 %	23 %	ND	ND	ND	ND
<i>C. tetricum</i>	81 %	15 %	13.9	6.75	50.4	7.5	
<i>G. gracilis</i>	Lot 1	85 %	14 %	9.74	ND	45.4	4.7
	Lot 2	88 %	5 %	ND	ND	ND	ND
<i>P. cartilagineum</i>	93 %	6 %	9.75	ND	10.4	1.06	
<i>Chloroquine</i>	100 %	100 %	0.004	0.076	ND	ND	

Tous les extraits algaux montrent une activité antipaludique. *Plocamium cartilagineum* possède une forte cytotoxicité. L'activité antipaludique des extraits a été évaluée *in vitro* soit par un test immunoenzymatique mesurant l'inhibition de la sécrétion de la lactate déshydrogénase de *Plasmodium* (pLDH) à deux concentrations (10 et 100 µg/mL), soit par un test radioisotopique mesurant l'inhibition de l'incorporation d'hypoxanthine tritiée, en présence et en absence de concentrations croissantes d'extrait algal (7 concentrations, dilution au 1/3 de 90 à 0,123 µg/ml). La cytotoxicité est mesurée par un test de viabilité cellulaire sur des cellules L6 (myoblastes de muscle squelettique de rat). Indice de sélectivité : CI₅₀ cytotoxicité L6/CI₅₀, *P. falciparum* souche K1.

Afin de poursuivre nos investigations, les extraits ont été fractionnés par MPLC par un gradient de polarité (Heptane-CH₂Cl₂-MeOH-lavage, Cf. partie expérimentale III. 1) afin d'obtenir 5 fractions. Ces fractions ont également été évaluées *in vitro* sur *P. falciparum* 7G8 par criblage à deux concentrations (Tableau 6).

On peut remarquer que la distribution de l'activité est sensiblement la même pour tous les extraits. Elle est le plus souvent concentrée dans les fractions 2 et 3, sauf pour l'extrait de *P. cartilagineum*. Cet extrait est également le plus cytotoxique. Une répartition de l'activité sur plusieurs fractions laisse supposer qu'il est possible d'avoir plusieurs molécules actives.

Tableau 6 : Activités antipaludiques *in vitro* des fractions obtenues par MPLC à partir des extraits actifs.

Espèce	Fractions	% d'inhibition	
		100 µg/mL (%)	10 µg/mL (%)
<i>M. stellatus</i>	F1	16	7
	F2	96	4
	F3	96	12
	F4	4	0
	F5	0	0
<i>D. carnosa</i> Lot 2	F1	26	10
	F2	98	9
	F3	82	9
	F4	82	6
	F5	11	3
<i>C. tetricum</i>	F1	26	8
	F2	98	24
	F3	98	17
	F4	53	2
	F5	32	12
<i>G. gracilis</i> Lot 2	F1	15	12
	F2	59	11
	F3	83	19
	F4	23	3
	F5	32	0
<i>P. cartilagineum</i>	F1	99	32
	F2	99	8
	F3	98	5
	F4	3	1
	F5	16	17

Le test d'activité employé est le test pLDH (*P. falciparum*, souche 7G8). Cinq fractions ont été obtenues par MPLC à partir de chaque extrait AcOEt actif (gradient Heptane-CH₂Cl₂-MeOH-lavage). En bleu, les fractions les plus actives. L'activité antipaludique est généralement distribuée sur les fractions 2 et 3 sauf pour l'extrait de *Plocamium cartilagineum*. F : fraction.

IV. 2. Inhibition de la voie des isoprénoïdes et de la prénylation des protéines

Une première tentative d'analyse différentielle des extraits d'algues rouges antipaludiques avait été entreprise par spectrométrie de masse, préalablement à ce travail et en collaboration avec Jean-Luc Wolfender (Université de Genève). Ces travaux, réalisés par le Dr Lise Bréant, suggéraient que les molécules communes à ces extraits et porteuses de l'activité antipaludique mesurée puissent appartenir à la famille des terpènes. En effet, dans ces extraits, la teneur d'une demi-douzaine de diterpènes et sesquiterpènes d'identité inconnue corrélait positivement avec l'activité antiplasmodiale¹²¹. Pour rappel, la voie de biosynthèse

des isoprénoïdes héritée de l'algue rouge ancestrale, est exprimée dans l'apicoplaste de *Plasmodium*. Ces résultats nous ont logiquement amenés à évaluer l'activité de nos extraits et fractions d'algues rouges sur la voie de biosynthèse des isoprénoïdes. L'hypothèse sous-jacente étant que ces terpènes d'algues puissent, par proximité structurale avec des substrats et/ou produits, interférer et perturber la voie de biosynthèse des isoprénoïdes exprimée chez *Plasmodium*.

Il existe deux voies de biosynthèse des isoprénoïdes chez les organismes vivants, la voie du mévalonate (voie du MVA) et la voie du méthylérythritol phosphate (voie du MEP). Les produits finaux de ces voies sont l'isopentényl diphosphate (IPP) et son isomère le diméthylallyl diphosphate (DMAPP). Ils sont ensuite impliqués dans la génération de nombreux produits : terpénoïdes, stérols, caroténoïdes, chlorophylles...¹²². Les terpénoïdes constituent notamment un signal d'adressage membranaire essentiel à la fonctionnalité de certaines protéines. Chez les algues, la distribution de ces voies est variable en fonction des espèces, mais elles peuvent aussi présenter les deux voies comme chez les végétaux¹²³. Chez *Plasmodium*, tout comme chez les autres apicomplexes possédant l'apicoplaste, seule la voie du MEP héritée de l'algue rouge ancestrale a été conservée. La voie de biosynthèse des isoprénoïdes chez *Plasmodium* est donc une cible de choix étant donné que cette voie diffère de celle exprimée dans les cellules animales.

Nous avons eu accès à un modèle de screening phénotypique permettant de mettre en évidence des inhibiteurs de la prénylation des protéines. Cette inhibition peut notamment être la conséquence de l'inhibition de la voie du MEP ou de l'étape de prénylation d'une protéine cible. Les tests ont été réalisés au sein de l'équipe *Metabolic Network* de l'UPR CNRS 9032 de l'Institut de Biologie Moléculaire des Plantes (IBMP) à Strasbourg par Quentin Chevalier en collaboration avec le Dr Andréa Hemmerlin. Ce modèle fait appel à des cellules de tabac BY-2 transformées qui expriment la calmoduline de riz couplée à la GFP (*Green Fluorescent Protein*). Cette protéine a été choisie car elle présente un motif d'adressage nucléaire en plus d'un site de prénylation.

Dans le cas où la protéine est prénylée, elle est exportée à la membrane plasmique. Dans le cas contraire (inhibition de la voie du MEP ou de la prényl-transférase) la protéine est adressée au noyau. L'observation en microscopie confocale du phénotype induit (fluorescence au niveau de la membrane ou du noyau) permet donc de mettre en évidence l'effet inhibiteur (noyau) ou l'absence d'effet (membrane) des composés évalués. Afin que ce phénotype ne soit exprimé que lorsque l'on traite les cellules, l'expression du gène codant pour la

calmoduline-GFP est sous le contrôle d'un promoteur inductible à la dexaméthasone qui est ajoutée après 3 heures d'incubation avec les composés à tester (Figure 23).

Il s'agit donc d'un modèle végétal qui reproduit la voie de biosynthèse des isoprénoïdes exprimée chez *Plasmodium*.

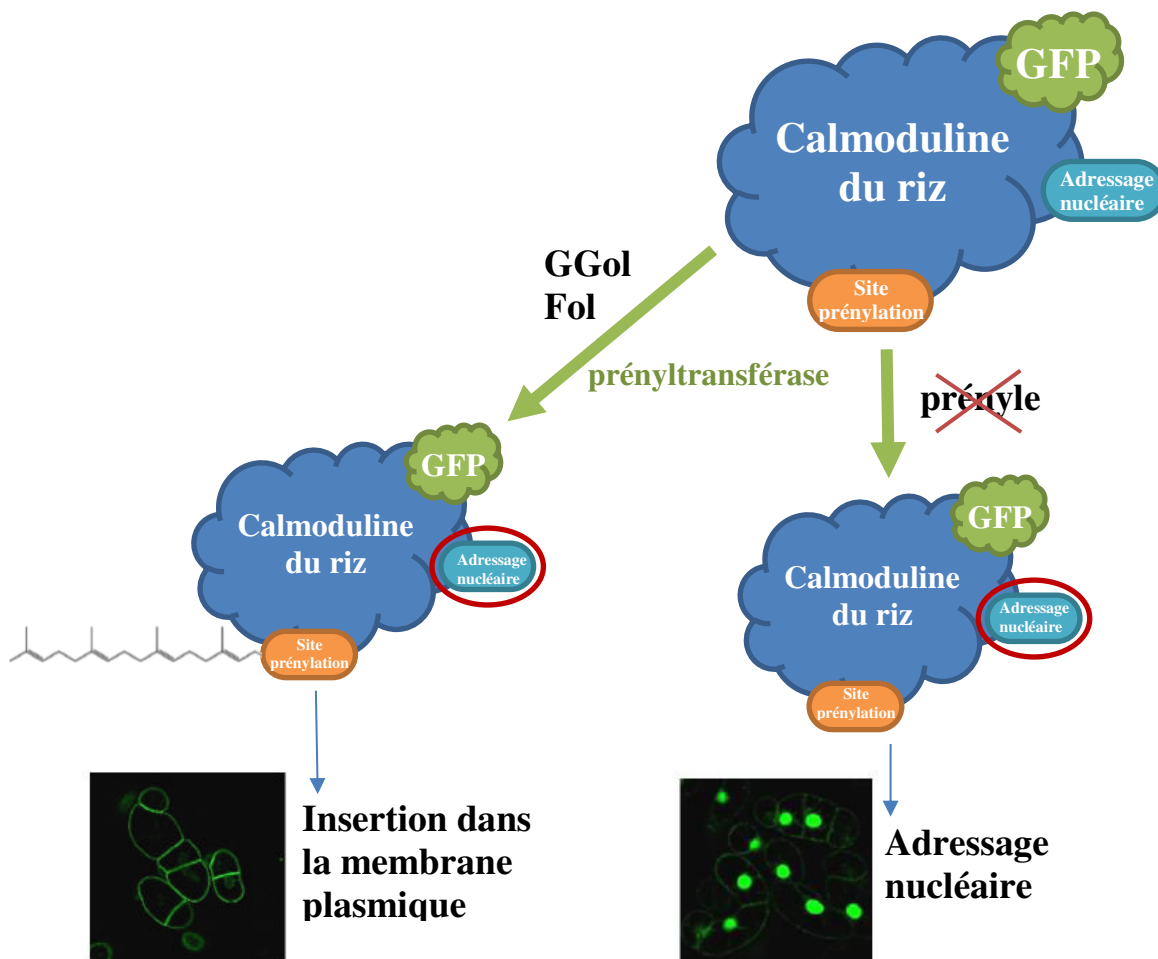


Figure 23 : Principe du test phénotypique permettant de mettre en évidence des inhibiteurs de la voie du MEP ou de la prénylation. GGol = Géranylgéranol (prényl en C₂₀), Fol = Farnésol (prényl en C₁₅), GFP : Green Fluorescent Protein. Si la prénylation est inhibée, l'adressage est nucléaire ; si la prénylation a lieu, l'adressage est membranaire.

Les extraits et les fractions actifs sur *Plasmodium* ont été évalués dans ce modèle (Tableau 7). Un phénotype intermédiaire est possible en cas d'inhibition partielle du processus.

Tableau 7 : Inhibition de la voie du MEP dans un modèle végétal (cellules BY-2 de tabac transformées) par les extraits et fractions les plus actives vis-à-vis de *P. falciparum*.

Espèce		Inhibition de la voie des isoprénoïdes		
		% phénotypes noyaux	% phénotypes membranes/noyaux	% phénotypes membranaires
<i>M. stellatus</i>	Extrait	ND	ND	ND
	F2	0.3	22.8	76.9
	F3	0.5	21.7	77.9
<i>D. carnosa</i> Lot 2	Extrait	0.8	5.5	93.7
	F2	0.8	16.1	83.1
	F3	0.6	8.7	90.7
	F4	ND	ND	ND
<i>C. tetricum</i>	Extrait	0.6	6.4	93.0
	F2	ND	ND	ND
	F3	0	5.7	94.3
<i>G. gracilis</i> Lot 2	Extrait	0.8	25.4	73.8
	F3	ND	ND	ND
<i>P. cartilagineum</i>	Extrait	0.6	19.1	80.3
	F1	0.1	24.4	75.5
	F2	0	5.3	94.7
	F3	0.5	17.0	82.5
Contrôle positif Fosmydomycine		98.6	1.4	0
Contrôle négatif Géranylgéranol		0	1.3	98.7

L'évaluation de l'inhibition de la voie du MEP fait appel à des cellules de tabac BY-2 transformées qui expriment la calmoduline de riz couplée à la GFP. Dans le cas où la protéine est prénylée, elle est exportée à la membrane plasmique. Dans le cas contraire (inhibition de la voie du MEP ou de la prényltransférase) la protéine est adressée au noyau. Les extraits AcOEt de chaque espèce, ainsi que les fractions obtenues à partir de ces extraits ont été testés à la concentration de 40 µg/mL. Contrôle positif : fosmydomycine (inhibiteur de la prényltransférase), contrôle négatif : Géranylgéranol (substrat de la prényltransférase). ND : Non déterminé. Des phénotypes intermédiaires, signifiant une inhibition partielle sont observés pour plusieurs fractions et extraits testés.

Les échantillons testés se sont montrés capables d'inhiber partiellement la prénylation de la protéine cible dans le modèle végétal employé. En effet, des phénotypes intermédiaires sont observés indiquant que la protéine cible n'est que partiellement adressée à la membrane plasmique en présence de ces échantillons. Des molécules algales sont donc capables d'interférer avec le processus de prénylation dans ce modèle végétal.

Il conviendrait d'identifier l'étape du processus mise en jeu en reproduisant ces expériences en présence de différents inhibiteurs ou substrats des enzymes potentiellement ciblées par nos échantillons (expérience de *rescuing*). La prényltransférase végétale est disponible sous forme recombinante. Cependant, à l'heure actuelle, l'enzyme plasmodiale correspondante n'est pas encore disponible (ni par purification, ni de façon recombinante).

Néanmoins, cette activité, bien que partielle, est encourageante et renforce l'hypothèse que des métabolites algaux puissent interférer dans les voies métaboliques exprimées dans l'apicoplaste.

V. Conclusion

Nous avons sélectionné huit nouveaux échantillons d'algues rouges qui ont été collectés et caractérisés botaniquement. Les extraits obtenus ont été évalués *in vitro* sur des cultures de *Plasmodium* dans des globules rouges humains. L'activité antipaludique des espèces précédemment actives a été retrouvée. Les fractions générées à partir de ces extraits en vue d'en identifier les molécules actives ont également été évaluées pour leur activité antipaludique. Celle-ci se répartit majoritairement entre deux fractions successives (fractions 2 et 3) sauf pour *P. cartilagineum*. Cette répartition suggère qu'il est possible que plusieurs molécules soient responsables de l'activité antipaludique mesurée. L'activité antipaludique des algues rouges est une nouvelle fois confirmée.

Les échantillons actifs sur *Plasmodium* ont été évalués pour leur capacité à interférer avec la voie du MEP, qui est la voie de biosynthèse des isoprénoïdes exprimée dans l'apicoplaste chez *Plasmodium*. Ces échantillons ont montré qu'ils étaient capables d'inhiber de façon partielle le processus de prénylation découlant de cette voie. Le modèle utilisé est un modèle végétal (cellules transformées de tabac) et ces résultats devront être reproduits chez *Plasmodium* pour être validés. Cependant, il n'existe pas à l'heure actuelle de modèle plasmodial reproduisant cette voie de biosynthèse. Ce résultat suggère que des métabolites des algues rouges inhiberaient la voie métabolique des isoprénoïdes, voie spécifique de l'apicoplaste chez *Plasmodium*, ou un processus en aval de cette voie, ce qui vient renforcer l'hypothèse de départ de notre projet, **à savoir que les algues rouges représentent une source privilégiée d'actifs antipaludiques susceptibles d'agir sur des voies métaboliques reliques héritées de l'algue ancestrale.**

L'identification de ces composés algaux doit être réalisée par une stratégie analytique innovante adaptée. Cette stratégie repose sur une analyse différentielle de spectres RMN convergeant vers l'empreinte spectrale de la ou des molécules porteuses de l'activité biologique (pharmacophore). Nous parlerons de la « méthode de déconvolution pharmacophorique ».



CHAPITRE III

Chapitre III : Automatisation de l'analyse différentielle par RMN

Ce chapitre porte sur la méthode de déconvolution pharmacophorique d'extraits algaux bioactifs et en particulier sur l'automatisation de l'analyse différentielle de spectres RMN de ces extraits.

Dans un premier temps, nous développerons les méthodes d'identification de molécules naturelles bioactives dans les extraits végétaux et la raison de développer une méthode alternative. Puis, nous présenterons les travaux d'analyse différentielle de spectres RMN du Dr F. Schroeder. L'application de sa stratégie à nos échantillons a été évaluée par une expérience préliminaire à partir de l'extrait de *Bifurcaria bifurcata*. Cette expérience a soulevé plusieurs limites qui ont conduit aux objectifs suivants : i) automatiser l'analyse différentielle par RMN, ii) diversifier le type d'analyses RMN pour enrichir l'information structurale.

Nous présenterons les outils informatiques et les expériences développés qui ont permis l'automatisation de l'analyse différentielle de données RMN, notamment par la création du programme Plasmodesma et du script '*Bucket Utilities*'.

Ces outils d'analyse différentielle automatique seront ensuite appliqués à deux échantillons bioactifs artificiels, recomposés à partir d'extraits algaux inactifs enrichis en une molécule antipaludique connue, soit la chloroquine, soit l'artémisinine.

Les travaux issus de ce chapitre ont donné lieu à un article dans la revue *Magnetic Resonance in Chemistry* en 2018 : *Automatic differential analysis of NMR experiments in complex samples*, Laure Margueritte, Petar Markov, Lionel Chiron, Jean-Philippe Starck, Catherine Vonthron-Sénécheau, Mélanie Bourjot, Marc-André Delsuc (**Annexe 7** : Publication).

I. Méthodes d'identification de molécules naturelles bioactives dans les extraits végétaux

Lors de leur évolution, les végétaux ont acquis de nombreux caractères adaptatifs pour pouvoir survivre et se développer dans leur environnement. Ces caractères peuvent être morphologiques et physiologiques. La production de métabolites secondaires résulte de cette évolution et joue un rôle essentiel à leur survie. Ces composés participent à la protection de la plante aux UV, à la sécheresse, aux attaques de pathogènes ou autres prédateurs. Au contraire, leurs influences sur la couleur et l'odeur de la plante peuvent jouer un rôle dans l'attraction des pollinisateurs. Ces métabolites particuliers présentent parfois des structures complexes et offrent ainsi une diversité chimique étendue. Ces molécules sont capables d'agir sur diverses cibles biologiques et représentent une source de composés à activité thérapeutique¹²⁴.

En effet, lors des 30 dernières années, sur la totalité des nouveaux médicaments approuvés et composés de petites molécules, 35% des médicaments étaient d'origine synthétique contre 32% d'origine naturelle ou dérivés (6% et 26% respectivement). A cela, s'ajoutent 27% d'une catégorie de produits de synthèse qui imitent des produits naturels. La place des molécules naturelles dans le domaine pharmaceutique reste importante¹²⁵. Cependant, celle-ci est en déclin en raison de la complexité d'isolement et de mesure de l'activité biologique des composés. Un autre facteur s'ajoute à cela, la mise en place de réglementations de protection des ressources (Protocoles de Rio et de Nagoya). Ces dernières années, la recherche en produits naturels s'est efforcée de pallier ces difficultés par la mise en place de nouvelles stratégies analytiques reposant notamment sur la dérégulation. Nous présenterons ici, les méthodes classiques et actuelles d'identification de composés naturels actifs.

I. 1. Techniques analytiques

Les extraits végétaux sont des mélanges complexes de molécules. L'identification de ces métabolites résulte de l'utilisation de techniques séparatives et de détection.

- **Techniques séparatives**

Les techniques chromatographiques sont souvent les plus utilisées pour la séparation de molécules d'un mélange complexe.

La chromatographie sur couche mince (CCM) est une technique simple et économique. Elle a l'avantage de permettre une visualisation rapide et simultanée des

composés de plusieurs échantillons. Elle peut être utilisée de manière analytique (qualitative) ou préparative (séparative). Elle peut être automatisée par un système appelé HPTLC (*High-performance thin-layer chromatography*). Elle est ainsi plus reproductible et précise. La CCM est une méthode classique recommandée par la pharmacopée européenne pour la validation de matière première végétale à destination de produits pharmaceutiques.

La chromatographie liquide moyenne pression (MPLC) est une méthode de chromatographie liquide à pression modérée (< 40 bars). Plusieurs types de phases stationnaires sont adaptables (silice, C¹⁸...). C'est une technique uniquement préparative qui est peu résolutive et permet des fractionnements grossiers.

La Chromatographie Liquide à Haute performance (HPLC) est largement répandue en raison de sa polyvalence, son efficacité et sa reproductibilité. La variété de phases stationnaires adaptables permet une séparation de composés aux propriétés physico-chimiques diverses¹²⁶. Une technique dérivée, l'UPLC (*Ultra Performance Liquid Chromatography*) est un système fonctionnant à très haute pression. Les phases stationnaires sont composées de particules plus fines qu'une phase pour HPLC, ce qui améliore la sensibilité, la résolution et la vitesse d'une analyse¹²⁷.

La chromatographie en phase gazeuse (GC) est souvent utilisée en analyse de produits naturels pour l'étude des composés volatils de petite masse moléculaire. Cette technique est particulièrement adaptée aux analyses des huiles essentielles et des parfums. Des dérivations chimiques (silylation, méthylation, estérification...) sont possibles pour étudier des composés non volatils trop polaires ou thermolabiles¹²⁸.

Par ailleurs, il existe des techniques de séparations liquide-liquide telle que la chromatographie de partage centrifuge (CPC). Elles ont l'avantage de limiter les dégradations ou les adsorptions définitives des composés sur les phases stationnaires et de pouvoir séparer ou purifier des quantités importantes de produit¹²⁹.

- **Techniques de détection**

Les techniques séparatives sont généralement reliées en ligne à des détecteurs.

Le détecteur le plus classique est le détecteur UV. Il mesure l'absorbance d'une molécule à une longueur fixe, variable ou multiple (détecteur à barrette de diode (*diode array detector*, DAD). Il détecte les molécules présentant un chromophore.

Le détecteur à diffusion de lumière (DEDL) mesure l'intensité de la lumière diffusée par une molécule. C'est une technique de détection quasi-universelle.

La spectrométrie de masse (MS) détecte les molécules par mesure de leur masse. Elle permet d'ioniser les molécules ce qui peut provoquer leur fragmentation. Un spectromètre de masse comprend en général, une source d'ionisation, un analyseur de masse qui sépare les ions selon leur m/z (masse/charge) et le détecteur, qui traduit les signaux électriques des ions en format final d'un spectre de masse.

Les principales sources d'ionisation utilisées sont l'ionisation par électrospray (*electrospray*, ESI), l'ionisation chimique à pression atmosphérique (APCI) et la désorption-ionisation assistée par matrice (MALDI). Cette dernière est plutôt réservée aux études des biomolécules ou des grosses molécules organiques. Les analyseurs les plus courants sont les analyseurs quadripolaires (Q), à temps de vol (TOF, *Time of Flight*) et les analyseurs à trappe ionique. C'est une méthode très sensible. La spectrométrie de masse haute résolution (HRMS) permet d'obtenir la formule brute d'un composé. Elle peut être réalisée en tandem (ou MS/MS ou MS²) par le couplage de deux analyseurs ou par l'utilisation d'une trappe ionique, ce qui permet d'accéder au m/z des ions parents et des fragments. L'étude des fragmentations et des réarrangements structuraux, ainsi que les signatures isotopiques caractéristiques de certains atomes sont essentielles à l'élucidation des structures chimiques. La comparaison de ces résultats avec les bases de données permet d'identifier rapidement des composés.

La spectroscopie à résonance magnétique nucléaire (RMN) est une technique quantitative et non destructive. Elle est très informative et obligatoire pour l'élucidation structurale d'une nouvelle molécule. Cette spectroscopie est fondée sur les propriétés magnétiques de certains noyaux atomiques (¹H, ¹³C, ¹⁷O, ¹⁹F, ³¹P...). Plusieurs analyses à une ou deux dimensions ont été développées, permettant de réaliser l'élucidation structurale d'une molécule. Les spectres des noyaux ¹H et ¹³C sont les plus étudiés.

Les spectres RMN ¹H fournissent quatre types d'informations :

- les déplacements chimiques en ppm. Ils représentent la fréquence de résonance des noyaux qui dépend de leur environnement électronique ;
- la multiplicité des signaux. Elle met en évidence le nombre de spins voisins ;

- Les constantes de couplage en Hertz. C'est la mesure de l'interaction entre deux spins, elle est influencée par le nombre, le type de liaison entre les noyaux et leurs environnements chimiques ;
- L'intégration : l'aire sous le signal est proportionnelle au nombre de noyaux (^1H).

Les analyses RMN ^{13}C sont moins informatives. Elles renseignent sur le déplacement chimique des noyaux ^{13}C . En général, les couplages $^1J_{\text{CH}}$ sont supprimés pour obtenir une raie unique correspondant à chaque carbone. Cependant, des séquences particulières appelées DEPT (*Distortionless Enhancement by Polarization Transfer*) permettent de déterminer le nombre de protons portés par le carbone. La RMN du ^{13}C est une expérience beaucoup moins sensible que le ^1H en raison de sa faible abondance naturelle (1.1%) et de son faible rapport gyromagnétique¹³⁰.

Ces analyses sont complétées par des analyses 2D qui enrichissent l'information pour mener à une élucidation structurale. Ces expériences permettent d'observer les couplages entre les noyaux. Il peut s'agir d'expériences homonucléaires ^1H - ^1H ou hétéronucléaires ^1H - ^{13}C (Tableau 8).

L'expérience ^1H - ^1H COSY (*CORrelation SpectroscopY*) est très répandue et montre les corrélations entre deux protons voisins de deux ou trois liaisons. Il existe également la DQF-COSY (*double quantum filter*). Cette expérience diminue l'intensité des signaux de la diagonale, elle est plus résolue et les multiplets sont visibles.

Les expériences ^1H - ^1H TOCSY (*TOTAL Correlation SpectroscopY*) permettent d'individualiser un système de spins. Les corrélations des protons sont visibles jusqu'à 5 liaisons. La présence d'hétéro-atomes tels que l'oxygène ou l'azote stoppe le transfert ^1H - ^1H TOCSY. Le nombre de liaisons visibles est modulable en ajustant le « spin-lock mixing time (D9) ». Elle est particulièrement utile pour l'étude de peptides ou de polysaccharides.

Les expériences ^1H - ^1H NOESY (*Nuclear Overhauser Effect SpectroscopY*) utilisent le couplage dipolaire. C'est un couplage direct à travers l'espace qui permet l'étude de la structure 3D d'une molécule.

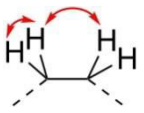
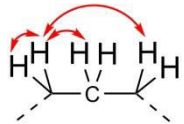
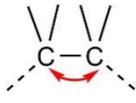
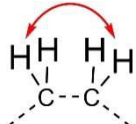
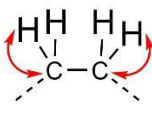
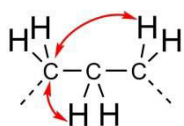
L'expérience ^{13}C - ^{13}C INADEQUATE (*Incredible Natural Abundance Double QUAntum Transfer Experiment*) montre les couplages C-C en 1J . Cette expérience bien que très informative est peu utilisée en raison de son manque de sensibilité.

Concernant les expériences hétéronucléaires, les plus utilisées sont la ^1H - ^{13}C HSQC (*Heteronuclear single quantum coherence spectroscopy*) qui met en évidence les couplages $^1J_{\text{HC}}$, ainsi que les expériences ^1H - ^{13}C HMBC (*Heteronuclear multiple-bond correlation spectroscopy*) qui montrent les couplages $^2J_{\text{CH}}$ à $^4J_{\text{CH}}$.

Ces méthodes sont parfois combinées entre elles, comme les expériences ^1H - ^{13}C HSQC-TOCSY ou ^1H - ^{13}C HSQC-NOESY.

Enfin, les expériences ^1H - δ DOSY (*Diffusion Ordered Spectroscopy*) sont des expériences 2D particulières adaptées aux mélanges. Elles informent sur le déplacement chimique ^1H et sur le coefficient de diffusion d'une molécule qui dépend de sa taille. Elle permet ainsi de distinguer plusieurs molécules en mélange¹³¹.

Tableau 8 : Expériences communes de spectroscopie RMN 2D.

	Expérience	Couplages observés	
Homonucléaire	COSY ^1H - ^1H		Couplage scalaire $^2J_{\text{HH}}$ à $^3J_{\text{HH}}$
	TOCSY ^1H - ^1H		Couplage scalaire système de spin ($^2J_{\text{HH}}$ à $^5J_{\text{HH}}$)
	INADEQUATE ^{13}C - ^{13}C		Couplage scalaire $^1J_{\text{CC}}$
	NOESY ^1H - ^1H		Couplage dipolaire (moins de 5 Å)
Hétéronucléaire	HSQC ^1H - ^{13}C		Couplage scalaire $^1J_{\text{CH}}$
	HMBC ^1H - ^{13}C		Couplage scalaire $^2J_{\text{CH}}$ à $^4J_{\text{CH}}$

I. 2. Le fractionnement bio-guidé

La méthode la plus classique pour isoler un composé naturel bioactif est le fractionnement bio-guidé. La procédure consiste en la réalisation d'un extrait du matériel végétal puis d'une étape de fractionnement validée par un test d'activité biologique. Cette étape de fractionnement est renouvelée jusqu'à l'isolement et la purification de la (les) molécule(s) active(s). La structure de la molécule doit ensuite être caractérisée (Figure 24).

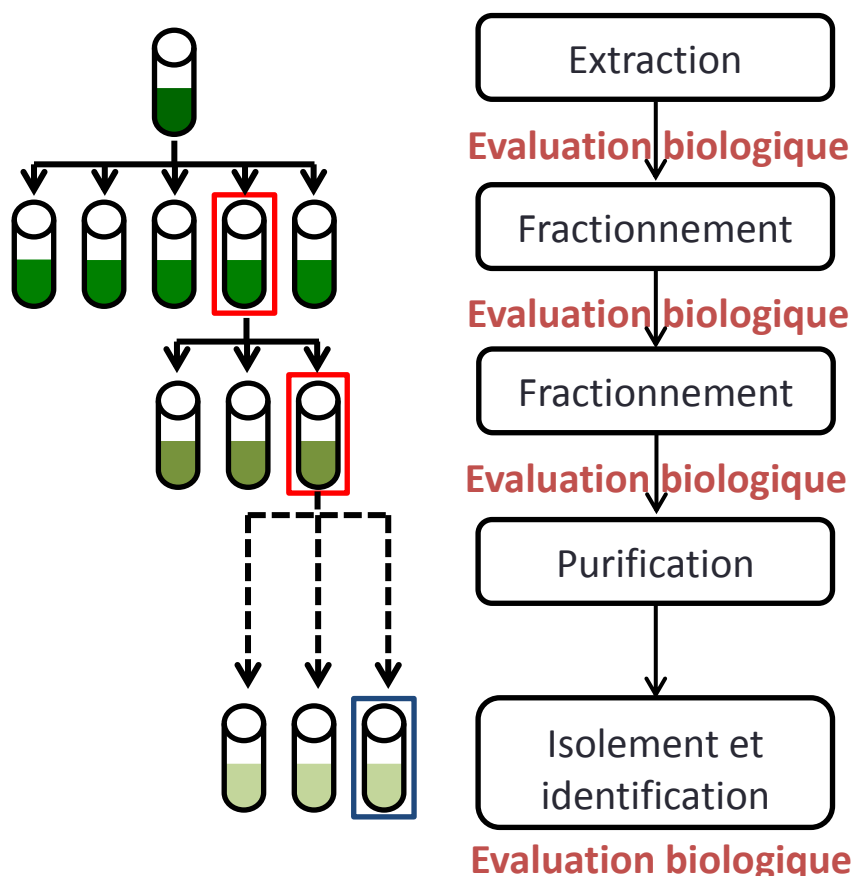


Figure 24 : Fractionnement bio-guidé. L'isolement de molécules actives par fractionnement bio-guidé est long et fastidieux.

Cette méthode est longue et onéreuse, en raison du nombre de tests d'activité biologique à réaliser. De plus, l'activité peut être perdue lors des étapes de fractionnement des composés bioactifs en raison de dégradation ou d'un effet synergique entre plusieurs composés. De plus, la masse de la molécule active est parfois trop faible pour valider l'activité biologique ou pour compléter les données de son activité (évaluation cytotoxique sur diverses cellules, évaluation *in vivo*...). La molécule isolée est parfois connue et disponible commercialement et une simple évaluation biologique de ce composé aurait pu valider sa

nouvelle activité. Aujourd'hui, cette méthode est le plus souvent évitée par des stratégies de déréplication.

I. 3. Les stratégies de déréplication

Depuis une quinzaine d'années, le développement de méthodes d'analyse des mélanges complexes est en plein essor, amorcé par les études protéomiques puis métabolomiques. Ce développement a été possible grâce à l'amélioration de la résolution et de la sensibilité des techniques analytiques, ainsi qu'aux couplages efficaces entre les instruments de séparation et les instruments spectroscopiques. Ainsi, un certain nombre de procédures de « déréplication » ont émergé¹³². Ce terme a été initialement défini dans les années 90, comme « **un processus d'identification rapide de chémotypes connus** »¹³³. Aujourd'hui, le concept de déréplication est élargi à l'identification rapide de substances connues dans un mélange. Ces méthodes reposent généralement sur la suppression ou la réduction des étapes de fractionnement, par l'utilisation d'outils analytiques couplés (LC/MS, GC/MS, LC/RMN) et sur l'utilisation de bases de données.

- **Déréplication par spectrométrie de masse en tandem (MS/MS) et réseaux moléculaires**

La spectrométrie de masse est une méthode sensible et rapide. La déréplication des produits naturels par spectrométrie de masse MS² se démocratise depuis la mise en place de la plateforme libre et accessible en ligne GNPS (*Global Natural Products Social Molecular Networking*) et la possibilité de créer des réseaux moléculaires¹³⁴. Cette stratégie a été développée par Pieter C. Dorrestein et Nuco Bandeira (Université de Californie, San Diego).

Cet outil repose sur le fait que des molécules structurellement proches se fragmentent de façon similaire, conduisant à des fragments ou des pertes de neutres communs. La plateforme GNPS permet d'aligner ces jeux de données et de calculer un indice de corrélation (*cosine score*) entre chaque paire de spectres. Ainsi, deux spectres MS² similaires auront un *cosine score* de 1 et deux spectres sans ion ou perte de neutre commun ont un *cosine score* de 0. Les résultats sont visualisés par le logiciel Cytoscape. L'utilisation de ces réseaux permet de détecter un ensemble d'analogues structuraux à partir d'un spectre MS² de référence, d'identifier des groupes spectraux MS² et de comparer finement les analyses MS de différents échantillons.

La puissance de la plateforme GNPS, en plus de contenir des bases de données de spectres MS², a l'avantage d'être communautaire. Les utilisateurs peuvent déposer leurs spectres expérimentaux et ainsi enrichir la base de données¹³⁵.

Cette méthode est très puissante et rapide mais présente quelques défauts liés aux variabilités d'ionisations dues aux différentes sources d'ionisation (*electrospray* ou ionisation chimique à pression atmosphérique) et aux différences de fragmentation entre les appareils Q-TOF (*quadripole-times of flight*) et trappes à ions. Le choix de l'énergie de collision permettant d'obtenir une fragmentation optimale peut également être une limite lorsque la classe de molécule à identifier est inconnue.

L'identification d'une molécule inconnue reste difficile bien que les données MS² permettent généralement de générer des données suffisantes pour déterminer la famille chimique. Il faut alors passer par des méthodes séparatives plus classiques pour pouvoir isoler le composé. De plus, bien qu'il soit possible de visualiser la provenance des ions en fonction des échantillons, il est impossible de prendre en compte la variation des activités biologiques par cette méthode car l'intensité des pics MS n'est pas prise en compte et ne serait pas représentative.

- **Déréplication par résonance magnétique nucléaire (RMN)**

La résonance magnétique nucléaire (RMN) est une technique informative et obligatoire pour l'élucidation structurale d'une nouvelle molécule. La RMN a longtemps été critiquée pour son manque de sensibilité. Aujourd'hui, elle est largement augmentée grâce aux avancées concernant les spectromètres à très haut champ (750 à 1.2 GHz) et les sondes cryogéniques et capillaires. Elle a également l'avantage d'être quantitative, de détecter des composés qui auraient pu être difficilement ionisables en MS. La RMN permet d'accéder à la structure d'une molécule inconnue non référencée dans une base de données.

L'approche la plus commune en métabolomique a été l'utilisation de spectres ¹H RMN. Or, cette technique présente plusieurs défauts tels que la variation des déplacements chimiques dépendant du solvant ou du pH de la solution, et le nombre de recouvrements des signaux dans un mélange complexe¹³⁶. Ces effets peuvent fortement influencer les procédures de déréplication. Cependant, ce point est corrigé par l'utilisation de spectres 2D et des expériences ¹H-δ DOSY¹³⁷.

Les analyses ^{13}C ont été moins utilisées pour l'analyse d'extraits naturels en raison de la plus faible sensibilité de ces expériences. Elles ont pourtant l'avantage de montrer moins de recouvrements de signaux compte tenu de leur fenêtre spectrale plus large (220 ppm). Les améliorations techniques des spectromètres permettent aujourd'hui d'aborder des études de déréplication par ce type d'expériences^{138,139}. En effet, une méthode de déréplication a été développée à l'Université de Reims. Premièrement, elle consiste en un fractionnement d'un extrait brut par chromatographie de partage centrifuge en plusieurs fractions qui sont analysées par RMN ^{13}C . Deuxièmement, une détection et un alignement automatique des signaux sont réalisés. Ensuite, à partir de la liste de signaux détectés, une matrice 2D est construite où les colonnes correspondent aux fractions et les lignes correspondent à une fenêtre de déplacements chimiques contenant au moins un signal ^{13}C dans au moins une fraction. Cette matrice est ensuite soumise à une analyse par classification ascendante hiérarchique. Cette méthode détecte des déplacements chimiques similaires entre les fractions et les regroupe. Cela permet une visualisation efficace des signaux de RMN partagés entre plusieurs fractions et appartenant à la même molécule. Enfin, ces groupes de déplacements chimiques sont comparés à une base de données interne de spectres RMN ^{13}C et de spectres ^{13}C prédits à partir des composés décrits dans la littérature sur l'espèce étudiée¹³⁸. Cette méthode est particulièrement efficace pour identifier des métabolites connus et majoritaires dans des mélanges complexes.

Les expériences RMN 2D sont en développement pour les études de déréplication et plus particulièrement, les expériences ^1H - ^{13}C HSQC et ^1H - ^1H TOCSY en raison de l'émergence de bases de données¹⁴⁰⁻¹⁴³. Une base de données appelée COLMAR a été développée par l'Université de l'Etat d'Ohio. Elle est spécialisée dans les analyses RMN ^{13}C - ^{13}C TOCSY, ^{13}C - ^1H HSQC-TOCSY et récemment ^1H - ^1H TOCSY^{140,144}. Elle a ajouté dernièrement les données spectrales de la *Biological Magnetic Resonance Data Bank* et la *Human Metabolome Database*¹⁴⁵⁻¹⁴⁷. Il est possible de comparer ses spectres expérimentaux à cette base de données, l'échantillon devant être préparé selon leur protocole et la liste des déplacements chimiques des signaux entrée manuellement. Cette base de données est plutôt orientée vers les biomolécules.

Par ailleurs, il existe plusieurs bases de données de spectres RMN qui ne permettent pas de déréplication. NMRShiftDB est une plateforme en ligne intéressante qui contient 42 000 structures et qui peut aider à l'identification de molécules organiques. Elle est libre d'accès et communautaire. Elle peut aussi prédire des déplacements chimiques en RMN du

^{13}C et ^1H . D'autres bases de données spectrales libres d'accès existent telles que CSEARCH, NAPROC-13 ou SpecInfo.

Aujourd'hui, il n'existe pas de plateforme équivalente à GNPS capable de faire des déréplications de mélanges complexes à partir d'analyses RMN.

- **Déréplication par RMN et MS**

Une nouvelle approche appelée SUMMIT MS/NMR (*Structure of Unknown Metabolomic Mixture components by MS/NMR*) a été développée¹⁴⁸. Cette méthode à l'originalité de combiner à la fois la spectroscopie de masse et la RMN. Elle consiste en l'analyse d'un échantillon par HRMS, pour conduire aux formules brutes de chaque composant. Toutes les structures possibles pour chaque formule sont ensuite générées, ce qui constitue une liste très importante de composés. Un spectre RMN prédit est généré pour chaque composé listé. Le spectre RMN expérimental du mélange est également acquis. Les déplacements chimiques des spectres prédits et du spectre expérimental sont comparés et les molécules sont ainsi identifiées.

Bien que cette approche soit intéressante, des centaines de structures possibles peuvent être proposées pour chaque formule brute. Générer un spectre prédit pour chacune de ces structures peut être long, en particulier, si le mélange est très complexe. De plus, les structures sont proposées par la base de données ChemSpider, ce qui nécessite que la structure cherchée soit connue¹⁴⁹. La prédiction de spectres RMN est réalisée par le logiciel MestReNova (Mestrelab Research, Santiago de Compostela, Spain) qui n'est pas libre d'accès.

I. 4. Les méthodes couplées

Les couplages des différentes techniques de séparation telles que la GC aux détecteurs MS ou UV existent depuis plusieurs années. Ils permettent aujourd'hui une déréplication rapide des mélanges complexes, comme nous avons pu le voir avec la construction de réseaux moléculaires. Cependant, ces techniques sont généralement insuffisantes pour accéder à la structure d'une molécule nouvelle.

Lors du premier couplage direct HPLC-RMN développé, l'HPLC était directement reliée à la cellule RMN¹⁵⁰. L'élution était stoppée quand le composé se trouvait dans la cellule et uniquement le spectre ^1H était enregistré. Cette technologie souffrait du coût de consommation de solvant deutéré et des signaux trop importants de l'acétonitrile non deutéré

et de l'eau résiduelle. Ensuite, a été commercialisé le système HPLC-SPE-RMN. Une cartouche SPE (*solide phase extraction*) a été ajoutée pour piéger et concentrer les métabolites. Cette méthode est devenue un outil efficace dans la recherche de produits naturels. Elle a plusieurs fois prouvé que l'élucidation structurale complète à partir d'un extrait brut était possible sans pré-purification¹⁵¹⁻¹⁵⁵. Les auteurs s'accordent sur le gain de temps et sur la très faible consommation de la masse de l'extrait. Néanmoins, le coût de l'appareillage n'est pas négligeable.

I. 5. Pourquoi développer une méthode alternative ?

L'évolution des techniques séparatives et analytiques ces dernières années ont rendu possible les études de déréplication. Ces méthodes sont efficaces, mais reposent toujours sur l'identification d'un composé via une base de données. Cela sous-entend qu'il a déjà été isolé et caractérisé. Identifier un composé original ou absent d'une base de données demande à retourner à des méthodes conventionnelles de fractionnement.

De plus, aucune de ces méthodes ne prend en compte les variations d'activité biologique entre un grand nombre d'échantillons.

A titre d'exemple, une analyse par UHPLC-TOF-MS (ESI+) des extraits AcOEt de 3 algues rouges (*Grateloupia turuturu*, *Mastocarpus stellatus* et *Dilsea carnosa*) et de 3 algues brunes (*Sargassum muticum*, *Bifurcaria bifurcata* et *Fucus serratus*) avait été réalisée précédemment au laboratoire par Dr. Lise Bréant (collaboration Pr. J-L. Wolfender, Université de Genève). Une analyse comparative des données de masse en fonction de l'activité antipaludique des extraits avait été établie. Celle-ci soulignait que l'activité était due à la présence d'un groupe de composés dont les activités s'additionneraient. Chez *B. bifurcata*, l'éléganolone, un composé connu pour son activité antipaludique, avait été identifié. Sept autres composés présents dans les extraits pouvaient contribuer à l'activité, 3 étaient de type diterpénique ($C_{20}H_{30}O_2$, $C_{20}H_{28}O_2$ ou $C_{20}H_{32}O_2$) et 4 composés non pas pu être identifiés.

Ainsi, deux options s'offraient à nous ; retourner à une méthode classique de fractionnement ou, **développer une stratégie alternative permettant d'accéder plus directement aux structures des composés actifs. C'est cette dernière que nous avons privilégiée et baptisée « déconvolution pharmacophorique ».** Elle consiste à :

- **Fractionner un extrait actif ;**
- **Evaluer précisément l'activité biologique de chaque fraction ;**
- **Générer leurs profils RMN ;**
- **Procéder automatiquement à l'analyse différentielle des spectres ;**
- **Aboutir à l'empreinte spectrale de la ou des molécules porteuses de l'activité, le pharmacophore de l'extrait.**

Cette approche est une extension des travaux précurseurs de F. Schroeder que nous avons étendus et automatisés.

II. Mise en place de la stratégie analytique (Dr M. Bourjot)

II. 1. Travaux du Dr Frank Schroeder

Il a été observé que la plupart des algues rouges évaluées montraient une activité antipaludique. Une stratégie de déréplication par spectrométrie de masse aurait pu être employée sur ces échantillons, mais des travaux préliminaires au laboratoire ont montré que les composés identifiés comme possiblement actifs ne pouvaient pas être identifiés ainsi.

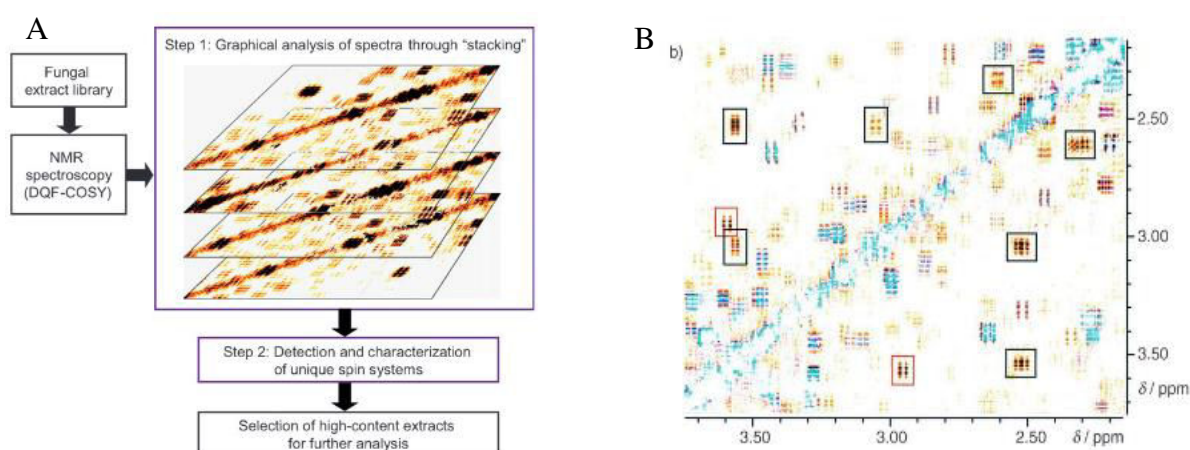


Figure 25 : Stratégie d'analyse différentielle d'après F. Schroeder. Les extraits fongiques sont analysés par ^1H - ^1H DQF-COSY-RMN. Les spectres ont été superposés via Photoshop, les signaux présents uniquement sur un spectre sont détectés (A). Les spectres ^1H - ^1H DQF-COSY sont superposés (B). Les signaux bleus sont les signaux superposés, les signaux marrons ne sont pas superposés, ils sont donc uniquement présents sur un seul spectre.

Dans ce cas, les analyses RMN sont une bonne alternative pour accéder aux structures. Aussi, le Dr M. Bourjot a trouvé intéressant d'appliquer les stratégies d'analyses différentielles de spectres RMN de mélanges développées par le Dr Frank C. Schroeder à notre situation¹⁵⁶. F. Schroeder a étudié les variations de métabolites secondaires du champignon *Tolypocladium cylindrosporum* cultivé selon 7 milieux et/ou protocoles de culture différents. Il a ensuite réalisé 25 analyses ¹H-¹H DQF-COSY (*Double Quantum Filtered*) de ces échantillons (7 extraits AcOET différents de ce champignon en triplicata et 4 contrôles des milieux). Les images du spectre ont été superposées via l'utilisation d'Adobe Photoshop (version CS2) (Figure 25).

Les pixels colorés et superposés changent alors de couleurs et suggèrent la présence du signal sur plusieurs spectres. Les signaux non superposés restent dans leurs couleurs d'origine. Ainsi, quelques signaux qui n'avaient pas subi de changement de couleur ont été détectés et appartenaient à un seul échantillon de champignon. Des analyses RMN ¹H-¹³C HSQC, ¹H-¹³C HMBC, et ¹H-¹H NOESY supplémentaires ont été réalisées et deux structures ont été suggérées. Les structures ont été confirmées à la suite de l'isolement de ces composés. Ils étaient connus mais n'avaient jamais été décrits dans cette espèce.

Ainsi, la technique de F. Schroeder a permis l'identification rapide de deux nouveaux composés. Il fait également part des limites de cette méthode en lien avec les limites de sensibilité du spectromètre RMN. Il propose cependant de la contourner par une étape de préfractionnement en amont de l'analyse RMN. Cela a d'ailleurs été appliqué dans une seconde étude portant sur la comparaison de métabolomes de *Caenorhabditis elegans* mutant et sauvage¹⁵⁷.

Ainsi, l'application de la méthode de F. Schroeder semble bien adaptée à nos échantillons. Cependant, elle ne permet pas forcément d'obtenir une identification structurale complète et doit souvent être complétée par d'autres techniques analytiques telles que la spectrométrie de masse ou un fractionnement classique pour confirmer les résultats. La plateforme analytique de la Faculté de Pharmacie de Strasbourg a la chance de posséder une méthode couplée HPLC-DAD/MS-SPE-RMN qui pourrait justement permettre de limiter le nombre d'analyses supplémentaires à effectuer pour caractériser la molécule identifiée par analyse différentielle et donc de caractériser directement une molécule inconnue. De plus, les analyses différentielles RMN de F. Schroeder reposent uniquement sur des spectres ¹H-¹H

DQF-COSY. Aussi, il nous a semblé intéressant d'appliquer l'analyse différentielle avec d'autres types d'analyses RMN afin d'enrichir l'information structurale.

II.2. Essai de l'analyse différentielle de F. Schroeder sur *Bifurcaria bifurcata* (Dr M. Bourjot)

Une application de cette méthode sur un extrait actif de *Bifurcaria bifurcata* précédemment travaillé au laboratoire a été réalisée.

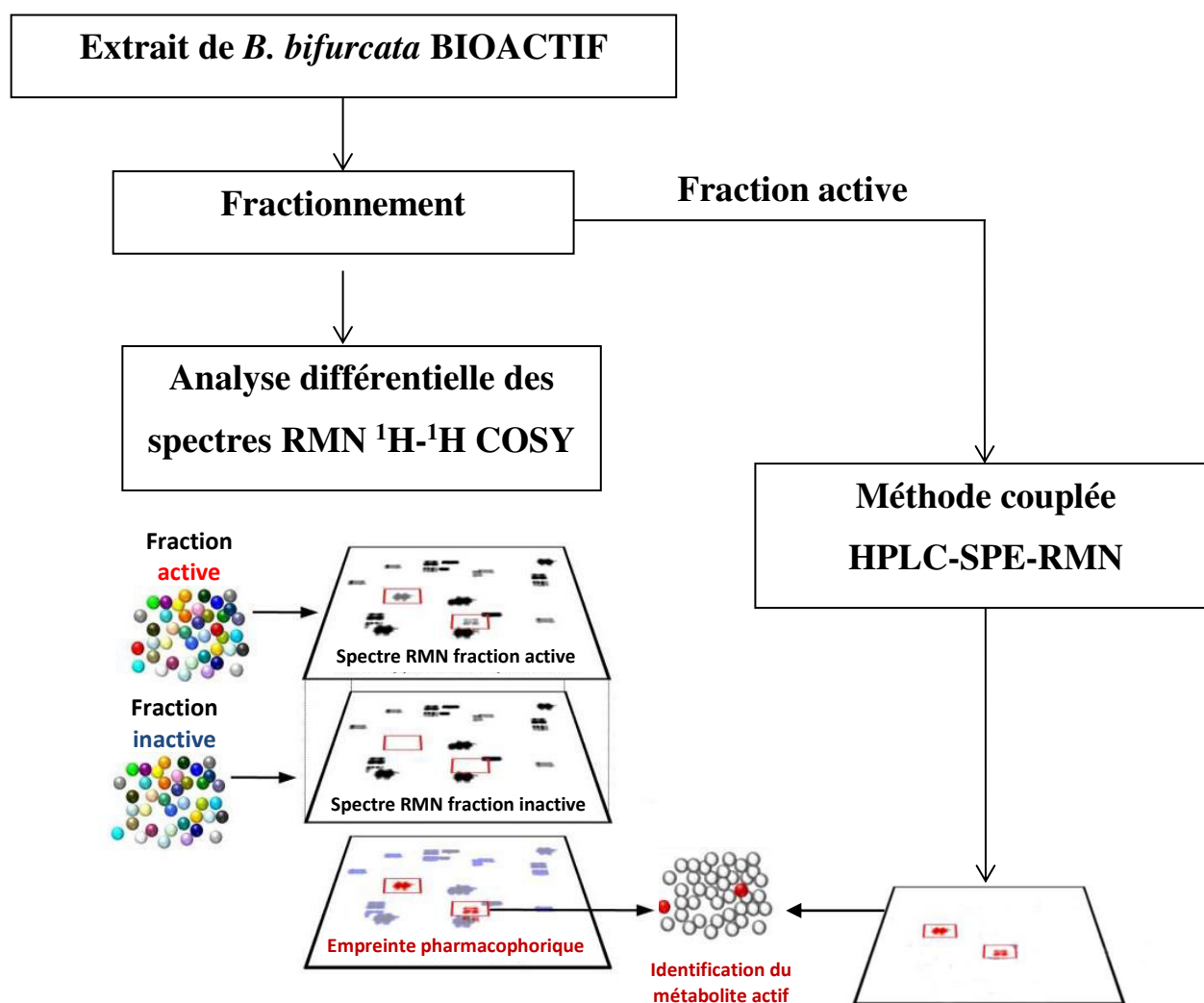


Figure 26 : Méthode d'analyse différentielle de F. Schroeder appliquée à *B. bifurcata*. Après un fractionnement de l'extrait bioactif, l'activité biologique des fractions a été évaluée. Les RMN ¹H-¹H COSY des fractions ont été réalisées et une analyse différentielle des spectres RMN a été effectuée de manière visuelle afin d'obtenir l'empreinte pharmacophorique du métabolite actif. La fraction active a été injectée dans la méthode HPLC-SPE-RMN permettant d'obtenir le spectre RMN d'un métabolite ciblé et de le comparer avec l'empreinte pharmacophorique.

Afin d'évaluer la faisabilité de cette méthode, une étude préliminaire a été réalisée par le Dr Mélanie Bourjot au laboratoire, à partir d'un extrait AcOEt bioactif de *Bifurcaria bifurcata*¹⁵⁸. La stratégie analytique de F. Schroeder est appliquée à nos échantillons mais une étape supplémentaire d'analyse par la méthode couplée HPLC-SPE-RMN est ajoutée (Figure 26).

Cet extrait avait précédemment été travaillé au laboratoire par fractionnement bio-guidé et avait conduit à l'isolement d'un diterpène linéaire, l'**éléganolone**¹²¹. Le but de cette étude est de vérifier qu'une analyse différentielle de cet extrait menait à l'identification de l'éléganolone. Dans cette situation, l'échantillon actif ne peut pas être comparé au même échantillon inactif, comme dans les cas d'étude de F. Schroeder. L'échantillon est donc fractionné pour obtenir des fractions actives et inactives.

L'extrait bioactif a été fractionné afin d'obtenir 5 fractions. Celles-ci ont été évaluées biologiquement par mesure de la pLDH et possédaient une activité antipaludique de F1 à F5 de 12 %, 31 %, 99 %, 96 %, 62 %, respectivement (à 100µg/mL).

Des analyses RMN-2D ¹H-¹H COSY des 5 fractions sont enregistrées. Une analyse différentielle de ces spectres a été réalisée de **manière visuelle** par la comparaison de spectres deux à deux entre les deux fractions actives F3 et F4.

Les taches de corrélation communes aux deux fractions actives ont été sélectionnées et forment l'empreinte pharmacophorique (Figure 27).

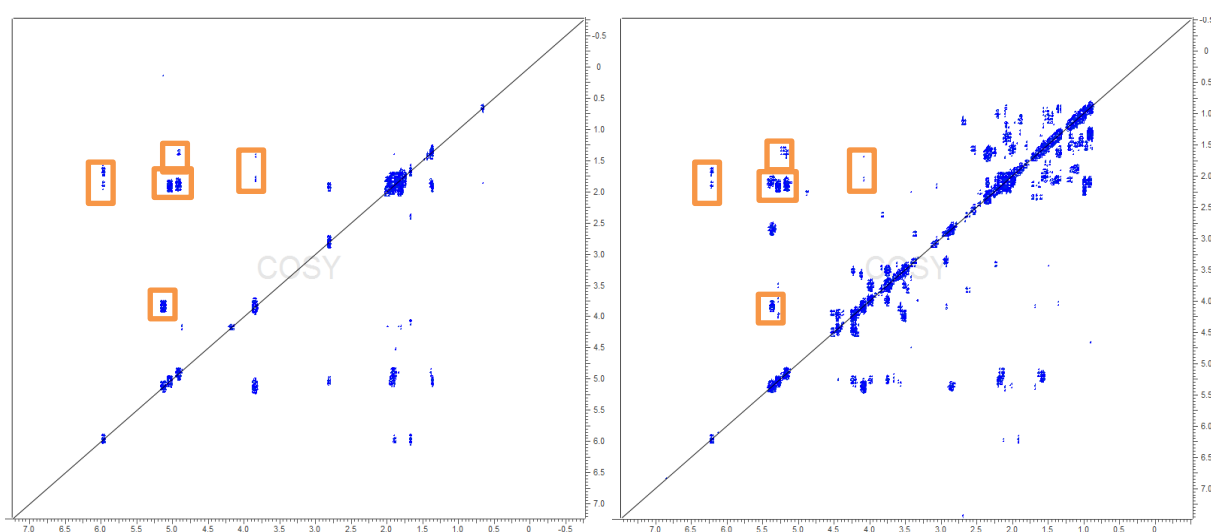


Figure 27 : Analyse différentielle par comparaison visuelle des spectres ¹H-¹H COSY des fractions 3 (A) et 4 (B) actives (CDCl₃, 500 MHz). Les taches de corrélation communes aux deux fractions sont sélectionnées (cadres orange) et représentent l'empreinte pharmacophorique de la molécule active

La fraction 3 de *B. bifurcata* la plus active, est analysée sur la méthode HPLC-SPE-NMR.

Cet équipement est présent à la plateforme analytique de la faculté de Pharmacie (PACSI). Il s'agit d'une HPLC couplée à un détecteur UV et à un spectromètre de masse. Deux chromatogrammes sont générés, un pour l'UV et un pour la masse. Il est ensuite possible de sélectionner les pics du/des chromatogramme(s) et de stocker individuellement chaque molécule correspondante sur une cartouche SPE. Ces cartouches sont ensuite désorbées avec un solvant deutéré directement dans la sonde capillaire du spectromètre RMN. Cela permet d'obtenir les spectres RMN homo- et hétéronucléaires de chaque molécule sélectionnée à partir du chromatogramme UV ou masse.

Le chromatogramme HPLC-UV-MS de la fraction 3 est présenté dans la **Figure 28**. Plusieurs pics ont été sélectionnés et les molécules correspondantes ont été stockées individuellement sur des cartouches SPE.

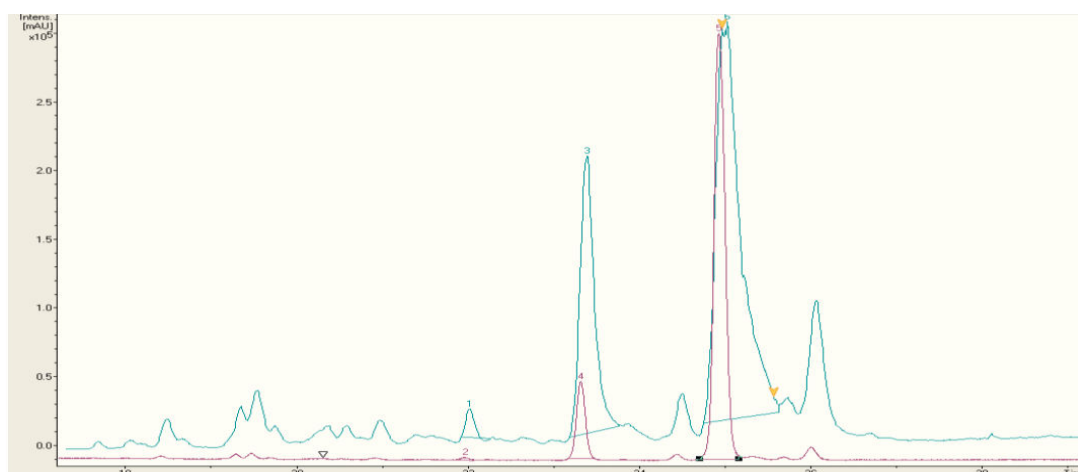


Figure 28 : Chromatogramme HPLC-UV (rose) et MS (bleu) de la fraction 3 de *B. bifurcata*.

Les spectres ^1H - ^1H COSY des différentes molécules stockées sont comparés à l'empreinte pharmacophorique (Figure 29). Un de ces spectres correspondait à l'empreinte pharmacophorique. Les déplacements chimiques des signaux correspondaient à ceux de l'**éléganolone** précédemment isolée. Cependant, *B. bifurcata* est riche en diterpènes linéaires et, il est difficile avec une seule expérience RMN de déterminer s'il s'agit de l'**éléganolone** ou d'un analogue structural¹⁵⁹.

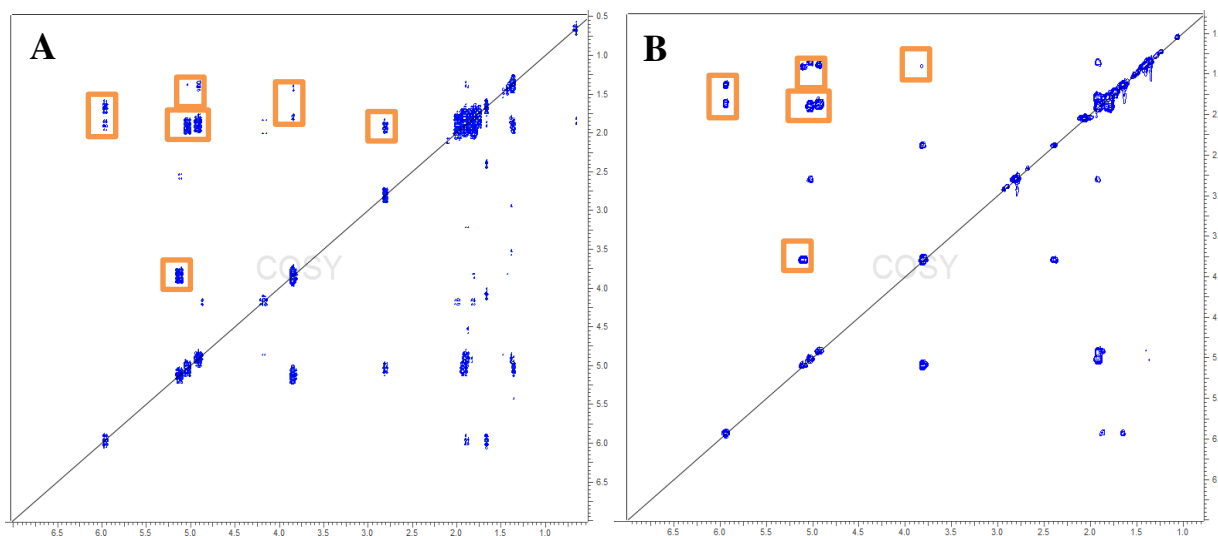


Figure 29 : Comparaison de l'empreinte pharmacophorique de *B. bifurcata* (A) avec le spectre ^1H - ^1H COSY d'une molécule stockée sur SPE (B). De nombreuses taches de corrélation du spectre sont similaires à l'empreinte, il pourrait donc s'agir de la molécule active.

Cette expérience préliminaire a montré les limites suivantes :

- la première, liée à la complexité des données à traiter, est la capacité restreinte de l'analyse différentielle **de manière visuelle** ;
- la seconde difficulté majeure est l'impossibilité de prendre en compte l'intensité des taches de corrélation en RMN-2D. En effet, les logiciels de RMN ne permettent pas de comparer facilement l'intensité de pics de plusieurs échantillons ;
- la troisième est la pauvreté de l'information structurale via un seul type d'analyse RMN.

Par conséquent, les objectifs de cette partie du travail ont été :

- i) d'automatiser l'analyse différentielle de spectres RMN ;
- ii) d'enrichir l'information structurale en diversifiant le type d'analyses RMN.

II. 3. Présentation générale de la méthode de déconvolution pharmacophorique

La méthode d'analyse différentielle appliquée à nos échantillons est détaillée dans la **Figure 30**.

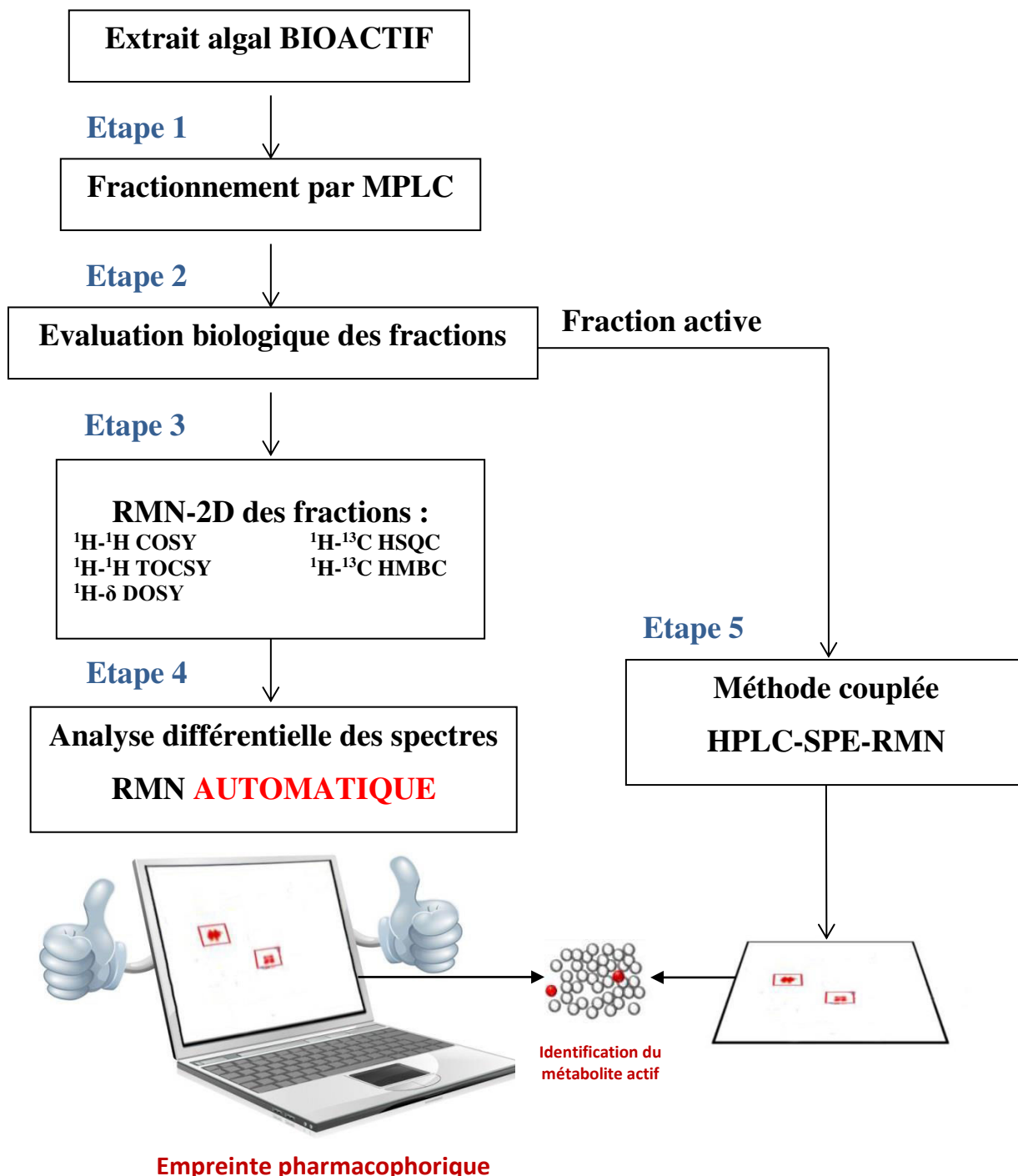


Figure 30 : Méthode de déconvolution pharmacophorique complète. Après un fractionnement grossier de l'extrait bioactif (**Etape 1**), les fractions sont évaluées biologiquement (**Etape 2**). Les analyses RMN-2D des fractions sont réalisées (**Etape 3**) et une analyse différentielle automatique des spectres RMN est effectuée afin d'obtenir l'empreinte pharmacophorique du métabolite actif (**Etape 3 et 4**). La fraction active est analysée par HPLC-SPE-RMN permettant d'obtenir le spectre RMN d'un métabolite ciblé et de le comparer avec l'empreinte pharmacophorique (**Etape 5**).

Etape 1 : Fractionnement préliminaire par MPLC

La première étape est le fractionnement de l'extrait algal bioactif par chromatographie liquide à moyenne pression (MPLC ou Flash) en utilisant un gradient de polarité croissante afin d'obtenir plusieurs fractions. Cette étape de fractionnement est obligatoire afin de concentrer l'activité biologique dans certaines fractions seulement et d'obtenir un échantillon actif et un échantillon inactif avec un recouvrement structural partiel. L'idéal serait d'avoir des fractions de même composition ne différant que par la présence ou non de la molécule porteuse de l'activité biologique recherchée. C'est bien sûr impossible. On peut néanmoins se rapprocher au mieux de cette situation grâce à un fractionnement grossier, qui couvre une large gamme de polarité. En effet, les fractions successives sont ainsi susceptibles de contenir des molécules communes ce qui permettra d'éliminer plus facilement un grand nombre de composants inactifs par analyse différentielle. Cela est d'autant plus intéressant dans le cas d'échantillon végétal où les composés sont souvent présents sous la forme de série.

Etape 2 : Evaluation biologique des fractions

La deuxième étape consiste en l'évaluation biologique de ces fractions. La mesure de l'activité antipaludique a été évaluée *in vitro* à deux concentrations (10 µg/mL et 100 µg/mL) par la mesure de la pLDH.

Etape 3 : Acquisition des spectres RMN-2D

Les analyses RMN-2D des fractions sont enregistrées. Elles doivent se réaliser selon les bonnes pratiques détaillées en partie expérimentale (p. 192). Réaliser une analyse différentielle sur plusieurs types d'analyses permet d'obtenir des informations structurales complémentaires et de multiplier le nombre de données avant une susceptible perte d'information provoquée par l'analyse différentielle.

Etape 4 : Analyse différentielle **automatique des spectres RMN**

Pour contourner les limites observées lors de l'analyse différentielle de *B. bifurcata*, l'analyse différentielle des spectres RMN de manière **automatique** est réalisée via l'utilisation de Plasmodesma, un programme que nous avons spécialement développé pour cette étape et qui sera détaillée à la suite de cette partie.

Les taches de corrélation conservées représentent l'**empreinte pharmacophorique** de la molécule active, c'est-à-dire la trace RMN de la molécule porteuse de l'activité biologique.

Étape 5 : Analyse par méthode HPLC-DAD/MS-SPE-RMN couplée

Il s'agit d'une HPLC couplée à un détecteur UV et à spectromètre de masse. La fraction active est injectée, et deux chromatogrammes sont générés, un pour l'UV et un pour la masse. Il est ensuite possible de sélectionner les pics du/des chromatogramme(s) et de stocker individuellement chaque molécule correspondante sur une cartouche SPE. Ces cartouches sont ensuite désorbées avec un solvant deutéré directement dans la sonde du spectromètre RMN. Cela permet d'obtenir les spectres RMN homo- et hétéronucléaires de chaque molécule sélectionnée à partir du chromatogramme UV ou masse.

Ces spectres peuvent alors être comparés à l'empreinte pharmacophorique et la caractérisation structurale de la molécule peut être réalisée en une fois (données masse, UV et RMN acquises).

III. Automatisation de l'analyse différentielle par RMN¹⁶⁰

Comme vu précédemment, la nécessité d'automatiser l'étape d'analyse différentielle s'est rapidement imposée comme une évidence. Aujourd'hui, les spectromètres RMN sont améliorés pour pouvoir acquérir de manière automatique des données RMN (passeur automatique, tuning de la sonde et acquisition automatique). Cependant, le traitement final de ces informations est moins évident : Transformée de Fourier, phasage, correction de la ligne de base, détection des pics et enfin l'interprétation des spectres. Les études métabolomiques génèrent des avalanches de données dont le traitement automatique est essentiel. Le but de ces études est généralement de vérifier la présence ou l'absence d'un composé ciblé, connu ou de le quantifier.

Dans cette partie, nous présentons le développement du programme informatique Plasmoderma qui permet de manière autonome le traitement d'expériences RMN 1D et 2D. Plasmoderma est relié au programme *Spectrometry Processing Innovative Kernel* (SPIKE) pour la plupart de ces opérations. Ces programmes sont rédigés en langage Python pour une meilleure portabilité.

L'automatisation a été développée en collaboration avec le Dr Marc-André Delsuc (UMR 7104, IGBMC, Strasbourg), dont l'expertise dans le domaine de la RMN, du traitement et de l'exploitation des données qu'elle génère, fait référence.

III. 1. Le langage Python et son environnement

Python est un langage de programmation apparu dans sa première version en 1991. Il a été créé par Guido van Rossum et tient son nom d'un hommage à la troupe de comiques « Monty Python ». Le succès de ce langage repose sur sa facilité d'apprentissage et ses nombreuses fonctionnalités qui lui sont intégrées grâce à son association à des **bibliothèques** (*libraries*). L'utilisation de ce langage varie de l'écriture de petits programmes simples (*script*) à des programmes complets (suites bureautiques...). Ses bibliothèques peuvent permettre de créer des interfaces graphiques, faire circuler des informations au travers d'un réseau ou d'accéder de manière avancée au système d'exploitation de l'ordinateur.

Python est un langage de programmation **interprété**, c'est-à-dire que les instructions qui lui sont envoyées sont directement « transcrites » au fur et à mesure de leur lecture. Cela diffère des langages « compilés », où un logiciel intermédiaire est chargé de transformer le code du programme en langage machine. L'avantage d'un langage interprété est donc de supprimer cette étape intermédiaire, ce qui permet une meilleure portabilité entre différents systèmes d'exploitation (Windows, Linux, Mac OS).

La programmation Python est une **programmation orientée objet (POO)**, c'est-à-dire qu'il crée des entités indépendantes (*objets*), chacun responsables de leurs propres attributs et méthodes, qu'il est possible de manipuler par la suite. Ce type de programmation s'oppose à la programmation procédurale et permet une simplification de l'écriture.

L'utilisation de bibliothèques qui permettent le calcul, la manipulation et la visualisation des données explique l'essor de ce langage dans le domaine scientifique.

Dans le développement de la méthode pharmacophorique, la suite de logiciels **Anaconda 5.1** a été utilisée. Elle comprend un environnement Python et de nombreuses bibliothèques. Plusieurs bibliothèques ont été utilisées :

- **NumPy**, qui permet la manipulation de grands ensembles de nombres, et les calculs numériques ;
- **Pandas**, qui permet la manipulation avancée de tableaux et matrices ;
- **SciPy**, qui concerne le calcul scientifique et contient des fonctions permettant des analyses statistiques, des interpolations, transformées de Fourier... ;
- **Matplotlib** qui permet de tracer et de visualiser des données sous formes de graphiques.

Pour une facilité d'écriture et une visualisation des résultats en simultanément, le *notebook Jupyter* a été utilisé. Un *notebook* est une interface web dans laquelle il est possible d'écrire du code, de l'exécuter, de voir directement les résultats et également de les visualiser par graphique. Il est également possible d'écrire des commentaires en format texte et d'exporter les scripts ou fichiers sous plusieurs formats (HTML, LaTeX...).

III. 2. SPIKE (*Spectrometry Processing Innovative Kernel*)

SPIKE représente le programme de base idéal pour manipuler des données RMN brutes et réaliser l'analyse différentielle. C'est un programme libre écrit en Python 2.7 et 3.5. Il est utilisable par *Jupyter Notebook* et peut supporter facilement l'ajout de *plug-in* sans modification de son code interne.

SPIKE est dédié aux spectroscopies à transformée de Fourier dont il peut réaliser le traitement et l'analyse. Il fournit plusieurs fonctionnalités telles que l'apodisation, le phasage de spectre RMN, la détection de signaux, des corrections de ligne de base... Il supporte les données provenant de différents constructeurs (Bruker, ThermoFisher...). Il utilise les bibliothèques externes *numpy*, *scipy*, *pandas*, *scikit-learn*, et *hdf5*.

III. 3. Plasmodesma et le *bucketing*

Le programme Plasmodesma a spécialement été développé pour ce projet. Il est relié à SPIKE pour la plupart de ces opérations. Il peut traiter de manière autonome de larges séries d'expériences RMN différentes provenant de plusieurs échantillons aux conditions d'acquisition variables. Ce traitement analytique implique la manipulation de séries complexes d'expériences RMN (1D et 2D homo- et hétéronucléaires) et la production automatique de rapports portant sur les données RMN analysées, des figures, des listes de pics et *buckets*. C'est un programme Python compatible Python 2.7 et 3.5. Il est implémenté à SPIKE comme un *plug-in*. Il est applicable aux analyses ^1H , ^1H - ^1H COSY, ^1H - ^1H TOCSY, ^1H - δ DOSY, ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC.

III. 3. a. Principe des opérations

Ce programme fonctionne sans intervention de l'homme.

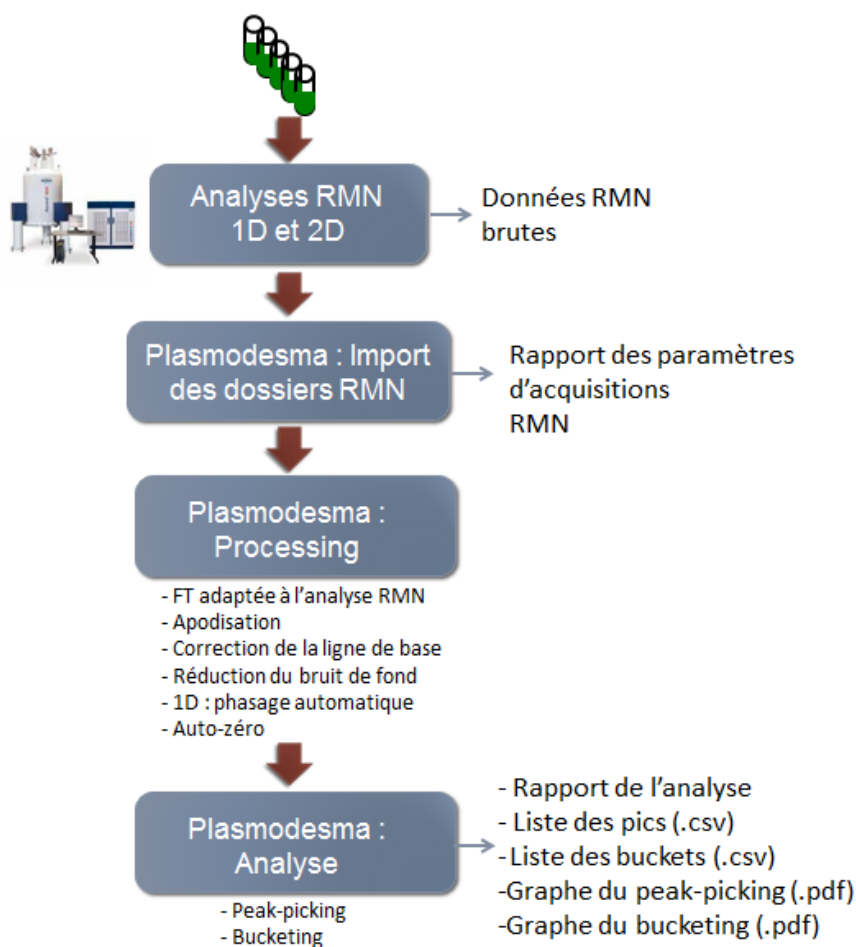


Figure 31 : Principe des opérations de Plasmodesma. Après l'acquisition des données RMN, celles-ci sont importées par Plasmodesma à SPIKE qui va les traiter de manière automatique. Cela implique l'application de transformées spécifiques de l'expérience RMN, une apodisation, une correction de la ligne de base, une réduction du bruit de fond, un phasage automatique des 1D et un calibrage du spectre. Ensuite, Plasmodesma va réaliser la détection automatique des pics et un *bucketing* des données. Plusieurs fichiers sont générés : des listes des signaux détectés et des *buckets* et des rapports sur les paramètres d'analyses et les traitements de données.

Lorsqu'il est appliqué à un dossier, tous les fichiers d'analyses RMN 1D et 2D qu'il contient vont être automatiquement importés, et un rapport global des analyses qu'il contient est produit. Tout le traitement des données et les étapes d'analyses ont été optimisés en fonction des paramètres d'acquisition des différentes analyses RMN. Les données RMN brutes sont apodisées, une transformée de Fourier est appliquée et la ligne de base corrigée, une réduction du bruit de fond est également réalisée. Une calibration du spectre est effectuée par la détection du pic de TMS à 0 ppm. Les spectres 1D sont également phasés de manière automatique. Plasmodesma réalise ensuite, un *peak-picking* et le *bucketing* des données, la liste des signaux et des *buckets* sont enregistrés sous un fichier `csv` pour chaque expérience

et des figures de chaque spectre sont créées. Enfin, un rapport résumant les analyses RMN et les fichiers issus de Plasmodesma est produit (Figure 31). Les développements spécifiques (auto-phasage, auto-zéro, *processing*, *peak-picking*) sont décrits dans la partie expérimentale (p. 193).

III. 3. b. Le bucketing

Le *bucketing* est l'étape la plus importante du programme Plasmodesma. C'est une méthode utilisée en métabolomique pour permettre l'alignement de spectre RMN 1D et repérer un métabolite connu dans un échantillon biologique. Cette option est d'ailleurs proposée par des logiciels de RMN tels que Mnova, ACD/Labs, NMRProcFlow. Le *bucketing* consiste en la division des données spectrales en petits **segments de taille égale appelés *bucket*** sur la largeur de la fenêtre spectrale et du calcul de l'**aire** sous le spectre de chaque segment (Figure 32).

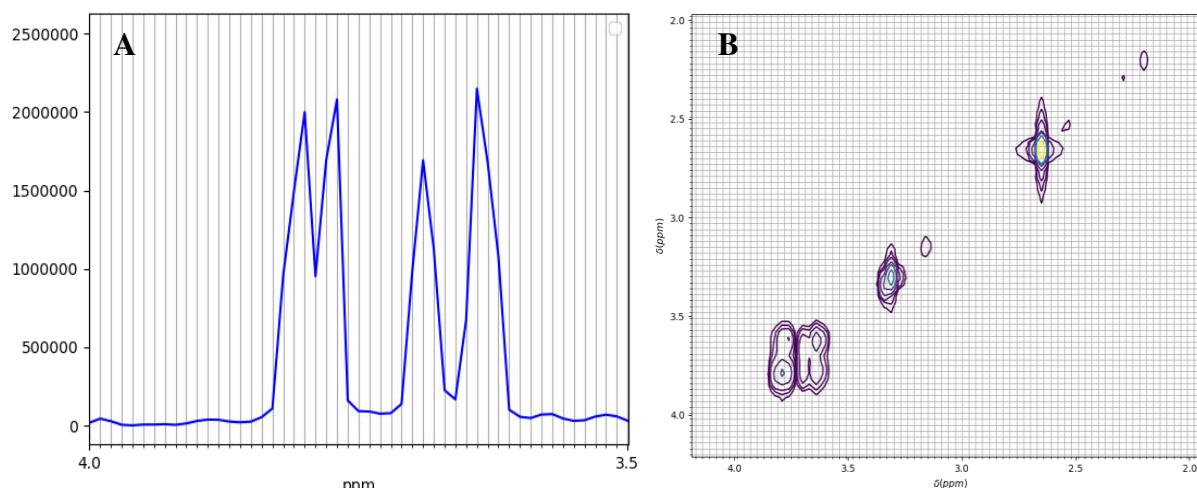


Figure 32 : Représentation de la division de spectres RMN 1D (A) et 2D (B) en *buckets*. Un signal est compris dans plusieurs *buckets*.

Les spectres RMN sont des données lourdes, une RMN 2D contient des millions de points d'acquisition. Le *bucketing* permet de réduire la masse de données, donc, de faciliter leurs manipulations et d'augmenter les vitesses de calculs pour les analyses différentielles à suivre. Le *bucketing* permet également de pouvoir comparer le même signal sur des spectres RMN de deux échantillons différents qui auraient subi un léger déplacement. Les signaux pourront être dans le même *bucket* et donc correctement alignés.

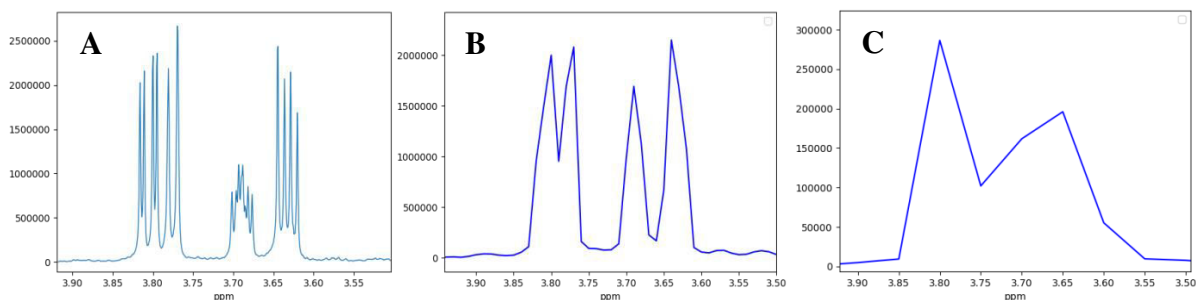


Figure 33 : Effet de la taille des *buckets* sur un spectre RMN ^1H . Spectre non *bucketé* (MeOD, 700MHz) (A), spectre *bucketé* à 0.01 ppm (B) et spectre *bucketé* à 0.05 ppm (C). Contrairement au spectre B, le spectre C ne permet plus de distinguer les pics entre eux. Les multiplicités sont perdues lors du *bucketing*.

La taille du *bucket* doit être suffisamment large pour recouvrir des légères erreurs de *shift* et permettre un bon alignement des spectres, et suffisamment petite pour ne pas perdre en résolution. Les tailles des *bucket* sélectionnées sont de **0.01 ppm pour les 1D**, de **0.03 ppm pour les 2D ^1H** et **1.0 ppm pour ^{13}C** (Figure 33).

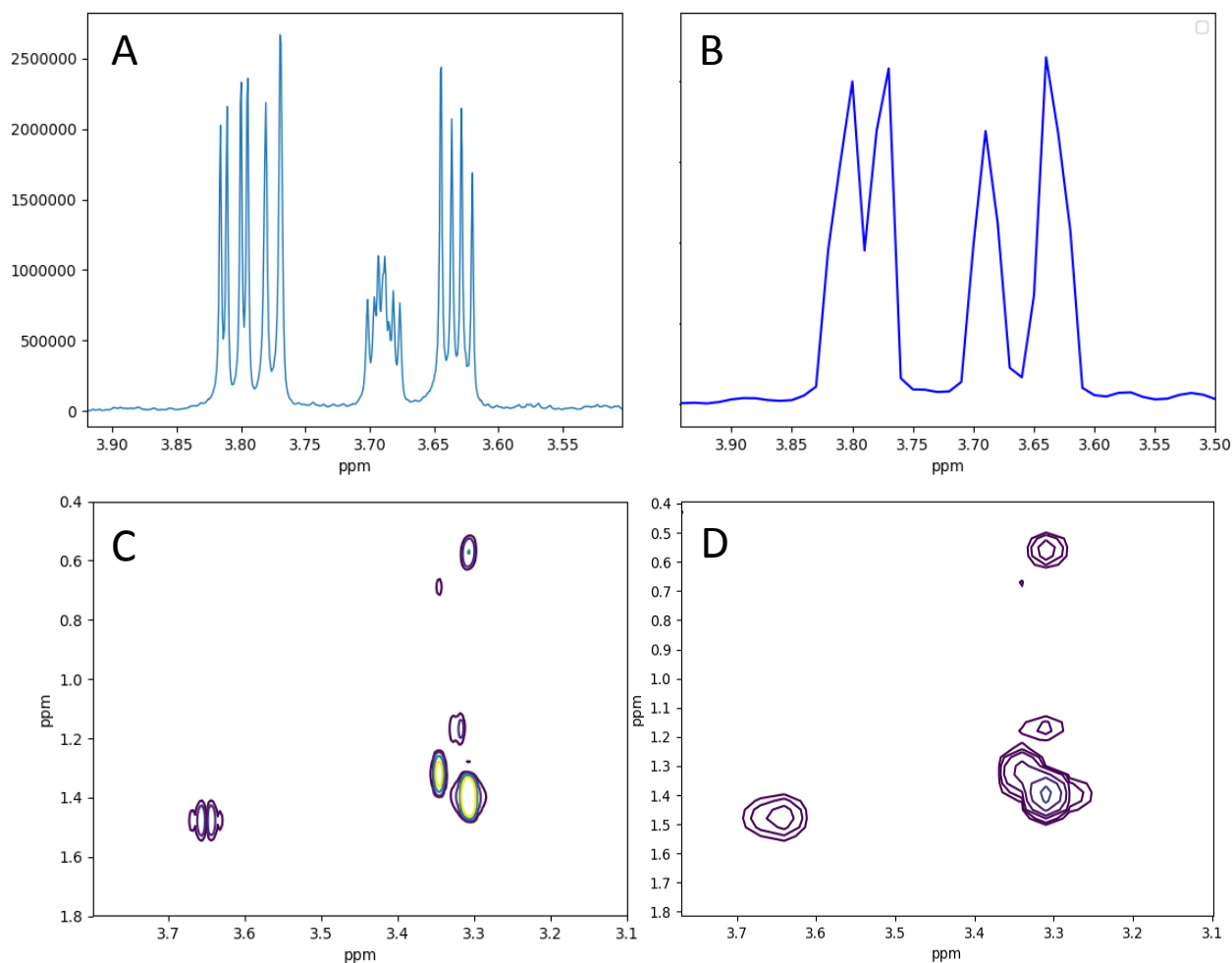


Figure 34 : Représentation d'un spectre RMN ^1H entre 3.5 et 3.9 ppm (MeOD, 700 MHz) (A) et le spectre *bucketé* correspondant (B) et un spectre ^1H - ^1H -TOCSY (C) et le spectre *bucketé* correspondant (D). Tous les signaux sont présents, il n'y a pas de perte d'information lors du *bucketing*.

D'après la **Figure 34**, avec les fenêtres de *buckets* choisies, nous pouvons observer qu'il n'y a pas de pertes de sensibilité ou de déplacements de signaux sur les spectres *bucketés*. Il y a une perte des multiplicités mais cela ne représente pas une perte d'information majeure dans le cas de comparaison d'empreintes.

Pour chaque *bucket* produit, le point **maximal et minimal, l'aire et l'écart-type** sont calculés.

III. 4. *Bucket Utilities*

Une fois les données RMN *bucketées*, elles sont prêtes à être utilisées pour les analyses différentielles. Il faut déterminer quels variables et calculs sont à réaliser pour produire la meilleure empreinte pharmacophorique. Un script appelé «*BucketUtilities*» contient plusieurs fonctions permettant le **pré-traitement** des spectres *bucketés*, les **importations** de ces données et l'**affichage** des résultats.

III. 5. **Pré-traitement des spectres : Nettoyage et symétrisation**

III. 5. a. *Nettoyage des spectres*

Lors de l'acquisition de spectres RMN, des artefacts tels que des bruits t_1 ou des anti-diagonales sont observés. Ces artefacts, habituellement de faibles intensités, sont exacerbés dans nos analyses par la volonté de détection de signaux de faible intensité. Ces artefacts peuvent altérer la production d'une empreinte pharmacophorique. Pour remédier à ce problème, deux solutions ont été apportées. La première consiste à une opération de «*nettoyage*» du spectre et vise à éliminer le bruit de fond et les raies t_1 .

Si les données RMN *bucketées* sont représentées sous formes d'histogramme, nous remarquons qu'elles suivent une distribution gaussienne terminée par une «*longue traîne*» (Figure 35, A). La majorité des données en RMN correspond au bruit de fond et cette longue traîne représente les signaux. Pour les éliminer, trois fonctions de «*nettoyage*» ont été essayées.

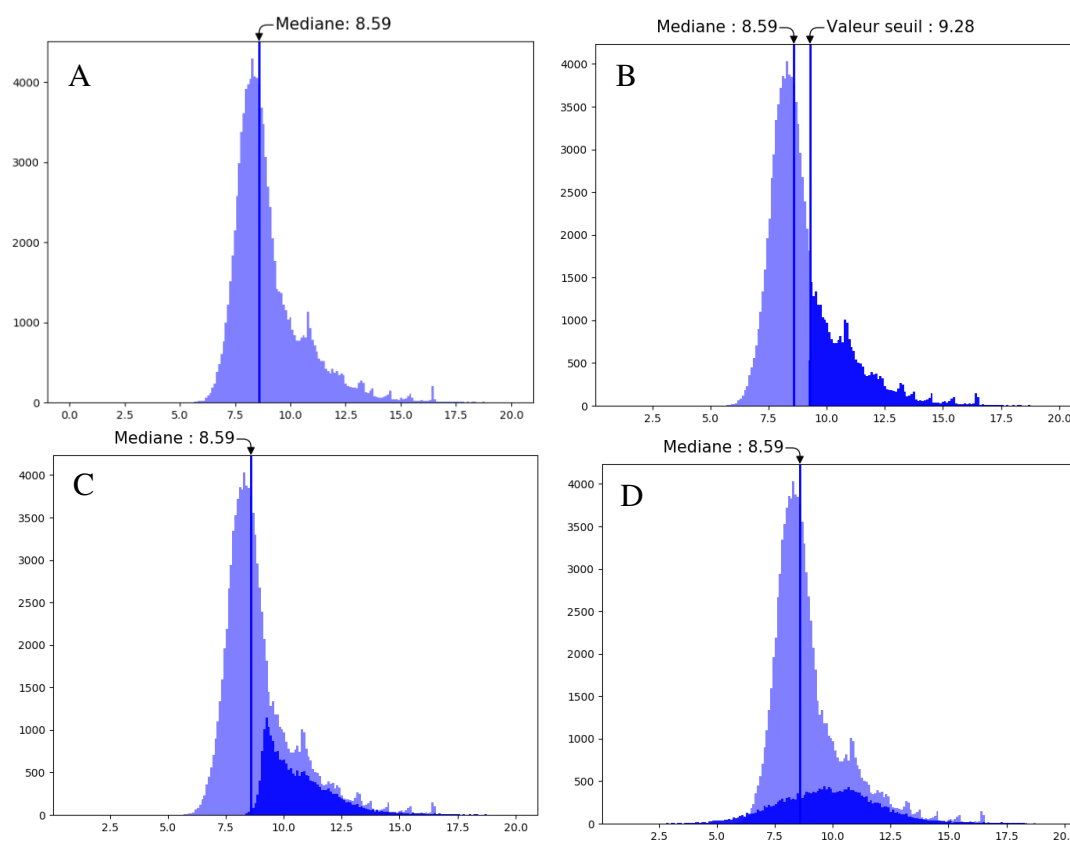


Figure 35 : Histogrammes des données RMN COSY *buckétées* d'ARTE 3 sans nettoyage (A, échelle logarithmique), avec fonction « nettoie » (B, bleu foncé), avec fonction « nettoie_mieux » (C, bleu foncé) et fonction « nettoie_encore_mieux » (D, bleu foncé). La fonction « nettoie_mieux » sera appliquée à nos données.

La première fonction appelée « nettoie » consiste en la suppression de toutes les valeurs supérieures à une valeur seuil. Cette valeur seuil est égale à la valeur logarithmique de la médiane des données multipliée par un facteur 2 (Figure 35, B).

La seconde fonction appelée « nettoie_mieux » repose sur le même principe que la fonction `nettoie` mais la valeur seuil est calculée pour chaque colonne de la matrice (ce qui correspond à chaque colonne de 0.03 ppm sur un spectre 2D par exemple (Figure 35, C)).

La dernière fonction de nettoyage testée est appelée `nettoie_encore_mieux` repose sur le même principe que la fonction `nettoie_mieux` mais redistribue les valeurs suivant une distribution normale (Figure 35, D).

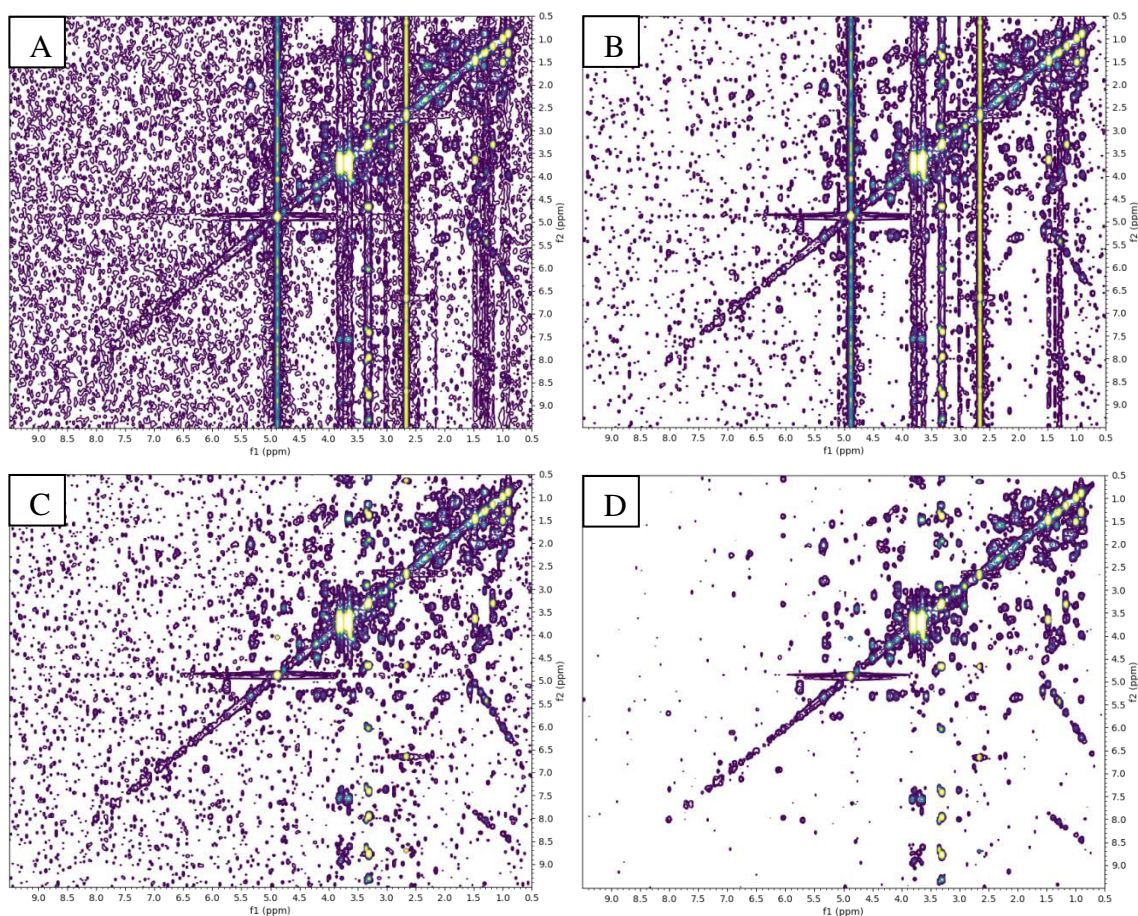


Figure 36 : Effet des fonctions de nettoyage sur les spectres *buckétés* (à la même échelle). Spectre sans nettoyage (A), spectre avec fonction "nettoie" (B), spectre avec fonction « nettoie_mieux », spectre avec fonction « nettoie_encore_mieux » (D). La réduction du bruit de fond par la fonction « nettoie_encore_mieux » est trop importante et conduit à la perte d'information.

L'effet des différentes fonctions de nettoyage est représenté dans la **Figure 36**. Nous pouvons observer qu'il y a bien une réduction du bruit de fond grâce aux fonctions de nettoyage mais celle-ci est trop importante dans le cas de l'utilisation de `nettoie_encore_mieux` et pourrait mener à une perte d'information. Les fonctions de nettoyage appliquées aux colonnes sont parfaitement efficaces pour l'élimination des raies t_1 . La fonction `nettoie_mieux` sera utilisée pour toutes les analyses.

III. 5. b. Symétrisation des spectres homonucléaires

Nous pouvons observer sur nos spectres la présence d'artefacts de type anti-diagonale (Figure 37). Pour les supprimer, une étape de symétrisation a été ajoutée. Celle-ci est seulement applicable aux données RMN symétriques donc les analyses RMN 2D ^1H - ^1H -COSY et ^1H - ^1H -TOCSY.

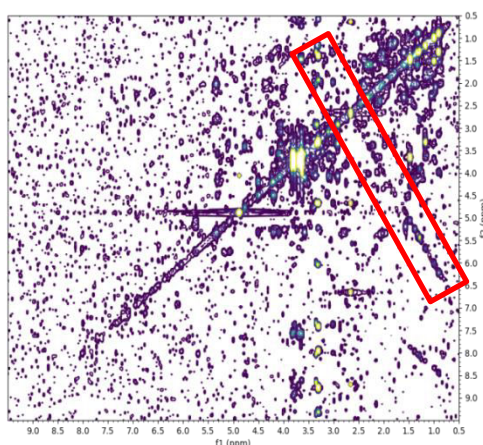


Figure 37 : Spectre ^1H - ^1H -COSY buckété. Présence d'artefacts de type antidiagonal (encadré rouge).

Pour réaliser une symétrisation de nos données, il suffit de les représenter sous forme d'une matrice, de réaliser son inverse, et de choisir la valeur minimum entre la matrice initiale et la matrice inverse. Ainsi, une tache de corrélation qui ne posséderait pas de symétrie par rapport à la diagonale (un artefact) est réduite à sa valeur minimale, elle est donc supprimée (Figure 38).

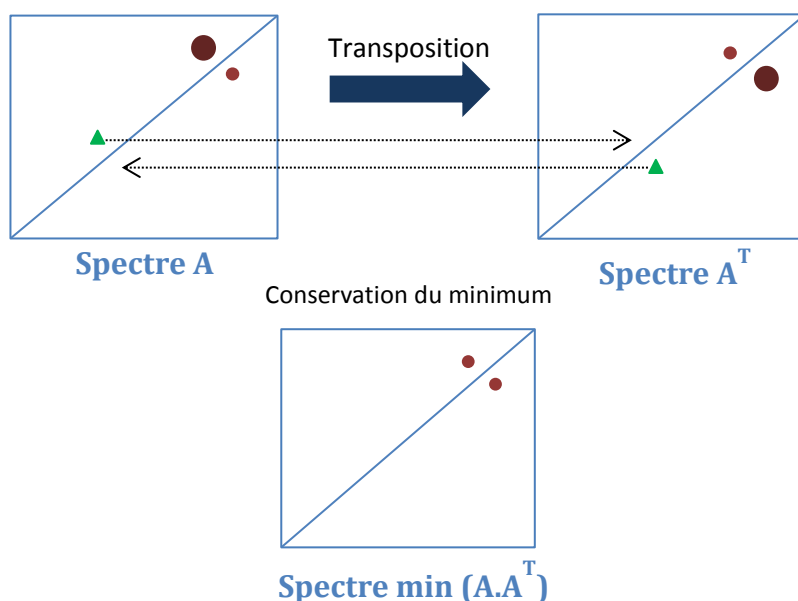


Figure 38 : Représentation schématique de la symétrisation des données RMN. L'artefact (triangle vert) ne possède pas de symétrie par rapport à la diagonale. La valeur minimale pour un point entre le spectre initial (spectre A) et le spectre « transposé » (spectre A^T) est sélectionnée et provoque une suppression de l'artefact.

Nous pouvons observer d'après la **Figure 39**, que l'anti-diagonale est bien supprimée, ainsi que tout autre signal qui n'aurait pas son symétrique.

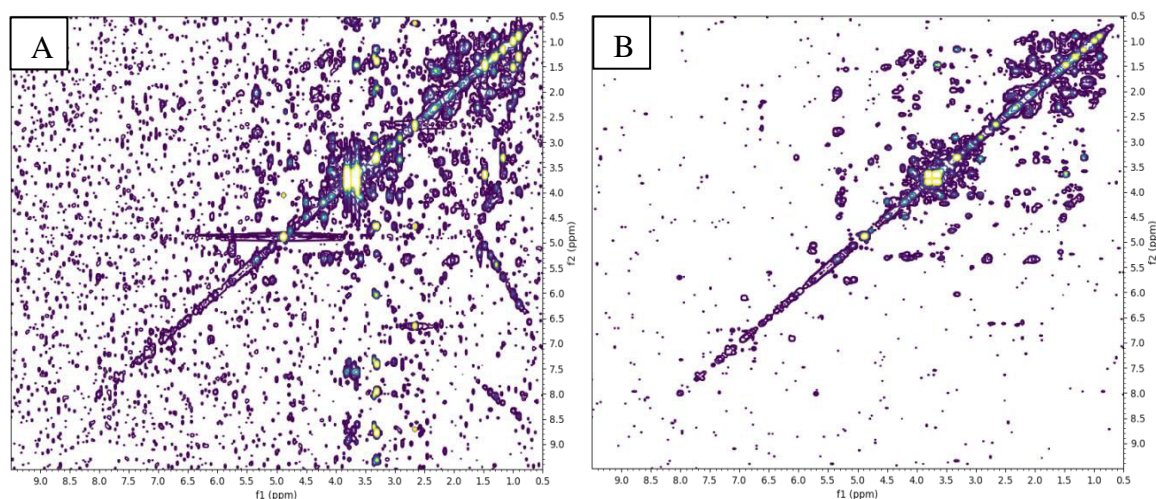


Figure 39 : Spectre ^1H - ^1H -COSY *bucketé*, nettoyé non symétrisé (A) et symétrisé (B). Les antidiagonales et artefacts ont été supprimés.

III. 6. Conclusion

Plasmodesma offre un lot d'outils pour le traitement et l'analyse de spectres RMN 1D et 2D complètement automatique sans aucune interaction. Il est capable d'agir sur un jeu de données comportant plusieurs types d'analyses et d'adapter leur traitement de manière optimisée pour chacune. Il est capable de créer des images de spectre, des listes de signaux et de *buckets* et de faire un rapport de ces données et analyses. *BucketUtilities* fournit entre autres des fonctions de réduction du bruit de fond et d'artefacts permettant de générer des données les mieux adaptées à une analyse différentielle.

IV. Exemplification de l'analyse différentielle sur des échantillons reconstitués

Pour permettre l'informatisation de l'analyse différentielle automatique, des simulations de cas simples ont été réalisées. Des échantillons « reconstitués », contenant un extrait algal inactif (extrait hydroalcoolique de *Sargassum muticum*) et une molécule active antipaludique ont été préparés (Figure 40). Deux séries de 5 échantillons ont été obtenues :

- une première série (série CHLORO) constituée de cinq échantillons d'extrait de *S. muticum* supplémentés par des concentrations croissantes de chloroquine, un alcaloïde quinoléique ;
- une seconde série (série ARTE) constituée de cinq échantillons supplémentés par des concentrations croissantes en artémisinine, une lactone sesquiterpénique.

Nous détaillerons tout d'abord la mise en place des calculs via la série CHLORO, leurs applications et la production d'empreintes spectrales qui en résultent. Les résultats issus des spectres ^1H - ^1H COSY, puis des spectres ^1H - ^1H TOCSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC et ^1H - δ DOSY seront présentés. Toutes les analyses RMN ont été réalisées dans **les mêmes conditions** (MeOD, 700MHz, Cf. Partie expérimentale III. 2).

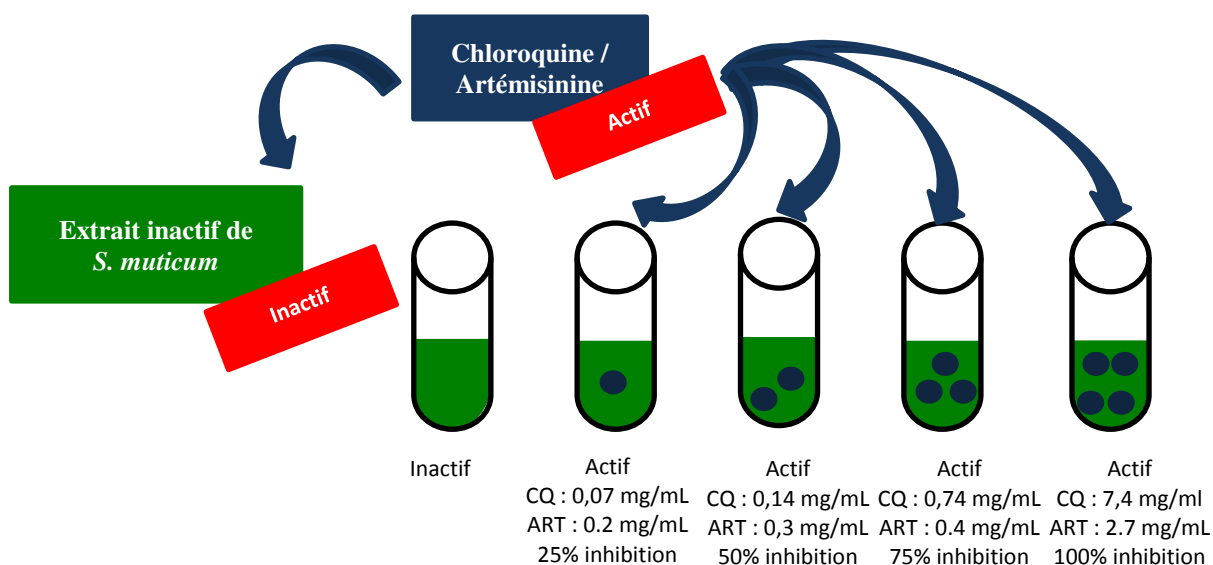


Figure 40 : Échantillons de la série CHLORO et ARTE. Cinq échantillons sont préparés par l'ajout d'une quantité croissante de chloroquine ou d'artémisinine correspondant à un pourcentage d'inhibition de croissance théorique de *P. falciparum* de 0, 25, 50, 75 et 100%. **CQ** : la concentration en chloroquine ; **ART** : concentration en artémisinine dans l'échantillon.

La concentration croissante en molécule active dans l'échantillon correspond à des concentrations inhibitrices théoriques de croissance *P. falciparum* de 0, 25, 50, 75 et 100% d'inhibition. Ces échantillons ont ensuite été évalués biologiquement. Les échantillons de la série CHLORO ont montré une inhibition de 0, 21, 65, 72 et 89 %, et la série ARTE de 13, 37, 81, 97, 100 %.

IV. 1. Présentation de la série CHLORO

L'ensemble des spectres de la chloroquine de référence et les conditions d'analyses RMN de cette série sont détaillées dans la partie Annexe (p.229).

La chloroquine appartient à la classe des 4-aminoquinoléines. Elle est donc constituée d'un noyau quinoléique (bicycle avec N en position 1) substitué par un atome de Cl en position 7, et d'une chaîne latérale basique en position 4 terminée par un groupement amino-alkyle (Figure 41).

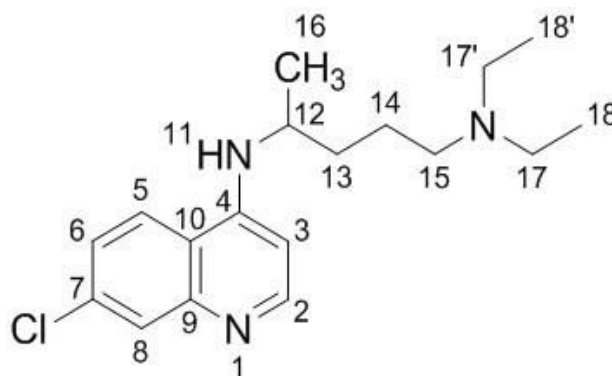


Figure 41: Structure chimique de la chloroquine

Un échantillon de référence contenant uniquement de la chloroquine pure a été analysé en RMN. Les déplacements ^1H et ^{13}C sont donnés dans le **Tableau 9**.

Tableau 9 : Déplacements chimiques ^1H et ^{13}C de la chloroquine (MeOD, 700MHz)

Position	δ_{H} (ppm), mult	δ_{C} (ppm)
2	8.39 (d, $J=7.2$ Hz, 1H)	-6.5 (142.2)
3	6.93 (d, $J=7.4$ Hz, 1H)	99.4
5	8.46 (d, $J=9.0$ Hz, 1H)	125.6
6	7.68 (dd, $J=9.0, 2.1$ Hz, 1H)	128.22
8	7.93 (d, $J=2.1$ Hz, 1H)	120.0
11	8.62	/
12	4.17 (m, 1H)	50.6
13	1.90 (m, 2H)	32.8
14	1.84 (m, 2H)	21.4
15	3.15 (m, 2H)	52.1
17, 17'	3.20 (q, $J=7.2$ Hz, 4H)	47.6
18, 18'	1.29 (t, $J=7.3$ Hz, 6H)	8.6
19	1.43 (d, $J=6.4$ Hz, 3H)	19.5

Ces signaux se distribuent dans la zone aromatique (6.5 et 9 ppm) et entre 0 et 4.5 ppm (Figure 42). La chloroquine de référence possède une impurété qui fait apparaître deux signaux supplémentaires notamment visibles en RMN 2D ^1H - ^1H COSY (3.66 ; 1.20 ppm et 3.70 ; 1.54, croix rouges sur les spectres).

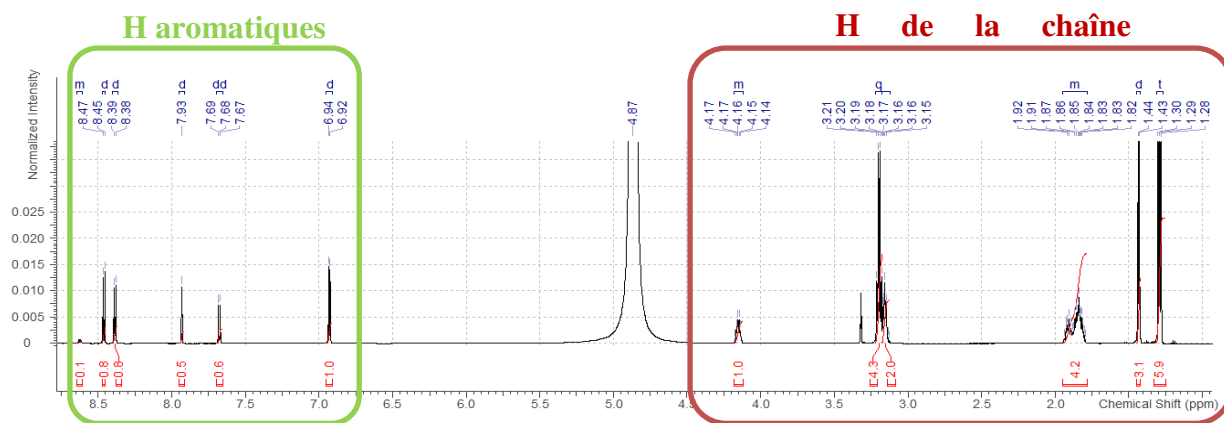


Figure 42 : Spectre ^1H de la chloroquine (MeOD, 700 MHz). Ces signaux se distribuent dans la zone aromatique (6.5 et 9 ppm) et entre 0 et 4.5 ppm.

Dans la série CHLORO, la chloroquine a été distribuée en concentration croissante de l'échantillon CHLORO 1 (sans chloroquine) à CHLORO 5 afin que sa concentration correspondent à une concentration inhibitrice de croissance du parasite de 0, 25, 50, 75 et 100%. CHLORO 1 est uniquement constitué de l'extrait hydroalcoolique de *S. muticum*. Son spectre ^1H - ^1H COSY donne un aperçu de la distribution des signaux RMN de l'extrait (Figure 43). En effet, la majorité des taches de corrélation sur le spectre ^1H - ^1H COSY se situe entre 0 et 4.5 ppm et il y a très peu de signaux dans la zone aromatique. L'empreinte de la chloroquine sera sûrement plus visible dans cette zone.

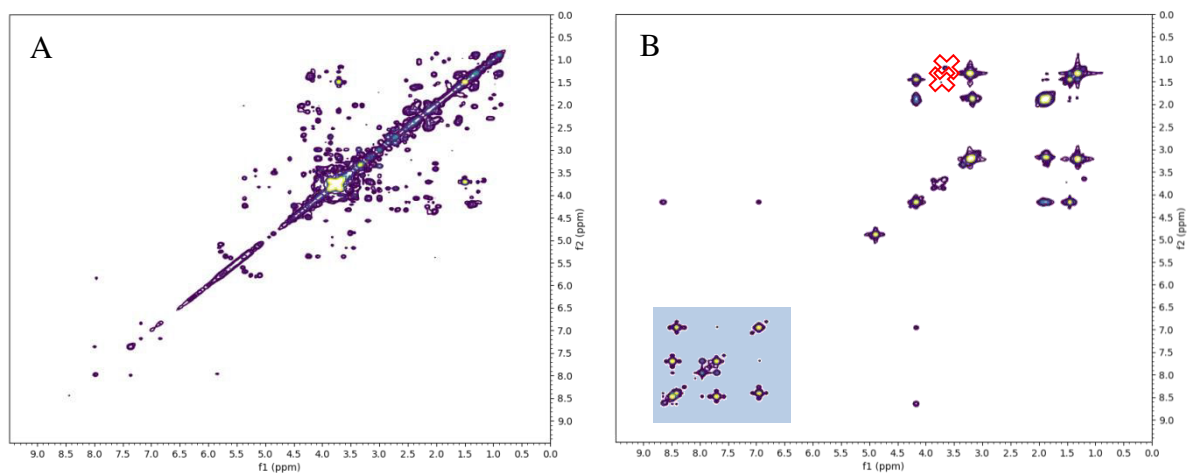


Figure 43 : Comparaison du spectre ^1H - ^1H COSY *buckété* de l'extrait de *S. muticum* (A) et de la chloroquine (B). L'extrait contient très peu de signaux dans la zone aromatique contrairement à la chloroquine (zone bleue). Une impureté est signalée par une croix rouge.

Les spectres ^1H - ^1H COSY *buckétés* des 5 échantillons de la série sont représentés dans **Figure 44**. La RMN est une technique quantitative, l'intensité des signaux augmente proportionnellement à la concentration d'un composé. L'augmentation de la concentration de la chloroquine est nettement visible au sein des échantillons.

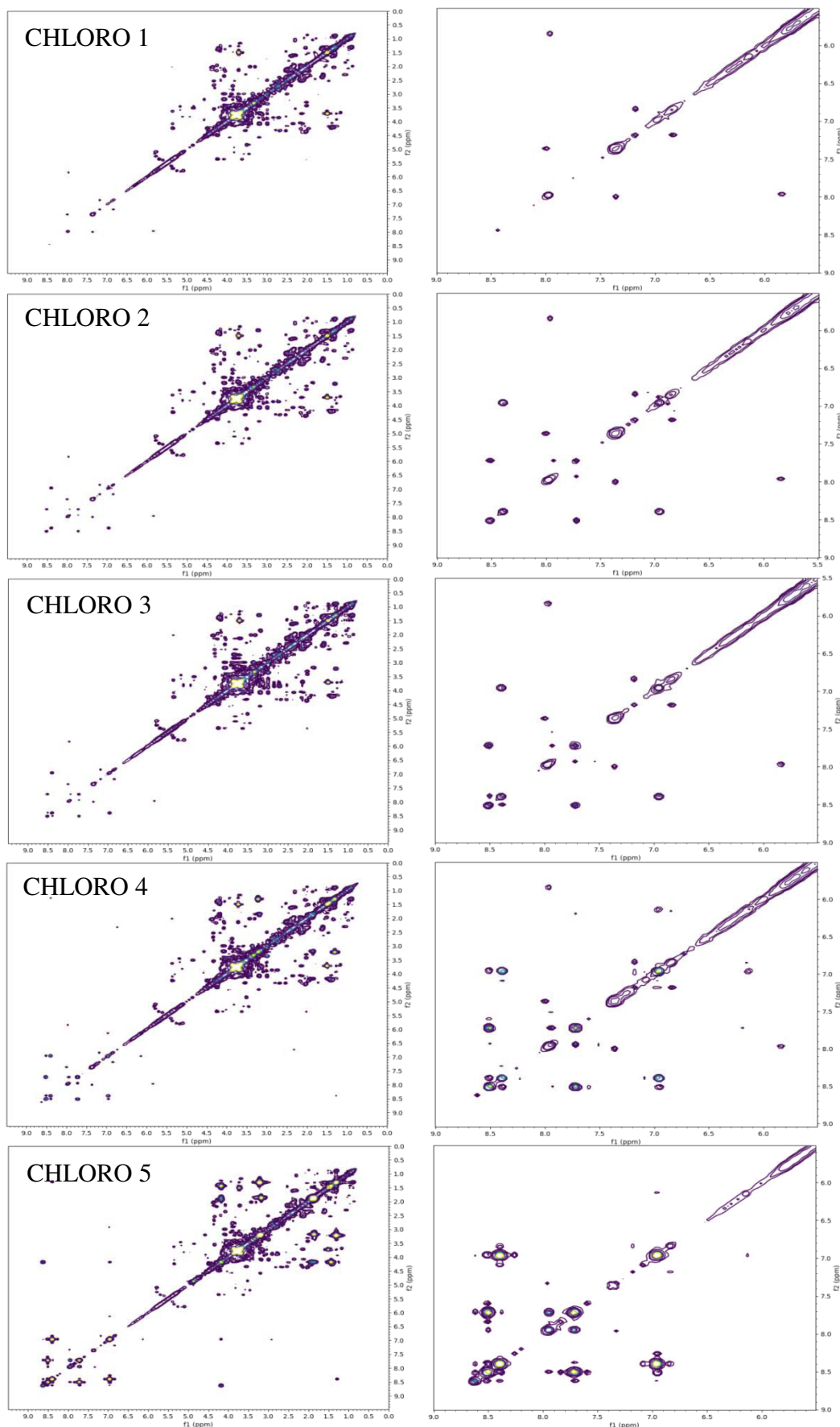


Figure 44 : Spectres COSY *buckétés* des échantillons CHLORO 1 à 5 et zoom sur la partie aromatique du spectre (colonne de droite). La chloroquine est visible dans les échantillons est l'intensité du signal croit en fonction de la concentration.

IV. 2. Présentation de la série ARTE

L'artémisinine est une lactone sesquiterpénique qui contient un inhabituel pont peroxyde (Figure 45).

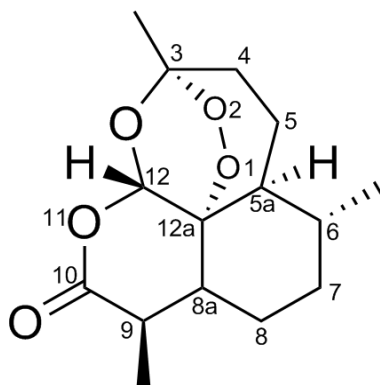


Figure 45 : Structure chimique de l'artémisinine

Elle a l'avantage d'avoir une fenêtre spectrale RMN plus réduite que la chloroquine. En effet, ces signaux se situent entre 0 et 3.5 ppm et un signal à 6 ppm. Il y a de nombreuses superpositions de signaux entre 0 et 2.5 ppm (Figure 46).

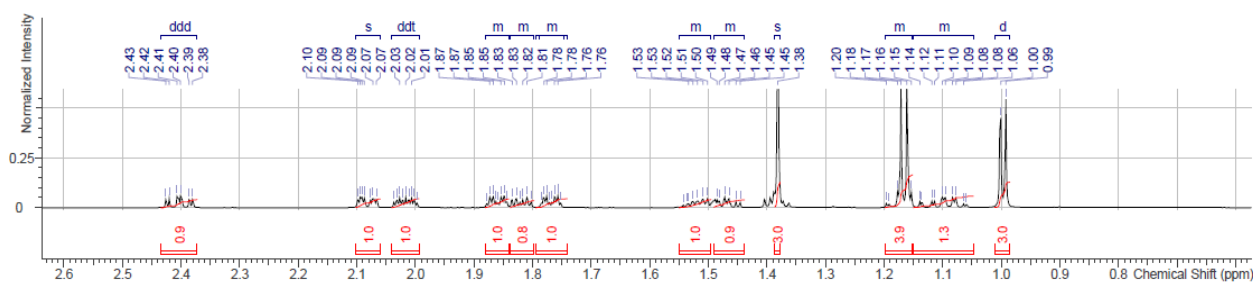


Figure 46 : Zoom du spectre ^1H de l'artémisinine entre 0 et 2.5 ppm. L'artémisinine possède de nombreux recouvrements de signaux et multiplets.

Ces déplacements sont plus intéressants à étudier car ils se situent dans la même zone que les signaux de l'extrait algal, et il y aura sûrement des recouvrements (Tableau 10). L'ensemble des spectres de l'artémisinine de référence et les conditions d'analyses RMN de cette série sont détaillées dans la partie Annexe (p. 234).

Tableau 10 : Déplacements chimiques ^1H et ^{13}C de l'artémisinine (700 MHz, MeOD)

Position	δ_{H} (ppm), mult	δ_{C} (ppm)
4 α	2.08 (m, 1H)	37.0
4 β	2.40 (ddd, $J=14.3*2, 4.1$ Hz, 1H)	"
5 α	2.02 (ddt, $J=14.3, 6.5, 3.5$ Hz, 1H)	26.2
5 β	1.47 (m, 1H)	"
5a	1.39 (m, 1H)	51.5
6	1.52 (m, 1H)	38.5
7 α	1.11 (m, 1H)	34.9
7 β	1.77 (m, 1H)	"
8 α	1.86 (m, 1H)	24.3
8 β	1.17 (m, 1H)	"
8a	1.82 (m, 1H)	46.1
9	3.31 (m, 1H)	34.3
12	6.03 (s, 1H)	95.8
3-Me	1.38 (s, 3H)	25.5
6-Me	1.01 (d, $J=6.2$ Hz, 3H)	20.3
9-Me	1.17 (m, 3H)	13.0

De la même manière que pour la série CHLORO, la série ARTE est constituée de 5 échantillons dont la concentration en artémisinine varie afin de correspondre à des concentrations inhibitrices de la croissance de *P. falciparum* de 0, 25, 50, 75 et 100 %. Expérimentalement, l'inhibition de croissance était respectivement de 13, 37, 81, 97, 100 %.

La masse de l'artémisinine était de 0, 1.5, 2, 3 et 20 % par rapport à la masse de l'extrait.

IV. 3. Analyse différentielle de spectres deux à deux

Les travaux de F. Schroeder s'appuyaient sur une analyse différentielle entre un échantillon actif et un échantillon inactif, il s'agissait dans ce cas d'une comparaison de spectres deux à deux. Cette situation de « active *versus* inactive » est la première approche à avoir été développée informatiquement.

IV. 3. a. Choix de la variable et du calcul

Plasmodesma calcule de manière automatique le maximum, le minimum, l'aire et l'écart-type des *buckets* et enregistre ces valeurs dans la *bucketlist*. Il faut sélectionner la variable et le calcul qui permettront d'obtenir la meilleure empreinte pharmacophorique avec

le minimum de pertes d'information. Ces quatre variables retransmettent quasiment la même image du spectre *buckété* qui est fidèle au spectre RMN brut. Ce sont donc les calculs via ces variables qui seront déterminants. Le choix des variables et des calculs sera présenté pour la série CHLORO puis exemplifié par la série ARTE.

- **Série CHLORO**

Le calcul le plus simple et intuitif est de réaliser une **soustraction** des données d'un spectre par rapport à l'autre. La stratégie est d'employer ce calcul sur l'échantillon qui contient le plus de chloroquine (CHLORO 5, 7.5 mg/mL), puis celui qui en contient le moins (CHLORO 2, 0.07 mg/ml) *versus* celui qui n'en possède pas (CHLORO 1). Les essais avec un échantillon peu concentré en la molécule active sont primordiaux afin d'anticiper les cas futurs où la molécule active sera présente en faible concentration.

Les essais de soustraction avec les 4 variables sont représentés dans la **Figure 47**, les carrés orange sont les signaux de la chloroquine, les croix rouges correspondent aux impuretés de la chloroquine et les croix noires, aux signaux d'origine algale.

Nous pouvons observer qu'il n'y a pas de différence de spectre entre les variables. Cependant, en raison de la présence de taches de corrélation n'appartenant pas à la chloroquine, la soustraction ne semble pas être un calcul très sélectif de l'empreinte mais est plutôt une image du spectre soustrait.

La soustraction n'est donc pas un calcul efficace pour produire une empreinte pharmacophorique.

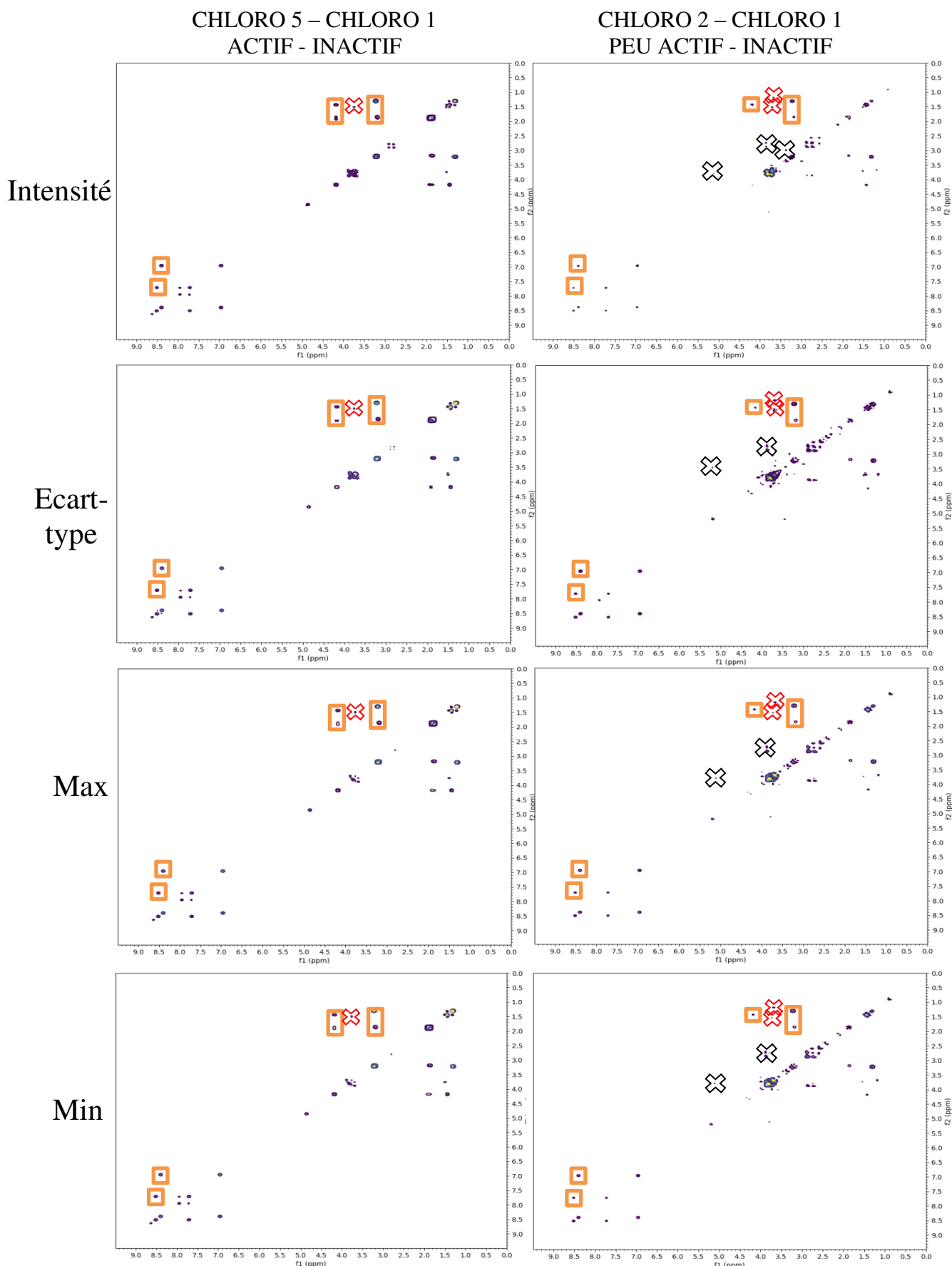


Figure 47 : Soustractions des différentes variables des *buckets* des expériences ^1H - ^1H COSY. Les résultats sont similaires entre les variables. La soustraction est une image du spectre soustrait et ne semble pas le calcul le plus efficace. Carré orange : signal de la chloroquine ; croix rouge : impureté de la chloroquine.

Le second calcul testé est **le rapport** des variables des *buckets* (Figure 48). Nous pouvons observer qu'il y a peu de différences de spectre entre les variables. Mais les spectres résultant du rapport des intensités ou des écarts-types sont légèrement meilleurs car la majorité des taches de corrélation appartiennent à la chloroquine et il y a moins de taches de corrélation algales.

Le rapport des intensités ou des écarts-types des *buckets* semble donc être un bon moyen de détection de l'empreinte dans les cas de présence/absence d'une molécule bioactive dans deux échantillons à comparer.

Il y a une suppression des taches de corrélation sur la ligne de symétrie entre 1 et 4.5 ppm. Cela s'explique par la forte intensité des taches de corrélation à la fois sur les spectres de CHLORO 5, CHLORO 2 et CHLORO 1.

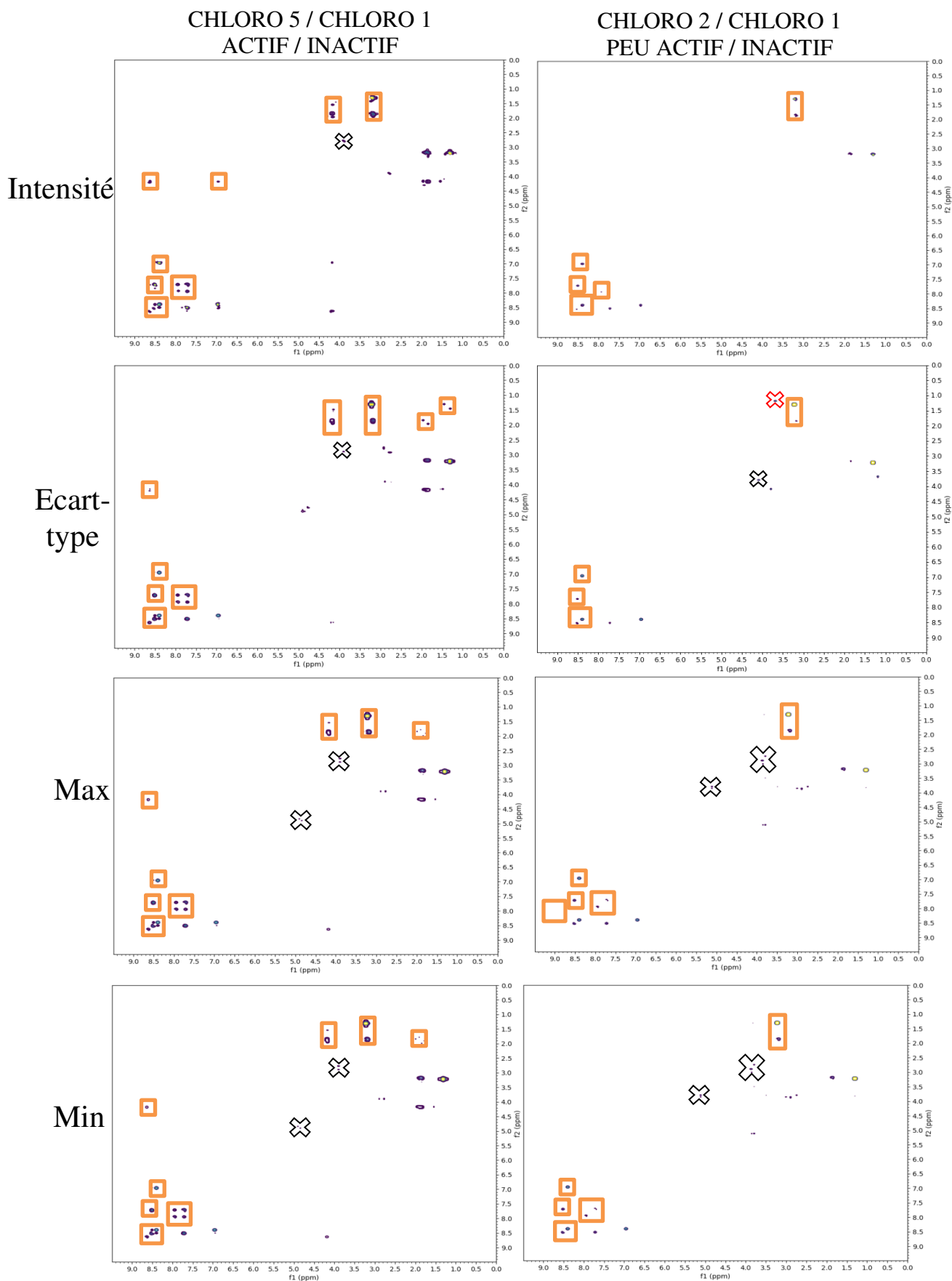


Figure 48 : Rapports des différentes variables des buckets des expériences ¹H-¹H COSY. L'utilisation des intensités et des écarts-types donnent les meilleures empreintes de la chloroquine. Carré orange = signal de la chloroquine ; croix rouge = impureté de la chloroquine ; croix noires = signal algal

Le rapport des écarts-types et des intensités des *buckets* des expériences ^1H - ^1H COSY est appliqué entre les échantillons (Figure 49).

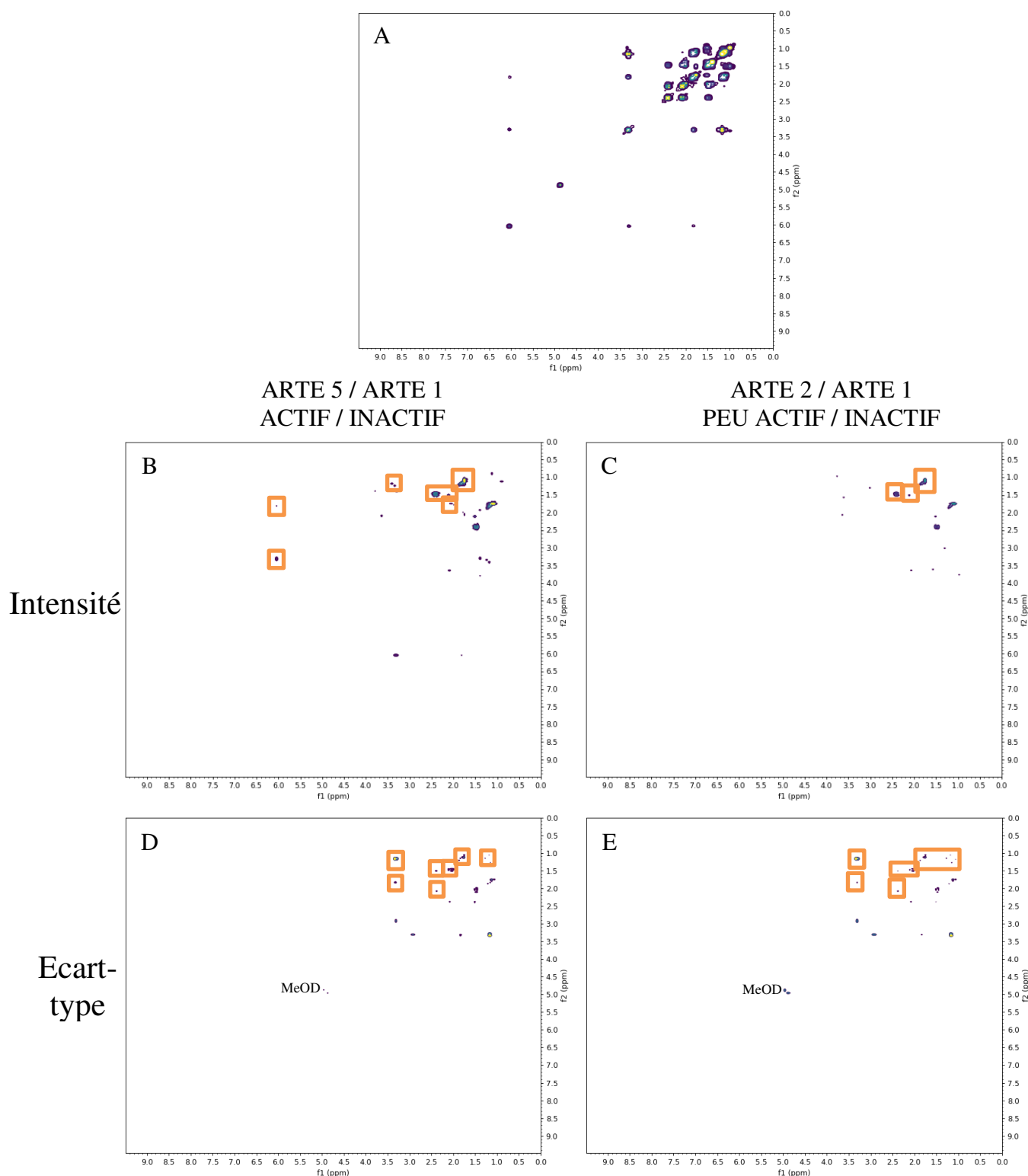


Figure 49 : Production automatique d'empreinte par le rapport des intensités ou des écarts-types des *buckets* des expériences ^1H - ^1H COSY des échantillons ARTE 5 (B et D) et ARTE 2 (C et E) versus ARTE 1 (inactif). Spectre COSY de l'artémisinine (A). L'empreinte de l'artémisinine est visible, également à partir de l'échantillon le moins concentré (ARTE 2). L'empreinte est de meilleure qualité en utilisant les écarts-types (D et E).

Le rapport des écarts-types des *buckets* a permis de produire l'empreinte spectrale ^1H - ^1H COSY de l'artémisinine à partir de chaque échantillon ARTE. L'artémisinine est facilement identifiable, notamment à partir de ARTE 2, l'échantillon le moins concentré en artémisinine (1% w/w). L'empreinte de l'artémisinine est nettement meilleure en utilisant les écarts-types.

Le rapport des intensités ou des écarts-types est appliqué aux autres analyses RMN enregistrées, ^1H - ^1H TOCSY, ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC des deux séries.

IV. 3. b. Analyse différentielle deux à deux de spectres ^1H - ^1H TOCSY

- **Série CHLORO**

Les analyses ^1H - ^1H TOCSY (*T*otal *C*orrelation *S*pectroscop*Y*), est une expérience RMN ^1H - ^1H « longue distance ». Elle crée des corrélations entre tous les protons qui sont dans le même système de spin. Dans notre cas, elle présente l'avantage de multiplier le nombre de signaux donc d'augmenter l'information pour un temps d'analyse équivalent ou inférieur à une ^1H - ^1H COSY.

Les résultats des rapports des intensités ou des écarts-types des analyses ^1H - ^1H TOCSY permettent d'obtenir l'empreinte de la chloroquine depuis les spectres CHLORO 5, 2 sur 1. L'empreinte est meilleure lorsque le rapport des écarts-types est effectué, notamment pour CHLORO 5 sur CHLORO 1 où aucune tache de corrélation de l'extrait n'apparaît (Figure 50).

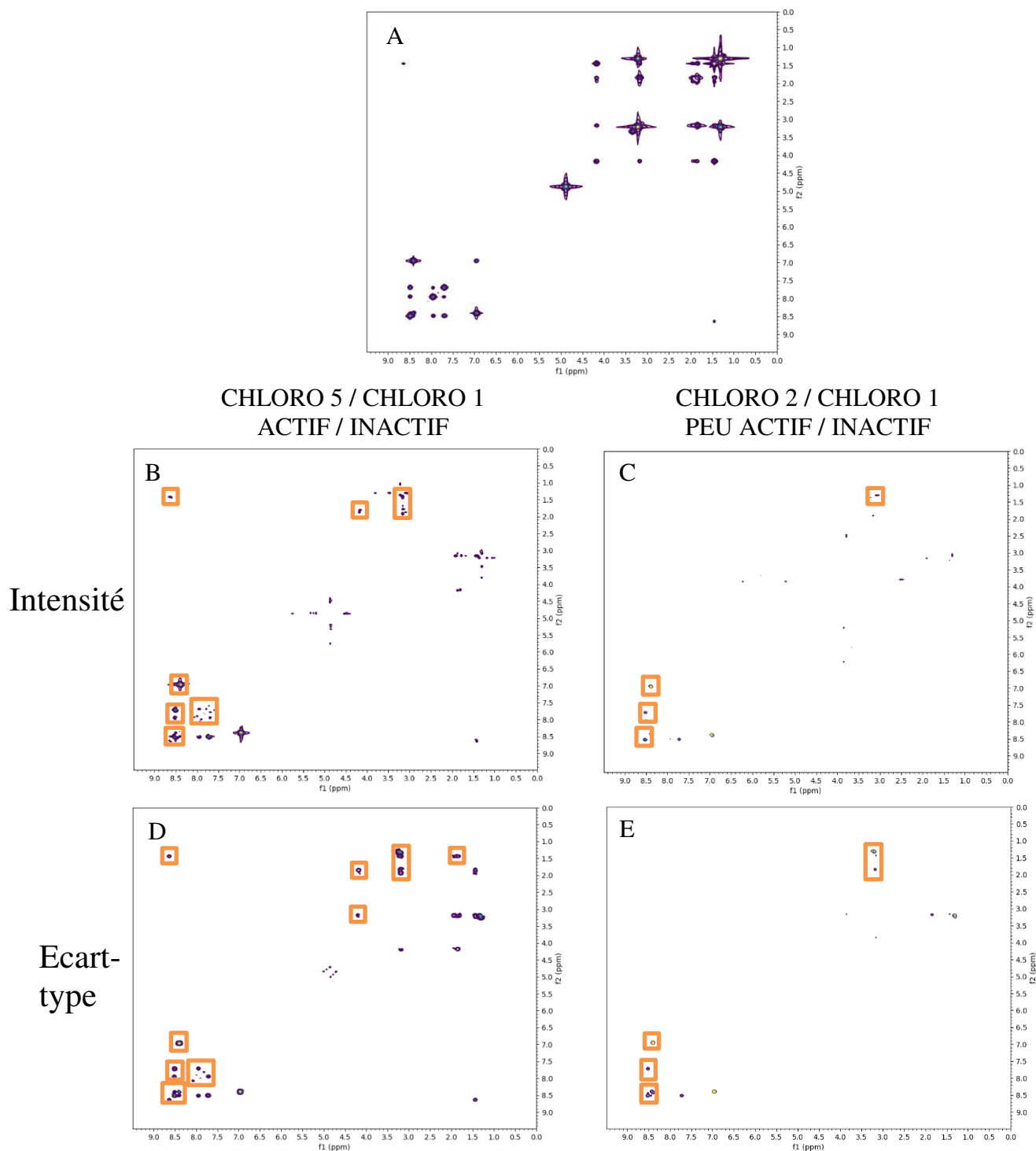


Figure 50 : Production automatique de l'empreinte par le rapport des intensités ou des écarts-types des buckets des expériences ^1H - ^1H TOCSY des échantillons CHLORO 5 (B et D) et CHLORO 2 (C et E) versus CHLORO 1 (inactif). Spectre ^1H - ^1H TOCSY de la chloroquine (A). L'empreinte de la chloroquine est visible dans chaque cas, également à partir de l'échantillon le moins concentré. Les meilleurs résultats sont obtenus par l'utilisation des écarts-types.

- Série ARTE :

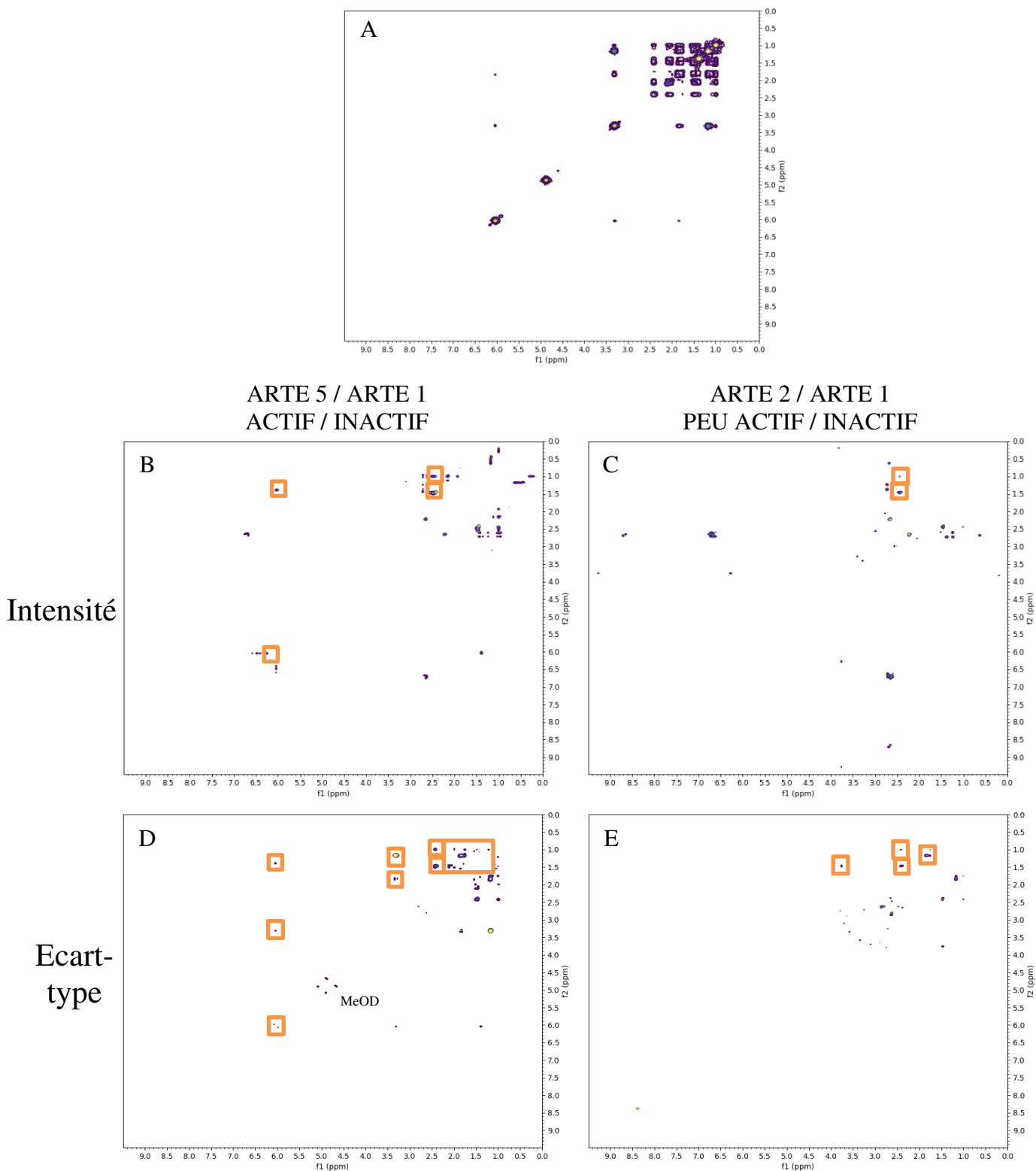


Figure 51 : Production automatique de l'empreinte par le rapport des intensités ou des écarts-types des buckets des expériences ^1H - ^1H TOCSY des échantillons ARTE 5 (B et D) et ARTE 2 (C et E) versus ARTE 1 (inactif). Spectre ^1H - ^1H TOCSY de l'artémisinine (A). L'empreinte de l'artémisinine est très peu visible par le rapport des intensités mais visible depuis le rapport des écarts-types.

Le rapport des écarts-types et des intensités des échantillons actifs *versus* inactif est appliqué à la série ARTE (Figure 51).

L'empreinte spectrale de l'artémisinine est nettement plus visible en utilisant les écarts-types. L'artémisinine n'est pas identifiable depuis les intensités de l'échantillon ARTE 2.

IV. 3. c. Analyse différentielle deux à deux de spectres ^1H - ^{13}C HSQC

Les expériences ^1H - ^{13}C HSQC (*Heteronuclear Single Quantum Correlation*) sont des analyses RMN 2D hétéronucléaire. Elle permet de corrélérer les déplacements chimiques ^1H et ^{13}C d'un même groupement ($^1J_{\text{CH}}$).

Nos expériences ^1H - ^{13}C HSQC possèdent des signaux dans le négatif. Cela est dû à l'utilisation d'une technique appelée « *aliasing* » développé par Damien Jeannerat à l'Université de Genève¹⁶¹. Cette technique consiste en le repliement d'un spectre 2D hétéronucléaire sur lui-même. Habituellement, les spectres ^1H - ^{13}C HSQC ont une fenêtre large de 200 ppm. Nous avons réduit celle-ci à 150, ce qui reste néanmoins large par rapport à ce qui est possible par le « *aliasing* ». Pour un temps d'analyse équivalent, cette technique permet d'augmenter la résolution d'un spectre, de diminuer le rapport signal/bruit et donc d'augmenter la sensibilité. En conséquence, il est possible de réaliser une expérience ^1H - ^{13}C HSQC moins longue pour une sensibilité équivalente.

- **Série CHLORO :**

Les rapports des écarts-types et des intensités sont appliqués aux analyses ^1H - ^{13}C HSQC (Figure 52).

L'empreinte de la chloroquine est produite de manière sélective grâce au rapport des écarts-types ou des intensités. Cependant, nous pouvons observer que l'empreinte est incomplète et peu sélective pour CHLORO 2 / CHLORO 1. La ^1H - ^{13}C HSQC, en raison de son couplage hétéronucléaire ^1H - ^{13}C et de la faible abondance naturelle du ^{13}C est une expérience moins sensible qu'une expérience homonucléaire ^1H - ^1H . La concentration minimum de chloroquine était de 0.07 mg/mL soit 0.5 % de la masse d'extrait algal. Les résultats sont légèrement meilleurs avec l'écart-type, les taches de corrélations de la chloroquine étant plus abondantes et plus intenses.

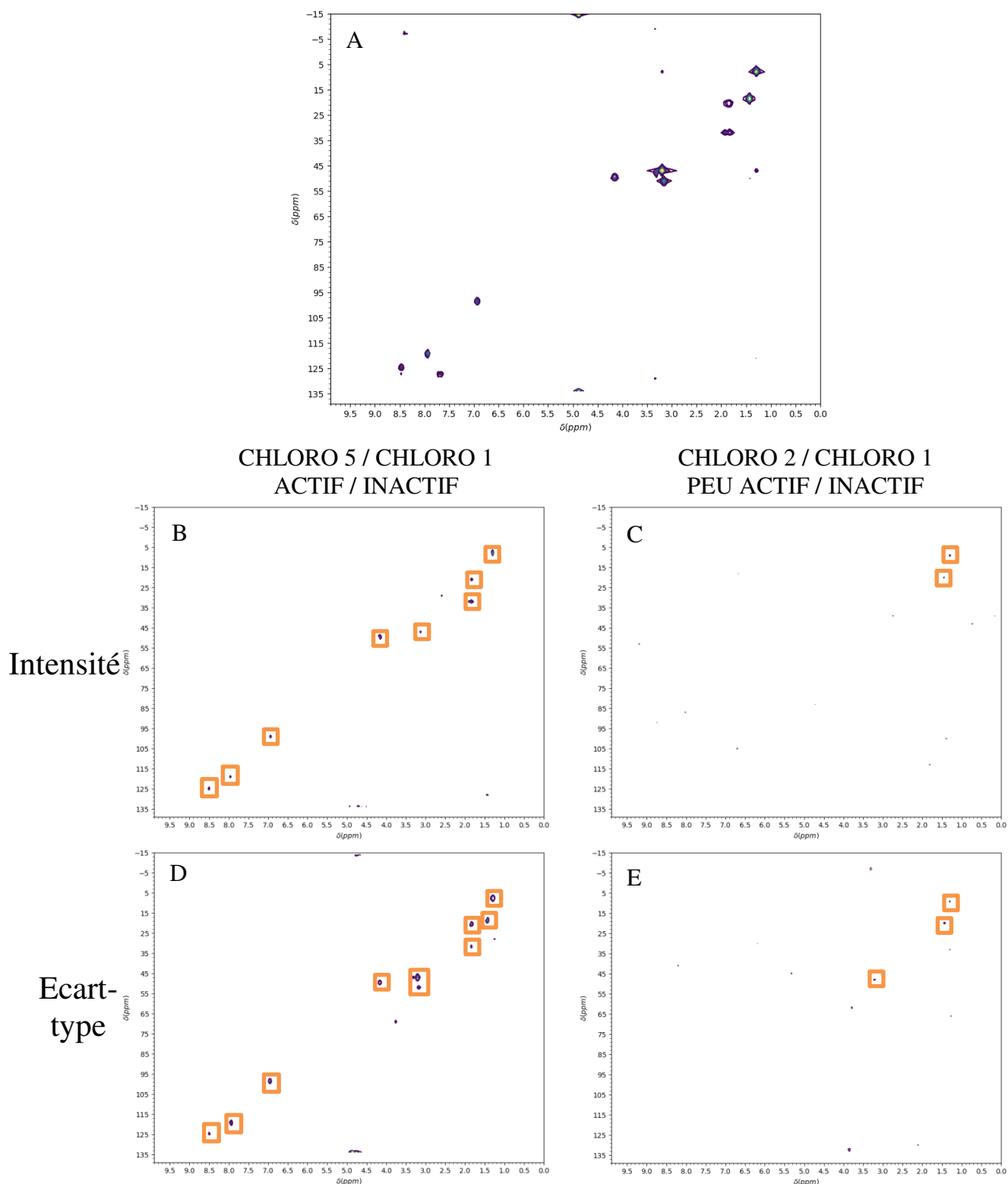


Figure 52 : Production automatique de l'empreinte par le rapport des intensités ou des écarts-types des buckets des expériences ^1H - ^{13}C HSQC des échantillons CHLORO 5 (B et D) et CHLORO 2 (C et E) versus CHLORO 1 (inactif). Spectre ^1H - ^{13}C HSQC de la chloroquine (A). L'empreinte est quasiment complète via CHLORO 5 mais très incomplète via CHLORO 2.

- Série ARTE :

Les rapports des écarts-types et des intensités sont appliqués aux analyses ^1H - ^{13}C HSQC de la série ARTE (Figure 53).

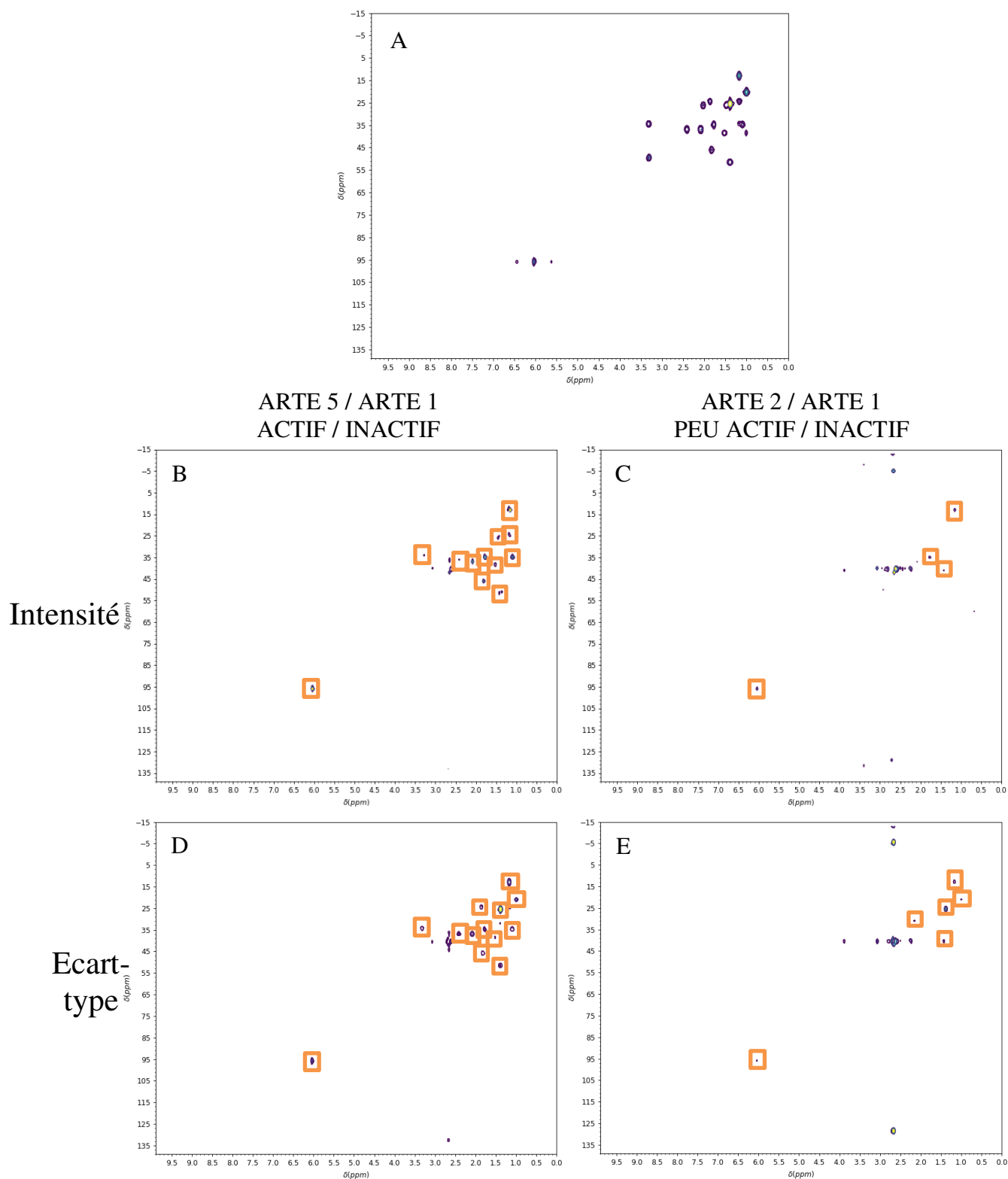


Figure 53 : Production automatique de l'empreinte par le rapport des intensités ou des écarts-types des buckets des expériences ^1H - ^{13}C HSQC des échantillons ARTE 5 (B et D) et ARTE 2 (C et E) versus ARTE 1 (inactif). Spectre ^1H - ^{13}C HSQC de l'artémisinine (A). L'empreinte est quasiment complète via ARTE 5 mais incomplète via ARTE 2 avec la présence de signaux de l'extrait algal.

L'empreinte spectrale de l'artémisinine est nettement visible depuis ARTE 5, elle est plus incomplète depuis ARTE 2 mais reste identifiable. Cependant, la présence d'une tache de corrélation de forte intensité est présente (δ_H 2,64 ; δ_C 40,4 ppm) sur chaque échantillon, mais isolée, elle ne perturbe pas l'identification de l'artémisinine. Les résultats sont également meilleurs par le rapport des écarts-types des *buckets*.

IV. 3. d. Analyse différentielle deux à deux de spectres 1H - ^{13}C HMBC

Les expériences 1H - ^{13}C HMBC (*Heteronuclear Multiple Bond Correlation*) sont des analyses RMN 2D hétéronucléaires. Elle met en évidence les couplages 1H - ^{13}C « longue distance », 2J , 3J et plus. L'*aliasing* a également été utilisé sur les expériences 1H - ^{13}C HMBC.

- **Série CHLORO :**

L'expérience 1H - ^{13}C HMBC est une expérience peu sensible en raison de son couplage ^{13}C et présente un temps d'acquisition très long dans notre cas. L'empreinte issue de CHLORO 2 / CHLORO 1 est incomplète mais les taches de corrélation les plus intenses appartiennent à la chloroquine. Les résultats sont meilleurs avec l'utilisation des écarts-types (Figure 54).

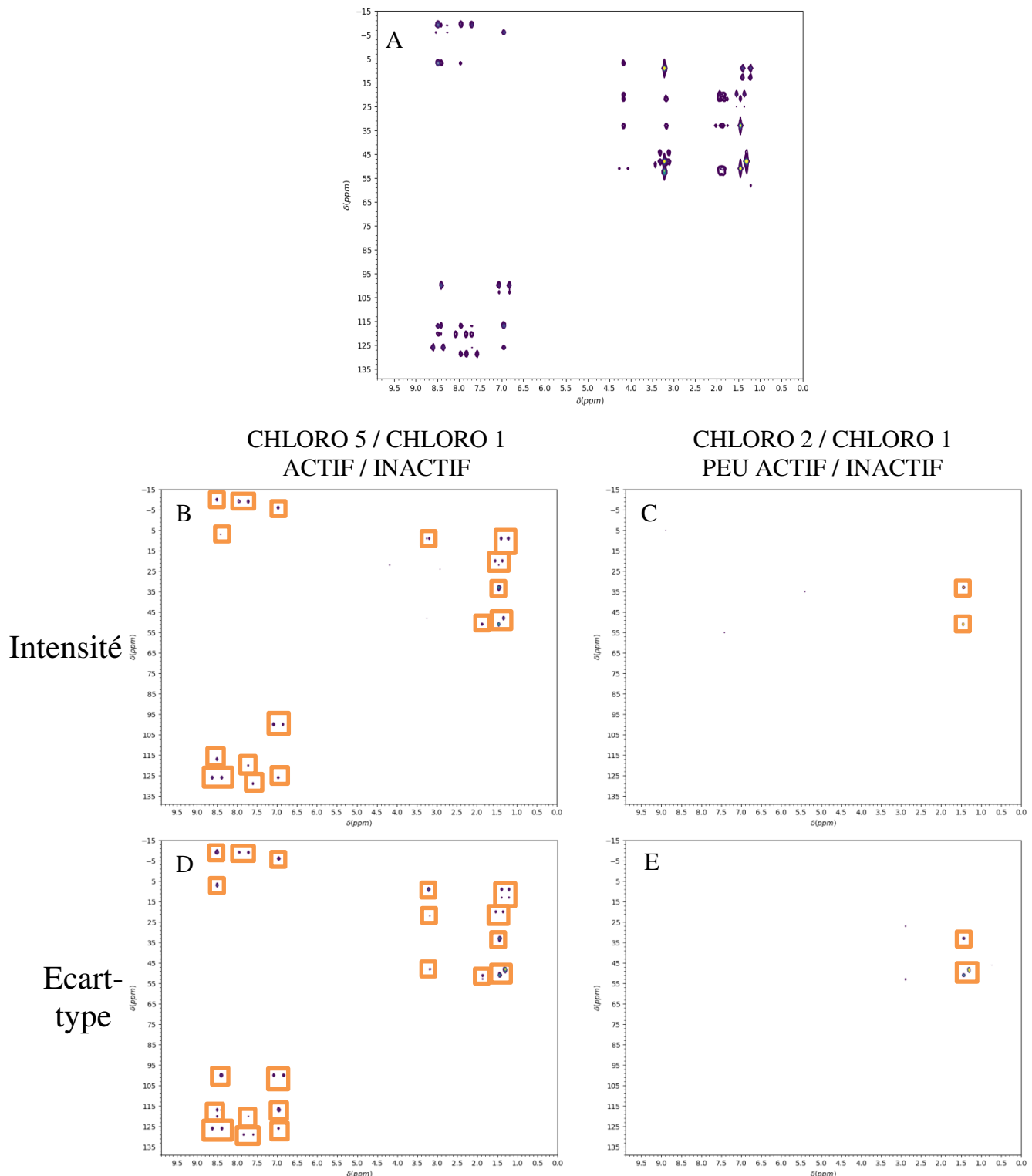
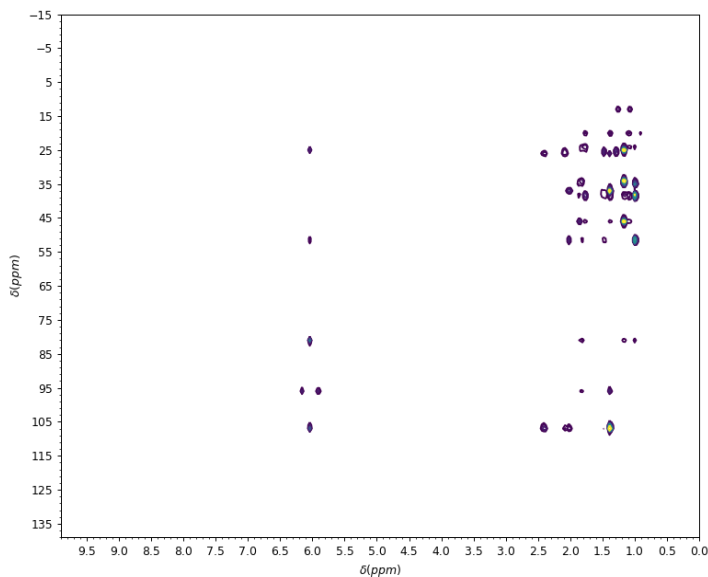


Figure 54 : Production automatique de l'empreinte par le rapport des intensités ou des écarts-types des buckets des expériences ^1H - ^{13}C HMBC des échantillons CHLORO 5 (B et D) et CHLORO 2 (C et E) versus CHLORO 1 (inactif). Spectre ^1H - ^{13}C HMBC de la chloroquine (A). L'empreinte de la chloroquine est quasiment complète pour CHLORO 5 et très incomplète depuis CHLORO 2.

• Série ARTE :



ARTE 5 / ARTE 1
ACTIF / INACTIF

ARTE 2 / ARTE 1
PEU ACTIF / INACTIF

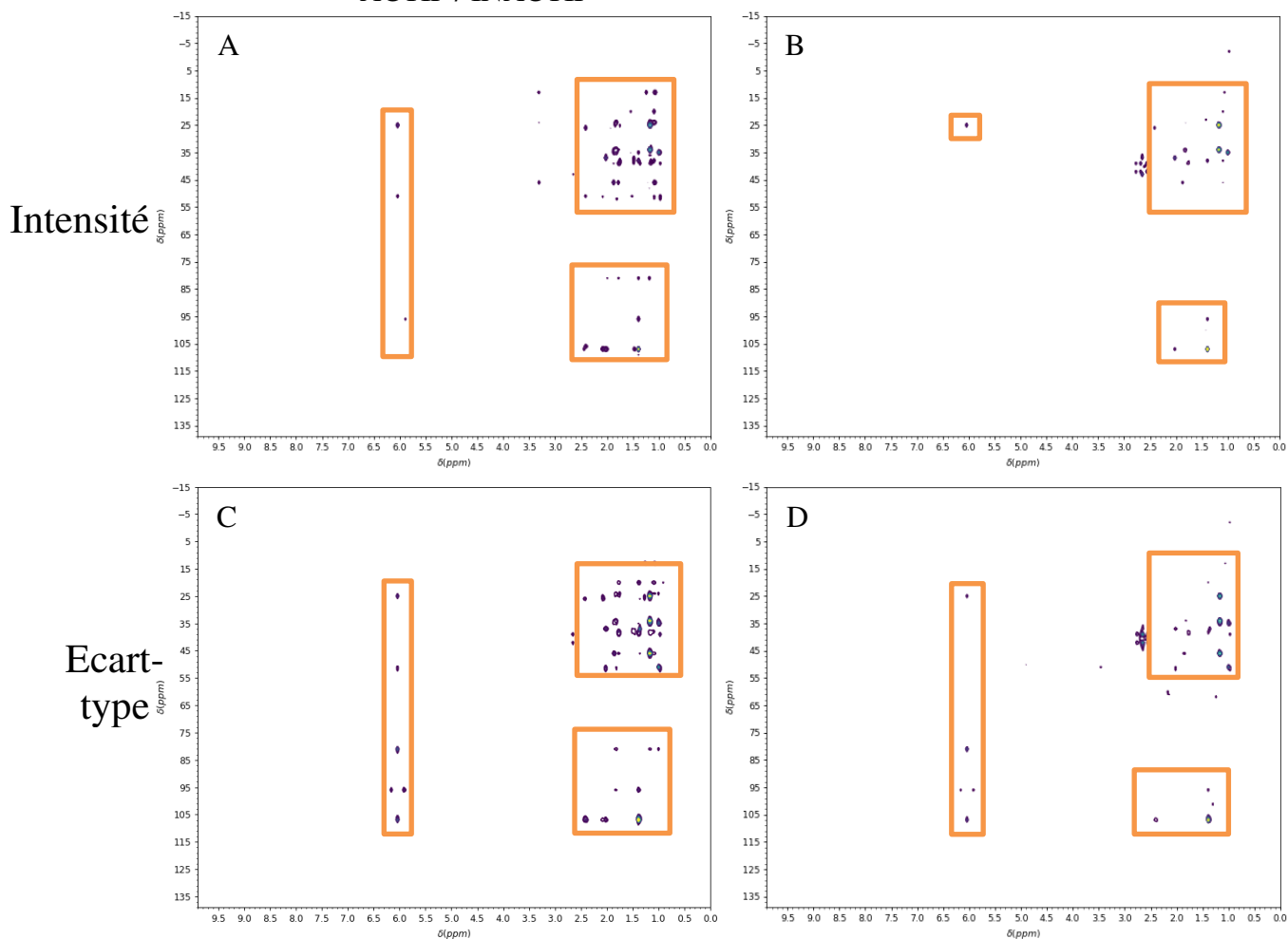


Figure 55 : Production automatique de l’empreinte par le rapport des intensités ou des écarts-types des buckets des expériences ^1H - ^{13}C HMBC des échantillons ARTE 5 (B et D) et ARTE 2 (C et E) versus ARTE 1 (inactif). Spectre ^1H - ^{13}C HMBC de l’artémisinine (A). L’empreinte est quasiment complète via ARTE 5, quelques signaux de l’extrait algal depuis ARTE 2.

L'empreinte spectrale de l'artémisinine est nettement visible depuis ARTE 5 et ARTE 2. Quelques signaux de l'extrait algal apparaissent mais l'artémisinine est clairement identifiable même à partir de l'échantillon le moins concentré (Figure 55).

IV. 3. e. Conclusion sur l'analyse différentielle de spectres de fractions actives versus inactives

L'analyse différentielle automatique de spectres RMN d'échantillons actifs *versus* inactifs a permis de produire une empreinte pharmacophorique permettant d'identifier la chloroquine pour les échantillons CHLORO 5 à 2 *versus* CHLORO 1. Cela a été possible plus particulièrement grâce aux rapports des **écarts-types des buckets** qui donnent les meilleurs résultats et cette méthode est applicable aux expériences homo- et hétéronucléaires.

Ces résultats ont été appuyés par la série ARTE qui était composée d'artémisinine. Cette molécule possède une fenêtre spectrale plus réduite que la chloroquine, la plupart de ces signaux étant compris entre 0 et 3.5 ppm. Les rapports des écarts-types des *buckets* ont également donné les meilleurs résultats pour la production de l'empreinte. L'empreinte était visible pour les échantillons ARTE 5 à 2.

Nous avons pu observer que les expériences hétéronucléaires ^1H - ^{13}C peuvent présenter une limite de détection des molécules minoritaires en raison de la faible abondance naturelle de ^{13}C .

La concentration en chloroquine dans CHLORO 2 était de 0.07 mg/mL soit 0.5 % de la masse de l'extrait. La concentration d'artémisinine dans ARTE 2 était de 0.2 mg/mL soit 1.3 % de la masse de l'extrait. Dans les deux cas, l'empreinte obtenue permet d'identifier la molécule active dans l'extrait, bien que 0.5% de molécule active semble être la limite de détection.

IV. 4. Analyse différentielle de l'ensemble des échantillons

Dans un premier temps l'analyse différentielle a été réalisée par comparaison d'échantillons deux à deux dans des situations de présence *versus* absence de la molécule bioactive. Mais, dans le cas où l'activité biologique est partagée entre plusieurs fractions, un autre type de calcul a été établi pour produire l'empreinte pharmacophorique.

La RMN étant une méthode quantitative, l'intensité des signaux de la molécule active sera proportionnelle à sa concentration. Cette propriété peut donc être exploitée pour distinguer les signaux de la molécule bioactive des autres signaux.

Pour cela, la bibliothèque `scikit-learn` a été utilisée. Elle est spécialisée dans le « *machine learning* » (apprentissage statistique) et contient une panoplie d'outils statistiques pour la classification, régression, « *clustering* », application de modèle... Elle a l'avantage d'être complètement compatible avec python, Jupyter et SPIKE.

Nous avons donc cherché à exacerber des signaux qui suivent une fluctuation de concentration massique ou d'activité biologique. Pour cela, des modèles de régression ont été appliqués à nos données. Nous avons utilisé la fonction de régression linéaire `linear_model.LinearRegression()` ainsi que la méthode d'élimination récursive des variables (*recursive feature elimination*, RFE). Ces outils permettent de sélectionner un petit nombre de paramètres qui corréleront le mieux avec la concentration de la molécule active. Les motifs sélectionnés devraient alors correspondre à l'empreinte spectrale de la molécule.

Dans le cas de notre analyse différentielle, la régression linéaire a été appliquée entre l'intensité des signaux et l'activité biologique des cinq échantillons.

Dans un premier temps, nous observerons les résultats de deux types de régression linéaire sur les analyses ^1H - ^1H COSY de la série CHLORO, puis sur l'ensemble des expériences RMN de la série CHLORO et ARTE.

IV. 4. a. Analyse par régression linéaire : Utilisation de la fonction `linear_model.LinearRegression`

Dans le cas des séries CHLORO et ARTE, nous connaissons les **concentrations** en molécules actives dans l'extrait pour chaque échantillon. Ces valeurs seront tout d'abord utilisées pour réaliser la régression linéaire en fonction de l'intensité des *buckets* de l'ensemble des expériences RMN, puis, dans un second temps, les **pourcentages d'inhibition** des échantillons seront utilisés. L'application de la régression linéaire sur nos données est détaillée dans la partie expérimentale. La régression linéaire est appliquée aux 5 échantillons en fonction des intensités des *buckets* et de la concentration en chloroquine. Elle permet d'obtenir une empreinte spectrale ^1H - ^1H COSY de la chloroquine quasiment complète sans la présence de signaux appartenant à l'extrait algal (Figure 56).

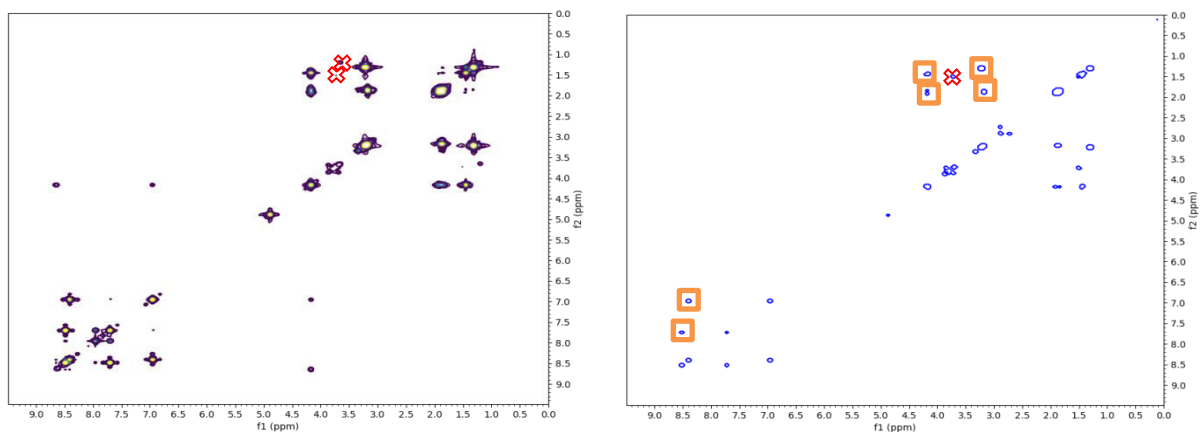


Figure 56 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des *buckets* des cinq échantillons CHLORO et les valeurs de concentration en chloroquine dans les fractions. L'empreinte de la chloroquine est quasiment complète.

Il est également possible de réaliser une régression linéaire entre les intensités des *buckets* et les pourcentages d'inhibition de croissance mesurés. Les résultats sont montrés dans la **Figure 57**.

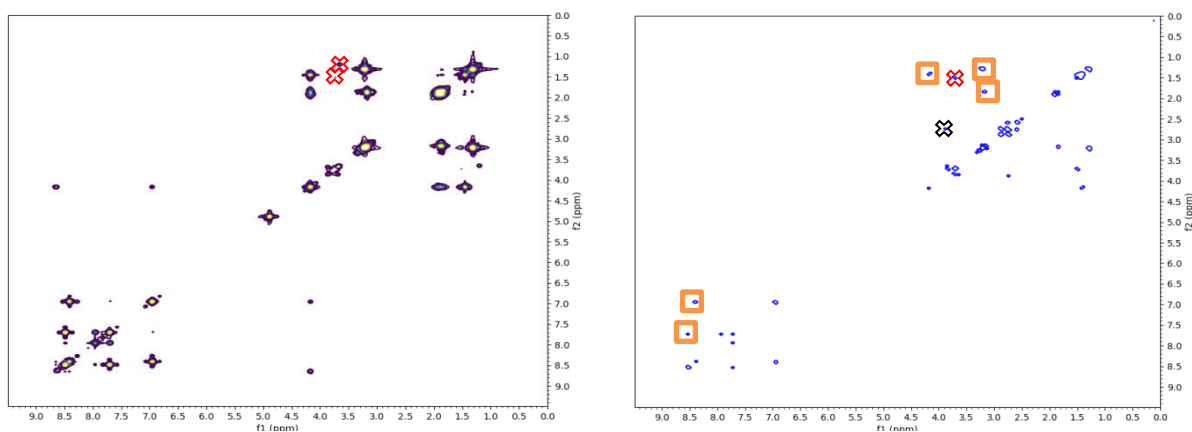


Figure 57 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des *buckets* des cinq échantillons CHLORO et les pourcentages d'inhibition. L'empreinte produite est quasiment complète, une tache de corrélation seulement provient de l'extrait algal.

L'empreinte produite permet d'identifier clairement la chloroquine, une tache de corrélation seulement provenant de l'extrait algal.

Le dernier échantillon de la série CHLORO, CHLORO 5, possède une concentration très forte en chloroquine (7.4 mg/mL). Une telle concentration pourrait influencer fortement les résultats d'empreinte. La régression linéaire est appliquée sans cet échantillon pour vérifier qu'il n'influe pas les résultats (la concentration de chloroquine dans CHLORO 1 à 4 est de 0 à 0.7 mg/mL soit 0 à 0.5% de chloroquine dans l'extrait, respectivement) (Figure 58).

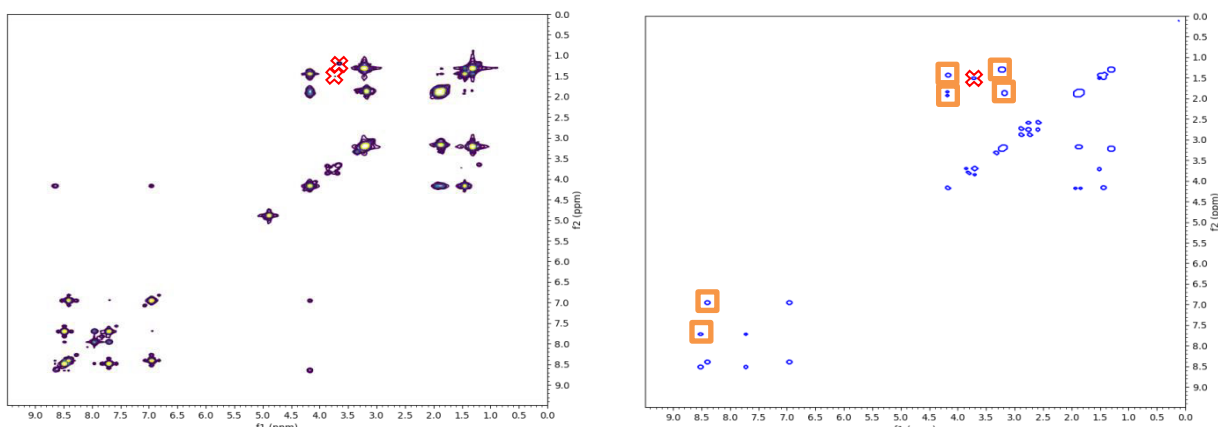


Figure 58 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des *buckets* des échantillons CHLORO 1, 2, 3 et 4 et les concentrations en chloroquine. L'empreinte obtenue est quasiment complète et dépourvue de signaux appartenant à l'extrait algal.

Sans l'échantillon CHLORO 5, l'empreinte obtenue est complète et dépourvue de signaux appartenant à l'extrait algal. La régression linéaire sélectionne de manière spécifique les signaux qui évoluent selon la gamme de concentration qu'on lui applique.

L'expérience est également réalisée en fonction de la gamme de pourcentage d'inhibition (Figure 59).

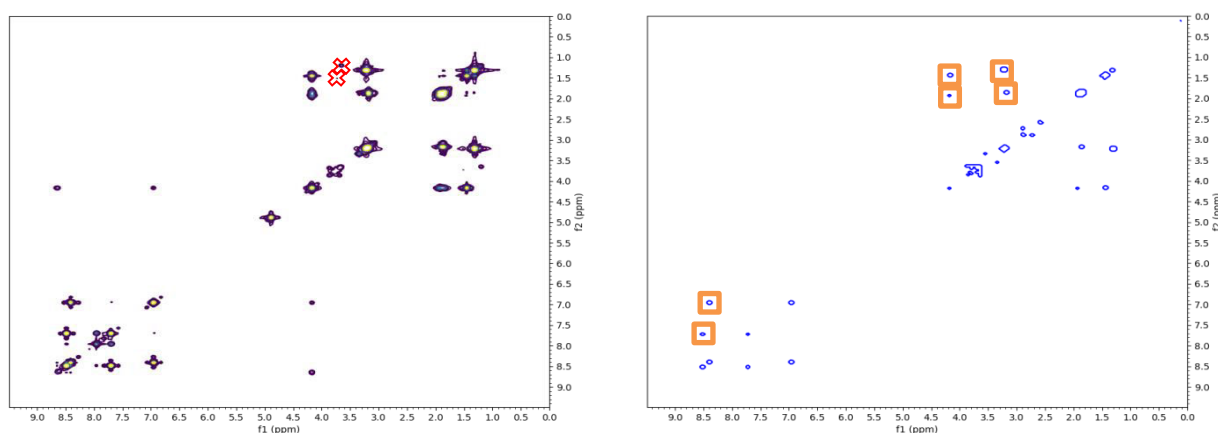


Figure 59 : Empreinte spectrale ^1H - ^1H COSY obtenue par régression linéaire entre les intensités des *buckets* des échantillons CHLORO 1, 2, 3 et 4 et les pourcentages d'inhibition. L'empreinte produite est quasiment complète et dépourvue de signaux de l'extrait algal.

L'empreinte produite est complète sans aucune tache de corrélation provenant de l'extrait algal.

La régression linéaire est donc un moyen puissant de produire une empreinte spectrale de la molécule active, que ce soit en fonction des variations de concentration en la molécule active ou des variations du pourcentage d'inhibition. Dans les situations où seule l'activité

biologique est connue, il serait préférable d'exprimer celle-ci de manière la plus précise possible en réalisant une mesure de CI_{50} .

Les mêmes résultats ont été observés pour la série ARTE (Figure 60). L'empreinte de l'artémisinine a bien été produite par régression linéaire, tous les signaux de l'extrait algal ont été supprimés.

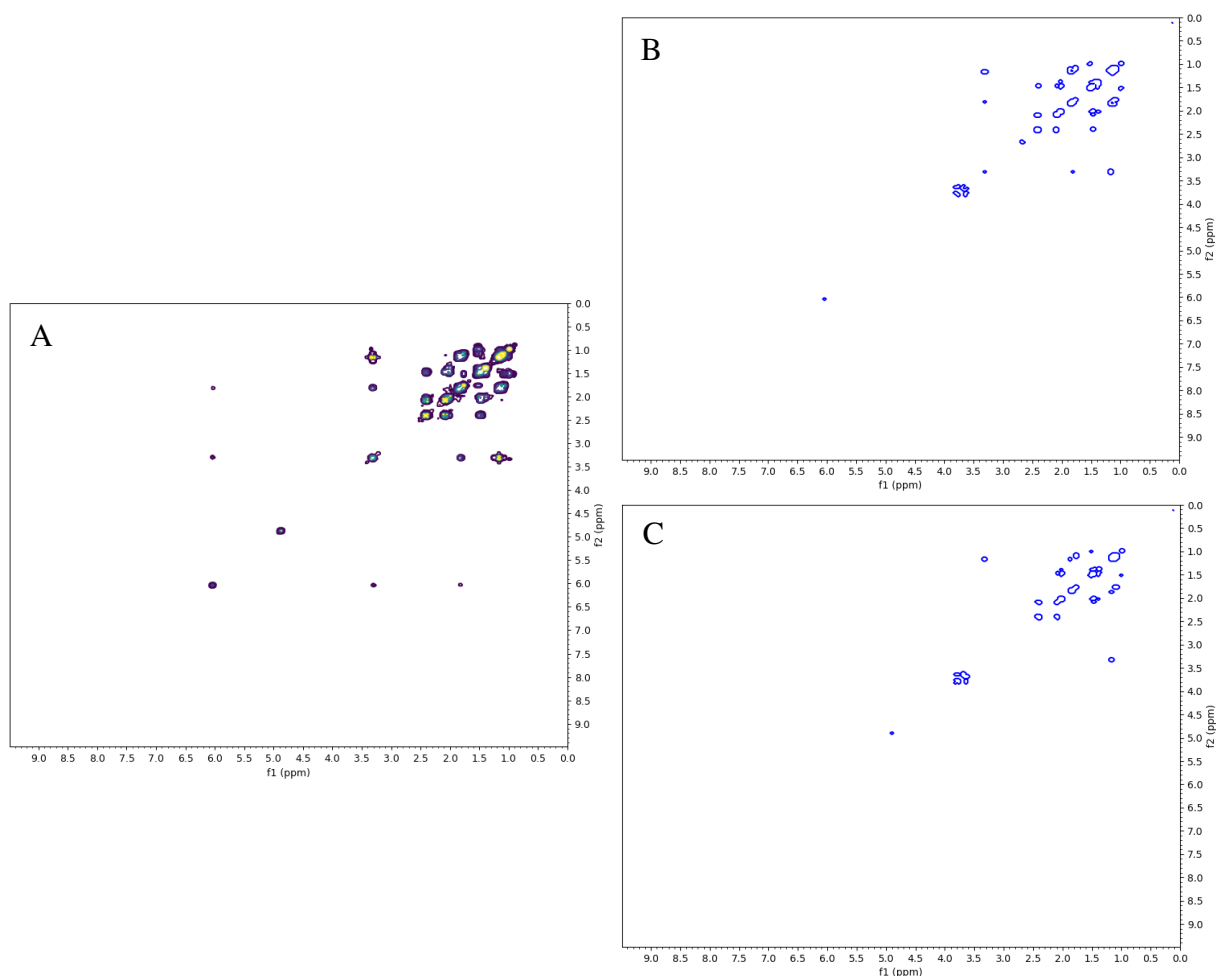


Figure 60 : Empreintes spectrales ^1H - ^1H COSY obtenues par régression linéaire entre les intensités des *buckets* des cinq échantillons ARTE et les concentrations en artémisinine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^1H COSY de l'artémisinine (A). L'empreinte de l'artémisinine est obtenue dans les deux cas, il n'y a pas de signaux appartenant à l'extrait algal.

IV. 4. b. Analyse par régression linéaire : Utilisation de l'outil RFE

La méthode d'élimination récursive des variables (RFE) est un outil de sélection de motif via une régression linéaire (*estimator*) ou un autre type de régression. L'estimateur va sélectionner un motif de données auxquelles il va assigner un poids. Le but de la RFE est de sélectionner ces motifs et les recalculer plusieurs fois en considérant des lots de motifs de plus en plus petits. Cet outil permettrait donc une seconde sélection, plus forte et précise que la sélection réalisée par la régression linéaire.

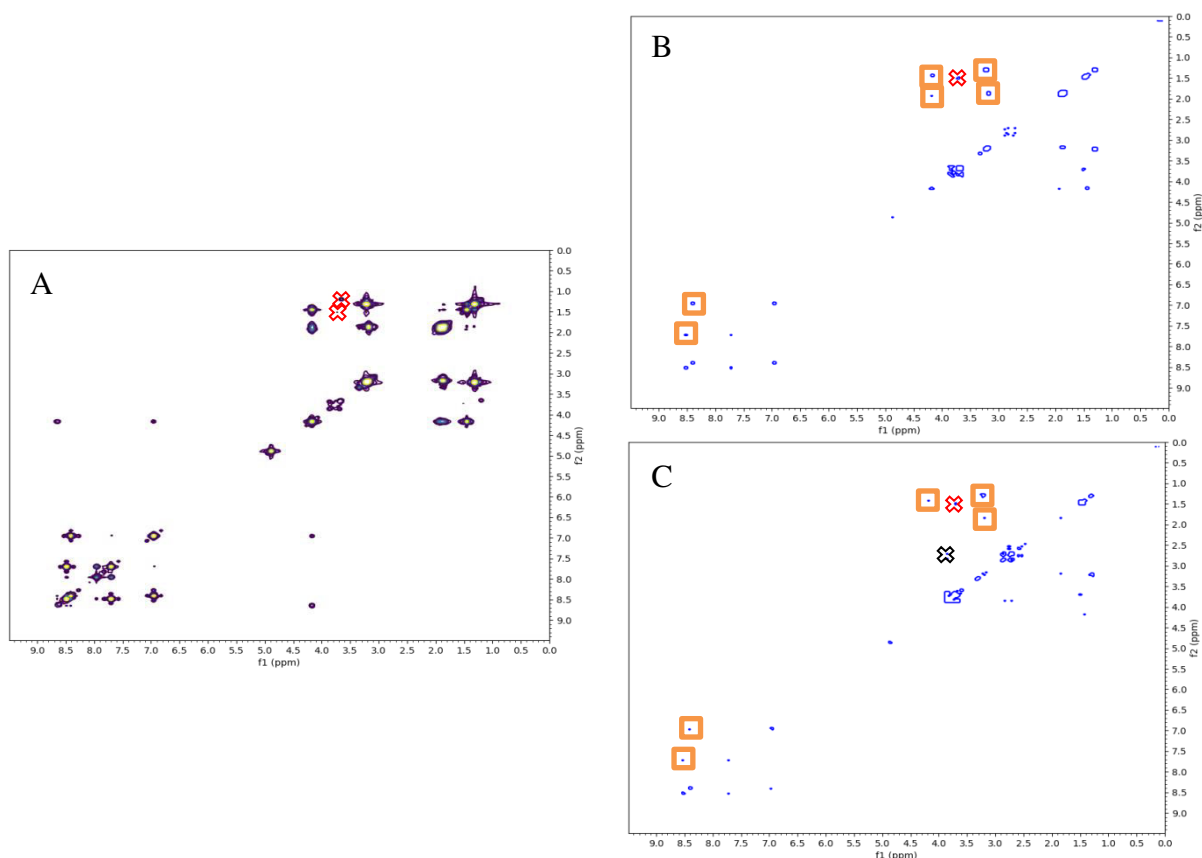


Figure 61 : Empreintes spectrales ^1H - ^1H COSY de la chloroquine obtenues par RFE entre les intensités des *buckets* des cinq échantillons CHLORO et les concentrations en chloroquine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^1H COSY de la chloroquine (A). L'empreinte de la chloroquine est visible et quasiment complète.

L'empreinte spectrale produite par la RFE est quasiment complète pour la série CHLORO en utilisant la gamme de concentrations ou de pourcentages d'inhibition. Cependant, quelques taches de corrélation provenant de l'extrait algal apparaissent sur l'empreinte en fonction des pourcentages d'inhibition sans pour autant interférer dans l'identification de la chloroquine (Figure 61, C).

La même analyse est utilisée sur les échantillons CHLORO 1 à 4 (Figure 62). L'empreinte spectrale produite est quasiment complète et correspond à la chloroquine.

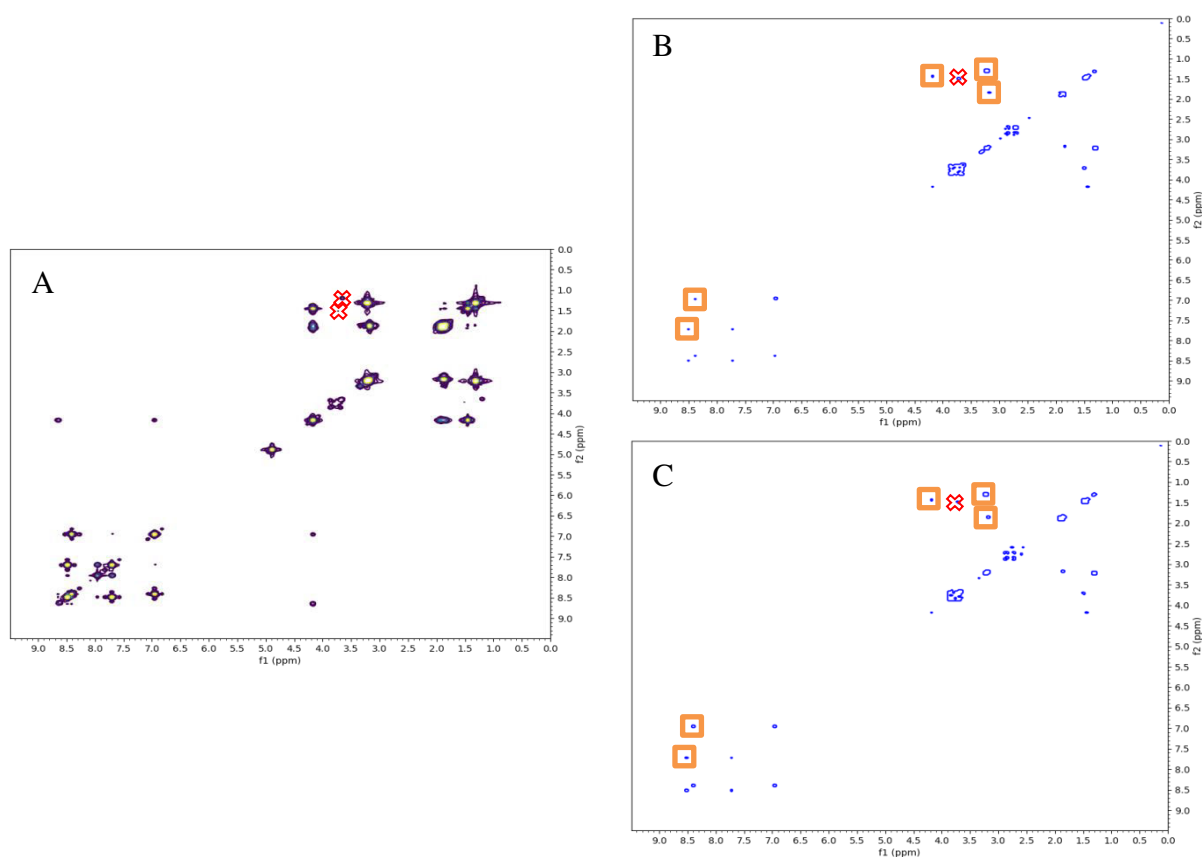


Figure 62 : Empreintes spectrales ^1H - ^1H COSY de la chloroquine obtenues par RFE entre les intensités des *buckets* des échantillons CHLORO 1 à 4 et les concentrations en chloroquine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^1H COSY de la chloroquine (A). L'empreinte de la chloroquine est visible et quasiment complète.

L'outil RFE permet de produire une empreinte spectrale correcte de la molécule active mais ne permet pas une amélioration notable de l'empreinte issue de la régression linéaire.

La même expérience est réalisée sur les autres expériences RMN et sur la série ARTE (données non montrées). Les conclusions sont les mêmes que pour les expériences ^1H - ^1H COSY de la série CHLORO.

Ainsi, la régression linéaire simple sera conservée pour la production automatique de l'empreinte dans les cas où l'activité biologique est partagée entre plusieurs fractions.

IV. 4. c. Analyse différentielle par régression linéaire sur les autres expériences RMN

L'analyse différentielle par régression linéaire est appliquée aux autres analyses RMN enregistrées, ^1H - ^1H TOCSY, ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC.

- **Résultats pour les expériences ^1H - ^1H TOCSY :**
 - **Série CHLORO**

L'analyse différentielle par régression linéaire est appliquée aux expériences RMN ^1H - ^1H TOCSY (Figure 63).

L'empreinte spectrale ^1H - ^1H TOCSY produite par régression linéaire entre les intensités et les concentrations en chloroquine est complète sans aucune autre tache de corrélation. Cependant, l'empreinte est incomplète par l'utilisation des pourcentages d'inhibition : il y a un signal appartenant à l'extrait algal, mais la chloroquine reste clairement identifiable.

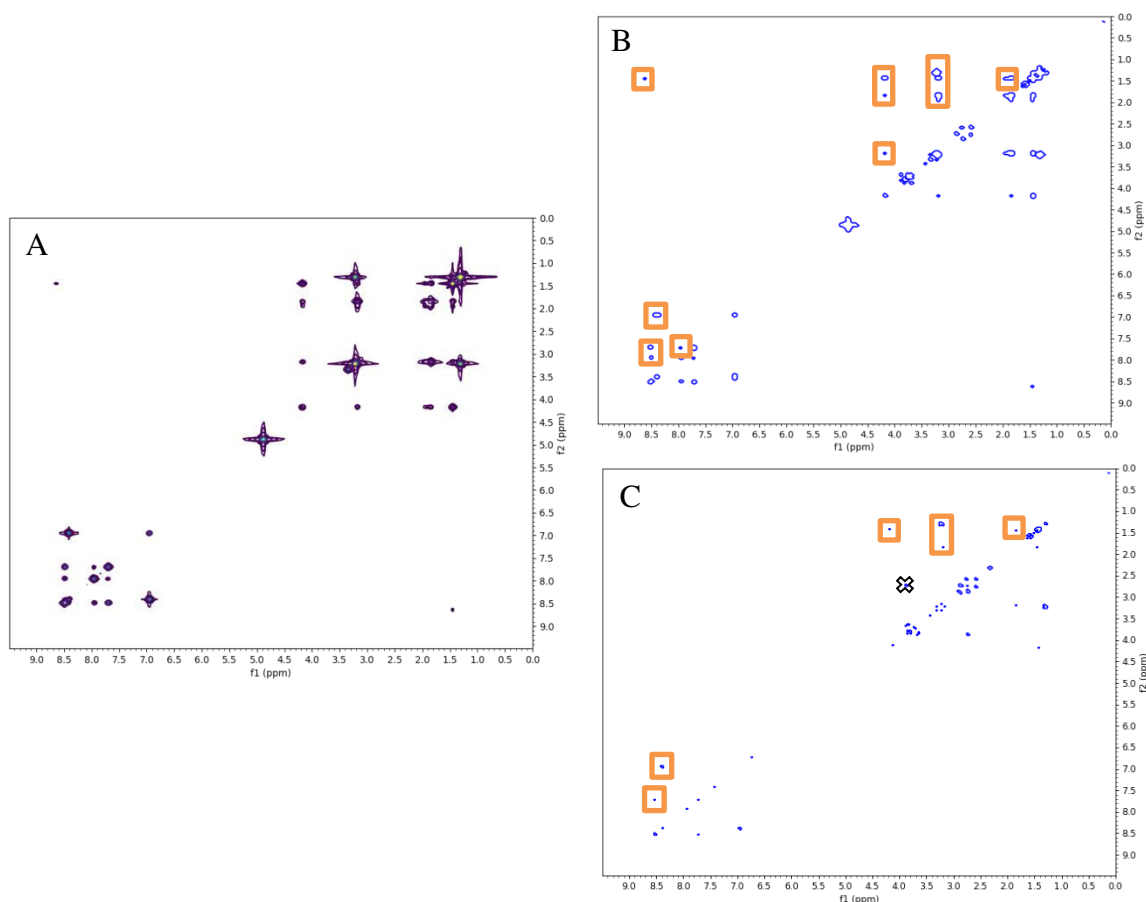


Figure 63 : Empreintes spectrales ^1H - ^1H TOCSY de la chloroquine obtenues par régression linéaire entre les intensités des *buckets* des échantillons CHLORO et les concentrations en chloroquine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^1H TOCSY de la chloroquine (A). L'empreinte de la chloroquine est visible et quasiment complète, une tache de l'extrait algal est visible sur le spectre C.

- Résultats pour les expériences ^1H - ^1H TOCSY :
 - Série ARTE

Les empreintes produites par régression linéaire des analyses ^1H - ^1H TOCSY correspondent bien à l'artémisinine. L'identification de l'artémisinine est évidente (Figure 64).

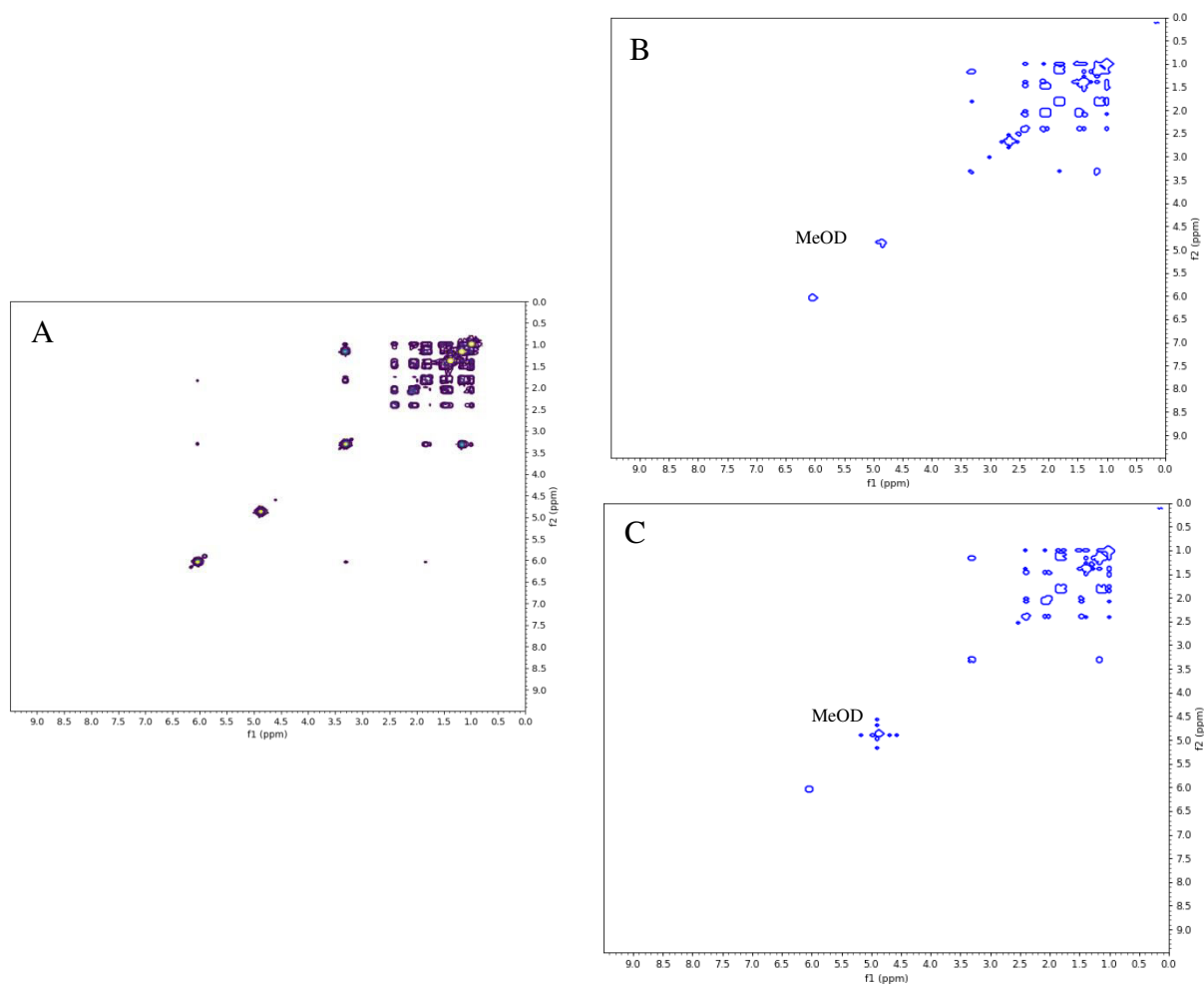


Figure 64 : Empreintes spectrales ^1H - ^1H TOCSY de l'artémisinine obtenues par régression linéaire entre les intensités des *buckets* des échantillons ARTE et les concentrations en artémisinine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^1H TOCSY de l'artémisinine (A). L'empreinte produite permet d'identifier clairement l'artémisinine. Toutes les taches de corrélation de l'extrait algal ont été supprimées.

- Résultats pour les expériences ^1H - ^{13}C HSQC :
 - Série CHLORO

L'analyse différentielle par régression linéaire est appliquée aux expériences RMN ^1H - ^{13}C HSQC (Figure 65).

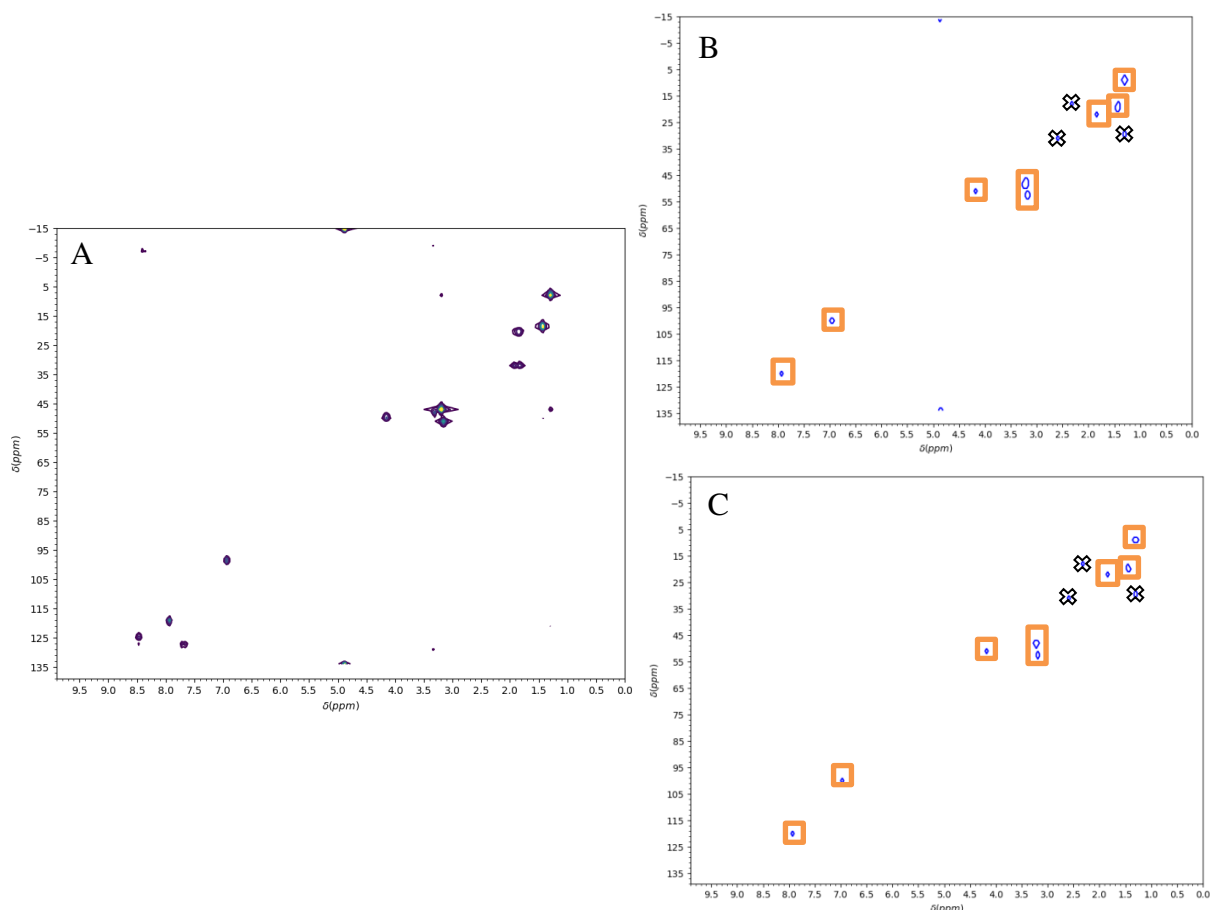


Figure 65 : Empreintes spectrales ^1H - ^{13}C HSQC de la chloroquine obtenues par régression linéaire entre les intensités des *buckets* des échantillons CHLORO et les concentrations en chloroquine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^{13}C HSQC de la chloroquine (A). L'empreinte de la chloroquine est visible et quasiment complète. Quelques taches de corrélations de l'extrait apparaissent.

L'empreinte spectrale ^1H - ^{13}C HSQC produite par la régression linéaire et les concentrations en chloroquine est quasiment complète, quelques autres taches de corrélation de l'extrait apparaissent.

- Résultats pour les expériences ^1H - ^{13}C HSQC :
 - Série ARTE

L'analyse différentielle par régression linéaire est appliquée aux expériences RMN ^1H - ^{13}C HSQC (Figure 66).

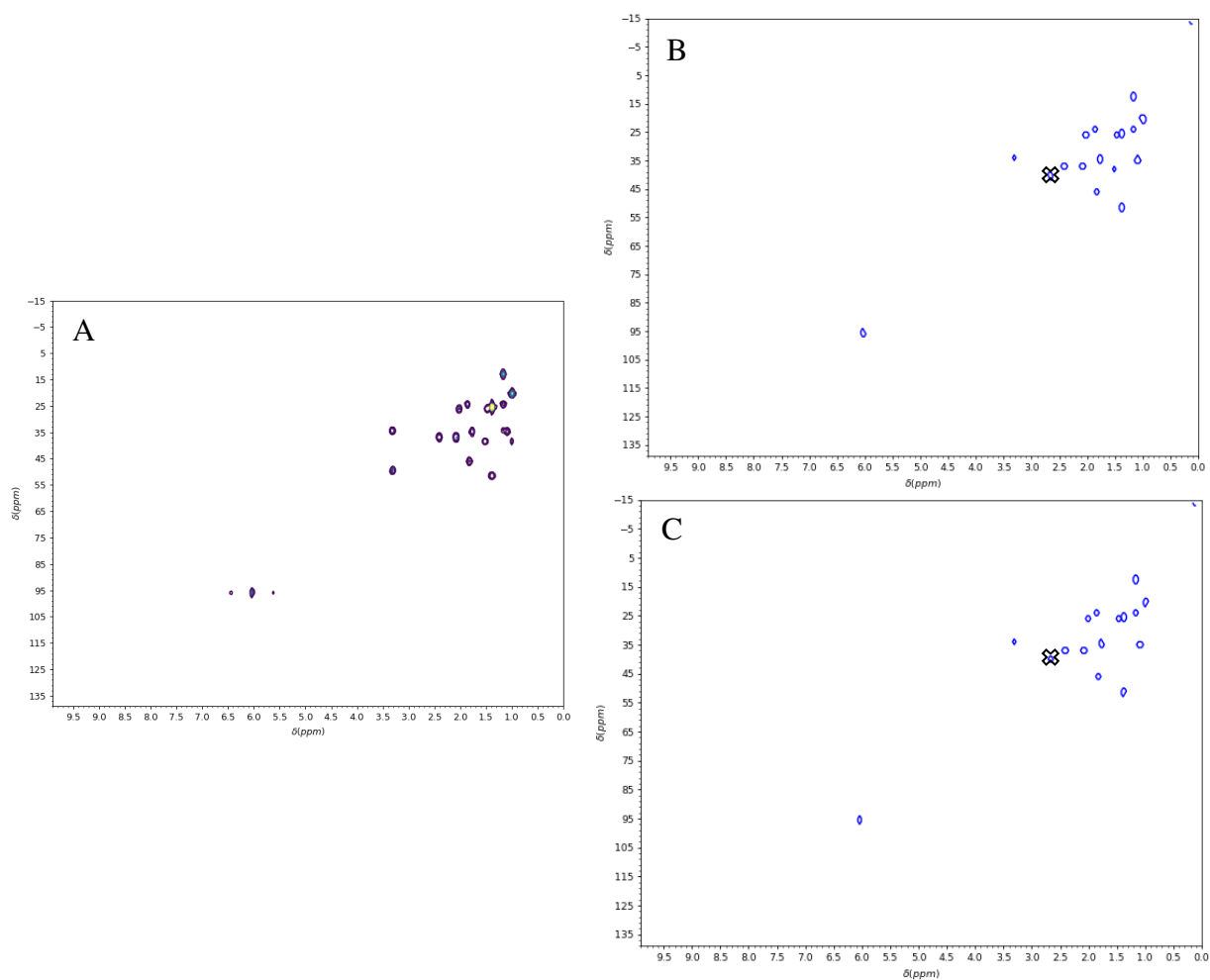


Figure 66 : Empreintes spectrales ^1H - ^{13}C HSQC de l'artémisinine obtenues par régression linéaire entre les intensités des *buckets* des échantillons ARTE et les concentrations en artémisinine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^{13}C HSQC de l'artémisinine (A). L'empreinte de l'artémisinine est quasiment complète, une seule tache de corrélation appartenant à l'extrait apparaît.

L'empreinte de l'artémisinine est bien produite par la régression linéaire en fonction des concentrations ou du pourcentage d'inhibition. Seule une tache de corrélation de l'extrait est présente.

- Résultats pour les expériences ^1H - ^{13}C HMBC :
 - Série CHLORO

Enfin, l'analyse différentielle par régression linéaire est appliquée aux expériences RMN ^1H - ^{13}C HMBC (Figure 67).

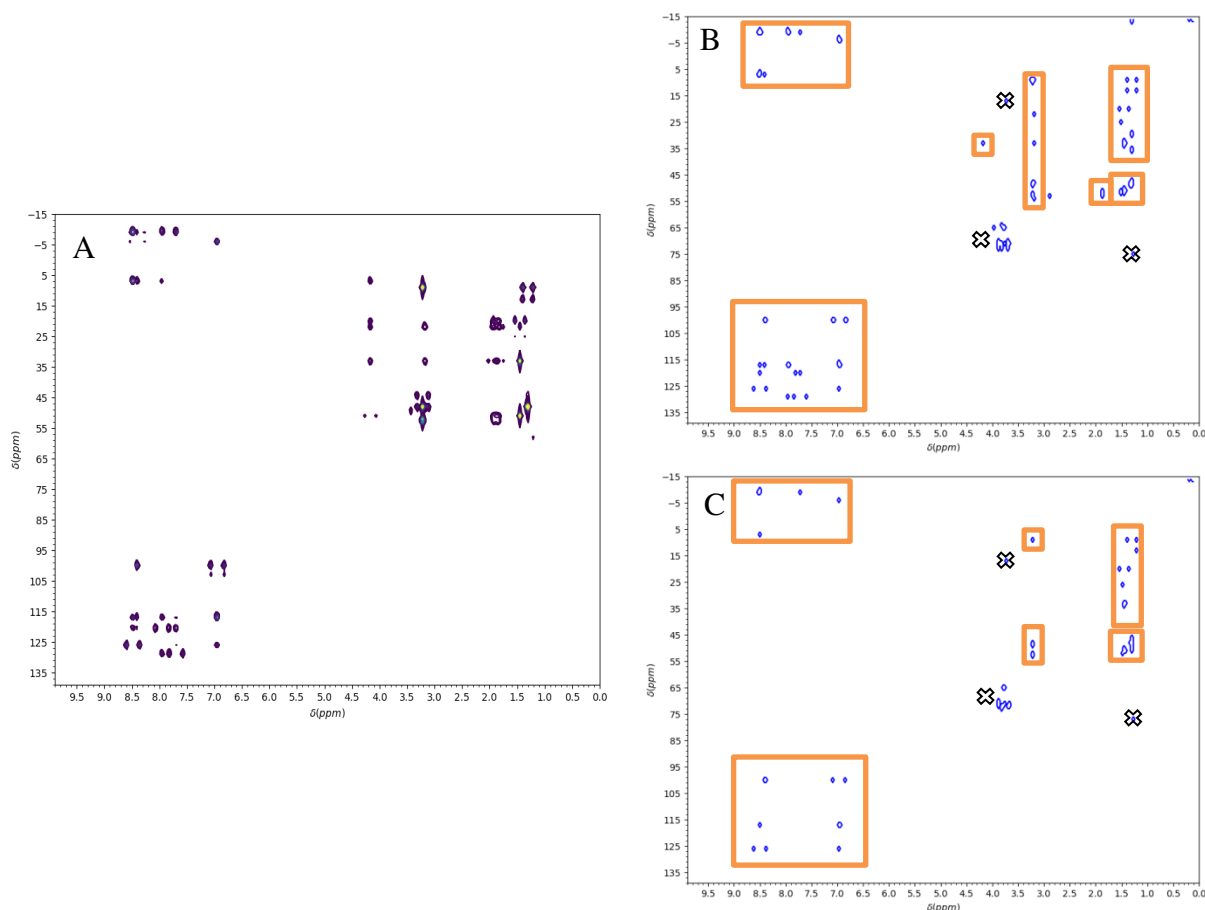


Figure 67 : Empreintes spectrales ^1H - ^{13}C HMBC de la chloroquine obtenue par régression linéaire entre les intensités des *buckets* des échantillons CHLORO et les concentrations en chloroquine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^{13}C HMBC de la chloroquine (A). L'empreinte de la chloroquine est visible et quasiment complète. Quelques taches de corrélations de l'extrait apparaissent.

L'analyse différentielle par régression linéaire sur les expériences ^1H - ^{13}C HMBC permet de produire une empreinte spectrale de la chloroquine quasiment complète. Quelques taches proviennent de l'extrait algal mais restent minoritaires et n'altèrent pas l'identification de la chloroquine.

- Résultats pour les expériences ^1H - ^{13}C HMBC :
 - Série ARTE

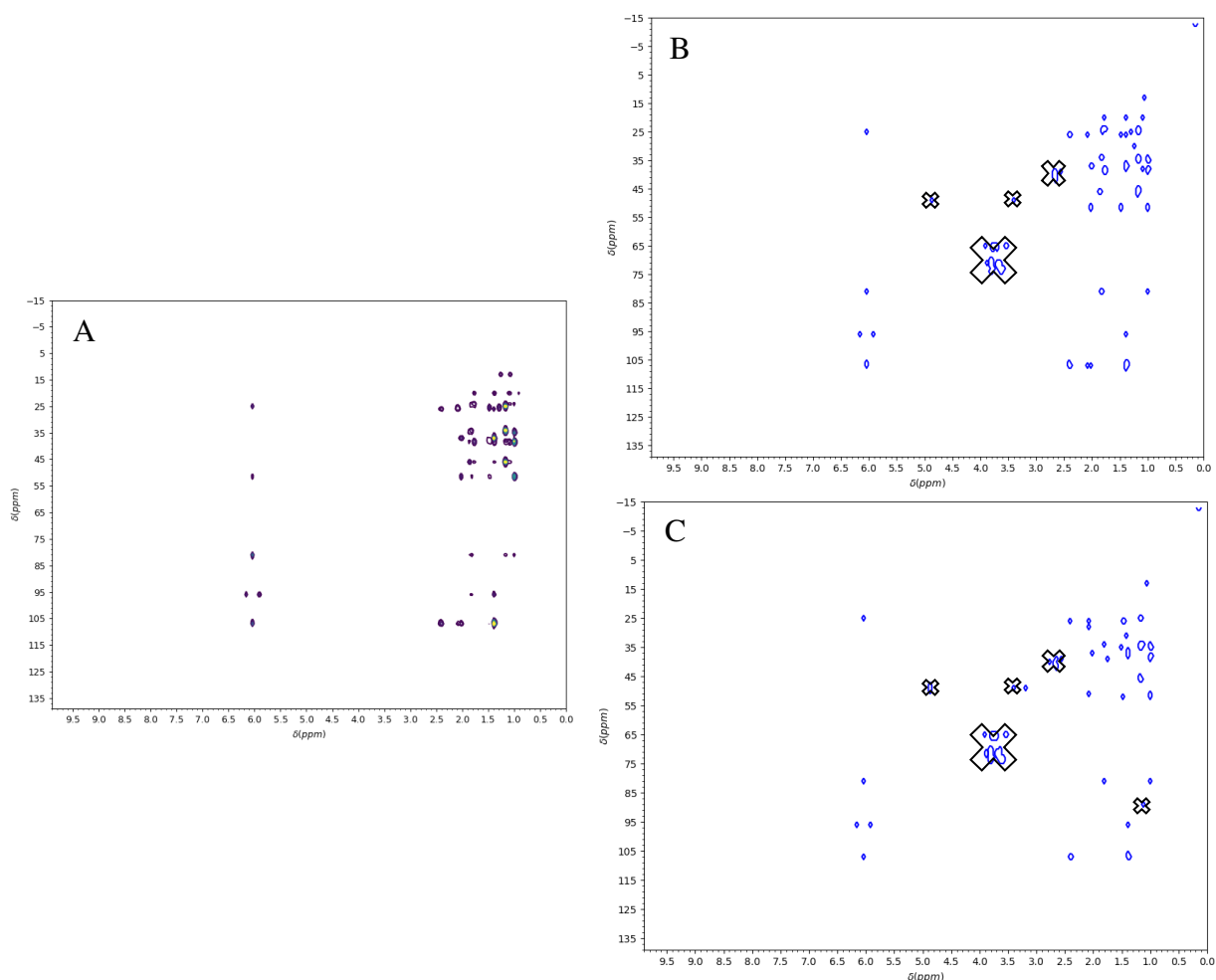


Figure 68 : Empreintes spectrales ^1H - ^{13}C HMBC obtenues par régression linéaire entre les intensités des *buckets* des échantillons ARTE et les concentrations en artémisinine (B) ou les pourcentages d'inhibition de croissance (C). Spectre ^1H - ^{13}C HMBC de l'artémisinine (A).

L'analyse différentielle par régression linéaire sur les expériences ^1H - ^{13}C HMBC permet de produire une empreinte spectrale de l'artémisinine quasiment complète. Quelques taches proviennent de l'extrait algal mais restent minoritaires et n'altèrent pas l'identification de la molécule active (Figure 68).

IV. 5. Conclusion sur l'analyse différentielle de l'ensemble des échantillons

L'analyse différentielle par régression linéaire est réalisable sur l'ensemble des échantillons lorsque l'activité biologique est répartie entre plusieurs fractions. Elle s'applique entre l'intensité des *buckets* et l'activité biologique (exprimée en concentration ou en pourcentage d'inhibition).

Deux types d'analyse ont été employées, une **régression linéaire classique** et une régression linéaire couplée à un outil statistique, la **RFE**. La régression linéaire simple permet d'obtenir de meilleurs résultats que la RFE.

En excluant l'échantillon le plus concentré en molécules bioactives, CHLORO 5 ou ARTE 5, la régression linéaire produit toujours une empreinte spécifique de la molécule active (testé pour chaque expérience, données non présentées).

Les empreintes spectrales ont permis d'identifier la chloroquine et l'artémisinine dans toutes les expériences RMN. Celles-ci ont donc été détectables avec une teneur variant **de 0.5 à 5 % dans l'extrait**.

IV. 6. Les expériences RMN ^1H - δ DOSY

Les expériences ^1H - δ DOSY (*Diffusion Order Spectroscopy*) sont des analyses RMN 2D. Une dimension correspond aux déplacements chimiques ^1H et l'autre aux coefficients de diffusion. Le coefficient de diffusion est caractéristique d'une molécule. La ^1H - δ DOSY est une expérience RMN de choix pour l'analyse de mélanges complexes¹⁶² (Figure 69).

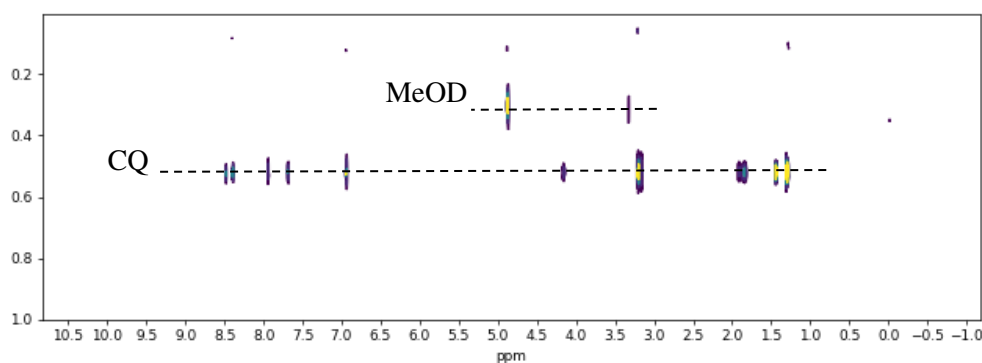


Figure 69 : Spectre ^1H - δ DOSY de la chloroquine (MeOD, 700MHz). La chloroquine (CQ) n'a pas le même coefficient de diffusion que le méthanol deutéré.

La ^1H - δ DOSY nous permet de mettre en évidence les zones de déplacements chimiques ^1H où des signaux RMN sont présents. Elle sera utilisée comme **un filtre** en application à d'autres expériences RMN.

Pour cela, une analyse différentielle des expériences ^1H - δ DOSY est réalisée par régression linéaire classique et par régression RFE entre les intensités des *buckets* et la concentration en chloroquine. Nous obtenons ainsi une empreinte spectrale ^1H - δ DOSY. Pour une meilleure lisibilité, les spectres sont projetés en 1D (Figure 70).

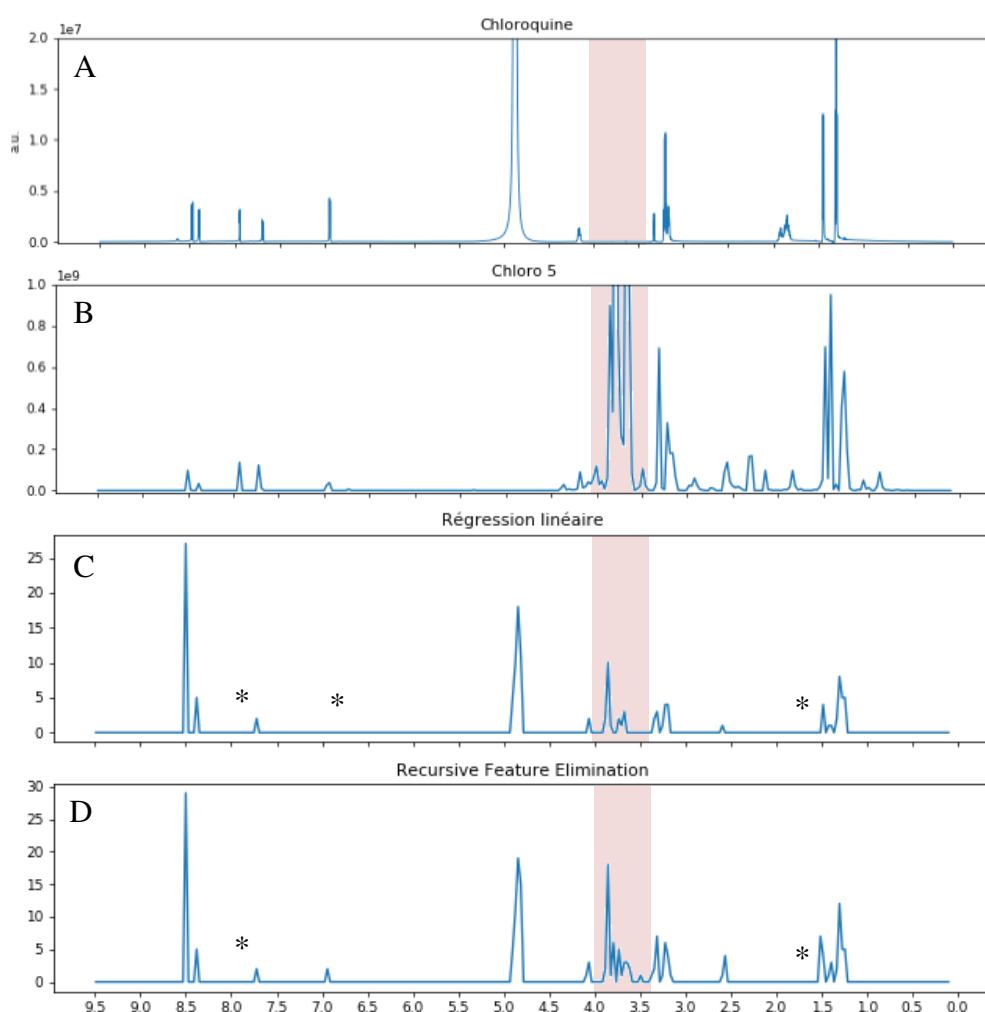


Figure 70 : Empreinte spectrale ^1H - δ DOSY obtenue par régression linéaire entre les intensités des *buckets* des échantillons CHLORO et les concentrations en chloroquine. Spectre ^1H de la chloroquine (A). Spectre ^1H - δ DOSY projeté en 1D de l'échantillon CHLORO 5 (B). Empreinte spectrale ^1H - δ DOSY projetée en 1D produite par régression linéaire entre les intensités des *buckets* et la concentration en chloroquine des échantillons CHLORO (C). Empreinte spectrale ^1H - δ DOSY projetée en 1D par RFE entre les intensités des *buckets* et la concentration en chloroquine des échantillons CHLORO (C). Les régressions linéaires atténuent mais ne suppriment pas les signaux algaux présents entre 3.5 et 4 ppm (rectangle rouge). La régression linéaire supprime trois signaux de la chloroquine et la RFE supprime deux signaux (étoiles noires).

L'empreinte ^1H - δ DOSY produite par régression linéaire classique et RFE en fonction de la concentration en chloroquine permet d'atténuer les signaux de l'extrait algal mais ne les supprime pas complètement (Figure 70, B). Trois signaux appartenant à la chloroquine sont supprimés par régression linéaire et deux sont supprimés par la RFE. La RFE produit le meilleur résultat pour les expériences ^1H - δ DOSY, c'est ce spectre qui sera conservé comme filtre.

En effet, ce spectre va être appliqué comme un filtre aux expériences ^1H - ^{13}C HSQC. Cela signifie que l'on conserve d'avantage les signaux ^1H - ^{13}C HSQC qui sont compris dans les mêmes zones que les signaux présents sur la ^1H - δ DOSY.

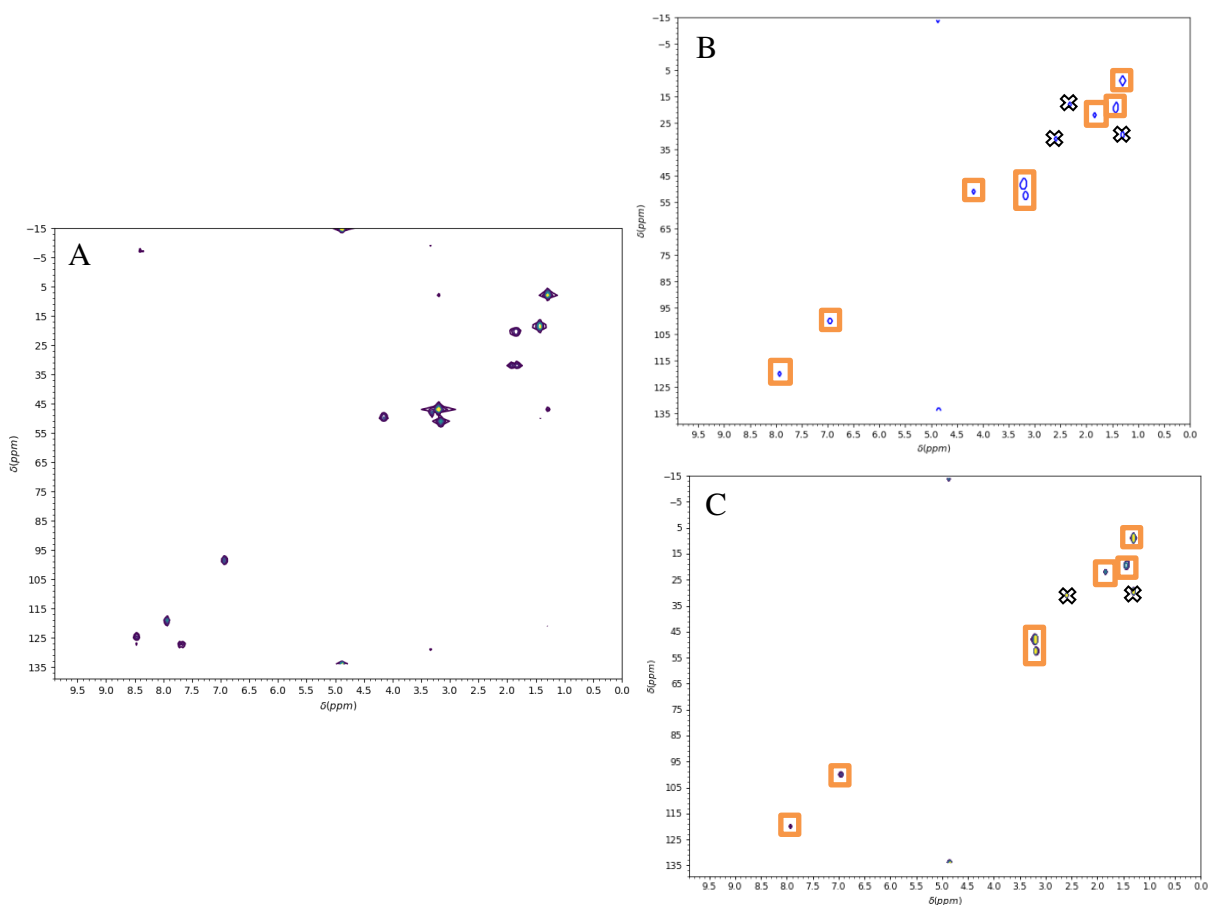


Figure 71 : Empreinte spectrale ^1H - ^{13}C HSQC produite par filtration de l'empreinte ^1H - ^{13}C HSQC obtenue par régression linéaire (B) par l'empreinte ^1H - δ DOSY (C). Spectre ^1H - ^{13}C HSQC de la chloroquine (A). Une seule tache de corrélation supplémentaire de l'extrait algal a été supprimée mais également une tache de corrélation de la chloroquine.

L'empreinte ^1H - ^{13}C HSQC de la chloroquine n'est pas nettement améliorée par cette méthode (Figure 71). Une seule tache de corrélation supplémentaire de l'extrait algal a été supprimée ainsi qu'une tache de corrélation de la chloroquine.

La filtration des empreintes des expériences RMN hétéronucléaires par l'empreinte ^1H - δ DOSY peut sembler être un outil performant pour affiner une empreinte spectrale. Cependant, celle-ci n'a pas été significativement améliorée par rapport à une régression linéaire classique. Des résultats équivalents ont été obtenus pour les expériences ^1H - ^{13}C HMBC et pour la série ARTE (données non montrées).

Il faut noter que les échantillons des séries ARTE et CHLORO sont exactement composés du même extrait algal. Si les signaux différaient plus entre les échantillons, la sélection des signaux spécifiques de l'échantillon actif serait peut-être meilleure.

V. Conclusion

Ce chapitre a porté sur le développement de l'analyse différentielle de manière automatique.

La stratégie initiale s'inspire des travaux de F. Schroeder¹⁵⁶. Une analyse différentielle préliminaire portant sur un extrait de l'algue brune *B. bifurcata* a montré plusieurs difficultés :

- la capacité restreinte de l'analyse différentielle réalisée **de manière visuelle** ;
- l'impossibilité de prendre en compte l'intensité des taches de corrélation en RMN-2D ;
- la pauvreté de l'information structurale via un seul type d'analyse RMN.

L'automatisation de la production d'empreintes a donc été développée afin de contourner ces difficultés, et dans un souci de gain de temps. De plus, il était intéressant d'élargir l'analyse différentielle automatique à différent type d'expériences RMN pour apporter de l'information structurale.

L'automatisation de l'analyse différentielle a pu être possible grâce au programme SPIKE développé par Marc-André Delsuc. Ce programme libre, écrit en Python permettant de traiter et d'analyser toutes les analyses à transformée de Fourier, était la ressource idéale pour pouvoir mettre au point cette analyse. Un programme, Plasmodesma, a été écrit et est implémenté sous la forme d'un *plug-in* dans SPIKE. Plasmodesma est capable d'importer les

données RMN brutes, de réaliser le *processing* et le *bucketing* des spectres RMN, de détecter des signaux et enfin, d'exporter sous la forme de rapport toutes les informations d'acquisition et de *bucketing*.

Les fichiers produits par Plasmodesma, sont ensuite utilisés par diverses fonctions du script *Bucket Utilities*. En effet, ce script contient des fonctions permettant le nettoyage et la symétrisation des spectres homonucléaires, ce qui améliorent nettement la qualité du spectre et de l'analyse différentielle. Il contient aussi des fonctions d'affichage des spectres par leurs données d'intensité ou d'écart-type.

Deux séries d'échantillons reconstitués ont été réalisées. Les séries CHLORO et ARTE contenaient chacune 5 échantillons composés d'un extrait de *S. muticum* inactif supplémentés d'une molécule antipaludique en concentration croissante, l'artémisinine ou la chloroquine.

Ces séries ont permis d'évaluer plusieurs calculs pour produire la meilleure empreinte de la molécule active. Le rapport **des écarts-types des buckets** fournit les meilleurs résultats pour l'analyse différentielle de spectres deux à deux, entre une fraction active *versus* inactive. Dans le cas d'une activité biologique partagée entre plusieurs fractions, une **régression linéaire** est appliquée sur l'ensemble des données RMN en fonction de l'intensité des *buckets* et de la concentration ou du pourcentage d'inhibition de la molécule active.

Une filtration des empreintes RMN hétéronucléaires par des empreintes ^1H - δ DOSY a été réalisée, mais cela n'améliore pas de manière significative les empreintes de la molécule active.

Le rapport des écarts-types des buckets d'une fraction active *versus* inactive et la régression linéaire classique entre les intensités des buckets et l'activité biologique seront les calculs conservés pour la suite des expériences d'analyse différentielle.

L'empreinte spectrale issue de ces calculs était correcte sur les deux séries CHLORO et ARTE, ainsi que pour toutes les analyses RMN enregistrées.

Cependant, le temps d'acquisition de chaque échantillon en RMN était d'environ de 15h, soit 3 jours d'analyse pour une série. Il a été décidé de réduire cette durée à 24h.

L'analyse ^1H - ^{13}C HMBC a une durée de 10h par échantillon alors que la ^1H - ^{13}C HSQC dure 56 min, l'expérience ^1H - ^{13}C HMBC ne sera donc pas conservée pour les prochaines analyses. Concernant les analyses RMN homonucléaires, pour un temps d'analyse équivalent (2h), l'expérience ^1H - ^1H TOCSY est beaucoup plus informative que la ^1H - ^1H COSY puisqu'elle conduit à l'obtention d'un nombre de taches de corrélation supérieur. Elle permet une meilleure identification de la molécule active, elle est donc conservée pour les prochaines séries.

Les analyses RMN ^1H , ^1H - ^1H TOCSY, ^1H - δ DOSY et ^1H - ^{13}C HSQC sont conservées pour la suite des expériences. La durée d'acquisition RMN d'une série de 5 échantillons est **réduite à 24h.**

Les deux séries ARTE et CHLORO étaient des cas simples adaptés au développement et à la validation de la méthode automatique. Il s'agit maintenant de l'éprouver sur des échantillons réels, plus complexes.



CHAPITRE IV

Chapitre IV : Identification de molécules bio-actives dans des extraits végétaux

Ce chapitre porte sur l'utilisation de la méthode automatique de déconvolution pharmacophorique sur différents extraits végétaux.

Dans un premier temps, la méthode d'analyse différentielle sera appliquée à un cas semi-réel, un extrait de *S. muticum* enrichi en artémisinine à une teneur équivalente à sa proportion naturelle dans *artemisia annua*. Contrairement aux séries précédentes, l'extrait « reconstitué » est fractionné, afin que l'artémisinine se répartisse aléatoirement entre les différentes fractions.

Ensuite, la méthode automatique sera appliquée à un cas réel mais connu, sur l'extrait actif *B. bifurcata* afin de vérifier si l'empreinte de l'éléganolone est correctement produite.

Dans un but de visualiser les limites et les capacités de l'analyse différentielle, celle-ci sera éprouvée sur un extrait de quinquina rouge, *Cinchona pubescens*. Cet extrait peut contenir jusqu'à 30 % d'alcaloïdes quinoléiques dont plusieurs possèdent une activité antipaludique. Il sera intéressant d'observer l'empreinte produite à partir de cet extrait.

Enfin, afin de répondre à notre problématique, une analyse différentielle de *Dilsea carnosa*, une algue rouge dont l'extrait AcOET possède une activité antipaludique, sera étudiée. L'empreinte obtenue sera également comparée avec l'empreinte spectrale d'une seconde algue rouge, *Gracilaria gracilis*.

I. Analyse différentielle « en se rapprochant du réel » - Série SMARTE

I. 1. Introduction

Les dernières séries ARTE et CHLORO étaient des échantillons « reconstitués » qui contenaient tous exactement le même extrait d'algue, seule la concentration en molécule bioactive variait.

La série SMARTE a été réalisée pour se rapprocher d'un extrait naturel. Pour cela, l'extrait hydro-alcoolique (60% EtOH) de *Sargassum muticum* inactif a été enrichi en artémisinine dans des proportions naturelles. Chez *Artemisia annua*, cette molécule peut varier entre 0.02 et 1.38% dans la plante sèche^{163,164}. Nous nous sommes appuyés sur les résultats de Klayman, qui a isolé l'artémisinine à partir d'*A. annua* cultivée aux États-Unis¹⁶⁵. Ainsi, nous avons enrichi notre extrait Répartition aléatoire à cette teneur, soit 2.6% m/m_{extrait} (équivalent à 0.06% m/m_{feuilles sèches}).

L'extrait enrichi a ensuite été fractionné par MPLC en phase normale en suivant un gradient de polarité (Figure 72, Cf. Partie Expérimentale III. 1).

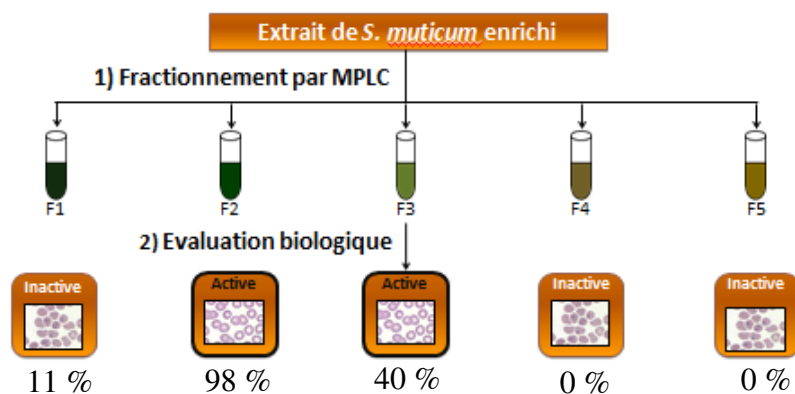


Figure 72 : Schéma des deux premières étapes de la méthode de déconvolution pharmacophorique appliquées à l'extrait SMARTE. Après un fractionnement grossier de l'extrait bioactif (1), les fractions obtenues sont évaluées pour leur activité anti-*Plasmodium* (2). Les fractions 2 et 3 montrent une activité antipaludique.

Les fractions ont été évaluées biologiquement sur *P. falciparum* à deux concentrations, 100 et 10 µg/ml. Elles ont montré une inhibition de croissance du parasite respectivement de 43, 97, 99, 12 et 0% à 100 µg/mL et de 11, 98, 40, 0 et 0% d'inhibition à 10 µg/mL.

Contrairement aux séries CHLORO et ARTE, la série SMARTE est issue d'un extrait fractionné, donc la composition de chaque fraction diffère et des molécules sont susceptibles

de se retrouver dans des fractions successives. C'est d'ailleurs le cas observé puisque les fractions 2 et 3 sont actives mais la fraction 3 en moindre mesure, donc l'artémisinine, seule molécule active de l'extrait, est répartie entre ces fractions. Les analyses RMN ^1H , ^1H - ^1H TOCSY, ^1H - δ DOSY et ^1H - ^{13}C HSQC de chaque fraction ont été enregistrées dans les mêmes conditions (MeOD, 700MHz). L'analyse différentielle par comparaison de spectres deux à deux sera donc appliquée entre les fractions actives et inactives successives et la régression linéaire sur l'ensemble des échantillons en fonction de l'intensité des *buckets* et du pourcentage d'inhibition de croissance de *Plasmodium*.

I. 2. Analyse différentielle : Fractions actives versus inactives

La comparaison deux à deux entre une fraction active et inactive est possible entre la fraction 1 inactive et la fraction 2 active, mais également entre la fraction 3 active et la fraction 4 inactive. Le rapport des écarts-types des *buckets* est appliqué entre ces fractions.

- Résultats pour les expériences ^1H - ^1H TOCSY :

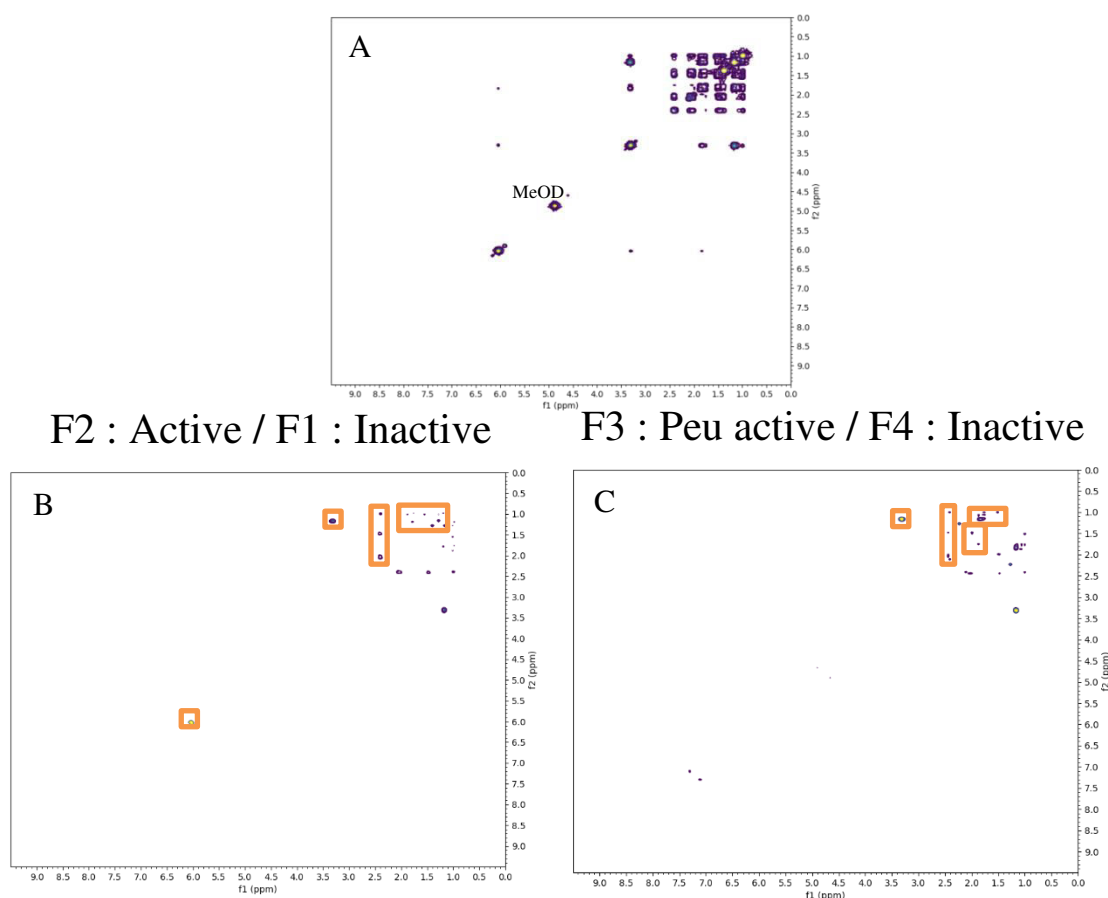


Figure 73 : Empreintes spectrales ^1H - ^1H TOCSY obtenues par rapport des écarts-types des *buckets* entre la fraction active F2 versus la fraction inactive F1 (B) et entre la fraction active F3 versus la fraction inactive F4 de la série SMARTE (C). Spectre ^1H - ^1H TOCSY de l'artémisinine (A). F : Fraction. Les deux empreintes possèdent des taches de corrélation communes qui correspondent aux déplacements chimiques de l'artémisinine.

Nous remarquons qu'il y a des taches de corrélations communes entre les deux empreintes. Le rapport des écarts-types entre les fractions 2 et 1 produit une empreinte spectrale où l'artémisinine est clairement identifiable. L'empreinte est moins marquée entre les fractions 3 et 4, mais le nombre de taches de corrélation est suffisant pour l'identifier (Figure 73).

- **Résultats pour les expériences ^1H - ^{13}C HSQC :**

La même opération est réalisée sur les expériences ^1H - ^{13}C HSQC.

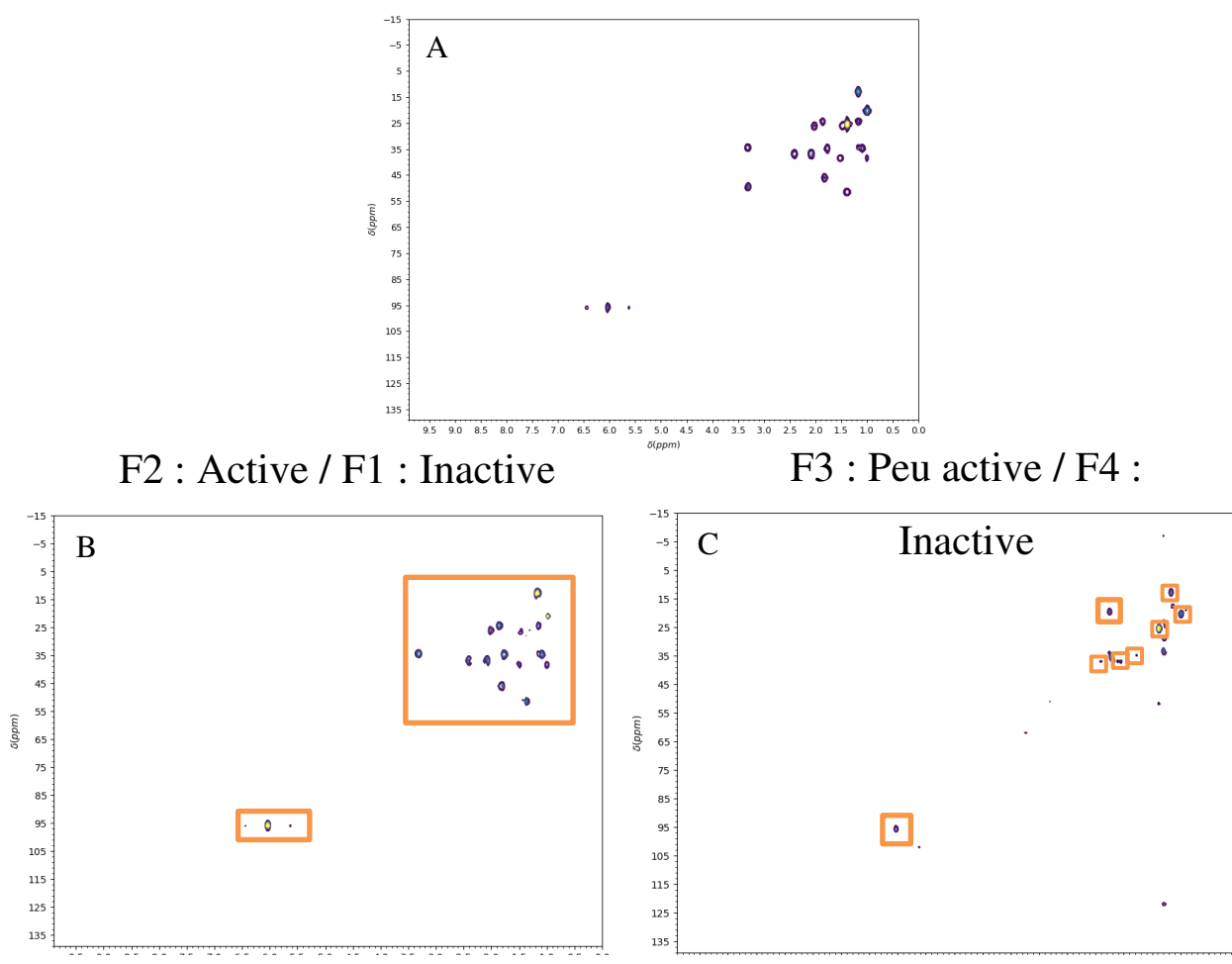


Figure 74 : Empreintes spectrales ^1H - ^{13}C HSQC obtenues par le rapport des écarts-types des *buckets* entre la fraction active F2 versus la fraction inactive F1 (B) et entre la fraction active F3 versus la fraction inactive F4 de la série SMARTE (C). Spectre ^1H - ^{13}C HSQC de l'artémisinine (A). F : Fraction. Les deux empreintes possèdent des taches de corrélation communes qui correspondent aux déplacements chimiques de l'artémisinine. L'empreinte issue du rapport de F2 et F1 est quasiment totale.

L'empreinte obtenue à partir de la fraction 2 est quasiment parfaite. Toutes les taches présentes correspondent à l'artémisinine et il n'y a aucune tache issue de l'extrait. L'empreinte obtenue à partir de la fraction 3 est moins complète mais l'artémisinine est facilement identifiable (Figure 74).

I. 3. Analyse différentielle par régression linéaire

Parmi les 5 fractions obtenues de l'extrait de *S. muticum* enrichi en artémisinine, les fractions 2 et 3 sont actives. Cependant la fraction 3 a une activité moindre puisqu'elle présente un pourcentage d'inhibition de 40 % à 10 $\mu\text{g/mL}$ contre 99 % pour la fraction 2. Cette variation d'activité est bien adaptée pour l'application de la régression linéaire sur les échantillons. Contrairement aux séries ARTE et CHLORO, la concentration exacte en artémisinine n'est pas connue entre les fractions, seule l'activité biologique est accessible.

- **Résultats pour les expériences ^1H - ^1H TOCSY :**

L'empreinte obtenue par régression linéaire sur la série SMARTE est presque similaire au spectre ^1H - ^1H TOCSY de l'artémisinine. La régression linéaire est très efficace sur nos échantillons (Figure 75).

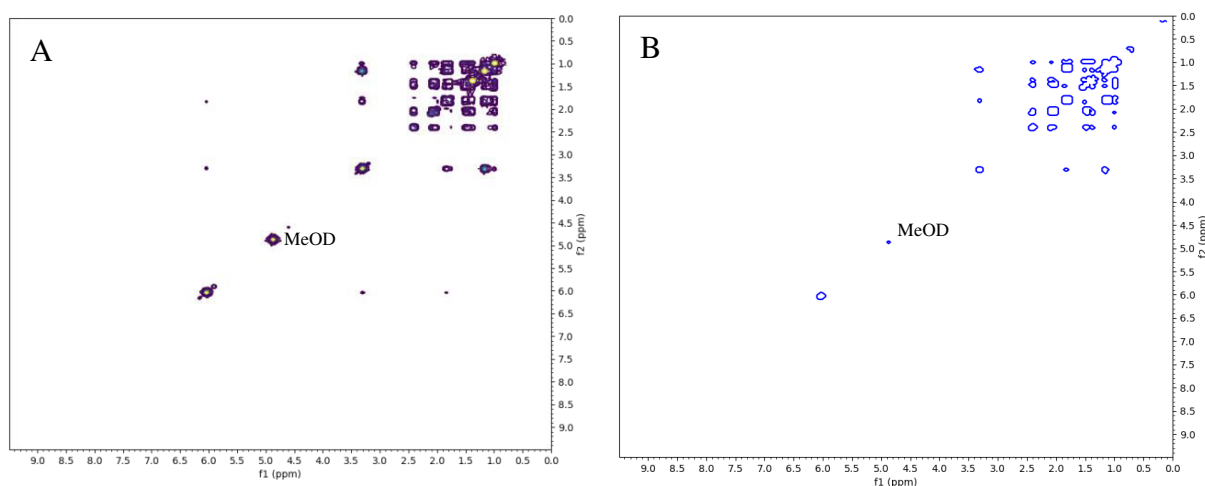


Figure 75 : Spectre ^1H - ^1H TOCSY de l'artémisinine (A). Empreinte spectrale ^1H - ^1H TOCSY obtenue par régression linéaire entre l'intensité des *buckets* et le pourcentage d'inhibition des fractions à 10 $\mu\text{g/mL}$ (B). L'empreinte pharmacophorique obtenue est quasiment similaire au spectre ^1H - ^1H TOCSY de l'artémisinine.

- **Résultats pour les expériences ^1H - ^{13}C HSQC :**

L'empreinte obtenue par régression linéaire sur la série SMARTE est pratiquement similaire au spectre ^1H - ^{13}C HSQC de l'artémisinine. Une tache supplémentaire seulement apparaît. La régression linéaire est efficace pour la production de l'empreinte (Figure 76).

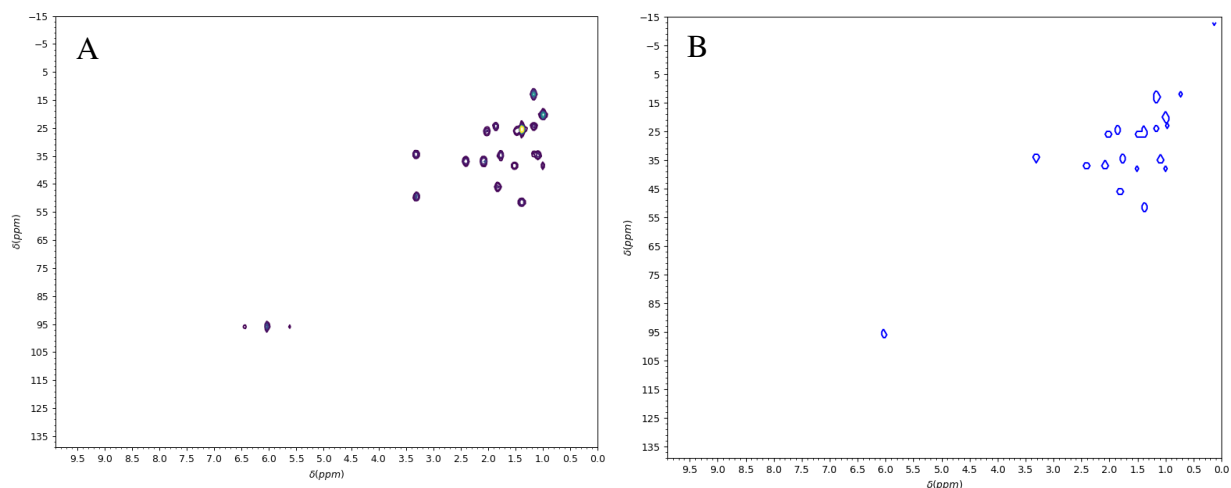


Figure 76 : Spectre ^1H - ^{13}C HSQC de l'artémisinine (A). Empreinte spectrale ^1H - ^{13}C HSQC obtenue par régression linéaire en fonction de l'intensité des *buckets* et le pourcentage d'inhibition des fractions à 10 $\mu\text{g/mL}$ (B). L'empreinte pharmacophorique obtenue est presque équivalente au spectre ^1H - ^{13}C HSQC de l'artémisinine.

I. 4. Filtration par les expériences ^1H - δ DOSY

Les analyses RMN ^1H - δ DOSY ont été enregistrées. Le spectre ^1H - δ DOSY de l'artémisinine est représenté dans la **Figure 77**.

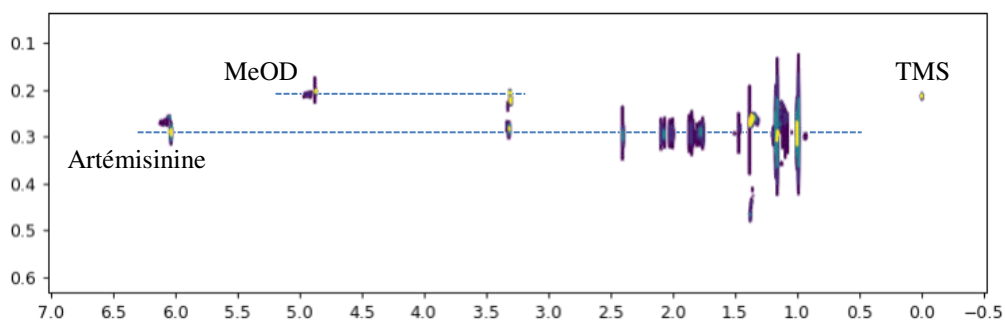


Figure 77 : Spectre ^1H - δ DOSY de l'artémisinine (MeOD, 700 MHz). L'artémisinine n'a pas le même coefficient de diffusion que le MeOD (TMS = triméthylsilane, référence)

Une régression linéaire classique et une RFE sont appliquées entre les intensités des *buckets* des expériences ^1H - δ DOSY et les pourcentages d'inhibition de croissance de *Plasmodium* (projection 1D, Figure 78). La RFE produit une empreinte de l'artémisinine plus fidèle.

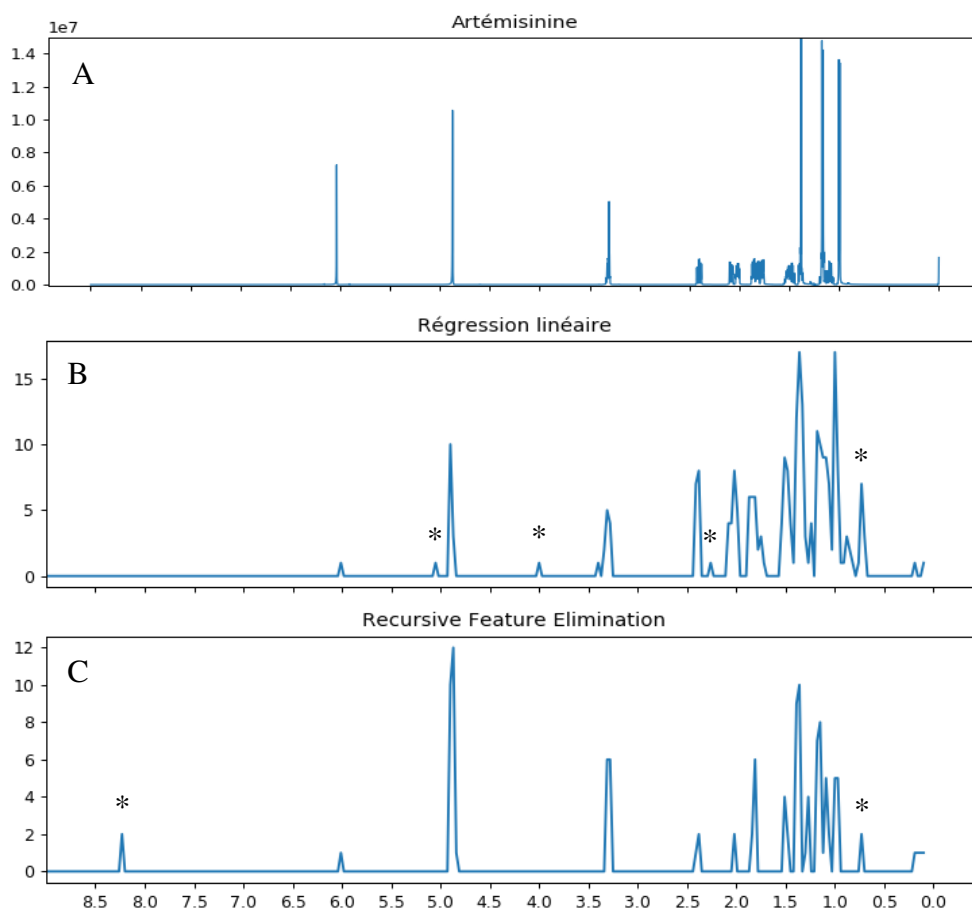


Figure 78 : Empreinte spectrale $^1\text{H}-\delta$ DOSY obtenue par régression linéaire ou RFE entre les intensités des *buckets* des échantillons SMARTE et les pourcentage d'inhibition des fractions. Spectre ^1H de l'artémisinine (A). Empreinte $^1\text{H}-\delta$ DOSY projetée en 1D produite par régression linéaire entre les intensités des *buckets* et le pourcentage d'inhibition des échantillons SMARTE (B). Empreinte $^1\text{H}-\delta$ DOSY projetée en 1D obtenue par RFE entre les intensités des *buckets* et le pourcentage d'inhibition des échantillons SMARTE (C). L'empreinte obtenue par RFE est plus proche du spectre ^1H de l'artémisinine (croix noires : signaux n'appartenant pas à l'artémisinine).

Une filtration de l'empreinte $^1\text{H}-^{13}\text{C}$ HSQC par l'empreinte $^1\text{H}-\delta$ DOSY issue de RFE est réalisée (Figure 79).

L'empreinte $^1\text{H}-^{13}\text{C}$ HSQC filtrée par l'empreinte $^1\text{H}-\delta$ DOSY permet une sélection des taches de corrélation de l'artémisinine et une diminution des taches de corrélation d'origine algale. Cependant, le nombre de signaux de l'artémisinine est également diminué. C'est un résultat supplémentaire qui permet d'enrichir l'information au niveau des taches de corrélation spécifiques de la molécule active mais n'apporte pas vraiment d'amélioration par rapport à la régression linéaire classique.

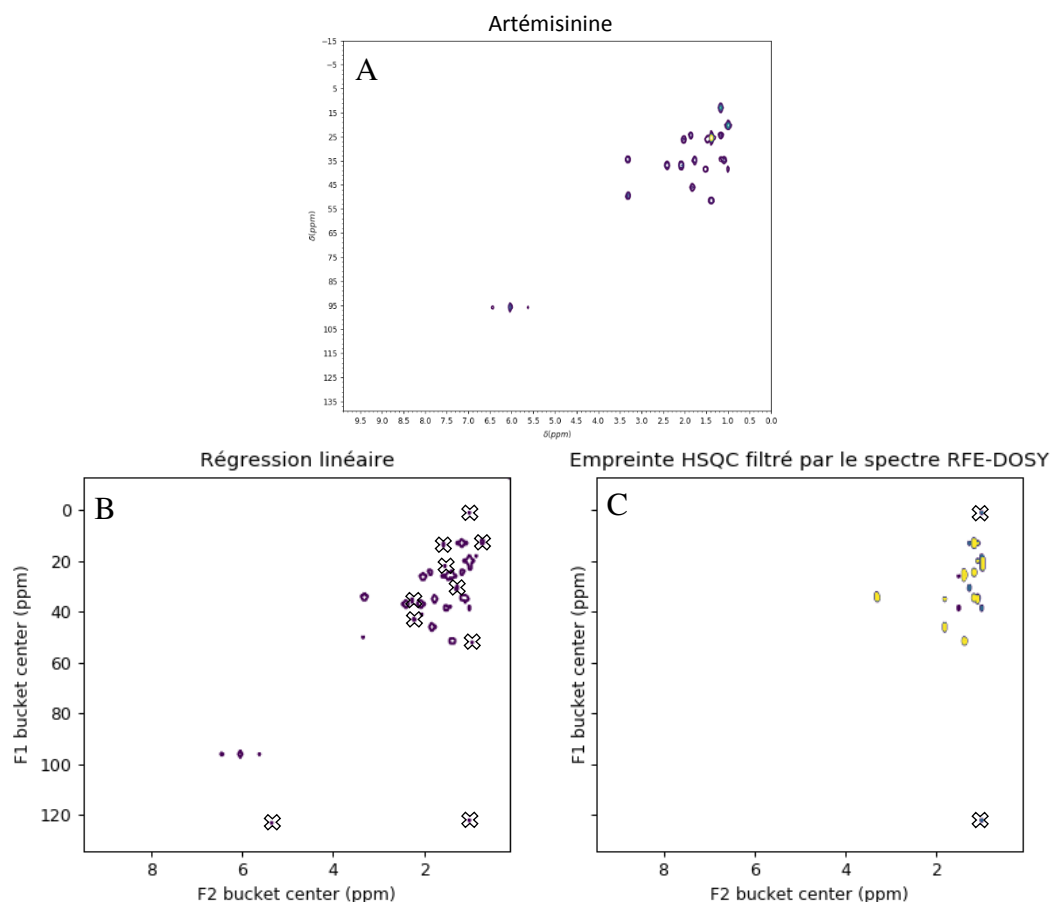


Figure 79 : Filtration de l'empreinte ^1H - ^{13}C HSQC obtenue par régression linéaire (B), par l'empreinte ^1H - δ DOSY obtenue par RFE (C). Spectre ^1H - ^{13}C HSQC de l'artémisinine (A). La filtration de l'empreinte ^1H - ^{13}C HSQC fonctionne mais un certain nombre de taches de corrélation de l'artémisinine ont été supprimées.

I. 5. Conclusion

L'artémisinine est facilement identifiable grâce à la production automatique de l'empreinte pharmacophorique obtenue par les différents calculs et pour chaque expérience RMN. La concentration en artémisinine n'était pas connue dans nos fractions mais l'activité biologique variait entre les fractions 2 et 3. Cela a été suffisant pour produire son empreinte par régression linéaire. Les différents calculs produisent des empreintes similaires qui permettent de sélectionner des taches de corrélation spécifiques et d'identifier la molécule active.

La série SMARTE est quasiment semblable à une situation réelle puisqu'elle contient un composé actif dans une **teneur similaire à un extrait naturel**. Cette série montre que les calculs sélectionnés grâce aux séries ARTE et CHLORO s'appliquent et fonctionnent même si la **composition totale en molécules diffère** entre chaque fraction.

II. Analyse différentielle de *Bifurcaria bifurcata* - Preuve de concept

L'extrait acétate d'éthyle de *Bifurcaria bifurcata* avait permis d'évaluer la faisabilité de la méthode d'analyse différentielle de F. Schroeder à nos extraits. Cette analyse différentielle avait été réalisée de manière visuelle (Chapitre 3, II. 2). L'éléanolone, molécule à activité antipaludique, avait été identifiée dans l'extrait par comparaison visuelle des spectres ^1H - ^1H COSY des deux fractions actives.

Cet extrait a donc été choisi pour poursuivre la preuve de concept de la méthode d'analyse différentielle **automatique**. En effet, il s'agit d'identifier une molécule active naturellement présente dans l'extrait, nous sommes dans la première situation réelle de l'application de la méthode.

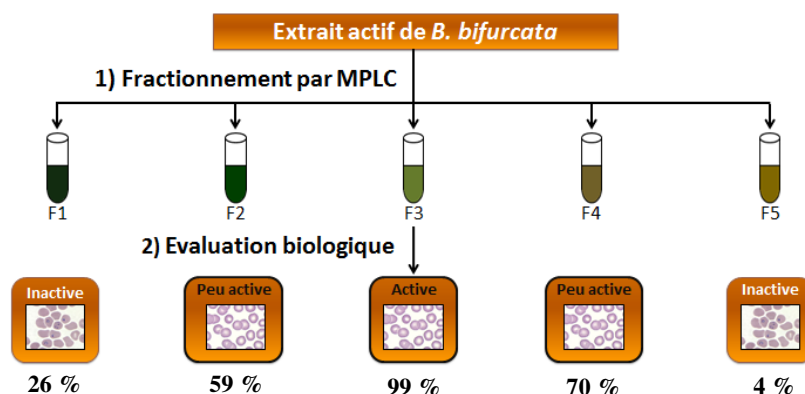


Figure 80 : Schéma des deux premières étapes de la méthode de déconvolution appliquées à l'extrait AcOEt de *B. bifurcata*. Après un fractionnement grossier de l'extrait bio-actif (1), les fractions obtenues sont évaluées pour leur activité anti-*Plasmodium* (2). Les fractions 2, 3 et 4 présentent une activité antipaludique.

L'extrait acétate d'éthyle a été fractionné par MPLC en phase normale pour obtenir 5 fractions. L'activité antipaludique de ces fractions a été évaluée à 100 µg/mL et à 10 µg/mL. Les pourcentages d'inhibition de croissance du parasite de F1 à F5 étaient respectivement de 26, 59, 99, 70, et 4 % à 100 µg/mL et de 4, 21, 34, 2, et 9 % à 10 µg/mL (Figure 80).

II. 1. Analyse différentielle automatique de *B. bifurcata*

II. 1. a. Analyse différentielle des spectres ^1H - ^1H TOCSY

L'analyse différentielle automatique entre les fractions actives et inactives successives, et par régression linéaire est réalisée. Le fractionnement de *B. bifurcaria* a conduit à une fraction 3 active et aux fractions 2 et 4 moins actives. En raison de cette faible activité, il est possible que l'empreinte issue de l'analyse différentielle automatique deux à deux soient erronée (Figure 81).

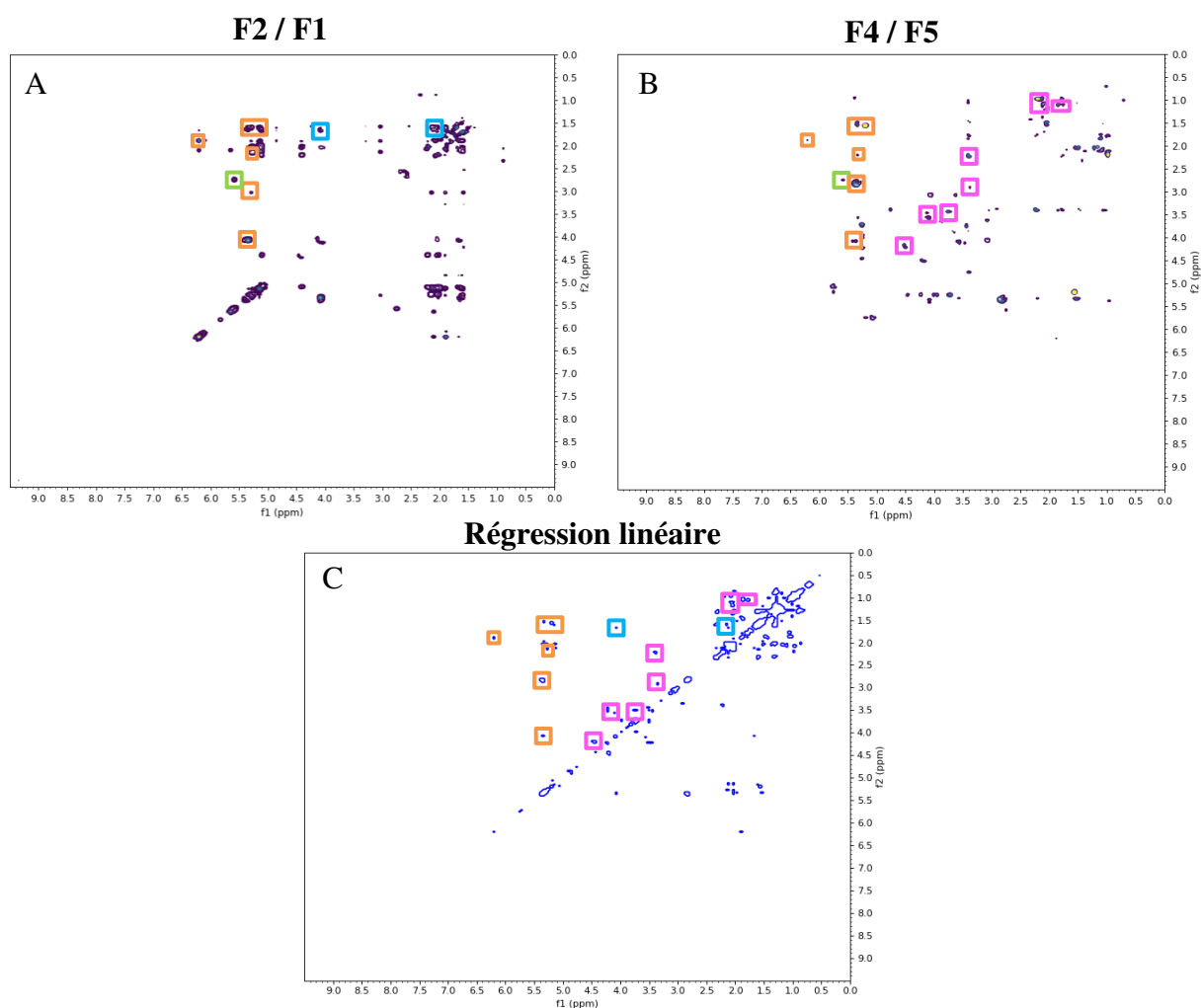


Figure 81 : Comparaison des différentes empreintes ^1H - ^1H TOCSY produites par le rapport des écarts-types des *buckets* des fractions 2 sur 1 (A) et des fractions 4 sur 5 (B) et par régression linéaire (C). Les taches de corrélation communes entre les spectres sont de la même couleur. Les taches de corrélation communes aux trois empreintes sont en orange. Les empreintes obtenues présentent de nombreuses taches de corrélation similaires, notamment partagées entre les trois empreintes (orange).

Les deux empreintes spectrales obtenues par le rapport des écarts-types des *buckets* des fractions 2 sur 1 et 4 sur 5 diffèrent mais possèdent des taches de corrélation communes. L'empreinte obtenue par régression linéaire recoupe partiellement les empreintes obtenues par

rapport des écarts-types. Des taches de corrélation sont partagées entre les trois empreintes, et appartiennent vraisemblablement à la molécule active (Figure 81).

II. 1. b. Analyse différentielle des spectres ^1H - ^{13}C HSQC

La même comparaison est produite pour les expériences ^1H - ^{13}C HSQC.

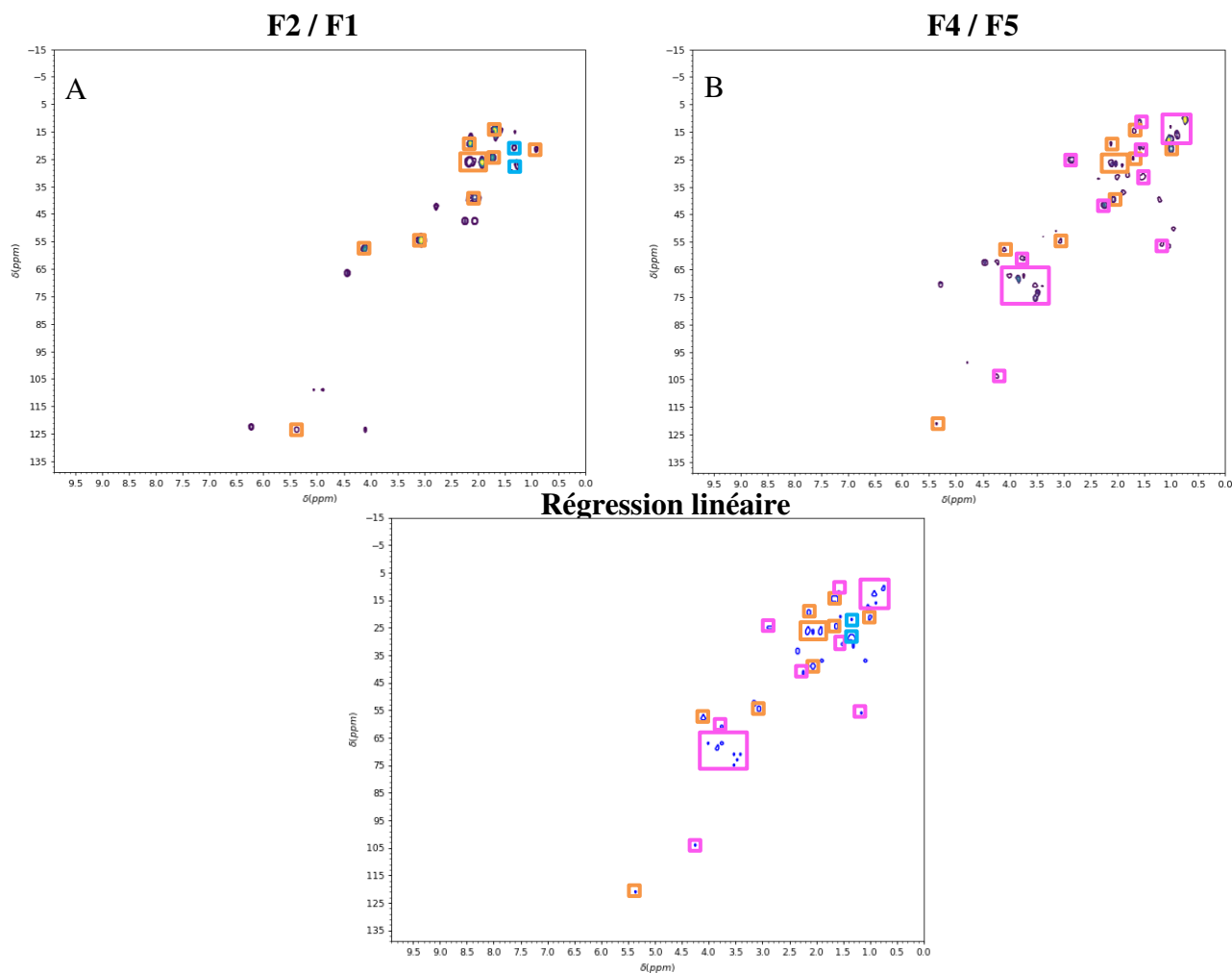


Figure 82 : Comparaison des différentes empreintes ^1H - ^{13}C HSQC produites par le rapport des écarts-types des *buckets* des fractions 2 sur 1 (A) et des fractions 4 sur 5 (B) et par régression linéaire (C). Les taches de corrélation communes entre les spectres sont de la même couleur. Les taches de corrélation partagées entre les trois empreintes sont en orange. Les empreintes obtenues présentes de nombreuses taches de corrélation communes, notamment entre les trois empreintes (orange).

Les empreintes obtenues par les différents calculs se recoupent entre elle. En effet, la régression linéaire regroupe les taches de corrélation des empreintes issues du rapport des écarts-types entre les fractions F4 et F5. Les taches de corrélations partagées entre les trois empreintes ont plus de probabilités d'appartenir à la molécule active (Figure 82).

II. 2. Comparaison des empreintes spectrales et du spectre RMN issu de la méthode couplée

Lors de la première étude sur *B. bifurcata*, une fraction active avait été injectée sur la méthode couplée HPLC-SPE-RMN. Une molécule avait spécifiquement été stockée sur une cartouche SPE et des acquisitions RMN avait été effectuées. Cette étape n'a pas été renouvelée avec le nouvel extrait. Seuls les spectres ^1H - ^1H COSY et ^1H - ^{13}C HSQC de cette molécule stockée ont été enregistrés.

II. 2. a. Comparaison de l'empreinte ^1H - ^1H TOCSY et du spectre ^1H - ^1H COSY de la molécule stockée sur la cartouche SPE

Nous pouvons comparer l'empreinte produite par régression linéaire (car elle rassemble les taches de corrélation communes aux trois calculs) et le spectre de la molécule stockée sur la cartouche SPE (Figure 83). Seul le spectre ^1H - ^1H COSY a été enregistré, les taches de corrélation probables en ^1H - ^1H TOCSY ont été représentées par des points rouges.

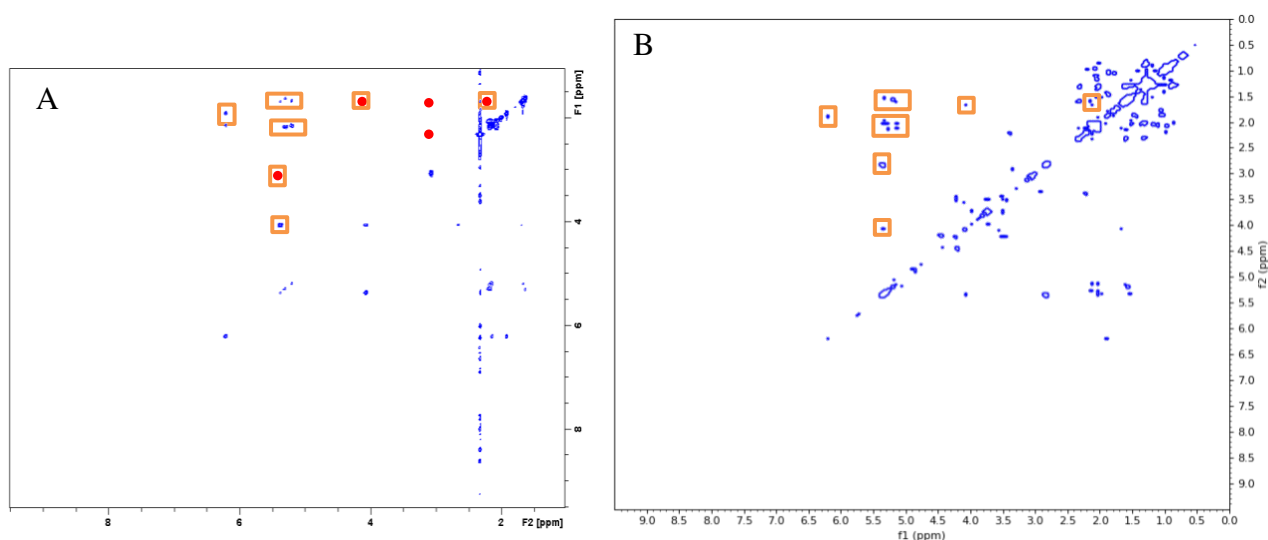


Figure 83 : Comparaison du spectre ^1H - ^1H COSY de la molécule stockée sur SPE (A, 500 MHz, acétone d-6, points rouges : taches de corrélation fictives correspondant à un spectre ^1H - ^1H TOCSY) et de l'empreinte spectrale ^1H - ^1H TOCSY obtenue par régression linéaire entre les intensités des *buckets* et le pourcentage d'inhibition de fractions (B). Les taches de corrélation communes sont encadrées en orange. L'empreinte obtenue par régression linéaire est quasiment identique au spectre de la molécule stockée.

L'empreinte produite par régression linéaire correspond sensiblement à la molécule stockée sur la cartouche SPE.

II. 2. b. Comparaison de l’empreinte ^1H - ^{13}C HSQC et du spectre ^1H - ^{13}C HSQC de la molécule stockée sur la cartouche SPE

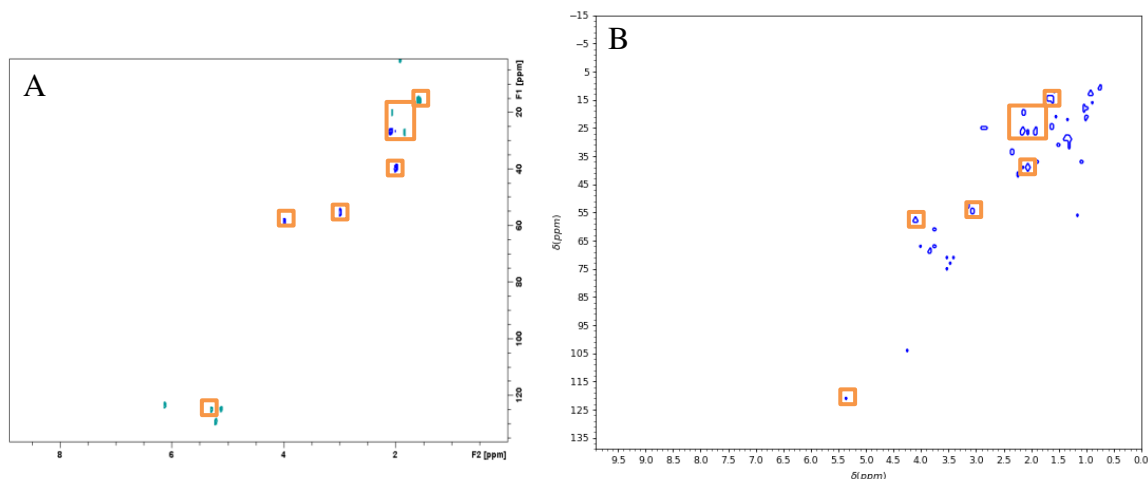


Figure 84 : Comparaison du spectre ^1H - ^{13}C HSQC de la molécule stockée sur SPE (A, 500 MHz, acétone d_6) et de l’empreinte spectrale ^1H - ^{13}C HSQC obtenue par régression linéaire entre les intensités des *buckets* et le pourcentage d’inhibition de fractions (B). Les taches de corrélation communes sont encadrées en orange. L’empreinte obtenue par régression linéaire ressemble au spectre de la molécule stockée.

Le spectre ^1H - ^{13}C HSQC de la molécule stockée sur la cartouche SPE est comparé à l’empreinte produite par régression linéaire. Plusieurs taches de corrélation présentes sur le spectre correspondent au spectre ^1H - ^{13}C HSQC de la molécule stockée. Cependant, de nombreuses autres taches sont observées (Figure 84).

II. 3. Conclusion

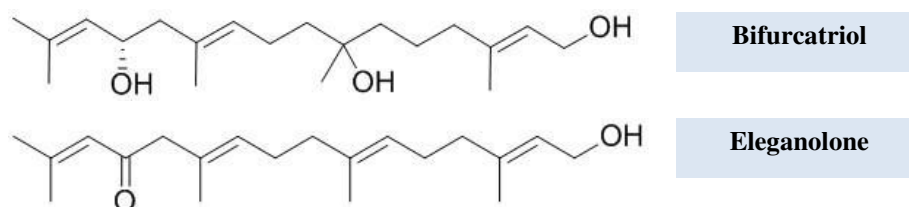
Nous avons pu observer que les empreintes issues des trois calculs se recoupaient entre elles. L’empreinte produite par régression linéaire recouvre celles produites par le rapport des écarts-types.

Une molécule a été stockée sur une cartouche SPE et ses spectres ^1H - ^1H COSY et ^1H - ^{13}C HSQC ont été enregistrés. Grâce à la méthode couplée, la masse de cette molécule a pu être déterminée et identifiée comme étant un diterpène linéaire, l’éléganolone. Cette molécule avait précédemment été isolée par un fractionnement bio-guidé et son activité antipaludique avait été démontrée ($\text{CI}_{50} = 7.9 \mu\text{M}$).

Les taches de corrélation communes aux trois empreintes correspondent au spectre RMN de l’éléganolone.

Cependant, les empreintes montrent des taches de corrélation supplémentaires qui pourraient signifier que l'éléganolone n'est pas la seule molécule active de cet extrait.

En effet, contrairement à l'expérience SMARTE (extrait enrichi en artémisinine), cette série est un extrait « naturel » qui est susceptible de contenir plusieurs molécules actives ou une classe de molécules partageant l'activité. D'après la publication du Dr. Gallé, l'éléganolone a été montré comme le principal constituant antipaludique isolé de l'extrait AcOEt de *B. bifurcata* mais la CI_{50} de cette molécule est équivalente à celle de la fraction active (2.6 $\mu\text{g/mL}$ et 3.0 $\mu\text{g/mL}$). Ce résultat suggère que l'activité n'est pas uniquement due à cette molécule mais qu'elle serait partagée entre plusieurs molécules¹²¹. Cela, est d'ailleurs vérifié, puisqu'un autre diterpène linéaire, le bifurcatriol a été isolé de *B. bifurcata* et a montré une activité antipaludique (test hypoxanthine tritiée, $CI_{50} = 0.65 \mu\text{g/mL}$)¹⁶⁶.



En RMN ^1H , le bifurcatriol ne diffère franchement de l'éléganolone que par un signal supplémentaire à 4.41 ppm (500 MHz, CDCl_3). Celui-ci semble seulement visible dans l'empreinte produite par le rapport des écarts-types des *buckets* de F2/F1. Le bifurcatriol a été isolé à partir de 12 g d'extrait, pour une masse de 7.7 mg. Cela représente une teneur de 0.06% dans l'extrait expliquant probablement son absence dans les empreintes.

L'empreinte obtenue par régression linéaire est quasiment identique à celle de l'éléganolone, la méthode automatique a donc été capable de produire l'empreinte de la molécule **majoritaire** de l'extrait **liée à l'activité antipaludique**.

III. Analyse différentielle de *Cinchona pubescens* – Cas complexe

L'écorce séchée de quinquina rouge contient entre 4.5 et 8.5% d'alcaloïdes totaux dont 30% à 60% sont constitués par des alcaloïdes quinoléiques. Une trentaine de ces alcaloïdes ont été décrits dans les quinquinas. Les noyaux quinoléiques sont éventuellement substitués en C-6', et reliés par un carbone porteur d'un hydroxyle secondaire à une quinuclidine. Il y a 4 alcaloïdes majoritaires, la quinine et la quinidine et leurs homologues déméthoxylés en C-6' (Figure 85). La quinine a été le premier médicament antipaludique mais les autres alcaloïdes quinoléiques présentent également une activité anti-*Plasmodium*^{167,168}.

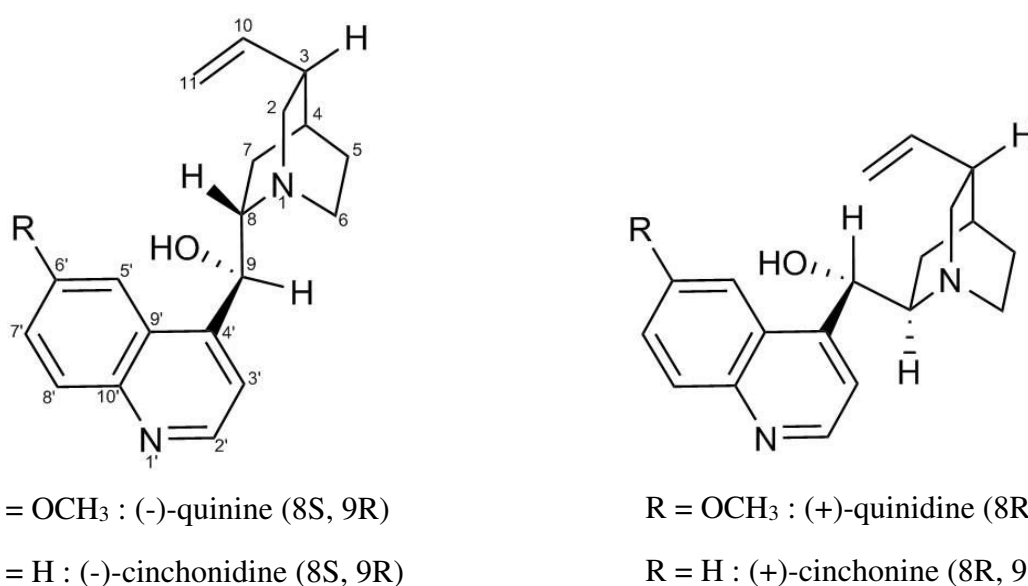


Figure 85 : Structures chimiques de la quinine, de la quinidine, de la cinchonine et de la cinchonidine.

Il est intéressant d'évaluer la méthode de déconvolution sur un extrait de quinquina. En effet, cet extrait représente un cas complexe où l'activité est partagée entre plusieurs molécules de la même classe.

Un extrait hydroalcoolique (60% EtOH) a été réalisé à partir de la poudre pulvérisée d'écorce de *Cinchona pubescens*. Cet extrait a été fractionné par MPLC en phase normale pour obtenir 6 fractions. Les analyses RMN ¹H-¹H TOCSY et ¹H-¹³C HSQC des fractions et des molécules de référence ont été réalisées. Les déplacements chimiques ¹H et ¹³C des références sont présentés dans le **Tableau 11**. La partie aromatique présente les déplacements chimiques qui varient le moins entre les alcaloïdes. Cependant, les déplacements chimiques des protons du noyau quinuclidine varient entre les molécules, et également entre les

stéréoisomères. Ces déplacements seront donc déterminants pour l'identification des alcaloïdes.

Tableau 11 : Déplacements chimiques des quatre alcaloïdes majoritaires, la quinine, la quinidine, la cinchonine et la cinchonidine (700MHz, MeOD).

N°	Quinine		Quinidine		Cinchonine		Cinchonidine	
	δ_C	δ_H , mult	δ_C	δ_H , mult	δ_C	δ_H , mult	δ_C	δ_H , mult
2'	-4.5 (150.9)	8.82 (d, 1H, 5.1)	-1.9 (150.9)	8.68 (d, 1H, 4.4)	1.4 (151.3)	8.82 (d, 1H, 5.1)	1.4 (151.3)	8.82 (d, 1H, 4.6)
3'	121.0	8.00 (d, 1H, 5.1)	120.5	7.77(d, 1H, 4.7)	119.8	7.73 (d, 1H, 4.1)	119.2	7.72 (d, 1H, 4.6)
5'	102.9	7.61 (d, 1H, 2.7)	102.6	7.51 (d, 1H, 2.6)	130.3	8.06 (d, 1H, 2.7)	130.3	8.05 (d, 1H, 2.7)
6'-R	57.9	4.09 (s, 3H)	57.3	4.03 (s, 3H)	128.4	7.66 (t, 1H, 7.2*2)	128.4	7.66 (t, 1H, 7.2*2)
7'	126.5	7.54 (dd, 1H, 9.3, 5.6)	124.0	7.39(dd, 1H, 9.1, 2.7)	130.9	7.77 (t, 1H, 7.7*2)	130.8	7.77 (t, 1H, 7.8*2)
8'	128.6	8.06 (d, 1H, 9.2)	131.4	7.93 (d, 1H, 9.2)	124.8	8.19 (d, 1H, 8.2)	124.9	8.23 (d, 1H, 8.2)
2a	55.8	3.34 (m, 2H)	50.6	3.52 (m, 2H)	50.9	2.78 (m, 1H)	57.9	2.64 (m, 1H)
2b	"	3.64 (m, 2H)	"	3.31 (m, 1H)	"	2.88 (m, 2H)	"	3.08 (m, 1H)
3	38.6	2.81 (m, 1H)	38.5	2.69 (q, 1H, 8.6)	41.5	2.31 (m, 1H)	41.1	2.34 (m, 1H)
4	28.4	2.10 (m, 1H)	29.2	1.97 (m, 1H)	29.8	1.73 (m, 1H)	29.5	1.79(m, 1H)
5a	25.3	1.95 (m, 1H)	24.3	1.82 (m, 1H)	27.2	1.54 (m, 1H)	28.3	1.57(m, 1H)
5b	"	2.19 (m, 2H)	"	1.91 (m, 2H)	"	1.59 (m, 1H)	"	1.86 (m, 2H)
6a	45.6	3.34 (m, 2H)	50.0	3.52 (m, 2H)	50.9	2.88 (m, 2H)	44.1	2.68 (m, 1H)
6b	"	4.30 (m, 1H)	"	4.26 (m, 1H)	"	3.51 (m, 1H)	"	3.61 (m, 1H)
7a	19.5	1.52 (m, 1H)	19.2	1.14 (m, 1H)	21.9	1.13 (m, 1H)	22.3	1.51(m, 1H)
7b	"	2.19 (m, 2H)	"	2.43 (m, 1H)	"	2.20 (m, 1H)	"	1.86 (m, 2H)
8	61.3	3.64 (m, 2H)	61.3	3.59 (m, 1H)	61.4	3.08(m, 1H)	61.8	3.12 (m, 1H)
9	68.4	6.25 (s, 1H)	68.5	6.31 (s, 1H)	72.5	5.70 (d, 4.1)	72.4	5.64 (d, 1H, 3.9)
10	-10.5 (139.4)	5.76 (ddd, 1H, 9.3, 2.6)	-11.3 (138.6)	6.10 (ddd, 1H, 9.3, 2.6)	-8.1 (141.8)	6.16 (m, 1H)	-7 (142.9)	5.76 (ddd, 1H, 17.4, 10.2, 7.6)
11a	117.3	5.02 (dt, 1H, 10.5, 1.1, 1.1)	117.8	5.25 (dt, 1H, 10.4, 1.2, 1.2)	115.4	5.07 (d, 1H, 10.3)	115.2	4.89 (m, 1H)
11b	"	5.12 (dt, 1H, 17.2, 1.2, 1.2)	"	5.22 (dt, 1H, 17.2, 1.3, 1.3)	"	5.11 (d, 1H, 17.8)	"	4.97 (dt, 1H, 17.1, 1.5, 1.5)

Les protons du noyau quinuclidine, de 2 à 8, sont ceux dont les déplacements chimiques varient le plus.

L'activité antipaludique des 6 fractions de l'extrait de quinquina et des molécules de référence, à savoir la quinine, la quinidine, la cinchonine, la cinchonidine a été évaluée. Ces alcaloïdes étant répartis sur l'ensemble des fractions, l'activité devait être exprimée de

manière précise. Aussi, une évaluation de la CI_{50} de chaque fraction et molécules de références a été réalisée.

Tableau 12 : Evaluation antipaludique des fractions issues de l'extrait hydroalcoolique de quinquina.

Echantillons	Références				Fractions de l'extrait de quinquina					
	Q	Qd	C	Cd	F1	F2	F3	F4	F5	F6
CI_{50} ($\mu\text{g/mL}$)	0.24	0.06	0.06	0.24	0.23	0.095	0.093	0.166	6.17	8.53

Q : quinine, Qd : quinidine, C : cinchonine, Cd : cinchonidine. Les molécules testées les plus actives sont la quinidine et la cinchonine. Les fractions les plus actives sont les fractions 2 et 3.

Toutes les fractions possèdent une activité élevée, les fractions 2 et 3 sont les plus actives. D'après nos tests d'activité, la quinidine et cinchonine sont les molécules les plus actives (Tableau 12).

III. 1. Analyse différentielle par régression linéaire

L'analyse différentielle de spectres deux à deux ne peut pas être appliquée dans cette situation car toutes les fractions sont actives. L'analyse par régression linéaire est donc le seul calcul utilisé.

III. 1. a. Analyse différentielle des expériences ^1H - ^1H TOCSY

L'empreinte spectrale obtenue est caractéristique d'un alcaloïde quinolérique (Figure 86). Il est possible de la comparer avec les spectres RMN des molécules de référence (Figure 87).

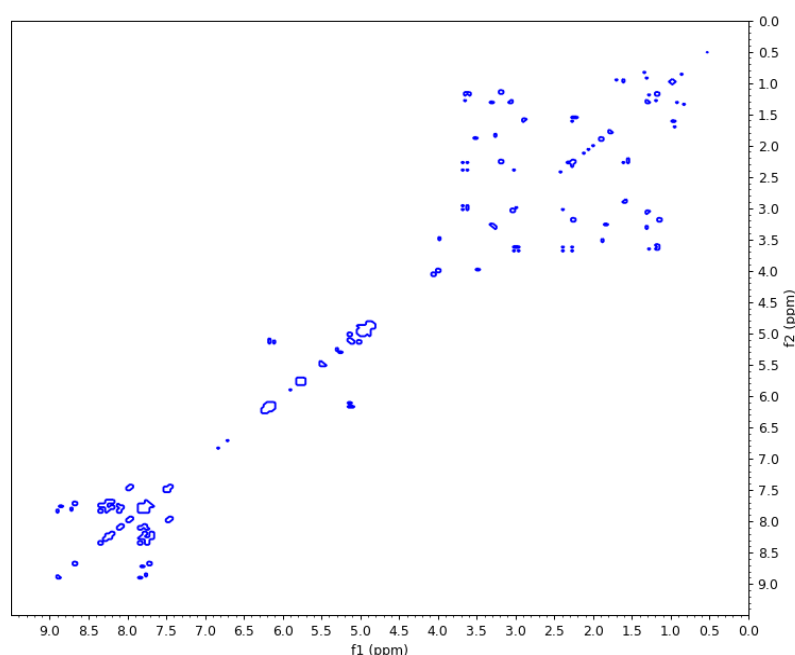


Figure 86 : Empreinte spectrale ^1H - ^1H TOCSY obtenue par régression linéaire des intensités des buckets des données ^1H - ^1H TOCSY des 6 fractions de quinquina et de leurs activités antipaludiques.

Ainsi, la majorité des taches de corrélation de l’empreinte semblent appartenir à la cinchonine (Figure 87, rouge) et à la quinidine (Figure 87, vert). Les autres taches de corrélation présentes n’ont pas été identifiées comme appartenant à un des alcaloïdes majoritaires.

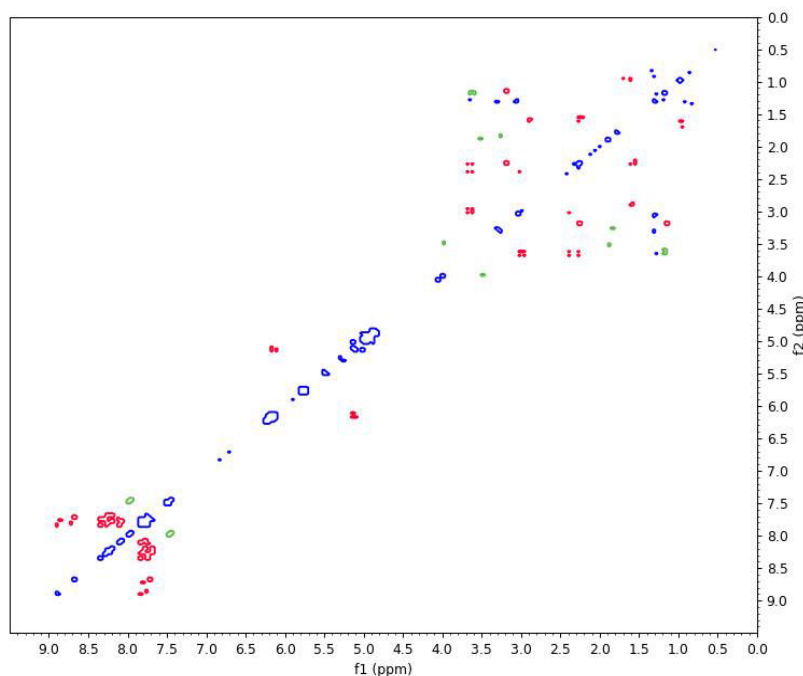


Figure 87 : Comparaison de l’empreinte ^1H - ^1H TOCSY et des spectres des 4 alcaloïdes quinoléiques majoritaires. Les taches de corrélation rouges appartiennent à la cinchonine et en vert, à la quinidine. La plupart des taches de corrélation appartiennent à la cinchonidine et à la quinidine, molécules évaluées comme les plus actives d’après nos tests.

III. 1. b. Analyse différentielle des expériences ^1H - ^{13}C HSQC

Les taches de corrélation de l’empreinte spectrale correspondent à des alcaloïdes quinoléiques. Cependant, en raison de la faible variation de déplacements chimiques sur l’axe ^{13}C , une tache de corrélation peut correspondre à plusieurs alcaloïdes sans qu’ils soient distinguables (Figure 88).

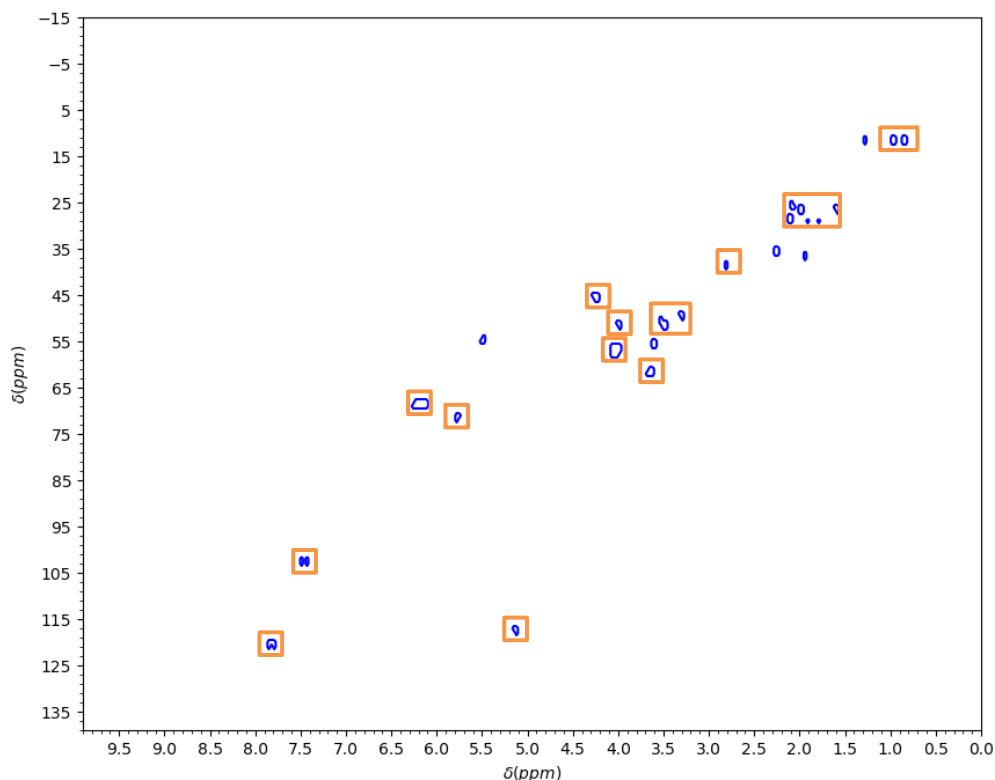


Figure 88 : Empreinte spectrale ^1H - ^{13}C HSQC obtenue par régression linéaire des données RMN ^1H - ^{13}C HSQC des 6 fractions de quinquina en fonction de leurs activités antipaludiques. Les taches de corrélation encadrées correspondent à aux moins un des alcaloïdes quinoléiques de référence.

III. 2. Conclusion

D'après nos tests d'activités, la cinchonine et la quinidine sont les deux molécules les plus actives contre *Plasmodium*, et ce sont les taches de corrélation de ces deux molécules qui apparaissent dans l'empreinte spectrale issue des analyses ^1H - ^1H TOCSY. Une sélection pertinente des taches de corrélation de la ou des molécules les plus actives a donc été réalisée parmi un mélange complexe d'analogues.

L'empreinte produite à partir des analyses ^1H - ^{13}C HSQC a permis de sélectionner les taches de corrélation appartenant à des alcaloïdes quinoléiques. Cependant, en raison de la faible variation des déplacements chimiques en ^{13}C , il n'est pas possible de distinguer les alcaloïdes entre eux.

Cet exemple a permis d'expérimenter la méthode dans le cas d'un mélange complexe composé de molécules de même classe chimique et dont l'activité biologique est répartie entre toutes les fractions. Les empreintes ^1H - ^1H TOCSY et ^1H - ^{13}C HSQC produites ont permis

d'identifier la classe de la molécule active mais seule l'analyse ^1H - ^1H TOCSY permet de faire apparaître **une empreinte mixte des molécules les plus actives**.

IV. Analyse différentielle de *Dilsea carnosa*

La méthode d'analyse différentielle automatique a été développée pour pouvoir identifier plus rapidement les molécules responsables de l'activité antipaludique partagée par les extraits d'algues rouges étudiées. Celle-ci a donc été appliquée à l'extrait AcOEt de *Dilsea carnosa* dans le temps imparti pour ce travail de thèse.

IV. 1. Introduction

Le genre *Dilsea* a été peu étudié. *D. carnosa* présente une activité contre *Staphylococcus aureus*¹⁶⁹, de même, une algue proche phylogénétiquement de *Dilsea*, *Neodilsea americana*, est active contre le virus *Herpes simplex*¹⁷⁰. Chez *Neodilsea yendoana* des molécules azotées telles que des acides aminés, des bétaines et sulfoniobétaines ont été isolées. Cette algue possède également une lipobétaine originale, la yendolipine¹⁷¹.

L'extrait AcOEt de *D. carnosa* a été étudié au laboratoire et il s'est montré capable d'inhiber de 23% et 98% la croissance de *P. falciparum* à 10 et 100 $\mu\text{g/mL}$, respectivement.

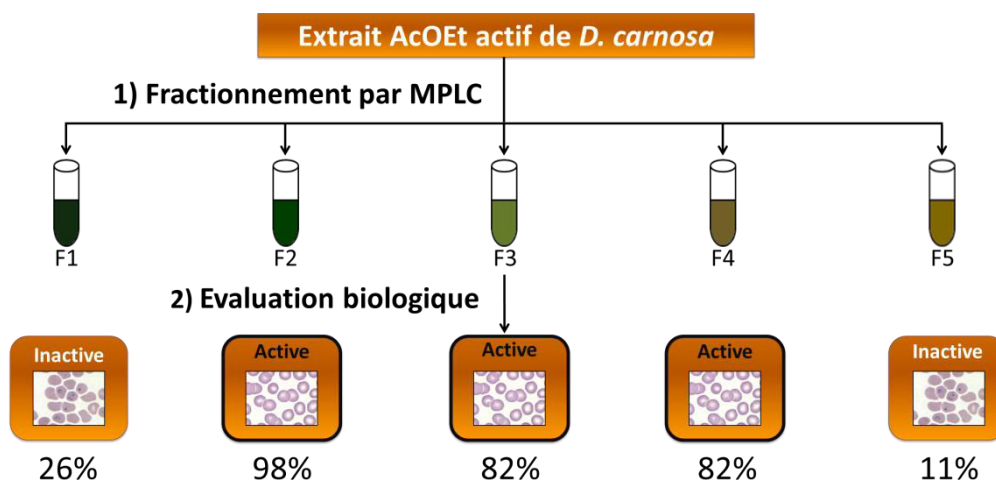


Figure 89 : Fractionnement par MPLC et évaluation biologique de l'extrait de *D. carnosa*. Les fractions F2, F3, F4 sont les plus actives.

L'extrait AcOEt a été fractionné par MPLC sur phase normale en 5 fractions. Ces fractions ont été évaluées biologiquement et ont montré, respectivement, un pourcentage d'inhibition de croissance de *Plasmodium* de 26, 98, 82, 82 et 11 à 100 µg/mL (Figure 89). L'analyse différentielle a ensuite été appliquée aux échantillons.

IV. 2. Analyse différentielle de *D. carnosa*

IV. 2. a. Analyse différentielle des spectres ^1H - ^1H TOCSY

L'activité biologique est répartie sur trois fractions, il est donc possible que les molécules actives soient différentes entre les fractions. Nous avons procédé à une analyse différentielle entre les fractions actives et les fractions inactives successives et par régression linéaire entre les intensités des *buckets* et l'activité biologique (Figure 90).

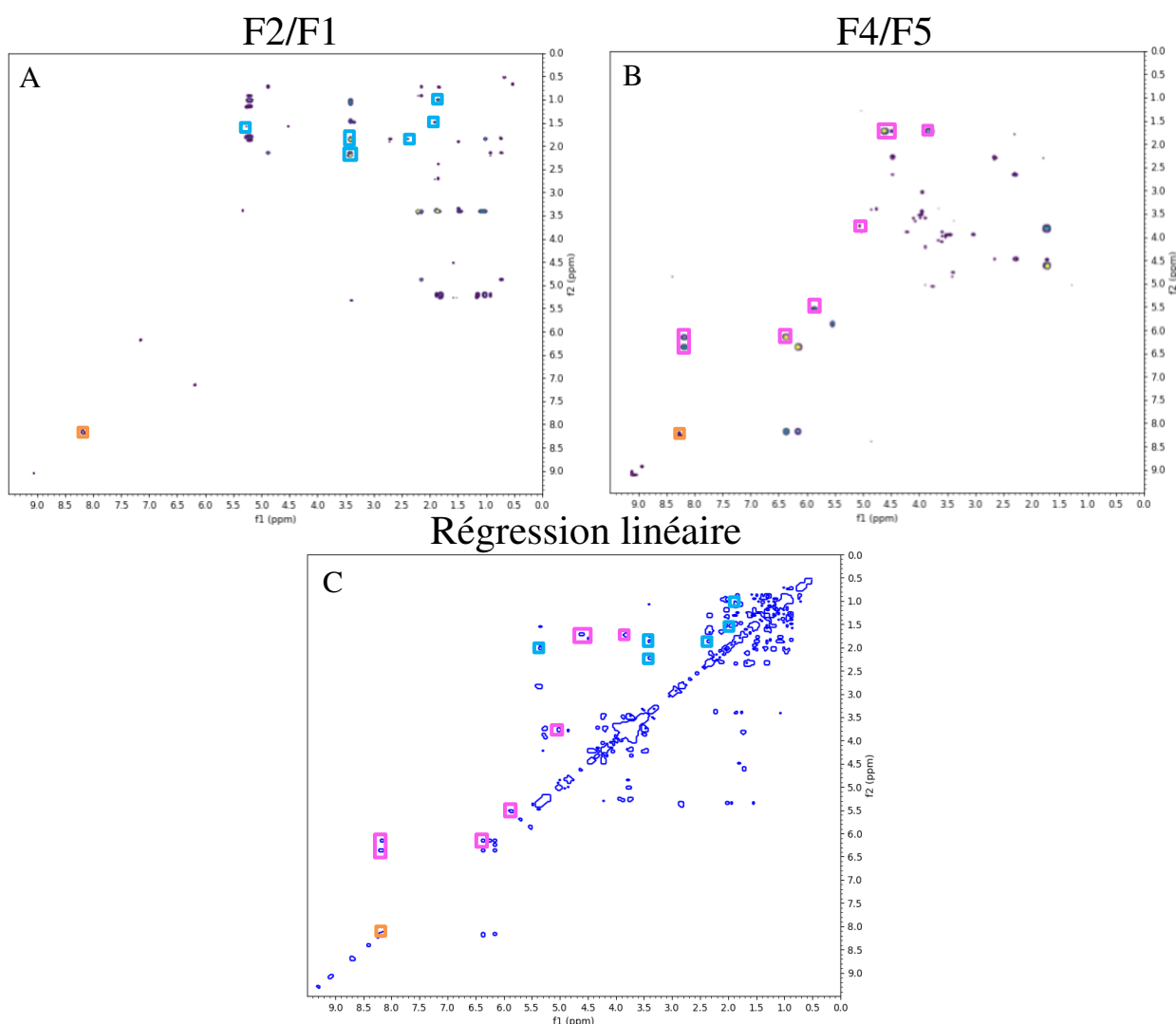


Figure 90 : Empreintes spectrales obtenues par le rapport des écarts-types des *buckets* des analyses ^1H - ^1H TOCSY des fractions 2 sur 1 (A) et des fractions 4 sur 5 (B), et par régression linéaire (C) Les empreintes obtenues par rapport des écarts-types sont très différentes, l'empreinte produite par régression linéaire recoupe les deux empreintes.

Le rapport des écarts-types des fractions actives *versus* inactives produisent deux empreintes très différentes l'une de l'autre (F2/F1 et F4/F5). Ces résultats suggèrent que l'activité serait portée par plusieurs molécules. De plus, les zones de déplacements chimiques sont également très différentes entre les deux empreintes, suggérant que les molécules actives n'appartiennent pas à la même classe.

IV. 2. b. Analyse différentielle des spectres ^1H - ^{13}C HSQC

L'analyse différentielle est appliquée aux analyses ^1H - ^{13}C HSQC.

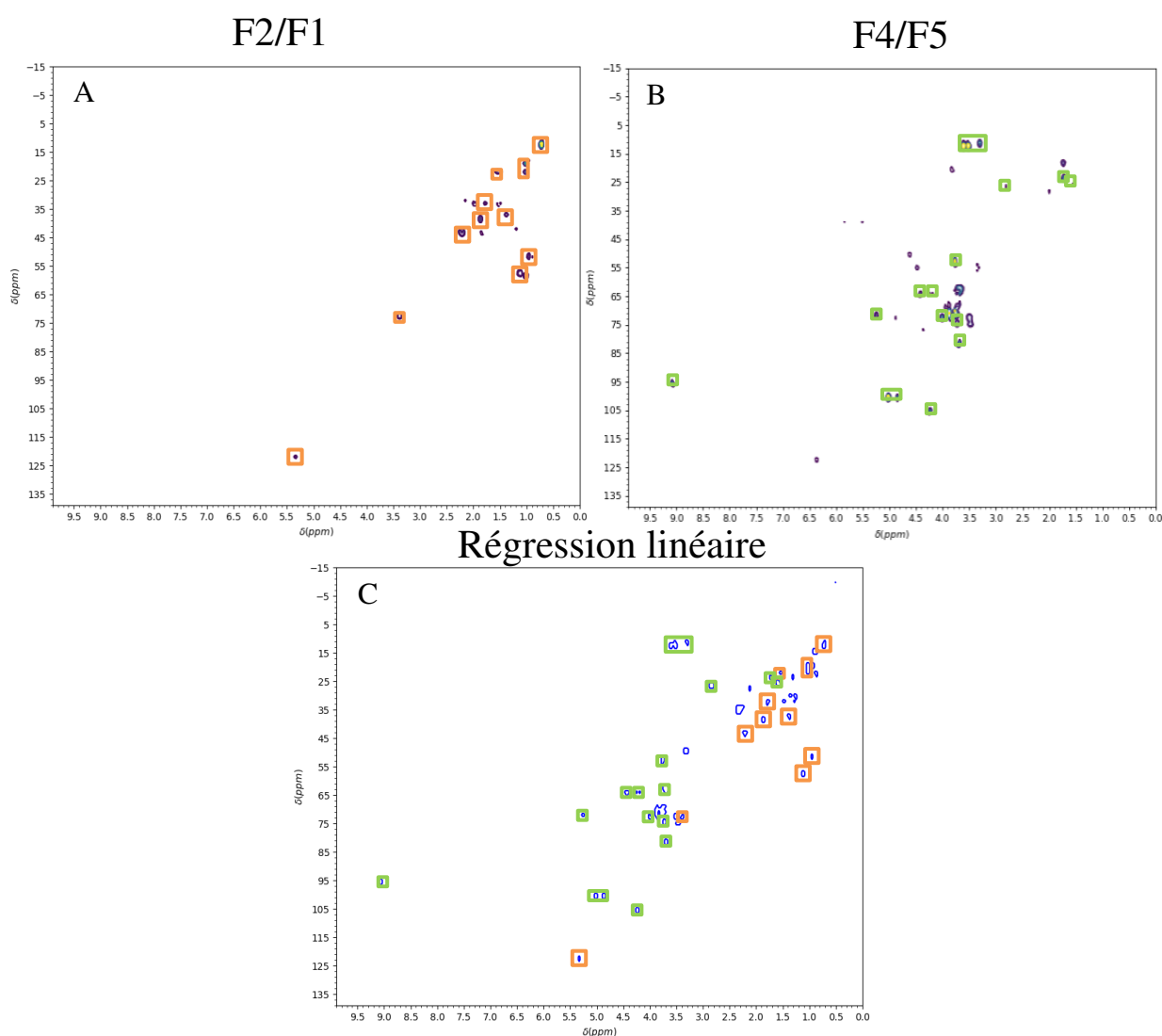


Figure 91 : Empreintes spectrales obtenues par le rapport des écarts-types des *buckets* des analyses ^1H - ^{13}C HSQC des fractions 2 sur 1 (A) et des fractions 4 sur 5 (B), et par régression linéaire (C). Les empreintes obtenues par rapport des écarts-types sont très différentes, l'empreinte produite par régression linéaire recoupe les deux empreintes.

Les résultats obtenus en ^1H - ^{13}C HSQC sont semblables aux résultats des ^1H - ^1H TOCSY. Le rapport des écarts-types des fractions actives *versus* inactives produisent deux

empreintes très différentes l'une de l'autre. Celle obtenue par régression linéaire les regroupe (Figure 91).

Ce résultat rappelle celui du quinquina. L'empreinte produite par régression linéaire mettrait en évidence deux molécules susceptibles de porter l'activité.

IV.3. Identification des composés par la méthode couplée HPLC-SPE-RMN

L'étape suivant la production de l'empreinte automatique RMN est l'analyse de la fraction la plus active par la méthode couplée HPLC-SPE-RMN. En raison d'un changement de personnel récent, de l'attente de leur formation, des réparations de plusieurs éléments de la méthode couplée, une analyse par fractionnement bio-guidé a été entamée.

IV.4. Identification des composés par fractionnement bioguidé

L'extrait AcOEt (2.6 g) a été fractionné par MPLC en utilisant un gradient heptane-CH₂Cl₂-MeOH (1:1:0 à 0:7:3) de polarité croissante, conduisant à l'obtention de 5 fractions (fractions destinées à l'analyse différentielle) notée F1 à F5. L'activité biologique de ces fractions a été évaluée sur *P. falciparum* 7G8. Les fractions F2 à F4 étaient actives avec un pourcentage d'inhibition de croissance du parasite de 98, 82, 82 %, respectivement.

La fraction F2 (240.5 mg), la plus active, a été sélectionnée et fractionnée sur colonne ouverte de silice en utilisant un gradient cyclohexane-AcOEt-MeOH de polarité croissante (8:2:0 à 0:8:2). Le fractionnement a été suivi par CCM et a permis de rassembler 6 fractions. Les fractions ont été testées sur *Plasmodium*, les fractions 2-3 à 2-6 étaient actives.

La fraction F2-3 active qui contenait le plus de masse a été fractionnée sur colonne ouverte de silice et a conduit à 3 composés majoritaires mais non purifiés. Le fractionnement a également été suivi par CCM. La fraction F2-3-3 inhibe la croissance de *Plasmodium* de 83% à 100 µg/mL (Figure 92).

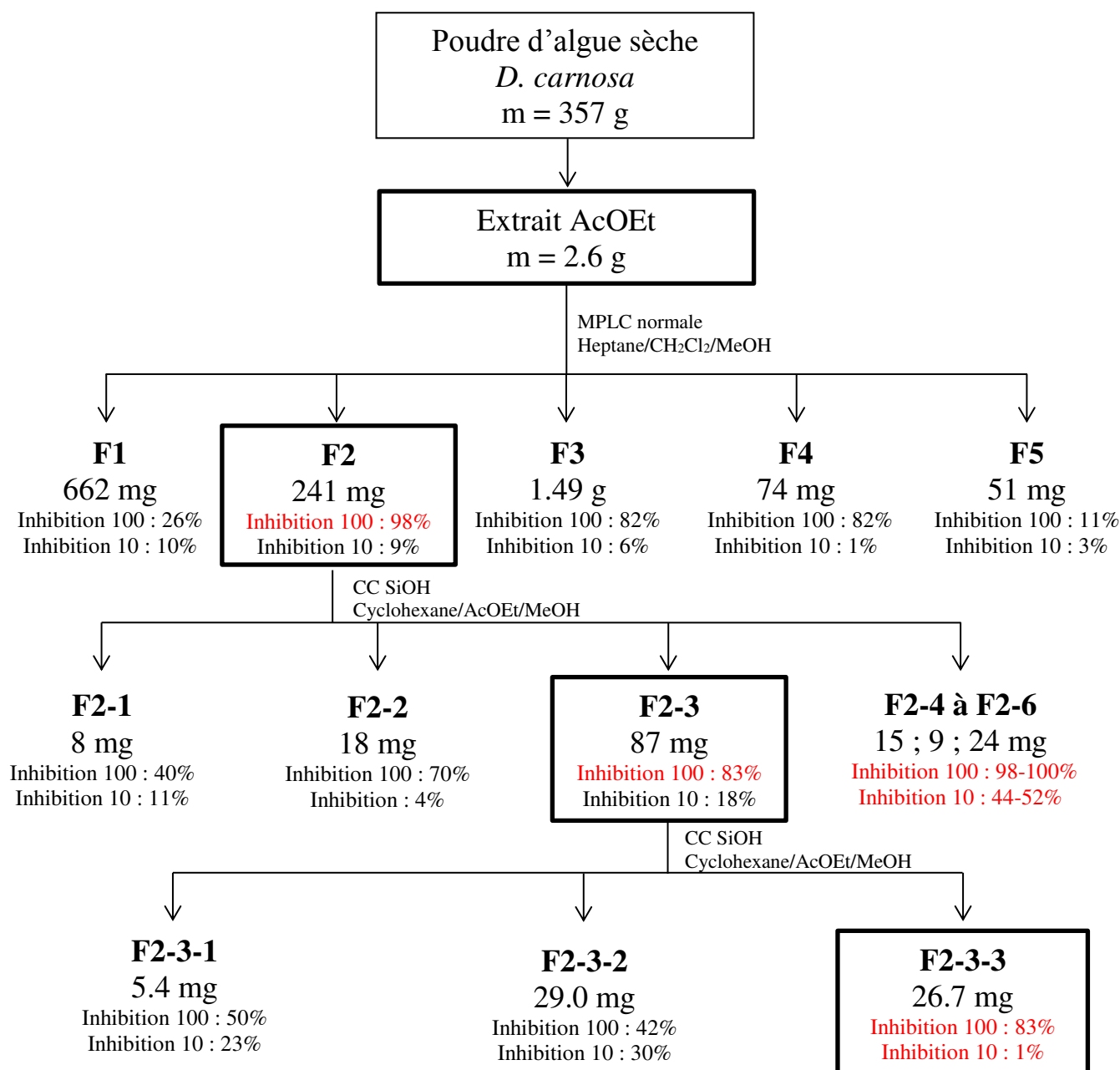


Figure 92 : Fractionnement de l'extrait AcOEt de *Dilsea carnosa*. Suite au premier fractionnement destiné à l'analyse différentielle (F1 à F5), la fraction la plus active a été fractionnée sur colonne ouverte. Les fractions F2-3 à F2-6 montrent une activité antipaludique. Le fractionnement de la fraction F2-3 a conduit à 3 composés non purifiés. La fraction F2-3-3 montre une activité antipaludique modérée. La fraction F2-4 est hémolytique. **Inhibition 100** : inhibition de croissance de *Plasmodium* à 100 µg/mL **Inhibition 10** : inhibition de croissance de *Plasmodium* à 10 µg/mL.

La répartition de l'activité biologique nous permet de nous conforter dans l'hypothèse formulée par la comparaison des fractions, que plusieurs molécules seraient porteuses de l'activité. La fraction F2-4 est hémolytique.

IV. 5. Comparaison de F2-3-3 à l’empreinte pharmacophorique

La fraction F2-3-3 est une fraction non purifiée mais composée d’une molécule majoritaire. Elle présente une activité de 83 % d’inhibition de croissance de *Plasmodium* à 100µg/mL, son activité est plus modérée que les fractions F2-4 à 6 mais il est intéressant de comparer son profil RMN aux empreintes précédemment produites.

Des analyses RMN ^1H - ^1H COSY et ^1H - ^{13}C HSQC de cette fraction sont réalisées puis comparées aux empreintes précédemment produites par régression linéaire (Figure 93).

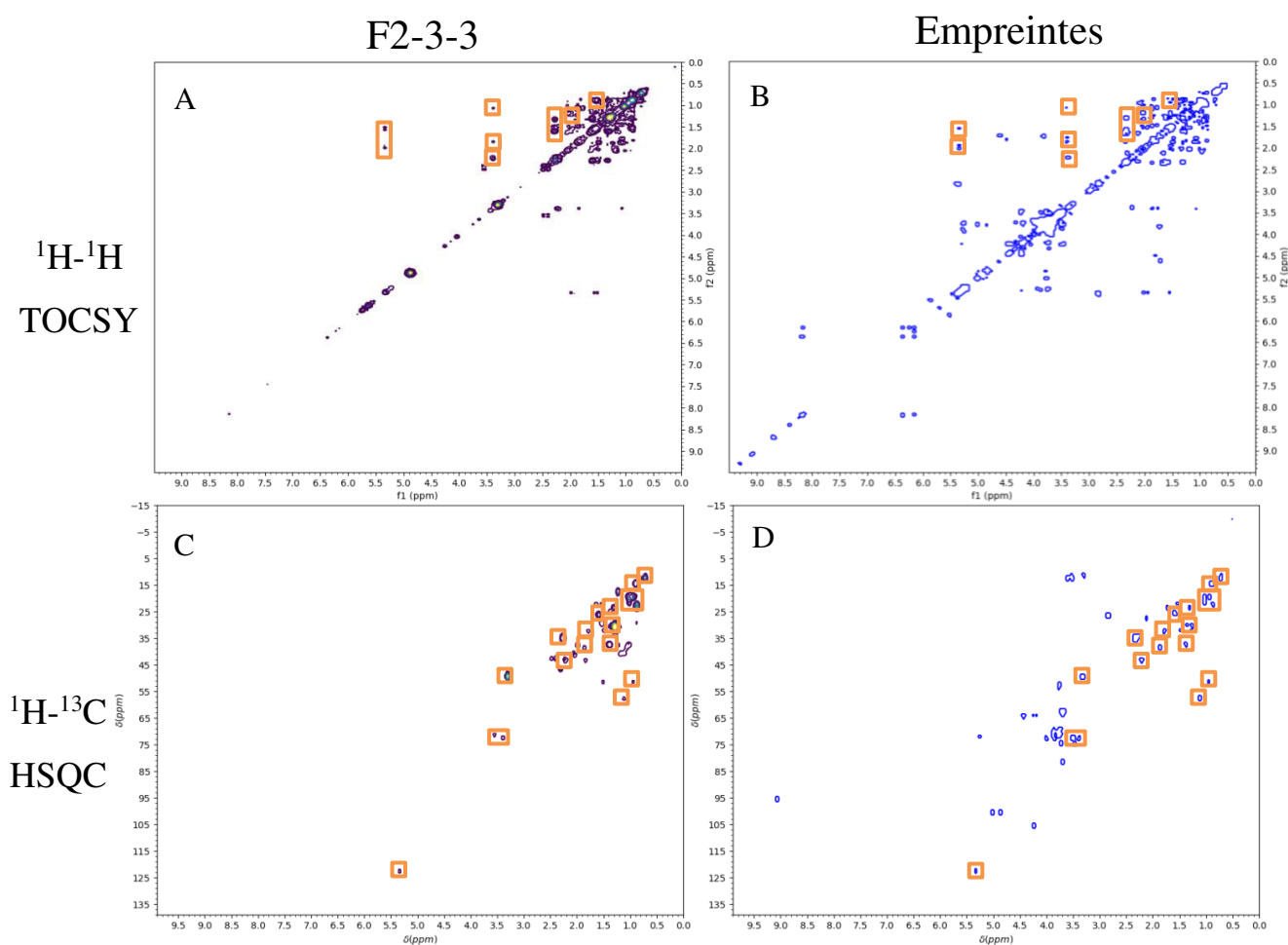


Figure 93 : Comparaison des spectres ^1H - ^1H COSY et ^1H - ^{13}C HSQC de F2-3-3 à l’empreinte pharmacophorique produite par régression linéaire. F2-3-3 possède de nombreuses taches de corrélation communes aux empreintes issues de la régression linéaire.

Il y a de nombreuses taches de corrélation communes entre les spectres RMN de F2-3-3 et les empreintes produites par régression linéaire, mais la correspondance n’est pas totale. De plus, la majorité des taches de corrélation de F2-3-3 sont également présentes dans l’empreinte spectrale produite par le rapport des écarts-types entre F2 et F1. Ce composé

semblerait correspondre pour moitié à l’empreinte pharmacophorique produite. Cela est d’ailleurs appuyé par les résultats d’activité biologique. En effet, cette fraction montre un pourcentage d’inhibition de 82 % à 100 µg/mL, ce qui est modéré par rapport à l’activité initiale de la fraction mais surtout par rapport aux fractions F2-4 à F2-6. Des analyses RMN ^1H , ^{13}C , ^1H - ^1H COSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC de ce composé ont été enregistrées. Par l’étude des déplacements chimiques et par comparaison avec les données de la littérature, nous pouvons affirmer qu’il s’agit d’un composé au squelette de type stérol présentant une insaturation dans la chaîne latérale^{172,173}. Cependant, les assignations exactes de cette chaîne restent incertaines en raison d’un trop grand nombre de recouvrement des déplacements chimiques. Les spectres RMN et les déplacements chimiques de cette fraction sont présentés en Annexe.

Une purification de ce produit reste à réaliser pour confirmer les résultats et son activité biologique. Il serait intéressant également de poursuivre le fractionnement pour isoler les autres molécules actives. Un fractionnement « RMN-guidé » pourrait être envisagé.

IV. 6. Comparaison avec l’empreinte de *Gracilaria gracilis*

Plusieurs extraits AcOEt d’algues rouges ont été testés au sein du laboratoire. Ils ont quasiment tous montré une activité antipaludique. Ces résultats ont permis d’émettre l’hypothèse que les algues rouges actives contre *Plasmodium* puissent partager des métabolites antipaludiques. Il est possible qu’une ou plusieurs classes de molécules communes à ces espèces soient responsables de cette activité.

Dans ce sens, une analyse différentielle de *Gracilaria gracilis* a été réalisée. Comme précédemment, l’extrait AcOEt a été fractionné en cinq fractions, qui ont été évaluées biologiquement. Ces fractions ont montré une inhibition de croissance de *Plasmodium* de 15, 59, 83, 23 et 32 % à 100 µg/mL, respectivement (Figure 94).

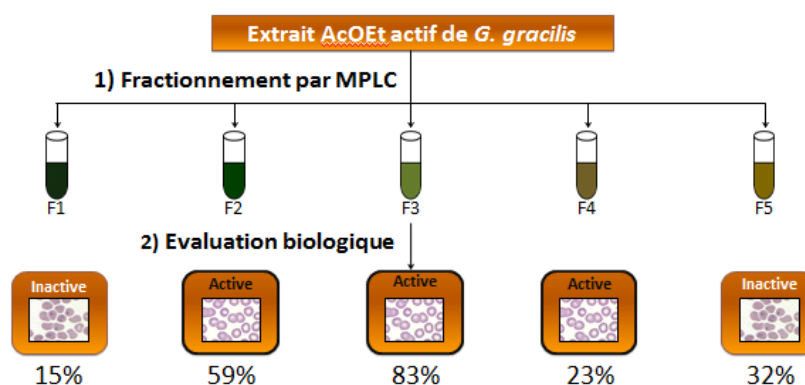


Figure 94 : Fractionnement et évaluation biologique de l'extrait de *G. gracilis*. La fraction 3 (F3) est la plus active, la fraction 2 (F2) possède une activité modérée.

IV. 6. a. Analyse différentielle des spectres RMN

La fraction 3 est la fraction la plus active, il est donc possible de réaliser une analyse différentielle de spectres deux à deux entre les fractions 2 et 3 et entre les fractions 3 et 4. La fraction 2 possède une activité très modérée, l'analyse différentielle à partir de cette fraction ne serait pas judicieuse.

- Analyse différentielle des expériences RMN ^1H - ^1H TOCSY

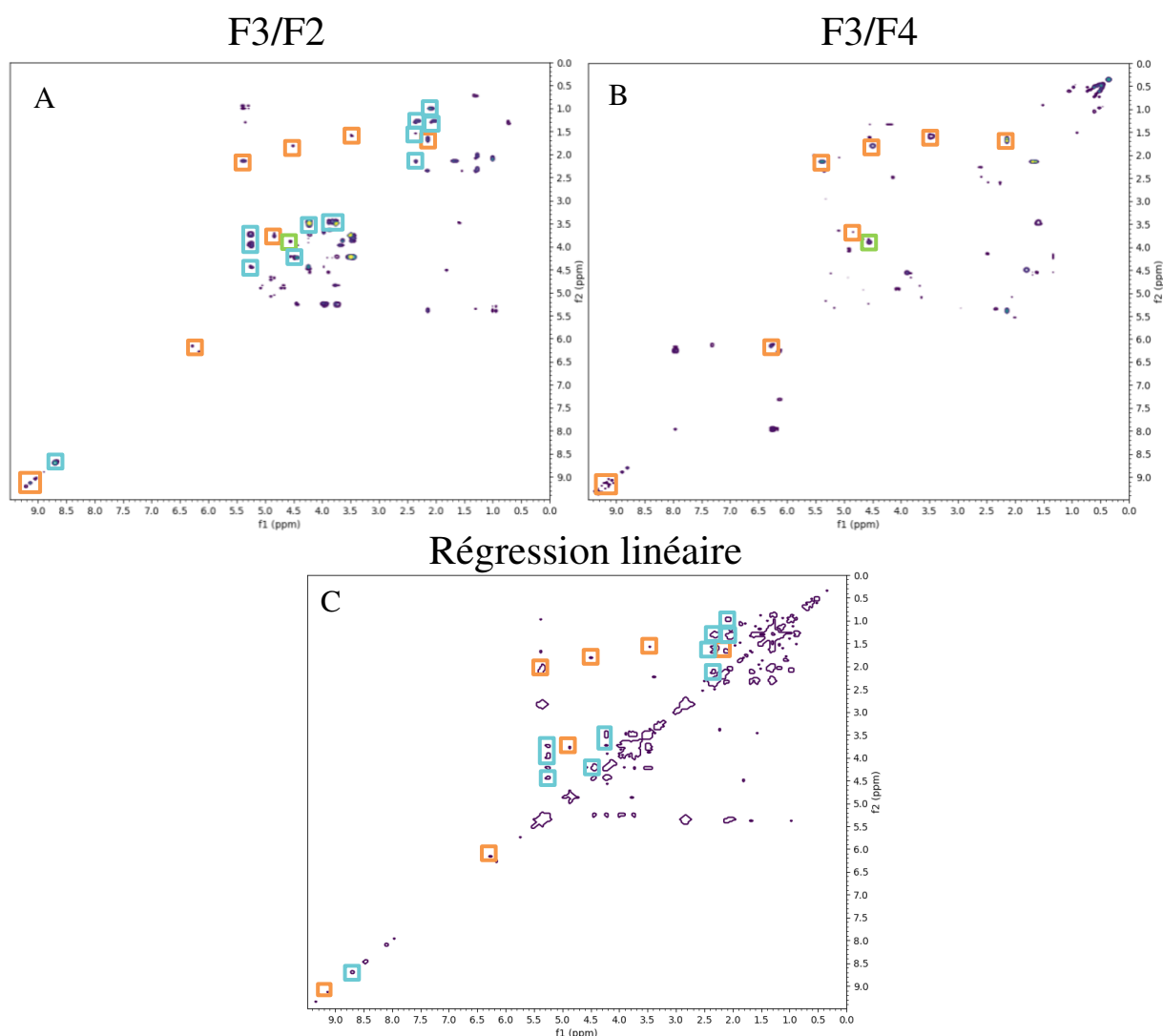


Figure 95 : Comparaison des empreintes spectrales ^1H - ^1H TOCSY de *G. gracilis*. Empreintes spectrales ^1H - ^1H TOCSY obtenues par rapports des écarts-types des *buckets* de F3 sur F2 (A) et de F3 sur F4 (B), et par régression linéaire des intensités des *buckets* en fonction du pourcentage d'inhibition (C). Les taches communes entre les spectres sont encadrées de la même couleur. Les carrés orange sont les taches communes aux trois empreintes.

La production d'empreinte par le rapport des écarts-types des *buckets* entre les fractions actives et inactives, ainsi que par la régression linéaire entre les intensités des

buckets et l'activité biologique a été réalisée. Les empreintes spectrales obtenues par les trois calculs se recourent entre elles pour quelques taches de corrélation (encadrées en orange). Il s'agit vraisemblablement de l'empreinte de la molécule active (Figure 95).

- **Analyse différentielle des expériences RMN ^1H - ^{13}C HSQC**

De la même manière que pour les empreintes ^1H - ^1H TOCSY, les empreintes spectrales obtenues par les trois calculs se recourent entre elles pour quelques taches de corrélation (Figure 96, encadrées en orange) qui appartiennent vraisemblablement à la molécule active.

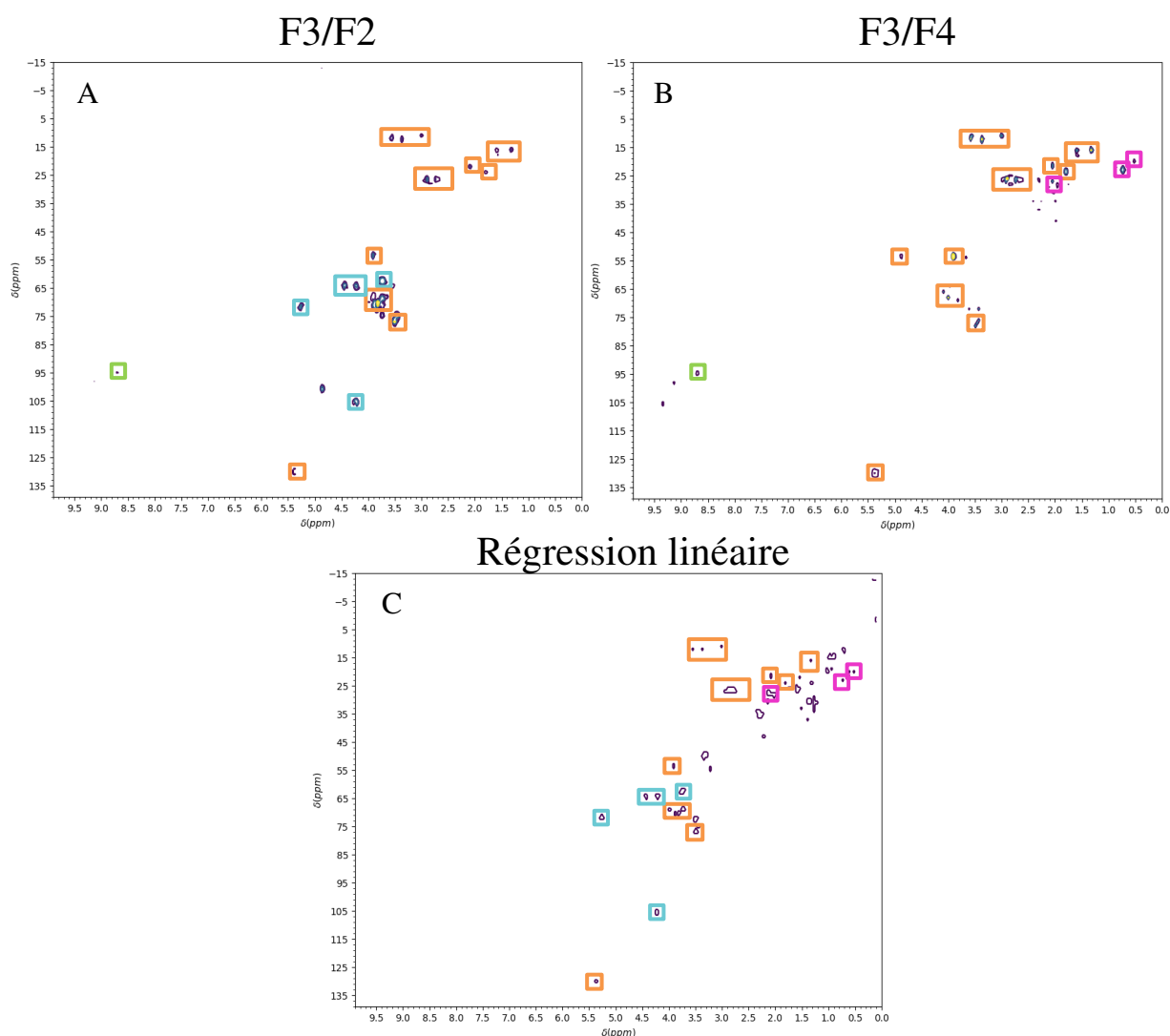


Figure 96 : Comparaison des empreintes spectrales ^1H - ^{13}C HSQC de *G. gracilis*. Empreintes spectrales ^1H - ^{13}C HSQC obtenues par rapports des écarts-types des *buckets* de F3 sur F2 (A) et de F3 sur F4 (B), et par régression linéaire des intensités des *buckets* en fonction du pourcentage d'inhibition (C). Les taches communes entre les spectres sont encadrées de la même couleur. Les carrés orange sont les taches communes aux trois empreintes.

IV. 6. b. Comparaison des empreintes de *D. carnosa* et de *G. gracilis*

Les empreintes produites par régression linéaire des deux algues ont été comparées entre elles (Figure 97).

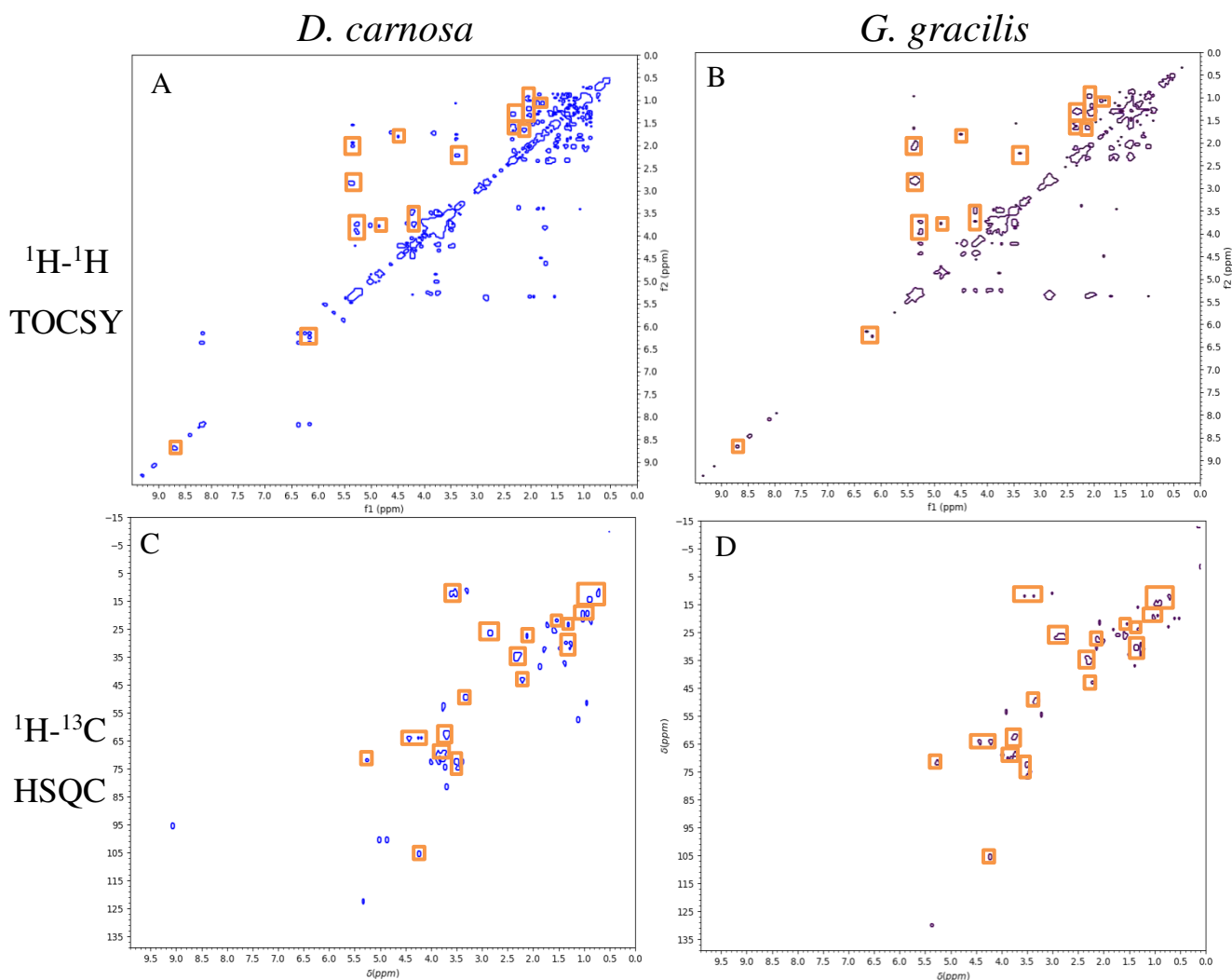


Figure 97 : Comparaison des empreintes ^1H - ^{13}C HSQC et ^1H - ^1H TOCSY obtenues par régression linéaire de *D. carnosa* et de *G. gracilis*. Empreinte ^1H - ^1H TOCSY de *D. carnosa* (A), empreinte ^1H - ^1H TOCSY de *G. gracilis* (B), empreinte ^1H - ^{13}C HSQC de *D. carnosa* (C), empreinte ^1H - ^{13}C HSQC de *G. gracilis* (D). Les empreintes montrent une similarité importante, elles ont de nombreuses taches de corrélation en commun.

Les empreintes de *D. carnosa* et de *G. gracilis* sont quasiment identiques. Cela suggère que la molécule active ou du moins sa classe chimique, est commune entre les algues rouges étudiées. Plusieurs taches de corrélation correspondent également à F-3-3.

Cependant, ces empreintes présentent également des similarités à celles obtenues pour *B. bifurcata* qui est une algue brune (Figure 98).

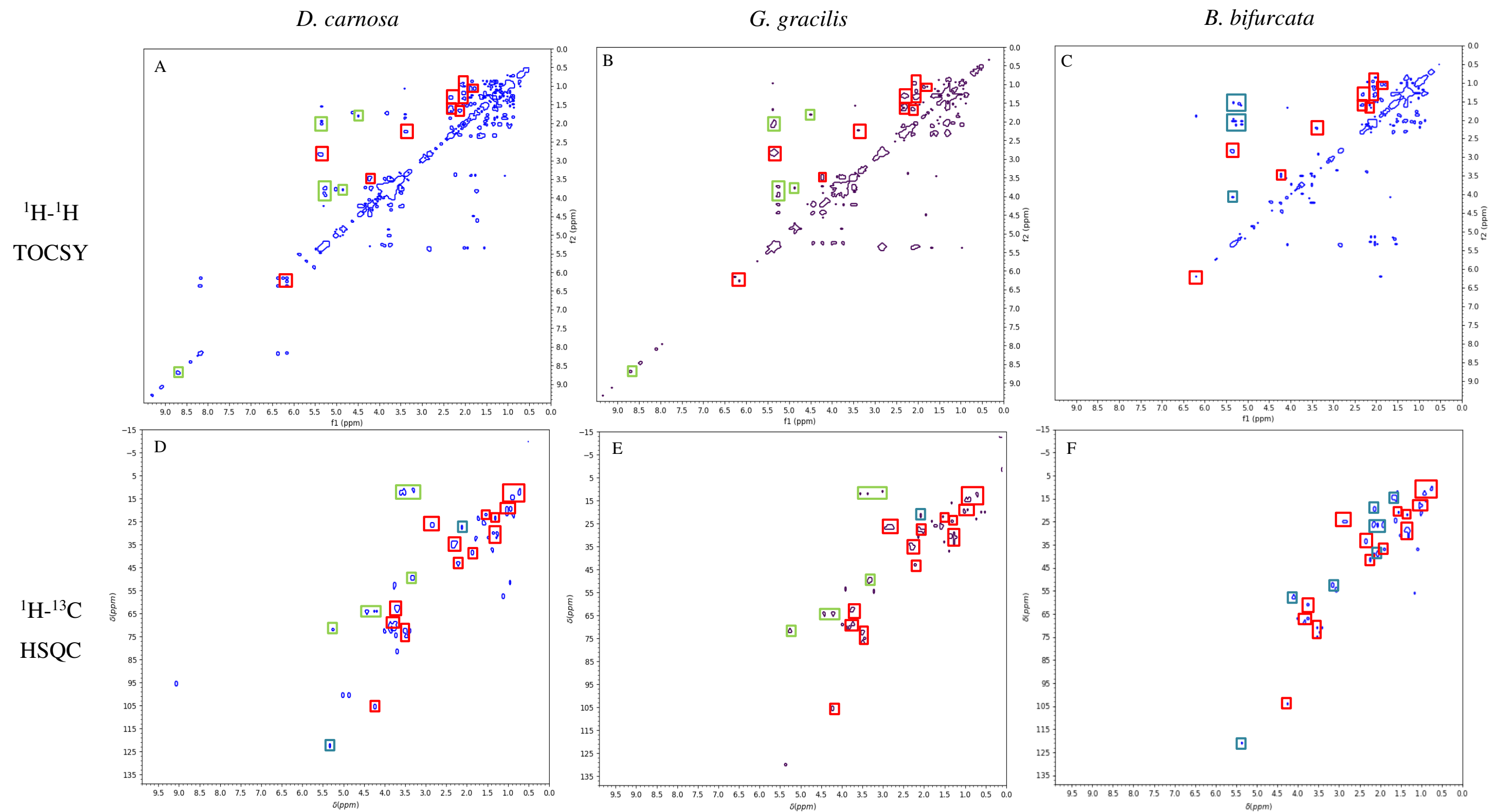
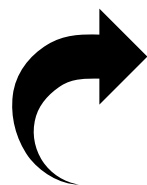


Figure 98 : Comparaison des empreintes ^1H - ^{13}C HSQC et ^1H - ^1H TOCSY obtenues par régression linéaire de *D. carnosa*, *G. gracilis* et *B. bifurcata*. Empreinte ^1H - ^1H TOCSY de *D. carnosa* (A), de *G. gracilis* (B), de *B. bifurcata* (C) empreinte ^1H - ^{13}C HSQC de *D. carnosa* (D), *G. gracilis* (E), de *B. bifurcata* (F). Les taches de corrélation encadrées en rouge sont communes aux trois algues, Les taches de corrélation vertes sont partagées entre les algues rouges, les taches bleues appartiennent à l'élégalonone.

Ouvrez la page S.V.P



Nous pouvons observer que les taches de corrélation présentes pour les empreintes de *B. bifurcata* sont également retrouvées chez les algues rouges (taches de corrélation encadrées en rouge). Cependant, bien que présentes chez *B. bifurcata*, les taches de corrélation appartenant à l'éléganolone (bleue), ne sont quasiment pas présentes dans les empreintes des algues rouges. La plupart de ces taches de corrélation correspondent au composé majoritaire de F2-3-3 de type stérol. Ces trois espèces possèdent donc un stérol en forte concentration dont la teneur évolue de la même manière que l'activité antipaludique et qui peut être relié à l'activité antipaludique.

En revanche, des taches de corrélation spécifiques des deux algues rouges apparaissent dans les empreintes de celles-ci, notamment des taches de corrélation aux déplacements chimiques compris entre 3 et 3.4 ppm en ^1H et environ 12 ppm en ^{13}C . Cette signature d'algue rouge suggère la présence d'une classe de molécule spécifique des algues rouges capable d'inhiber la croissance de *Plasmodium* (Figure 98).

IV. 7. Conclusion

L'analyse différentielle automatique de *Dilsea carnosa* était le premier cas d'analyse réel sur un échantillon inconnu dont la ou les molécule(s) active(s) étai(en)t inconnue(s). L'activité biologique était répartie sur les fractions 2, 3 et 4. La fraction 2 était la plus active. Cette répartition large de l'activité peut laisser penser que l'activité puisse être portée par plusieurs molécules.

Les deux empreintes produites par les rapports des écarts-types entre les fractions actives *versus* inactives ne se ressemblent pas. En revanche, leurs taches de corrélation sont présentes dans l'empreinte issue de la régression linéaire.

Malheureusement, du fait de contingences techniques, l'identification de composés par la méthode couplée HPLC-SPE-RMN n'a pas pu être réalisée. Ainsi, un fractionnement bio-guidé a été entrepris. Celui-ci a conduit à une fraction dont le composé majoritaire présente un squelette de type stérol.

L'empreinte pharmacophorique de *D. carnosa* correspond en partie à ce composé qui a montré une activité antipaludique modérée (83% d'inhibition à 100 $\mu\text{g/mL}$, 1% à 10 $\mu\text{g/mL}$).

Afin de répondre à la problématique de cette étude qui portait sur la possibilité que les algues rouges possèdent des métabolites secondaires spécifiques capables d'inhiber les voies

métaboliques de *Plasmodium*, une analyse différentielle d'une autre algue rouge, *G. gracilis*, a été entamée.

Ses empreintes RMN ^1H - ^1H TOCSY et ^1H - ^{13}C HSQC se sont montrées très similaires à celles de *D. carnosa* et plusieurs taches de corrélation correspondaient au composé de type stérol. Cependant, ces empreintes ont également été comparées à *B. bifurcata*, une algue brune dont le principal composé actif est un diterpène linéaire, l'éléganolone. Cette comparaison a révélé que l'empreinte du stérol était commune aux trois algues. Cependant, l'ensemble des taches de corrélation formant l'empreinte pharmacophorique de l'éléganolone n'a pas été retrouvé dans les algues rouges mais surtout, les deux algues rouges possèdent d'autres taches de corrélation spécifiques qui leur sont communes. S'agirait-il d'une signature des algues rouges ?

La production d'empreinte doit être étendue à l'ensemble des échantillons algaux pour vérifier si cette signature est retrouvée. En attente de la mise en fonctionnement de la méthode HPLC-SPE-RMN, un fractionnement RMN-guidé doit être poursuivi sur les fractions les plus actives de *D. carnosa* ou de *G. gracilis*. De plus, il serait intéressant de comparer ces empreintes avec un échantillon de rhodophytes inactif tel que *Palmaria palmata*, qui servirait alors de filtre.

V. Conclusion

Dans ce chapitre, l'analyse différentielle automatique a été évaluée sur plusieurs séries d'échantillons de difficulté croissante. La première série, SMARTE, était composée des fractions issues d'un extrait algal enrichi en artémisinine dans des teneurs naturelles. La concentration en artémisinine des fractions était inconnue, seule l'activité biologique avait été évaluée. Les empreintes spectrales ont été produites par le rapport des écarts-types des *buckets* entre les fractions actives *versus* inactives, et par régression linéaire des intensités en fonction de l'activité biologique. Ces empreintes ont permis d'identifier clairement l'artémisinine. Cette série valide la possibilité d'appliquer la méthode d'analyse différentielle automatisée sur des extraits algaux.

Puis l'analyse différentielle automatique a été appliquée au premier extrait algal naturel, l'extrait AcOEt de *B. bifurcata*. Précédemment, cet extrait avait été travaillé de manière classique par fractionnement bio-guidé. Il avait conduit à l'isolement d'un diterpène linéaire actif contre *Plasmodium*, l'éléganolone. A la suite de l'analyse différentielle

automatique de cet extrait, l'empreinte de ce composé a bien été retrouvée. Celle issue de l'analyse différentielle par régression linéaire a également fait apparaître d'autres taches de corrélation n'appartenant pas à cette molécule. Cela suggère que d'autres molécules sont responsables de l'activité, ce qui est appuyé par l'activité antipaludique de l'éléanolone équivalente à celle de l'extrait, et par l'isolement d'un second diterpène linéaire actif, le bifurcatriol. Cependant, ce composé ne correspond pas aux taches de corrélation de l'empreinte spectrale obtenue par régression linéaire. Il est possible qu'une autre classe chimique contribue à l'activité.

Ensuite, l'analyse différentielle a été évaluée sur un cas complexe : les alcaloïdes du quinquina. Toutes les fractions issues de l'extrait hydroalcoolique étaient actives, la mesure des CI_{50} s'est avérée nécessaire. Dans ce cas, seule la régression linéaire était applicable. L'empreinte spectrale produite correspondait aux deux molécules de référence les plus actives par nos tests biologiques, la quinidine et la cinchonine.

Enfin, afin de répondre à l'hypothèse de recherche qui suppose que **des métabolites secondaires des algues rouges pourraient interférer dans les voies métaboliques de *Plasmodium* et plus particulièrement de l'apicoplaste**, l'analyse différentielle automatique de l'algue rouge, *D. carnosa* a été réalisée. L'empreinte pharmacophorique obtenue par le rapport de l'écart-type des *buckets* des fractions actives *versus* inactives a produit deux empreintes différentes. L'empreinte issue de la régression linéaire regroupait ces deux dernières empreintes. La méthode couplée HPLC-SPE-NMR n'étant pas opérationnelle, un fractionnement bio-guidé a été entamé. Celui-ci a conduit à une molécule majoritaire partiellement purifiée correspondant en partie à l'empreinte. Les déplacements chimiques 1H et ^{13}C ont permis de proposer une structure de type stérol. Cependant, l'activité biologique de ce composé n'explique que partiellement l'activité de l'extrait. Puis, l'analyse différentielle a été appliquée à *G. gracilis*. L'empreinte obtenue a été comparée à celle de *D. carnosa* et de *B. bifurcata*. Les trois empreintes montrent des taches de corrélations communes de type stérol.

Particularité intéressante, **les empreintes des deux algues rouges partagent par ailleurs des taches de corrélation spécifiques**, dont certaines aux déplacements chimiques inhabituels (entre 3 et 3.5 ppm en 1H et environ 12 ppm en ^{13}C). Il pourrait s'agir des composés responsables de l'activité, et cela signifierait que les algues rouges possèdent une signature spectrale caractéristique. Ces résultats sont à approfondir par l'application de la méthode automatique d'analyse différentielle aux autres extraits d'algues rouges, et par

l'identification de ces métabolites, soit par l'utilisation de la méthode couplée HPLC-SPE-NMR, soit en poursuivant le fractionnement de manière RMN-guidée et bio-guidée.



Discussion

Discussion

Ce projet doctoral avait pour but de développer et valider une nouvelle stratégie de recherche de métabolites antipaludiques par une **approche phylogénétique**. En effet, les algues rouges et l'apicoplaste, organite relique de *Plasmodium*, sont liés par un passé évolutif commun. L'objectif scientifique était donc de montrer la présence **de métabolites secondaires actifs contre *Plasmodium*, le parasite responsable du paludisme, dans plusieurs espèces d'algues rouges.**

Pour y répondre, il était essentiel de développer **une nouvelle stratégie analytique** efficace et rapide d'identification de **molécules actives et inconnues dans des extraits végétaux.**

Cette stratégie développée repose sur l'analyse différentielle de spectres RMN produisant une empreinte pharmacophorique et l'utilisation d'une méthode HPLC-SPE-RMN couplées. L'analyse différentielle s'appuie sur les travaux initiaux du Dr. Schroeder. Après une première application de sa méthode à nos échantillons, plusieurs limites se sont révélées. Pour les contourner, l'objectif premier a été d'informatiser et d'automatiser cette analyse différentielle.

Cela a été possible grâce à une collaboration étroite avec le Dr. Marc-André Delsuc de l'IGBMC. Nos travaux conjoints ont permis de valoriser le programme libre SPIKE qui permet de traiter les données de spectroscopie à transformée de Fourier. Un programme, nommé Plasmodesma, a spécialement été conçu pour faire le lien entre nos données RMN, SPIKE et l'utilisation finale de ces données en pharmacognosie. La première version de ce programme a été développée par le stagiaire Petar Markov, dont le travail a été remarquable. Plasmodesma est inspiré de la métabolomique et réalise entre autres, une étape essentielle, le *bucketing*. Cette étape est primordiale pour aligner des spectres entre eux et ainsi, pouvoir réaliser une analyse différentielle. Cette version de Plasmodesma a ensuite fait l'objet d'améliorations continues en lien avec nos retours d'expériences.

Ce programme préparait nos données à l'analyse différentielle mais il fallait ensuite trouver le moyen de la réaliser et de produire la meilleure empreinte pharmacophorique. A partir de cette étape, un stage intensif de Python chez M-A. Delsuc a été obligatoire. Ce séjour a été particulièrement efficace pour échanger entre « l'informaticien-RMNiste » et moi-

même... la « détentrice de l’empreinte ». Ainsi, un outil a été mis au point, *BucketUtilities*, qui a également fait l’objet d’améliorations, notamment dans l’affichage de spectres RMN adapté aux phytochimistes.

Le développement informatique des analyses différentielles s’est appuyé sur deux séries d’échantillons « reconstitués », ARTE et CHLORO, composés d’un extrait algal inactif et d’une molécule antipaludique connue, la chloroquine ou l’artémisinine. Ainsi, les meilleures empreintes des molécules actives ont été obtenues par plusieurs calculs :

- le rapport des écarts-types des *buckets* des analyses RMN entre une fraction active et une fraction inactive ;
- l’application d’une régression linéaire entre les intensités des *buckets* des analyses RMN et l’activité biologique.

La méthode automatisée a donc été mise au point et validée sur ce premier cas d’école.

La méthode a ensuite été évaluée sur des séries d’échantillons de difficulté croissante. La première série, SMARTE, était constituée d’un extrait algal inactif supplémenté en artémisinine dans une teneur naturelle, puis fractionné en 5 fractions. L’analyse différentielle a été appliquée et grâce aux empreintes spectrales obtenues, l’artémisinine a pu être clairement identifiée. Cette série d’expériences valide donc la possibilité d’appliquer avec succès la méthode d’analyse différentielle automatisée à des fractions d’extraits végétaux en retrouvant une molécule active connue dans ce système complexe plus « naturel ».

L’expérience suivante a été réalisée sur un extrait algal actif de *B. bifurcata*. Bien qu’il s’agisse d’une algue brune, cet extrait a été choisi car il avait été précédemment travaillé au laboratoire par un fractionnement bio-guidé. Celui-ci avait conduit à l’isolement de son principe actif majoritaire, l’éléganolone. En effet, ce diterpène linéaire présentait une activité antipaludique ($CI_{50} = 7.9 \mu M$). Notre stratégie a, cette fois, été appliquée aux fractions issues de cet extrait actif. L’empreinte spectrale automatiquement générée a permis de retrouver et d’identifier l’éléganolone. En raison de son activité équivalente à l’activité de l’extrait, il était certain que d’autres molécules contribuaient à l’activité. Récemment, le bifurcatriol a été isolé de *B. bifurcata* et montre une activité antipaludique supérieure ($2.0 \mu M$) mais il n’était présent dans l’algue à une teneur de 0.06 % contre 0.1% pour l’éléganolone.

Ce résultat montre les limites de l'analyse différentielle. En effet, elle sélectionne les signaux les plus intenses qui évoluent de la même manière que l'activité biologique. Ainsi une molécule possédant une activité moins importante mais présente dans une concentration plus élevée qu'une autre serait sélectionnée. Cependant, une molécule dont la concentration évolue dans les fractions successives de la même manière que l'activité biologique mais qui ne serait pas liée à l'activité serait également sélectionnée. Pour cette raison, l'activité doit être évaluée de la manière la plus précise possible. Nous avons effectué la plupart de nos mesures d'activité biologique en deux points, mais dans les cas où l'activité biologique était identique entre les fractions, une mesure de la CI_{50} a été réalisée.

Le fractionnement a une importance dans la simplification de la composition moléculaire de l'échantillon et dans la distribution de l'activité biologique. Jusqu'ici, il s'est effectué d'une manière très standardisée entre les différents échantillons algaux. Ce fractionnement standardisé était nécessaire dans le cas des algues rouges puisque l'étude porte sur une analyse différentielle sur plusieurs espèces. En revanche, dans le cas d'analyses différentielles sur un extrait individuel, il pourrait être intéressant de coupler deux types de fractionnement orthogonaux, soit en utilisant une phase stationnaire différente, soit en utilisant une autre technique de fractionnement par chromatographie de partage centrifuge (CPC) par exemple. Cela multiplierait les échantillons et les profils de variation de l'activité biologique, ce qui permettrait une analyse différentielle d'autant plus spécifique et efficace.

Nous avons ensuite éprouvé la méthode d'analyse différentielle en étudiant les alcaloïdes quinoléiques du quinquina. Dans ce cas, toutes les fractions obtenues à partir de l'extrait brut se sont montrées biologiquement actives. Leurs CI_{50} ainsi que celles des 4 alcaloïdes de référence ont donc été mesurées. L'empreinte spectrale obtenue par régression linéaire correspondait aux mélanges des empreintes des deux alcaloïdes les plus actifs dans nos tests d'activité, la quinidine et la cinchonine. Cela prouve une nouvelle fois, l'efficacité de l'approche en situation réelle, permettant de ressortir les composants actifs de métabolites structurellement proches.

Enfin, but ultime de notre étude, l'analyse différentielle des algues rouges a été entreprise. *Dilsea carnosa* a été choisie en raison de la quantité d'extrait disponible et de son activité antipaludique. L'activité était répartie sur 3 fractions, ce qui suggérait que plusieurs composés étaient probablement responsables de l'activité de l'extrait. Les empreintes

spectrales issues du rapport des écarts-types différaient entre elles, et celle issue de la régression linéaire réunissait ces deux dernières.

L'étape suivante était normalement l'utilisation de la méthode HPLC-SPE-RMN à partir de la fraction active pour caractériser directement la structure de la molécule active correspondant à l'empreinte. Malheureusement, cette étape n'a pas pu être réalisée pour des raisons de logistique. L'HPLC-SPE-RMN est-elle une méthode miracle ? L'équipement est disponible sur la plateforme analytique de la faculté de Pharmacie de Strasbourg et partagé, tout comme les autres équipements du site, entre tous les utilisateurs internes à la faculté de pharmacie mais aussi avec les entreprises externes. Depuis mon arrivée, le personnel affecté à cette plateforme de service varie d'une à deux personnes maximum. Elles sont en charge d'injecter, analyser, réparer, commander et de se former sur tous ces appareils. Premièrement, mettre en œuvre une méthode couplée implique d'avoir un personnel formé à chaque élément couplé, c'est-à-dire du spectromètre de masse à la RMN, mais cela implique surtout, un personnel disponible pour le développement analytique. Deuxièmement, ce n'est pas un appareillage utilisé en routine : la remise en fonctionnement d'un appareillage qui a été arrêté pendant plusieurs semaines est long et provoque son lot de réparations et paramétrages. La méthode couplée n'additionnerait-elle pas les limites de chaque élément ? En effet, l'HPLC équipée d'une colonne C18 exclut les éléments très polaires et très apolaires. Ensuite, le détecteur DAD montre seulement les molécules visibles en UV mais le spectromètre de masse agit comme un second détecteur et peut contourner cette limite. En ce qui concerne le piégeage de la molécule sur les cartouches SPE, la molécule doit être correctement piégée et retenue sur la cartouche et plusieurs injections seront nécessaires pour accumuler suffisamment de masse afin de réaliser une analyse RMN. La molécule doit ensuite être correctement désorbée avec le solvant deutéré et en quantité suffisante pour réaliser l'acquisition RMN. Malgré tout, les cartouches SPE donnent l'avantage de pouvoir stocker les molécules, de sécher le solvant provenant de l'HPLC, de le remplacer par un solvant deutéré, et d'obtenir une masse suffisante pour l'analyse RMN, d'autant plus qu'il s'agit d'une sonde capillaire de 30 μ L. Malgré les limites qu'elle peut présenter, l'HPLC-SPE-RMN demeure une méthode efficace et prometteuse pour la déréduplication de produits naturels comme le montrent de nombreuses publications et notre expérience, lorsque les conditions opérationnelles optimales sont réunies.

Ainsi, un fractionnement bio-guidé a été entrepris. Cela a permis d'identifier une partie de l'empreinte spectrale pharmacophorique qui est de type stéroïdique.

En effet, les déplacements chimiques ^1H et ^{13}C , ainsi que les corrélations observées sur les expériences RMN 2D ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC permettent facilement de retrouver le squelette stérane dont le carbone 3 est porteur de la fonction hydroxyle (δ_{H} 5,29 ppm ; δ_{C} 71.9 ppm). Ensuite, trois multiplets sont visibles entre 5.22 ppm et 5.35 ppm. Sur la base des corrélations ^1H - ^{13}C HMBC, il s'agit de l'insaturation en position C-5,6 et d'une insaturation présente sur la chaîne latérale (Annexe). La fraction a été analysée par LC-BRMS (ESI+), plusieurs composés étaient visibles, notamment, un pic $[\text{M}+\text{H}]^+$ m/z 413.29. Cette masse ainsi que les déplacements chimiques observés en RMN pourraient correspondre au stigmastérol, stérol commun des algues rouges^{172,173}. Cela reste à être confirmé par la purification de ce produit ainsi que par des analyses HR-MS/MS.

Cette partie d'empreinte est commune aux espèces d'algues rouges *D. carnosa* et *G. gracilis* mais également à l'algue brune *B. bifurcata*. Cependant, une autre partie de l'empreinte est uniquement commune aux deux algues rouges et présente des taches de corrélation particulières comprises entre 3 et 4 ppm en ^1H -RMN et 9 et 12 ppm en ^{13}C -RMN. Cette empreinte renforce l'hypothèse de l'existence d'une classe commune de composés responsables de l'activité antipaludique dans les algues rouges.

Les composés responsables de l'activité doivent maintenant être isolés et caractérisés. Cependant, la production d'empreinte pourrait permettre un fractionnement RMN-guidé allégé de tests biologiques.

Ainsi, ce travail de thèse a permis de développer la production d'empreintes RMN des produits actifs de manière automatique. L'ensemble du processus a été validé dans diverses conditions expérimentales de complexité croissante. La méthode se révèle efficace, rapide et sensible. En effet, les empreintes produites pour chaque série ont toujours permis d'identifier la molécule active, même quand celle-ci ne représentait que 0.5 % de la masse de la fraction. La durée de la méthode entre le fractionnement de l'extrait jusqu'à la production de l'empreinte est estimée à 4 jours. Les points limitants sont l'évaluation biologique et l'analyse par la méthode HPLC-SPE-RMN.

Pour la première fois, une méthode analytique allie la déréplication (plutôt la déconvolution) et l'activité biologique. Celle-ci ne se réfère pas à des bases de données et peut permettre de déterminer la structure d'un composé inconnu. Elle permet d'économiser le bioguidage long, cher et fastidieux. L'obtention d'une empreinte RMN du composé actif permet ensuite de l'identifier beaucoup plus rapidement grâce à l'utilisation d'une méthode

couplée de type HPLC-SPE-RMN ou par la mise en place d'un fractionnement RMN-guidé. De plus, cette méthode est universelle car elle peut être applicable à n'importe quelles sources d'échantillons ou d'activités biologiques. Plasmodesma est en constante progression pour s'adapter aux différents types d'analyses et spectromètres RMN. Enfin, tous les programmes sont en lignes et complètement libres d'accès. La création d'un tutoriel permettant de diffuser la méthode est à réaliser en priorité.

De plus, Plasmodesma contient tous les outils qui pourraient servir au développement d'une véritable plateforme de déréplication de spectre RMN comme GNPS pour la spectrométrie de masse. Une base de données libre et riche de spectres RMN-2D pourrait permettre de réaliser des comparaisons automatisées en ligne, comme cela existe pour le génotypage ou sur... GNPS.

Récemment, une nouvelle procédure appelée « *bioactive molecular networking* » a été développée par Dr. L. F. Nothias en 2018¹⁷⁴. Elle permet d'allier la déréplication par réseaux moléculaires et l'activité biologique. Cette méthode repose tout d'abord sur le fractionnement de l'échantillon actif et l'évaluation biologique de ces fractions. Ces fractions sont ensuite analysées par LC-MS/MS. Un score de bio-activité est calculé en utilisant la corrélation de Pearson entre l'intensité des ions et l'indice de sélectivité des fractions. Ce score de bioactivité est directement visualisé dans le réseau moléculaire par la taille du nœud. L'abondance relative de la molécule dans la fraction est également visible. Cette étude permet aujourd'hui de créer des réseaux moléculaires où l'activité biologique des échantillons est reliée à l'abondance relative des molécules dans les échantillons. Cette étude a aussi montré que l'identification complète d'une molécule non répertoriée dans une base de données est impossible sans passer par une étape de purification.

Par ailleurs, la totalité des extraits et fractions d'algues rouges a été analysée par LC-MS² (collab. Dr David Touboul, ICSN). Une analyse de déréplication par réseaux moléculaires sera appliquée dans la continuité de ces travaux.

Qu'apporte notre méthode par rapport à celles existantes ? Lesquelles devons-nous privilégier ?

Les apports de chacune des méthodes actuelles d'identification de produits naturels sont résumés dans le tableau 13. La méthode de déconvolution pharmacophorique est la seule méthode qui permet d'accéder rapidement à la structure complète d'une molécule active

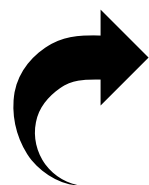
connue ou non. Cependant, elle n'est pas la plus efficace par exemple pour rechercher un métabolite connu dans plusieurs échantillons.

Le choix de la méthode repose principalement sur l'objectif ultime de l'étude. Par exemple, dans le cadre d'une recherche industrielle des domaines cosmétiques ou nutraceutiques, ou, si l'objectif est d'enrichir les connaissances phytochimiques d'un échantillon, il nous semble préférable de se tourner vers des méthodes de dérégulation. Si l'objectif premier est d'identifier un composé responsable d'une activité biologique au sein d'un mélange complexe il sera préférable de se tourner vers la méthode de déconvolution pharmacophorique.

Tableau 13 : Avantages et inconvénients des méthodes actuelles les plus utilisées pour l'identification de produits naturels dans des mélanges complexes.

	Déréplication par réseaux moléculaires	Déréplication par ^{13}C et ^1H - ^{13}C RMN	Déréplication par RMN ^1H et ^1H - ^1H	Méthode couplée HPLC-SPE-RMN	Déconvolution Pharmacophorique (Empreinte RMN + HPLC-SPE-RMN)
Point faible technique	Ionisation	Sensibilité	Shift des solvants deutérés	Echantillon unique Limites des appareils de chromatographie/détection	Limites de la chromatographie/détection de la HPLC-SPE-RMN
Référence aux bases de données	OUI	OUI	OUI	NON	NON
Identification de composés connus	OUI	OUI	OUI	OUI	OUI
Identification de composés inconnus	NON	NON	NON	OUI	OUI
Identification des composés bioactifs	OUI (2018)	NON	NON	NON	OUI
Points forts	Sensibilité Accessibilité	Structure	Structure	Structure complète	Structure complète Bio-activité

Ouvrez la page SVP



Sur le plan scientifique, les dernières activités biologiques obtenues, notamment l'inhibition de la voie de biosynthèse des isoprénoïdes par plusieurs extraits d'algues rouges, renforcent notre approche phylogénétique car cette voie métabolique est également présente dans l'apicoplaste chez *Plasmodium*. Cette nouvelle stratégie thérapeutique reste à être explorée et élargie à un plus grand nombre d'espèces d'algues rouges, et le sera dans la continuité de ces travaux notamment dans le cadre de collaborations (D. Tasdemir au GEOMAR-Biotech, ResNet NPND, Algaia™). Par ailleurs, l'approche phylogénétique pourrait être étendue au-delà des Rhodophytes. Par exemple, de nombreuses molécules antipaludiques avec une bonne sélectivité ont été isolées d'éponges marines¹⁷⁵⁻¹⁸¹. Or, *Chromera velia*, l'organisme qui a permis de résoudre l'énigme de l'origine algale de l'apicoplaste, représente le chaînon manquant entre les apicomplexes et les dinoflagellés. Ces organismes possèdent eux aussi un chloroplaste issu d'une endosymbiose avec une algue rouge. Ils sont connus pour vivre en symbiose avec des organismes marins comme les coraux ou les éponges marines. Il apparaît intéressant d'explorer cette source comme support à l'approche phylogénétique.

Pour ma part, ce projet de thèse a bénéficié d'un environnement complexe et enrichissant, partagé entre la programmation, la chimie analytique et la parasitologie. Cela a été une véritable expérience de gestion de projet et de relation entre les différents acteurs de cette étude collaborative. Il est difficile de partir vers d'autres horizons lorsque l'on sent enfin la pelote se dénouer. La pharmacognosie est une spécialité formidablement multidisciplinaire... une cuillère à soupe de chimie saupoudrée d'une pincée de biologie, mais dans mon cas, une louche de Python a été nécessaire. Cet ingrédient amer pour un palais inexpérimenté devient rapidement essentiel au bon équilibre de la recette. Ces trois années de potage m'ont permis, je l'espère, de scientifiquement bien croire et de pouvoir affronter la suite avec sérénité.



Partie

Expérimentale

Partie expérimentale

I. Procédures générales

I. 1. Solvants

Des solvants de pureté $\geq 99\%$ ont été utilisés pour les extractions solide/liquide du matériel végétal ainsi que pour les étapes de fractionnement.

I. 2. Fractionnements, purifications et suivi

I. 2. a. Chromatographie sur couche mince (CCM)

Les extraits et fractions issues des séparations chromatographiques ont été analysés par chromatographie sur couche mince (CCM) sur plaque de gel de silice 60 F254 de 0,2 mm d'épaisseur sur support d'aluminium (Fluka, Sigma-Aldrich). Les plaques ont été visualisées sous UV à 254 et 365 nm avant d'être révélées par pulvérisation d'un révélateur. Les révélateurs principalement utilisés étaient la vanilline sulfurique 10%, ainsi que le réactif à l'iodoplatinate (Sigma-Aldrich).

I. 2. b. Chromatographie sur colonne (CC)

Les CC ont été réalisées soit dans des colonnes en verre de différents diamètres en utilisant de la silice de granulométrie (40-63 μm) où l'éluant est poussé par moyenne pression, soit par l'utilisation d'un appareil de chromatographie à moyenne pression automatique (CLMP), (*Spot Liquid Chromatography Flash*, Armen Instrument) sur colonne en phase normale PuriFlash PF-30 SiHP/40G (Interchim). L'appareil est contrôlé par le logiciel Chromeleon. Les échantillons sont déposés en dépôt solide en les adsorbant sur de la célite (VWR).

I. 2. c. Chromatographie liquide haute performance (CLHP)

La chromatographie haute performance (CLHP) en phase inverse analytique a été réalisée sur une chaîne Agilent 1200 équipée d'un détecteur à barrette de diode (DAD) et d'une colonne analytique C18 Ascentis express (2,7 μm , 4,6 x 75 mm).

I. 3. Analyses physiques et spectrales des composés

I. 3. a. Spectrométrie de masse (SM)

Les analyses de spectrométrie de masse basse résolution (BRMS) ont été réalisées sur un spectromètre de masse à trappe ionique HCT Ultra (Bruker), équipé d'une source d'ionisation par électrospray (ESI), couplé à une chaîne LC (*liquid chromatography*) Agilent 1200. Les solvants d'élution étaient de l'eau milliQ (solvant A), et de l'acétonitrile (B). L'élution a consisté en un gradient de 2% à 100% de B pendant 2,5 min, et un maintien à 100% de B pendant 5 minutes à un débit de 800 $\mu\text{L}/\text{min}$.

Les analyses de spectrométrie de masse haute résolution (HRMS) ont été réalisées sur un spectromètre de masse quadripôle-*time-of-flight* (Q-TOF) Agilent Accurate 6520 (Agilent Technologies) équipé d'une source d'ionisation ESI et couplé à une chaîne LC Agilent 1200 équipée d'une colonne Zorbax SB-C18 (2.1 x 50 mm ; 1.8 μm , Agilent Technologies). Les solvants d'élution étaient de l'eau milliQ (solvant A), et de l'acétonitrile (B) à 0.05 % d'acide formique. L'élution a consisté en un gradient de 2% à 100% de B pendant 8 min, et un maintien à 100% de B pendant 4,5 minutes à un débit de 500 $\mu\text{L}/\text{min}$.

Des analyses LC-ESI-HRMS² ont été réalisées à l'ICSN (Gif-sur-Yvette). Les analyses LC ont été réalisées sur une chaîne Dionex Ultimate 3000 RSLC équipée d'une colonne C18 Accucore (2.1 x 100 mm ; 2.6 μm , ThermoScientific). Les solvants d'élution étaient de l'eau milliQ (solvant A), et de l'acétonitrile (B). L'élution a consisté en une étape isocratique à 20% de B pendant 5 min, puis un gradient de 20% à 100% de B pendant 20 min, et un maintien à 100% de B pendant 4 minutes à un débit de 350 $\mu\text{L}/\text{min}$. Le système LC est couplé à un spectromètre de masse Q-TOF Agilent 6540 (Agilent Technologies) équipé d'une source d'ionisation ESI.

I. 3. b. Résonance magnétique nucléaire (RMN)

Les expériences de RMN effectuées à la plateforme analytique de la Faculté de Pharmacie de Strasbourg, ont été réalisées sur des appareils Bruker Avance III 400 MHz équipé d'une sonde BBFO+ ou sur 500 MHz, Ultrashield+ équipé d'une sonde BBFO+. Les déplacements chimiques sont exprimés en ppm (partie par million) et calibrés par rapport au solvant de référence interne. Les valeurs des δ des solvants sont celles publiées par Cambridge Isotope Laboratories, Inc. (Tableau 14). Les constantes de couplage (J) sont exprimées en Hertz (Hz). La multiplicité des signaux est donnée par les abréviations suivantes : s (singulet), d (doublet), t (triplet), m (multiplet), q (quadruplet), qt (quintuplet).

L'attribution des signaux des protons et des carbones a été effectuée à partir des expériences ^1H , ^{13}C , ^1H - ^1H COSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, en utilisant les logiciels ACD Spectrum Processor ou TopSpin 3.5. Les solvants deutérés utilisés sont le chloroforme-d, le méthanol- d_4 ou l'acétone- d_6 (Sigma-Aldrich ou Euriso-Top).

Tableau 14 : Tableau des solvants deutérés de référence.

Solvants	δ_{C} en ppm (multiplicité)	δ_{H} en ppm (multiplicité)
Chloroforme-d	77,23 (3)	7,24 (1)
Méthanol- d_4	49,15 (7)	3,31 (5)
		4,87 (1)
Acétone- d_6	206.68 (1)	2.05 (1)
	29.92 (7)	

Les analyses RMN réalisées sur le spectromètre RMN 700 MHz sont détaillées dans la partie méthode de déconvolution (Analyses RMN).

II. Matériel végétal

II. 1. Matériel algal

II. 1. a. Collection et identification

Toutes les espèces d'algues utilisées sont originaires de Normandie. La détermination taxonomique a été réalisée par Dr. A.-M. Rusig et un échantillon de référence a été déposé à l'Herbier de l'Université de Caen-Normandie (Tableau 2).

II. 1. b. Préparation des extraits bruts

Le thalle de chaque espèce a été rincé à l'eau de mer filtrée (EMF) puis rapidement à l'eau douce avant d'être lyophilisé puis réduit en poudre. Celle-ci a été extraite par l'AcOEt (10% m/v pendant 3 x 24h) sous agitation et à l'abri de la lumière. Les extraits sont séchés à une température inférieure ou égale à 40°C à l'évaporateur rotatif puis par évaporateur à centrifugation (Genevac EZ-2) en enfin conservés à 4°C (Tableau 4).

II. 2. Matériel végétal non algal

- *Cinchona pubescens* L.

Quinze grammes de poudre d'écorce de *Cinchona pubescens* L. (Rubiaceae, Herboristerie CAILLEAU) ont été extraits par une solution hydroalcoolique à 60% (10% m/v, pendant 3 x 24h sous agitation et à l'abri de la lumière). La masse de l'extrait est de 3.6 g (rendement de 24%).

III. Méthode de déconvolution pharmacophorique

III. 1. Fractionnement par MPLC

III. 1. a. Extraits algaux

Tous les extraits AcOEt algaux utilisés pour la méthode différentielle ont été fractionnés par chromatographie MPLC selon la méthode suivante (Tableau 15).

Tableau 15 : Méthode MPLC de l'étape de fractionnement des extraits algaux.

Appareil	Spot Liquid chromatography Flash, Armen Instrument				
Temps (min)	0	60	100	120	150
Solvant A (%) : Heptane	50	0	0	0	0
Solvant B (%) : CH ₂ Cl ₂	50	100	70	70	0
Solvant C (%) : MeOH	0	0	30	30	0
Colonne	PuriFlash PF-30 SiHP/40gr Interchim				
Débit	18 ml/min				
Détection (nm)	254 nm et 365 nm				

Les fractions sont rassemblées toutes les 30 min, soit **F1** (heptane:CH₂Cl₂ 50:50 à 25:75), **F2** (heptane:CH₂Cl₂ 25:75 à 0:100), **F3** (CH₂Cl₂:MeOH 100:0 à 77.5:22.5), **F4** (CH₂Cl₂:MeOH 77.5:22.5 à 70:30), **F5** (lavage de la colonne par *Magic system*, (CH₂Cl₂:Acétone:MeOH:Acide formique 5:3:0.5:0.5) (Figure 99).

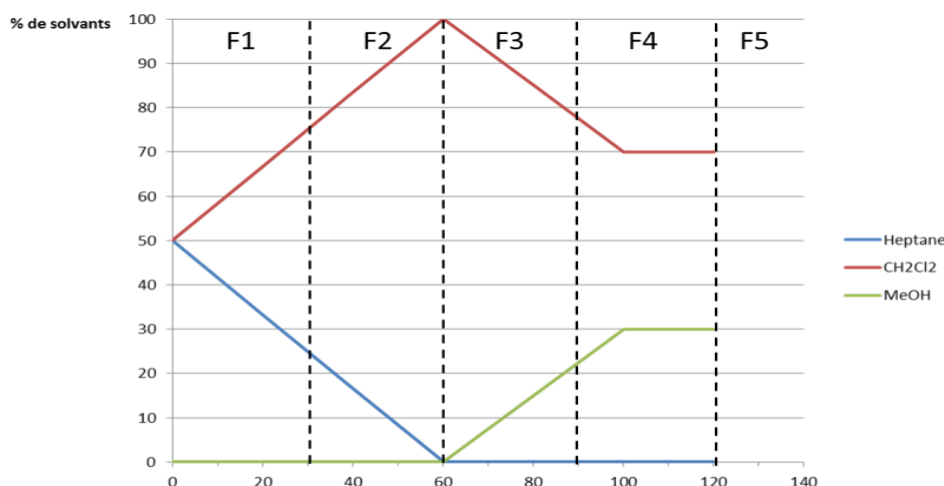


Figure 99 : Gradient des solvants de la méthode MPLC pour le fractionnement des extraits algaux.

III. 1. b. Extrait non algal : *Cinchona pubescens* L.

L'extrait hydroalcoolique de *C. pubescens* a été fractionné par chromatographie MPLC selon la méthode décrite dans le Tableau 16. 1% de diéthylamine sont ajoutés aux solvants d'éluion.

Tableau 16 : Méthode MPLC de l'étape de fractionnement de l'extrait de quinquina.

Appareil	Spot Liquid chromatography Flash, Armen Instrument				
Temps (min)	0	15	30	60	90
Solvant A (%) : CH ₂ Cl ₂	98	98	96	85	70
Solvant B (%) : MeOH	2	2	4	15	30
Colonne	PuriFlash PF-30 SiHP/40G Interchim				
Débit	18 ml/min				
Détection (nm)	254 nm et 365 nm				

Les fractions sont rassemblées en fonction de leurs profils CCM.

III. 2. Analyses RMN

Les analyses RMN destinées aux analyses différentielles ont été réalisées sur la plateforme de l'Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC) en collaboration avec Marc-André Delsuc. Elles ont été réalisées sur un spectromètre Bruker Avance-III à 700 MHz équipé d'une cryo-sonde TCI et d'un passeur d'échantillons Bac60. Les tubes RMN sont centrifugés avec une petite centrifugeuse manuelle pour culoter la partie insoluble au fond du tube. Chaque échantillon est automatiquement inséré dans le spectromètre, *tunné* et *shimmé* après un délai de stabilisation de 120 secondes. Toutes les

expériences RMN d'une série d'échantillons sont automatiquement lancées par une macro Topspin. Les paramètres spectraux (pulses $\pi/2$, gain...) sont optimisés sur l'échantillon le moins soluble et sont utilisés pour les autres échantillons de la série. Toutes les analyses sont réalisées à 295 K.

Les largeurs spectrales sont de 12 ppm pour les ^1H et de 150 ppm pour les ^{13}C .

Les spectres ^1H sont acquis sur 64 scans, 16 384 points et un délai de relaxation de 1.5 s, pour une durée d'analyse de 3 minutes.

Les expériences ^1H - ^1H COSY (*CORrelation SpectroscopY*) sont acquises sur 8 scans, avec 512 incréments de 4 096 points chacun, pour une durée totale d'analyse de 2 heures.

Les expériences ^1H - ^1H TOCSY (*TOTAL Correlation SpectroscopY*) sont acquises sur 8 scans avec 400 incréments de 4 096 points chacun et une séquence de mélange DIPSI-2 de 80 ms. La durée totale de l'acquisition ^1H - ^1H TOCSY est de 1 heure et 40 minutes.

Les expériences ^1H - δ DOSY (*Diffusion-Ordered SpectroscopY*) sont réalisés sur 32 scans avec 50 incréments de 4 096 points chacun pour une durée d'analyse de 50 minutes.

Les expériences ^1H - ^{13}C HSQC (*Heteronuclear Single Quantum Correlation*) sont acquises sur 4 scans avec 512 incréments de 2 048 points pour une durée d'acquisition d'une heure.

Les expériences ^1H - ^{13}C HMBC (*Heteronuclear multiple-bond Correlation*) sont acquises sur 48 scans avec 400 incréments de 4 096 points pour une durée d'acquisition de 10 heures.

III. 3. Bonnes pratiques d'acquisition RMN

Pour réaliser une analyse différentielle la plus performante possible, il est important que les échantillons ainsi que l'acquisition des données RMN soient réalisés dans les meilleures conditions. En effet, les analyses RMN du proton peuvent être plus délicates à utiliser en analyse différentielle en raison des shifts que peuvent subir les signaux d'un proton, d'où l'importance de préparer les échantillons et les analyses de manière la plus reproductible possible. Pour cela, quelques bonnes pratiques ont été établies :

Préparation des échantillons :

- L'échantillon doit être dépourvu de solvants et eau résiduels
- Le solvant deutéré doit contenir une référence interne, le tétraméthylsilane (TMS)
- Le solvant deutéré utilisé doit être identique entre les échantillons
- Le volume de solvant deutéré doit être identique entre les échantillons
- La concentration en produits doit être identique entre les échantillons

Acquisitions RMN :

- Température stable et identique des échantillons (295 K).
- Réglage du *lock*, accord de sonde, *shim* et impulsion à 90° (P1) sur l'échantillon qui est le moins soluble
- Le nombre de points d'acquisition doit être identique
- La fenêtre spectrale doit être identique entre les échantillons
- Un spectre ^1H de l'échantillon est réalisé en début et fin d'analyse

III. 4. Traitement des données

Le programme Plasmodesma est relié à SPIKE pour la plupart de ces opérations. Il peut traiter de manière autonome de large série de différentes expériences RMN provenant de différents échantillons aux conditions variables. Ce traitement analytique implique la manipulation de séries complexes d'expériences RMN (1D et 2D homo- et hétéronucléaires) et produit automatiquement des rapports portant sur les données RMN analysées, des figures, des listes de pics et *buckets*. C'est un programme Python compatible Python 2.7 et 3.5. Il est implémenté à SPIKE comme un plug-in. Il est applicable aux analyses ^1H , ^1H - ^1H COSY, ^1H - ^1H TOCSY, ^1H - δ DOSY, ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC.

Le programme Plasmodesma est disponible sur : <https://github.com/delsuc/plasmodesma>

Le programme SPIKE est disponible sur : <https://bitbucket.org/delsuc/spike>

III. 4. a. Principe des opérations

Ce programme fonctionne sans intervention de l'homme. Lorsqu'il est appliqué à un dossier, tous les fichiers d'analyses RMN 1D et 2D qu'il contient vont être automatiquement importés et, un rapport global des analyses qu'il contient est produit. Tout le traitement des données et les étapes d'analyses ont été optimisés en fonction des paramètres d'acquisition des différentes analyses RMN. Les données RMN brutes sont apodisées, une transformée de Fourier est appliquée et la ligne de base est corrigée. Une réduction du bruit de fond est

également réalisée. Une calibration du spectre est effectuée par la détection du pic de TMS à 0 ppm. Les spectres 1D sont également phasés de manière automatique. Ensuite, Plasmodesma réalise un peak-picking et le *bucketing* des données. La liste des pics et des *buckets* est enregistrée sous un fichier *csv* pour chaque expérience et des figures de chaque spectre sont créées. Enfin un rapport résumant les analyses RMN et les fichiers issus de Plasmodesma est produit (Figure 100).

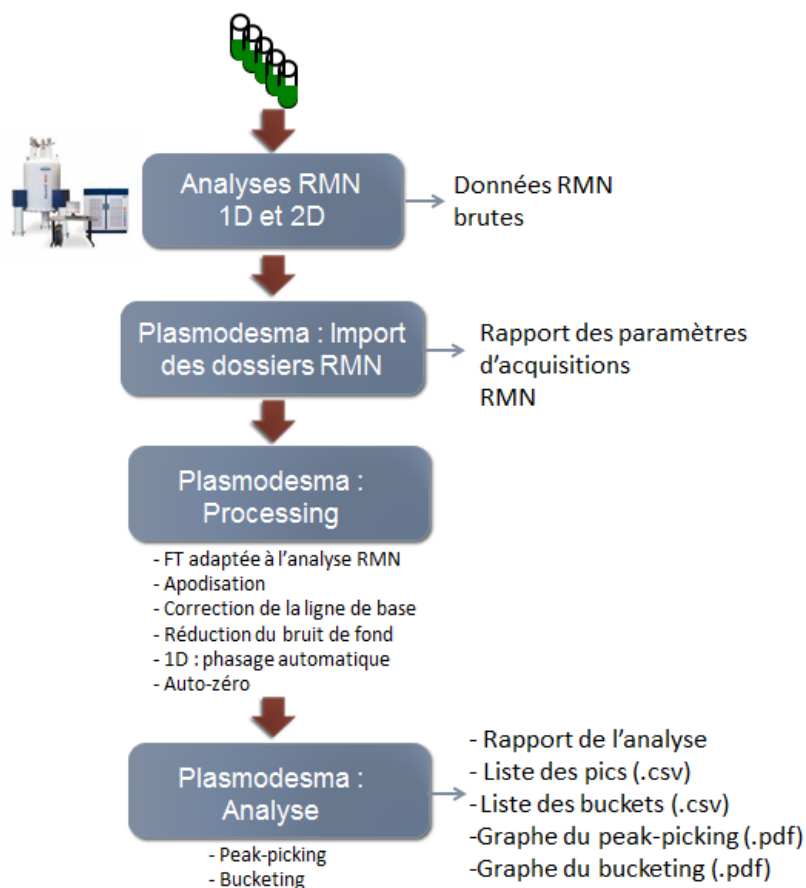


Figure 100 : Principe des opérations de Plasmodesma. Après l'acquisition des données RMN, celles-ci sont importées par SPIKE qui va les traiter de manière automatique. Cela implique l'application de transformées spécifiques de l'expérience RMN, une apodisation, une correction de la ligne de base, une réduction du bruit de fond, un phasage automatique des 1D et un calibrage du spectre. Ensuite, Plasmodesma va réaliser la détection automatique des pics et un *bucketing* des données. Plusieurs fichiers sont générés : les listes de pics et de *buckets* et des rapports sur les paramètres d'analyses et de traitement de données.

III. 4. b. Auto-phasage des analyses RMN-1D (fonction *FT1D*)

Le phasage correspond à un traitement du signal post-acquisition pour qu'il soit positif et symétrique. Il s'agit donc de minimiser le pied négatif du signal par correction sur l'ordre 0 (correction de la phase sur le signal le plus intense et ceux proches de celui-ci) puis sur l'ordre 1 (signaux plus éloignés) Cette étape peut avoir un fort impact dans les cas d'étude de

mélange complexe où l'intérêt porte parfois sur la détection de signaux de faible intensité, et cela peut être d'autant plus critique s'ils sont proches de signaux très intenses. Cela peut également mener à des erreurs de calcul des maximums d'intensité. Le phasage de manière automatique représente une certaine difficulté dans l'étude d'extrait en raison des signaux intenses du solvant et du nombre important de signaux de faible intensité. Le phasage automatique a pu être réalisé en s'appuyant sur le programme NPK (*NMR Processing Kernel*)¹⁸² qui est aussi implémenté comme un *plug-in* dans SPIKE.

III. 4. c. Auto-zéro

L'auto-zéro est une fonction définie dans Plasmodesma. Il permet de calibrer le spectre à 0 ppm. Cela est possible grâce à la présence d'une référence interne, le tétraméthylsilane (TMS). Le programme détecte le signal le plus intense entre -0.1 et 0.1 ppm et le calibre à 0.

III. 4. d. Processing

Le *processing* des données RMN est adapté à chaque analyse RMN, il est possible pour les ^1H - ^1H COSY, ^1H - ^1H TOCSY, ^1H - ^{13}C HSQC, ^1H - ^{13}C HMBC, ^1H - δ DOSY. La transformée de Fourier appliquée aux données RMN est choisie en fonction de l'analyse.

Les expériences ^1H - δ DOSY ne requièrent pas le même système de *processing* que les autres RMN-2D car elles mesurent une constante de diffusion. La transformée de Fourier classique n'est pas utilisée pour le traitement de ces données, c'est une transformée de Laplace inverse. Le *processing* des ^1H - δ DOSY est possible grâce à l'utilisation de l'algorithme PALMA qui implémente une rapide analyse par transformée de Laplace. PALMA est un *plug-in* de Spike.

III. 4. e. Peak-picking

Une détection des signaux RMN est automatiquement réalisée par Plasmodesma. Cela est possible grâce au *plug-in* Peaks implémenté dans Spike. Un fichier est produit contenant la liste des signaux détectés par Plasmodesma sous le nom de « nom de l'analyse RMN_numéro de l'expérience_peaklist.csv ». Il contient un numéro attribué au signal, sa position (ppm) et son intensité. Une image des signaux détectés est également produite automatiquement en format pdf, donnant un aperçu des signaux du spectre.

III. 4. f. Bucketing

Bucketing-Plasmodesma est le script permettant le *bucketing*. Il est implémenté comme *plug-in* dans Spike et contient les fonctions de *bucketing* des données RMN 1D et 2D appelées *bucket1d* et *bucket2d*. Plasmodesma fait appel à ces deux fonctions pour réaliser le *bucketing*.

Pour chaque *bucket* produit, le point **maximal et minimal**, l'**aire et l'écart-type** sont calculés. Un fichier `csv` contenant la valeur du centre du *bucket* (ppm), l'aire cumulée, la valeur maximal et minimal, l'écart-type et la taille du *bucket* (ppm) est produit sous le nom de « nom de l'analyse RMN_numéro de l'expérience_bucketlist.csv ». Pour chaque analyse RMN *buckétée*, une image correspondante (pdf) est produite et permet de donner un aperçu du spectre *buckété*.

III. 4. g. Rapport

A la fin du *bucketing* deux fichiers `csv` sont créés. Le premier appelé « *report* » résume tous les paramètres d'acquisition des analyses RMN et le second nommé « *files* » résume le *processing*, par échantillon et analyse. Il contient le nombre de *buckets* ou de pics par analyse.

III. 5. Bucket Utilities

Une fois les données RMN *buckétées*, elles sont prêtes à être utilisées pour les analyses différentielles. Il faut déterminer les variables et calculs qui sont à réaliser pour produire la meilleure empreinte pharmacophorique. Un script appelé « *BucketUtilities* » contient plusieurs fonctions permettant le **pré-traitement** des spectres *buckétés*, les **importations** de ces données et l'**affichage** des résultats. Pour l'utiliser, il suffit de l'importer en amont de la page du notebook.

La totalité du *script* de *BucketUtilities* est disponible sur : <https://github.com/delsuc/plasmodesma> et la dernière version (*BucketUtilities_LM*) est disponible dans la partie annexe.

III. 5. a. Nettoyage des spectres buckétés

La fonction « *nettoie_mieux* » est utilisée pour le nettoyage des spectres sauf pour la régression linéaire des expérience $^1\text{H-}\delta$ DOSY où la fonction « *encore_mieux* » est utilisée.

```

def nettoie(ZZ, factor=2.0):
    " clean noise in matrix - hard thresholding"
    ZZr = ZZ.copy()
    thresh = factor*np.median(ZZ)
    print( thresh)
    ZZr[ZZ<thresh] = 1.0 # 1.0 allows to display log(Z) !
    return ZZr

def nettoie_mieux(ZZ, factor=2.0):
    " clean noise in matrix - hard thresholding columnwise"
    ZZr = ZZ.copy()
    for i in range(ZZ.shape[1]):
        iZZ = ZZ[:,i]
        thresh = factor*np.median(iZZ)
        ZZr[iZZ<thresh,i] = 1.0
    return ZZr

def nettoie_encore_mieux(ZZ, factor=2.0):
    " clean noise in matrix - soft thresholding columnwise"
    ZZr = ZZ.copy()
    for i in range(ZZ.shape[1]):
        iZZ = ZZ[:,i]
        thresh = factor*np.median(iZZ)
        ZZr[:,i] = np.where(iZZ<thresh, 1, iZZ-thresh+1)
    return ZZr

```

III. 5. b. Symétrisation des spectres bucketés homonucléaires

Il s'agit de la fonction pour supprimer les artefacts de type anti-diagonale. Elle est applicable seulement sur les spectres homonucléaires.

```

def symetrise(ZZ):
    " symetrisation of spectra - simple minimum operation"
    return np.minimum(ZZ, ZZ.T)

```

III. 5. c. Importation et affichage des spectres bucketés

Bucketutilities contient les fonctions d'importation et d'affichage des données RMN *bucketées*, c'est-à-dire les données contenues dans le fichier *bucketlist.csv*. Il contient notamment les fonctions :

- loadStd2D : Importation des valeurs des écarts-types des *buckets* d'expériences RMN-2D contenu dans le fichier *bucketlist*
- loadInt2D : Importation des valeurs des intensités des *buckets* d'expériences RMN-2D contenu dans le fichier *bucketlist*
- affiche : affichage des données chargées sous l'image d'un spectre
- affratio : Calcule et affiche le spectre issu du ratio de l'intensité ou du ratio des *buckets* d'expérience RMN

- `affsub` : Calcule et affiche le spectre issu de la soustraction de l'intensité ou du ratio des *buckets* d'expérience RMN.

III. 6. Analyse différentielle par régression linéaire

Le *script* de la régression linéaire n'est pas inclus dans *Bucket Utilities*, il est à écrire directement dans le Jupyter Notebook. Un exemple est donné sur : <https://github.com/delsuc/plasmodesma>.

Lorsque la concentration massique dans les échantillons de la molécule bioactive est connue ou lorsque que la valeur de l'activité biologique mesurée est une concentration, celles-ci sont utilisées pour réaliser la régression linéaire sur l'ensemble des données des expériences RMN.

Les données des intensités des *buckets* des cinq échantillons sont chargées :

```
BucketUtilities.NETMODE = 'mieux'
X = []
for i in range(1,6):
    tit=op.join(results, r"name{}/2D/cosydfph_10_bucketlist.csv".format(i))
    print (tit)
    Int = loadInt2D(tit, net=True, sym=True)
    tInt = Int[2].ravel() # [tmask]
    X.append(tInt)
X = np.array(X)
print ("X.shape", X.shape)
```

Puis les concentrations de la molécule active dans les échantillons (valeurs doivent être comprises entre 0 et 1) :

```
Y = np.array([0, 0.15, 0.24, 0.32, 2])/2 # divide by 2 to be in 0..1 range
```

Chargement du modèle de régression linéaire de la bibliothèque Scikit-learn, définition de la valeur seuil et affichage :

```
import sklearn
from sklearn import linear_model
reg = linear_model.LinearRegression()
reg.fit(X,Y)
```

```
Threshold = 3E-14
coefmask = (reg.coef_) > Threshold
coef = reg.coef_[coefmask]
tcoefmask = np.where(coefmask.ravel() > 0) [0]
```

```
N = len(Im1[2].ravel())
m = np.zeros(N)
m[tcoefmask] = 1.0
m = m.reshape(Im2[2].shape)
affiche(Im1[0],Im1[1],m, scale=0.1)
```

Si modèle RFE :

```
from sklearn.feature_selection import RFE
NFeatures = 300 # we want a small number
estimator = linear_model.LinearRegression()
selector = RFE(estimator, step=0.5, n_features_to_select=NFeatures)
selector = selector.fit(X, Y)
```

```
N = len(Im2[2].ravel())
m = np.zeros(N)
m[selector.support_] = 1.0
m = m.reshape(Im2[2].shape)
affiche(Im1[0],Im1[1],m, scale=0.1)
```

Dans le cas où l'activité biologique s'exprime en pourcentage (pourcentage d'inhibition de croissance par exemple), il faut ajouter une étape supplémentaire permettant de passer d'une valeur de pourcentage à une concentration. Les activités biologiques ou les courbes dose-réponse sont des courbes sigmoïdes et il faut donc passer par leur inverse. Cette étape s'ajoute après le chargement des intensités des échantillons et des pourcentages d'inhibition.

```
def ty(x, A=1):
    "A defines the slope of the sigmoid"
    val = 0.5 + A*np.arctanh(1.5*x-1)
    val = np.minimum(val,1.0)
    val = np.maximum(val,0.0)
    return val
x = np.linspace(0.0,1.0,100)
Y = ty(Y)
Y
```

III. 7. Filtration des empreintes par les empreintes ¹H-δ DOSY

Il est possible de filtrer les empreintes issues d'analyses hétéronucléaires par les analyses ¹H-δ DOSY. Pour cela, une régression linéaire entre les intensités des *buckets* des expériences ¹H-δ DOSY et de l'activité biologique est réalisée. Le *script* n'est pas inclus dans *BucketUtilities*.

Chargement des intensités des *buckets* des expériences $^1\text{H-}\delta$ DOSY :

```

BucketUtilities_LM.NETMODE='encore'
X = []
for i in range(1,6):
    tit = op.join(results,r"name{}/2D/hsqcetgp_40_bucketlist.csv".format(i)
    print(tit)
    Int = loadInt2D(tit, net=True, sym=False)
    tInt = Int[2].ravel() # [tmask]
    X.append(tInt)
X = np.array(X)
print ("X.shape", X.shape, Int[2].shape)

```

Chargement de l'activité biologique :

```

Y = np.array([0.0, 0.2, 0.3, 0.4, 2.7])/10

```

Chargement des modèles de régression :

```

import sklearn
from sklearn import linear_model
from sklearn.feature_selection import RFE

```

Comparaison de la régression linéaire et de RFE :

```

# LinearReg
Threshold = 3E-12 # keep only points above the threshold
reg = linear_model.LinearRegression()
reg.fit(X,Y)
coefmask = (reg.coef_)> Threshold
coef = reg.coef_[coefmask]
tcoefmask = np.where(coefmask.ravel()>0)[0]
N = len(Int[2].ravel())
dosy = np.zeros(N)
dosy[tcoefmask] = 1.0
dosy = dosy.reshape(Int[2].shape)

# RFE
NFeatures = 150 # keep only this number of points
estimator = linear_model.LinearRegression()
selector = RFE(estimator, step=0.5, n_features_to_select=NFeatures)
selector = selector.fit(X, Y)
dosyrfe = np.zeros(N)
dosyrfe[selector.support_] = 1.0
dosyrfe = dosyrfe.reshape(Int[2].shape)

```

Affichage des ^1H - δ DOSY en projection 1D :

```
dosyl1D = dosy.sum(axis=0)
dosyrfel1D = dosyrfel.sum(axis=0)
f1, (ax1,ax2) = plt.subplots(nrows=2, sharex=True, figsize=(12,6))
ax1.set_title('Régression linéaire')
ax1.plot(Int[0][0], dosyl1D)
ticks = np.arange(0, 7, 0.5)
ax2.set_title('Recursive Feature Elimination')
ax2.plot(Int[0][0], dosyrfel1D)
ax2.set_xticks(ticks)
ax2.set_xlim(xmin=-0.5, xmax=7)
ax2.invert_xaxis()
f1.savefig("FigDOSY1D.pdf")
```

Chargement de intensités des *buckets* des expériences ^1H - ^{13}C HSQC :

```
BucketUtilities.NETMODE = 'encore'
# load data
H = []
for i in range(1,6):
    tit=op.join(results, r"name{}/2D/hsqcetgp_40_bucketlist.csv".format(i))
    print (tit)
    Int = loadInt2D(tit, net=True, sym=False)
    tInt = Int[2].ravel() # [tmask]
    H.append(tInt)
H = np.array(H)
print (H.shape, Int[2].shape)
```

Application de la régression linéaire :

```
reg = linear_model.LinearRegression()
reg.fit(H,Y)

coefmask = (reg.coef_)>1E-13
coef = reg.coef_[coefmask]
tcoefmask = np.where(coefmask.ravel()>0)[0]
len(coefmask),len(tcoefmask),len(coef)
N = len(Int[2].ravel())
mreg = np.zeros(N)
mreg[tcoefmask] = 1.0
#m[coef] = 1.0
mreg = mreg.reshape(Int[2].shape)
f1, ax = plt.subplots( figsize=(6,5))
ax.set_title('Empreinte HSQC')
affiche(Int[0],Int[1],mreg, scale=0.01, ax=ax)#, cmap=cm.spring)
```

Filtration de l'empreinte ^1H - ^{13}C HSQC par l'empreinte ^1H - δ DOSY :

```
hsqc_filtered = mreg*dosyrfel1D
f1, (ax1,ax2) = plt.subplots(ncols=2, sharex=True, sharey=True,
figsize=(14,6))
ax1.set_title('Régression linéaire')
affiche(Int[0], Int[1], mreg, scale=0.01, ax=ax1)
ax2.set_title('Empreinte HSQC filtré par le spectre RFE-DOSY')
affiche(Int[0], Int[1], hsqc_filtered, scale=1, reverse=False, ax=ax2)
```

III. 8. Méthode couplée HPLC-DAD-MS-SPE-RMN

Le travail est réalisé sur une chaîne Agilent 1200 équipée d'une colonne analytique C18 ProntoSIL (4mm x 120 mm, 5µm, BISCHOFF Chromatography) et d'un détecteur à barrette de diode (DAD). Le spectromètre de masse couplé est spectromètre à trappe ionique HCT Ultra (Bruker), équipé d'une source d'ionisation par electrospray (ESI). Le système est possiblement dérivé par une unité Prospect-2 (Spark Holland) équipée de cartouche SPE HLB (*Hydrophilic-Lipophilic Balance*) Oasis (2 x 10 mm, 30 µm, Waters). Ces cartouches sont désorbées avec un solvant deutéré et le produit est conduit à la sonde capillaire RMN LC-SEI 3 mm à gradient Z d'un volume de 30 µL. Le spectromètre RMN est un Bruker Avance III 500 MHz.

IV. Essais biologiques *anti-Plasmodium*

IV. 1. Culture de *P. falciparum*

La culture du parasite a été réalisée dans les locaux de l'Institut de Parasitologie et de Pathologie Tropicale de Strasbourg selon la méthode de *Trager et Jensen*, et adapté localement^{119,183}. *P. falciparum* est un parasite intracellulaire obligatoire, il a besoin pour sa survie d'être en présence d'hématie, d'un milieu de culture adapté fournissant tous les nutriments nécessaires, d'une température physiologique de 37 °C et d'un milieu en anaérobiose (atmosphère pauvre en oxygène). Celui-ci est obtenu grâce un système appelé « candle jar », qui est un dessiccateur étanche dans lequel une bougie est allumée afin de consommer l'oxygène présent. Toutes les manipulations concernant la culture cellulaire sont réalisées dans des conditions stériles sous hotte à flux laminaire.

IV. 1. a. Souche plasmodiale

Les extraits et fractions ont été testés sur la souche de *P. falciparum* 7G8. Il s'agit d'une souche multirésistante (chloroquine, pyriméthamine et cycloguanil), d'origine brésilienne¹⁸⁴.

IV. 1. b. Préparation des hématies

L'emploi d'hématies est nécessaire pour recréer artificiellement le cycle intra-érythrocytaire du parasite. Les hématies sont issues du sang de donneurs O+ n'ayant jamais été au contact du parasite (Etablissement Français du Sang, Strasbourg). Le contenu d'une poche de culot globulaire est transféré dans une flasque F175 et celui-ci est complété avec du

milieu de conservation SAGM (*Saline Adenine Glucose Medium*) à raison de 1 volume de SAGM pour 2.5 volumes d'hématies. Les hématies sont ainsi conservées 1 mois à + 4 °C.

Avant utilisation, les hématies sont lavées avec du RPMI (*Roswell Park Memorial Institute medium*, milieu de culture) et l'hématocrite (pourcentage d'hématies par rapport au volume total) est ajustée à 50 %. Les échantillons ainsi reconstitués peuvent être conservés 5 jours à + 4 °C.

IV. 1. c. « Milieu de culture Malaria » MCM

Le milieu est constitué d'un milieu de base RPMI 1640 additionné d'hypoxanthine (1 mg/l ; nécessaire à la croissance du parasite qui est incapable de réaliser la synthèse de novo de ce précurseur des bases puriques), de L-glutamine (1 mM ; source énergétique pour la glycolyse anaérobie), de pyruvate de sodium (110 mg/l) et est tamponné à 7,4 (pH physiologique) par un tampon HEPES (10 mM). Un antibiotique, la gentamicine, y est ajouté afin d'éviter une éventuelle contamination bactérienne (20 mg/l). Enfin une fraction de sérum de veau (α -calf) est ajoutée à hauteur de 10 %. Ce milieu est appelé « Milieu de Culture Malaria » ou MCM complet.

IV. 1. d. Décongélation de la souche et mise en culture

Les érythrocytes infectés par *Plasmodium* sont conservés dans l'azote liquide dans une solution qui permet de conserver l'intégrité des cellules sanguines qui les contiennent (solution d'Alsever). Afin de les remettre en culture, ces souches doivent être décongelées de façon rapide (1 à 2 minutes dans un bain-marie à 37 °C). Le volume de sang décongelé (V) est ensuite additionné goutte à goutte de 0,1 x V de solution de NaCl à 12 %. Ce mélange est incubé 5 minutes à température ambiante. Une deuxième solution moins concentrée de NaCl (1,6 %) est ensuite ajoutée goutte à goutte à raison de 10 x V. Cette étape permet d'éviter un choc osmotique qui entrainerait la lyse des globules rouges et donc la mort du parasite. Le tube est ensuite centrifugé pendant 5 minutes à 425 g à 20 °C et le surnageant est éliminé. Les hématies du culot sont remises en suspension dans 10 x V de MCM doucement et goutte à goutte puis centrifugées à nouveau. Le culot est alors repris dans 5 ml de MCM et cette suspension parasitaire est transférée dans une flasque de culture F25. Des hématies fraîches et du milieu de culture sont ensuite ajoutés de manière à obtenir un hématocrite de 2 % (0,2 ml d'hématies dans 10 ml de MCM). La flasque est mise à incuber en « candle jar » placé dans un incubateur à 37 °C. La parasitémie (pourcentage d'hématies infectées par rapport au

nombre total de globules rouges) est évaluée régulièrement par la lecture d'un frottis sanguin afin d'apprécier la croissance de la souche.

IV. 1. e. Frottis sanguin

Le frottis sanguin est une couche monocellulaire de cellules sanguine qui permet entre-autres de calculer la parasitémie de *Plasmodium*. Il est réalisé par étalement d'une goutte de sang sur une lame à l'aide d'une lamelle. Le frottis doit ensuite être séché rapidement afin d'éviter la formation d'échinocytes (globules rouges déformés se présentant sous une forme épineuse). Il est ensuite fixé au méthanol et coloré au Giemsa 10 % pendant 20 à 30 min. Cette technique de coloration permet de donner une couleur caractéristique au parasite (noyau rose et cytoplasme bleu). La lame est ensuite lue au microscope à l'objectif à immersion (X 100). Le nombre d'hématies parasitées est alors compté sur 30 champs homogènes et rapporté au nombre total d'hématies comptées dans ces champs.

IV. 1. f. Division

En fonction de la parasitémie et des stades présents, une alternative se présente :

- Entre 1 et 4 % de parasitémie dont majoritairement des trophozoïtes, le milieu est changé et ramené à **un hématocrite de 1 %**.
- Si la forme schizonte est majoritaire et la parasitémie supérieure à 2 %, ou que la parasitémie est supérieure à 4%, il faut réaliser une division de la culture afin de la ramener à une **parasitémie de 0,5 à 1 %**.

IV. 1. g. Congélation

Les hématies doivent être parasitées majoritairement par des formes rings présentes à une parasitémie d'environ 3 à 4 %. Le culot d'hématies parasitées est repris volume à volume dans une solution de congélation (solution d'Alsever) contenant du glycérol (28 % v/v), du sorbitol (3 % m/v) et du NaCl (0,65 % m/v), préalablement filtrée sur une membrane de 0,22 µm.

IV. 2. Essais biologiques

IV. 2. a. Screening des extraits

Les extraits sont premièrement testés à 100 et 10 µg/ml en duplicat. Une solution mère à 20 mg/ml est préparée dans du DMSO pur (composés apolaires ou moyennement polaire) ou directement dans du MCM à 1 % de DMSO sans α -calf (composés polaire) afin de

permettre une solubilisation optimale des composés. Celle-ci est ensuite diluée au 1/100^{ème} avec du MCM sans sérum de façon à obtenir une solution à 200 µg/ml présentant une concentration finale de DMSO de 1 %. Cette première solution fille est diluée une seconde fois au 1/10^{ème} avec du MCM à 1 % de DMSO pour obtenir une solution à 20 µg/ml. L'ajout de 100 µl/puits de suspension parasitaire diluée au ½ ces deux solutions filles et ramène à une concentration allant de 100 µg/ml et 10 µg/ml dans 0,5 % de DMSO. Chaque concentration est testée en triplicat.

IV. 2. b. Réalisation des gammes de concentration des composés à tester

Les gammes de concentrations sont réalisées dans une solution de MCM à 1% de DMSO (diméthyl sulfoxyde). Une solution mère des extraits ou composés à tester à 20 mg/ml est préparée dans du DMSO pur (composés apolaires ou moyennement polaire) ou directement dans du MCM à 1% de DMSO sans α -calf (composés polaire) afin de permettre une solubilisation optimale des composés. Celle-ci est ensuite diluée au 1/100^{ème} avec du MCM sans sérum de façon à obtenir une solution à 200 µg/ml présentant une concentration finale de DMSO de 1 %.

Une dilution en cascade au 1/3 de cette solution à 200 µg/ml est réalisée directement dans la plaque 96 puits : 50 µl sont transférés en cascade du premier puits (contenant 150 µl de solution à 200 µg/ml) dans les puits suivants contenant 100 µl de MCM sans sérum à 1 % de DMSO et ce jusqu'au puits G où les 50 µL en excès sont jetés ce qui ramène tous les puits à un volume final de 100 µl. Les puits de la ligne H servent de puits contrôle. Cette dilution en cascade permet d'obtenir une gamme de 7 concentrations allant de 200 à 0,276 µg/ml. L'ajout de 100 µl/puits de suspension parasitaire diluée au ½ les concentrations et ramène la gamme à une concentration allant de 100 µg/ml à 0,138 µg/ml dans 0,5 % de DMSO. Les gammes sont adaptées si besoin afin que la CI50 soit proche de la concentration médiane de la gamme. La chloroquine, molécule de référence, sert de contrôle positif sur une gamme de concentration allant de 500 ng/ml à 0,69 ng/ml. Chaque concentration est testée en triplicata.

IV. 2. c. Réalisation de la suspension parasitaire

La culture de *P. falciparum* est ajustée à une parasitémie de 1 % et 4 % d'hématocrite dans du MCM complet en fonction du volume nécessaire pour la réalisation des tests. Une fois dilué au ½ dans le puits (volume de solution du composé à évaluer), l'hématocrite aura une valeur de 2 %. Le pourcentage de la forme ring du parasite doit être supérieur à 50 %.

Dans chaque plaque deux contrôles sont ajoutés :

- Un puits de suspension globulaire à 2 % d'hématocrite dans 0,5 % de DMSO (contrôle négatif de croissance)
- Trois puits de suspension parasitaire dans 0,5 % de DMSO (contrôles positifs de croissance).

Ces contrôles serviront à déterminer le pourcentage d'inhibition de chaque puits, calculé comme suit :

$$\text{Pourcentage d'inhibition (\%)} = 100 - \left(\frac{DO_{\text{moyenne}_{\text{produit}}} - DO_{\text{moyenne}_{\text{GRS}}}}{DO_{\text{moyenne}_{\text{GRP}}}} \times 100 \right)$$

Les plaques sont mises à incuber pendant 96 heures à 37 °C en candle jar.

IV. 2. d. Quantification de la Plasmodium Lactate DeHydrogenase (pLDH)

La production de pLDH par le parasite est un reflet de sa **viabilité**. Cette enzyme est quantifiée par une méthode immuno-enzymatique à l'aide du kit ELISA Malaria Antigen test (ApDia®, Turnhout, Belgique). Dans ce test, c'est la souche multi-résistante 7G8 qui est employée (Figure 101).

Après 96 heures d'incubation, 5 µl de surnageant des puits sont transférés dans les puits correspondants des plaques 96 puits du kit ELISA, préalablement remplis avec 100 µl de tampon de lyse. Le fond des puits est recouvert d'anticorps monoclonaux anti-pLDH pan-spécifiques (qui reconnaissent toutes les espèces de *Plasmodium*). Les contrôles du test ELISA sont également ajoutés dans quatre puits :

- un puits contenant le contrôle positif (sérum humain contenant une quantité détectable de protéine recombinante pLDH)
- trois puits contenant le contrôle négatif (sérum humain contenant une quantité non détectable de recombinant pLDH)

La plaque est alors mise à incuber pendant 60 minutes à 37 °C sous agitation. A l'issue de l'incubation, les puits sont lavés 3 fois à l'aide d'une solution de lavage à base de tampon phosphate puis 100 µl d'une solution d'anticorps monoclonaux anti-pLDH (pan-spécifique) et biotinylés sont ajoutés dans chaque puits. La plaque est incubée 30 minutes à 37 °C avant d'être soumise à une nouvelle opération de lavage. 100 µl d'une solution de streptavidine-horseradish peroxydase sont alors ajoutés dans chaque puits et la plaque est mise à incuber 15 minutes à 37 °C. La plaque est ensuite lavée à nouveau et 100 µl de solution chromogène « 2 en 1 » [peroxyde d'hydrogène (substrat) + 3,3',5,5'-tétraméthylbenzidine, TMB, molécule chromogène] sont ajoutés. Une dernière incubation de 15 minutes à 37 °C, à l'abri de la lumière permet l'oxydation du TMB (incolore) par l'O₂ produit par la catalyse du peroxyde d'hydrogène, en produit de couleur bleue (3,3',5,5'-tetramethylbenzidine diimine). La réaction est ensuite stoppée par adjonction de 50 µl d'une solution d'acide sulfurique (0,5 M) qui induit une coloration jaune du produit et l'absorbance est lue dans les 15 minutes au spectrophotomètre à 450 nm. Chaque expérimentation est renouvelée une fois, voire deux si nécessaire (n= 2 à 3).

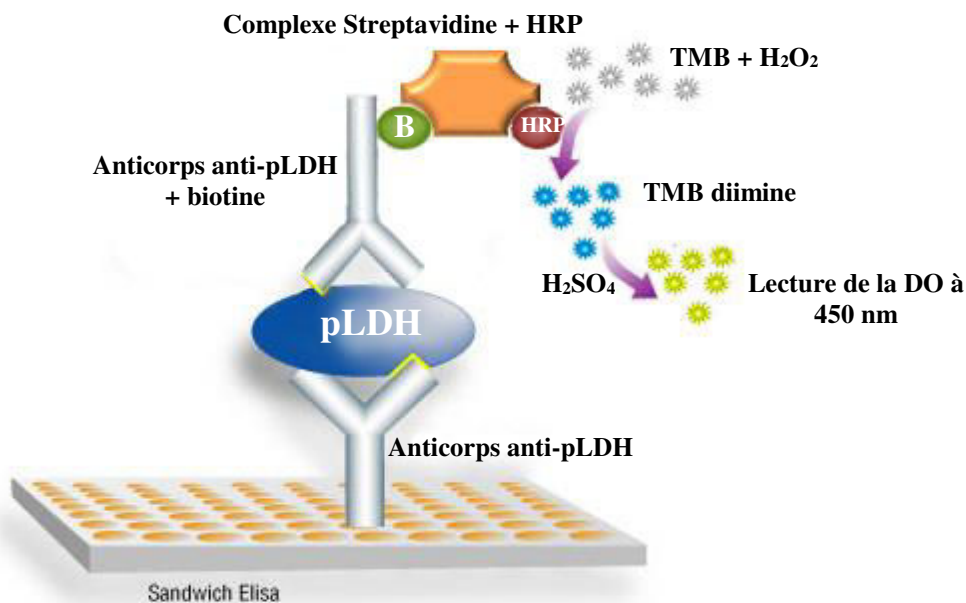


Figure 101 : Test d'activité anti-*Plasmodium* par la mesure de la sécrétion de pLDH via le kit ELISA Antigen ApDia®. Après 72h d'incubation de la suspension parasitaire avec l'échantillon à tester, le surnageant contenant la pLDH est prélevé et déposé dans les puits recouverts de l'anticorps anti-pLDH. Un anticorps anti-pLDH secondaire couplé à la biotine, puis une solution de streptavidine-horseradish peroxydase sont ajoutés. Une solution chromogène de 3,3',5,5'-tétraméthylbenzidine (TMB) et de peroxyde d'hydrogène est ajoutée. Le TMB va être oxydé et prendre une couleur bleue. La réaction est stoppée par ajout d'une solution d'acide sulfurique. L'absorbance est lue au spectrophotomètre à 450 nm.

IV. 2. e. Détermination de la CI_{50}

La détermination de la concentration inhibitrice 50 se réalise par une méthode graphique à partir des courbes effet-dose qui sont des sinusoïdales (l'activité est très faible à une certaine dose puis augmente rapidement avec celle-ci pour atteindre un plateau d'inhibition). La CI_{50} est la concentration qui permet d'inhiber 50 % de la croissance du parasite. Elle est calculée précisément à l'aide de GraphPad Prism 6 qui calcule l'interpolation linéaire entre les deux concentrations de drogues encadrant le 50 %.

IV. 2. f. Test d'inhibition de l'incorporation de l'hypoxanthine tritiée (collab. M. Kaiser, Swiss TPH)

Le test d'inhibition de l'incorporation de l' 3H -hypoxanthine est une technique radioisotopique décrite initialement par Desjardins *et al.* en 1979¹⁸⁵. Ce test, contrairement au test pLDH permet de mettre en évidence une inhibition de la **division plasmodiale** étant donné que l' 3H -hypoxanthine est un précurseur de base purique indispensable à la synthèse nucléique du parasite. Ainsi l' 3H -hypoxanthine est transformée en 3H -adénine ou 3H -guanine qui sont incorporées à l'ADN ou l'ARN du parasite. Le test consiste, comme pour le test pLDH, à mettre en présence une gamme de concentration du composé à tester avec une suspension parasitaire. Dans ce test, c'est la souche de *P. falciparum* K1 qui est employée. Après 48 heures d'incubation à 37 °C et en atmosphère anaérobie, l' 3H -hypoxanthine est additionnée à chaque puits et l'incubation est poursuivie pendant encore 24 heures dans les mêmes conditions. L'incorporation de radioactivité est évaluée après lyse des hématies à l'aide d'un scintillateur (coups par minute, cpm). La substance de référence utilisée est la chloroquine. Ce test est réalisé en collaboration avec le Dr M. Kaiser du *Swiss Tropical and Public Health Institute* de Bâle, plateforme de référence mondiale pour l'évaluation de la chimiosensibilité de parasites protozoaires, qui dispose des autorisations nécessaires pour la détention et l'utilisation de produits radioactifs. Chaque concentration est testée en triplicat (7 concentrations, dilution au 1/3 de 90 à 0,123 µg/ml). Chaque expérimentation est renouvelée 1 fois (n= 2).

IV. 2. g. Intérêt de combiner les deux techniques

L'intérêt de mesurer l'activité anti-*Plasmodium in vitro* en utilisant parallèlement les deux techniques est que l'on mesure deux paramètres différents. Or, les tests antiplasmodiaux *in vitro* discriminent entre les différents composés criblés en fonction de leur mécanisme d'action antiplasmodiale¹⁸⁶. Le test hypoxanthine permet de mesurer l'impact du composé sur la division parasitaire tandis que le test ELISA pLDH permet de déterminer l'impact sur la

survie du parasite. La pLDH étant une enzyme présentant une demi-vie longue par rapport au temps d'incubation du parasite avec les composés (demi-vie d'environ 30 heures vs test de 72 heures), les composés entraînant une mort lente du parasite pourraient ne pas être détectés car la CI_{50} mesurée serait plus élevée que la CI_{50} réelle¹⁸⁶. De plus, le milieu de culture utilisé à Strasbourg ayant été développé par un partenaire industriel pour augmenter la croissance parasitaire et obtenir une parasitémie élevée, les valeurs de CI_{50} peuvent également être augmentées. Il était donc important de confirmer les résultats obtenus à Strasbourg avec le test de référence utilisé à Bâle. En effet, dans nos mains, les activités mesurées par ces deux techniques ne conduisent pas toujours aux mêmes résultats.

V. Inhibition de la voie de synthèse des isoprénoides (MEP) ou de la prénylation des protéines

En collaboration avec Andréa Hemmerlin de l'équipe *Metabolic Network*, de l'Institut de Biologie Moléculaire des Plantes (IBMP), quelques extraits ou fractions ont pu être évalués sur un modèle de *screening* phénotypique permettant de mettre en évidence des inhibiteurs de la voie du MEP ou de l'enzyme responsable de la prénylation des protéines (prényl-transférase).

3 mL de suspension cellulaire de cellules transformées de tabac BY-2 sont mis en présence des extraits algaux à une concentration de 40 $\mu\text{g/mL}$. Ces cellules expriment la calmoduline de riz couplée à la GFP. Cette protéine a été choisie car elle présente un motif d'adressage nucléaire en plus d'un site de prénylation. Le gène codant pour la calmoduline-GFP est sous le contrôle d'un promoteur inductible à la dexaméthasone. Un puit de contrôle négatif (cellules non traitées), et un contrôle positif (Fosmidomycine, inhibiteur de la déoxyxylulose phosphate réductase de la voie du MEP) sont déposés comme témoins. Les cellules sont mises à incuber pendant 3 heures à température ambiante, puis 1.5 μL de dexaméthasone sont ajoutés (15 μM) ainsi que les composés à tester. Ensuite, les cellules sont mises à incuber pendant 18 heures, et une lecture est effectuée au microscope confocal à une longueur d'onde d'excitation de 446 nm pour observer le phénotype induit.

En effet, dans le cas où la protéine est prénylée, elle est exportée à la membrane plasmique, dans le cas contraire (inhibition de la voie du MEP ou de la prényl-transférase) la protéine est adressée au noyau.

Les cellules sont comptées pour chaque phénotype, membranes, membranes/noyaux, et noyaux par rapport au nombre de cellules totales présentes. Les images sont ensuite acquises avec le logiciel ZEN V2.2.



Bibliographie

Bibliographie

1. WHO | World malaria report 2017. *WHO* Disponible sur : <http://www.who.int/malaria/publications/world-malaria-report-2017/en/>. (Consulté le : 31 July 2018)
2. Prevention, C.-C. for D. C. and. CDC - Malaria - About Malaria. Disponible sur : <https://www.cdc.gov/malaria/about/index.html>. (Consulté le : 13th April 2017)
3. Hoppe, H. C., van Schalkwyk, D. A., Wiehart, U. I. M., Meredith, S. A., Egan, J. & Weber, B. W. Antimalarial quinolines and artemisinin inhibit endocytosis in *Plasmodium falciparum*. *Antimicrob Agents Chemother* **48**, 2370–2378 (2004).
4. Dubar, F., Khalife, J., Brocard, J., Dive, D. & Biot, C. Ferroquine, an ingenious antimalarial drug – Thoughts on the mechanism of action. *Molecules* **13**, 2900–2907 (2008).
5. Ebstie, Y. A., Abay, S. M., Tadesse, W. T. & Ejigu, D. A. Tafenoquine and its potential in the treatment and relapse prevention of *Plasmodium vivax* malaria: the evidence to date. *Drug Des Devel Ther* **10**, 2387–2399 (2016).
6. Noedl, H., Se, Y., Schaefer, K., Smith, B. L., Socheat, D., Fukuda, M. M. & Artemisinin resistance in Cambodia 1 (ARC1) Study Consortium. Evidence of artemisinin-resistant malaria in western Cambodia. *N. Engl. J. Med.* **359**, 2619–2620 (2008).
7. Wells, T. N. Natural products as starting points for future anti-malarial therapies: going back to our roots? *Malar J* **10**, S3 (2011).
8. Carvalho, L. H., Rocha, E. M., Raslan, D. S., Oliveira, A. B. & Krettli, A. U. *In vitro* activity of natural and synthetic naphthoquinones against erythrocytic stages of *Plasmodium falciparum*. *Braz. J. Med. Biol. Res.* **21**, 485–487 (1988).
9. Wright, C. W., Phillipson, J. D., Awe, S. O., Kirby, G. C., Warhurst, D. C., Quetin-Leclercq, J. & Angenot, L. Antimalarial activity of cryptolepine and some other anhydronium bases. *Phytotherapy Research* **10**, 361–363 (1996).
10. Mesia, K., Cimanga, R. K., Dhooghe, L., Cos, P., Apers, S., Totté, J., Tona, G. L., Pieters, L., Vlietinck, A. J. & Maes, L. Antimalarial activity and toxicity evaluation of a quantified *Nauclea pobeguini* extract. *J Ethnopharmacol* **131**, 10–16 (2010).
11. Graz, B., Willcox, M. L., Diakite, C., Falquet, J., Dackuo, F., Sidibe, O., Giani, S. & Diallo, D. *Argemone mexicana* decoction versus artesunate-amodiaquine for the management of malaria in Mali: policy and public-health implications. *Trans. R. Soc. Trop. Med. Hyg.* **104**, 33–41 (2010).
12. MacKinnon, S., Durst, T., Arnason, J. T., Angerhofer, C., Pezzuto, J., Sanchez-Vindas, P. E., Poveda, L. J. & Gbeassor, M. Antimalarial activity of tropical Meliaceae extracts and gedunin derivatives. *J. Nat. Prod.* **60**, 336–341 (1997).

13. Compound for use in the treatment of parasites mediated diseases European patent EP18305121.8. *ResearchGate* Disponible sur : https://www.researchgate.net/publication/323152755_Compound_for_use_in_the_treatment_of_parasites_mediated_diseases_European_patent_EP183051218. (Consulté le : 8 Aout 2018)
14. Bero, J., Frédérick, M. & Quetin-Leclercq, J. Antimalarial compounds isolated from plants used in traditional medicine. *J. Pharm. Pharmacol.* **61**, 1401–1433 (2009).
15. Rukunga, G. M., Muregi, F. W., Omar, S. A., Gathirwa, J. W., Muthaura, C. N., Peter, M. G., Heydenreich, M. & Mungai, G. M. Anti-plasmodial activity of the extracts and two sesquiterpenes from *Cyperus articulatus*. *Fitoterapia* **79**, 188–190 (2008).
16. Vongvanich, N., Kittakoop, P., Charoenchai, P., Intamas, S., Sriklung, K. & Thebtaranonth, Y. Antiplasmodial, antimycobacterial, and cytotoxic principles from *Camchaya calcarea*. *Planta Med.* **72**, 1427–1430 (2006).
17. Pillay, P., Vleggaar, R., Maharaj, V. J., Smith, P. J. & Lategan, C. A. Isolation and identification of antiplasmodial sesquiterpene lactones from *Oncosiphon piluliferum*. *Journal of Ethnopharmacology* **112**, 71–76 (2007).
18. Okasaka, M., Takaishi, Y., Kashiwada, Y., Kodzhimatov, O. K., Ashurmetov, O., Lin, A. J., Consentino, L. M. & Lee, K. H. Terpenoids from *Juniperus polycarpus* var. *seravschanica*. *Phytochemistry* **67**, 2635–2640 (2006).
19. Jullian, V., Bonduelle, C., Valentin, A., Acebey, L., Duigou, A.-G., Prévost, M.-F. & Sauvain, M. New clerodane diterpenoids from *Laetia procera* (Poepp.) Eichler (Flacourtiaceae), with antiplasmodial and antileishmanial activities. *Bioorg. Med. Chem. Lett.* **15**, 5065–5070 (2005).
20. Thiem, D. A., Sneden, A. T., Khan, S. I. & Tekwani, B. L. Bisnortriterpenes from *Salacia madagascariensis*. *J. Nat. Prod.* **68**, 251–254 (2005).
21. Lee, S.-E., Kim, M.-R., Kim, J.-H., Takeoka, G. R., Kim, T.-W. & Park, B.-S. Antimalarial activity of anthothocol derived from *Khaya anthotheca* (Meliaceae). *Phytomedicine* **15**, 533–535 (2008).
22. Linn, T. Z., Awale, S., Tezuka, Y., Banskota, A. H., Kalauni, S. K., Attamimi, F., Ueda, J.-Y., Asih, P. B. S., Syafruddin, D., Tanaka, K. & Kadota, S. Cassane- and norcassane-type diterpenes from *Caesalpinia crista* of Indonesia and their antimalarial activity against the growth of *Plasmodium falciparum*. *J. Nat. Prod.* **68**, 706–710 (2005).
23. Andrade-Neto, V. F. de *et al.* *In vitro* inhibition of *Plasmodium falciparum* by substances isolated from Amazonian antimalarial plants. *Memórias do Instituto Oswaldo Cruz* **102**, 359–366 (2007).
24. McFadden, G. I. & Waller, R. F. Plastids in parasites of humans. *Bioessays* **19**, 1033–1040 (1997).
25. Okamoto, N. & McFadden, G. I. The mother of all parasites. *Future Microbiology* **3**, 391–395 (2008).
26. Moore, R. B., Oborník, M., Janouskovec, J., Chrudimský, T., Vancová, M., Green, D. H., Wright, S. W., Davies, N. W., Bolch, C. J. S., Heimann, K., Slapeta, J., Hoegh-

- Guldberg, O., Logsdon, J. M. & Carter, D. A. A photosynthetic alveolate closely related to apicomplexan parasites. *Nature* **451**, 959–963 (2008).
27. Low, L. M., Staniscic, D. I. & Good, M. F. Exploiting the apicoplast: apicoplast-targeting drugs and malaria vaccine development. *Microbes Infect.* (2017).
28. Ralph, S. A., Dooren, G. G. van, Waller, R. F., Crawford, M. J., Fraunholz, M. J., Foth, B. J., Tonkin, C. J., Roos, D. S. & McFadden, G. I. Tropical infectious diseases: Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nature Reviews Microbiology* **2**, 203–216 (2004).
29. Dahl, E. L. & Rosenthal, P. J. Multiple antibiotics exert delayed effects against the *Plasmodium falciparum* apicoplast. *Antimicrob. Agents Chemother.* **51**, 3485–3490 (2007).
30. *Treatment and Prevention of Malaria: Antimalarial Drug Chemistry, Action and Use.* (Springer Basel, 2012).
31. The Heme Biosynthesis Pathway Is Essential for *Plasmodium falciparum* Development in mosquito stage but not in blood stages. Disponible sur : <http://www.jbc.org/content/289/50/34827.full>. (Consulté le : 1 Aout 2018)
32. Nagaraj, V. A., Arumugam, R., Prasad, D., Rangarajan, P. N. & Padmanaban, G. Protoporphyrinogen IX oxidase from *Plasmodium falciparum* is anaerobic and is localized to the mitochondrion. *Mol. Biochem. Parasitol.* **174**, 44–52 (2010).
33. Gisselberg, J. E., Dellibovi-Ragheb, T. A., Matthews, K. A., Bosch, G. & Prigge, S. T. The iron-sulfur cluster synthesis pathway is required for apicoplast maintenance in malaria parasites. *PLoS Pathog* **9**, (2013).
34. Charan, M., Singh, N., Kumar, B., Srivastava, K., Siddiqi, M. I. & Habib, S. Sulfur mobilization for Fe-S cluster assembly by the essential SUF pathway in the *Plasmodium falciparum* apicoplast and its inhibition. *Antimicrob. Agents Chemother.* **58**, 3389–3398 (2014).
35. Guggisberg, A. M., Anthor, R. E. & Odom, A. R. Isoprenoid biosynthesis in *Plasmodium falciparum*. *Eukaryotic Cell* **13**, 1348–1359 (2014).
36. Wiesner, J., Ziemann, C., Hintz, M., Reichenberg, A., Ortmann, R., Schlitzer, M., Fuhst, R., Timmesfeld, N., Vilcinskas, A. & Jomaa, H. FR-900098, an antimalarial development candidate that inhibits the non-mevalonate isoprenoid biosynthesis pathway, shows no evidence of acute toxicity and genotoxicity. *Virulence* **7**, 718–728 (2016).
37. Goulart, H. R., Kimura, E. A., Peres, V. J., Couto, A. S., Duarte, F. A. A. & Katzin, A. M. Terpenes arrest parasite development and inhibit biosynthesis of isoprenoids in *Plasmodium falciparum*. *Antimicrob. Agents Chemother.* **48**, 2502–2509 (2004).
38. No, J. H., de Macedo Dossin, F., Zhang, Y., Liu, Y.-L., Zhu, W., Feng, X., Yoo, J. A., Lee, E., Wang, K., Hui, R., Freitas-Junior, L. H. & Oldfield, E. Lipophilic analogs of zoledronate and risedronate inhibit *Plasmodium* geranylgeranyl diphosphate synthase (GGPPS) and exhibit potent antimalarial activity. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 4058–4063 (2012).

39. Fleurence, J. & Levine, I. Seaweed in Health and Disease Prevention. *Academic Press Elsevier Ltd*, (2016).
40. Burki, F. The Eukaryotic Tree of Life from a Global Phylogenomic Perspective. *Cold Spring Harb Perspect Biol* **6**, (2014).
41. Guiry, M.D. and Guiry, G.M. Algaebase. *World-wide Electron. Publ. Natl. Univ. Ireland, Galway* (2018).
42. Vonthron-Sénécheau, C., Kaiser, M., Devambe, I., Vastel, A., Mussio, I. & Rusig, A.-M. Antiprotozoal activities of organic Extracts from French marine seaweeds. *Mar Drugs* **9**, 922–933 (2011).
43. Süzgeç-Selçuk, S., Meriçli, A. H., Güven, K. C., Kaiser, M., Casey, R., Hingley-Wilson, S., Lalvani, A. & Tasdemir, D. Evaluation of Turkish seaweeds for antiprotozoal, antimycobacterial and cytotoxic activities. *Phytother. Res.* **25**, 778–783 (2011).
44. Lategan, C., Kellerman, T., Afolayan, A. F., Mann, M. G., Antunes, E. M., Smith, P. J., Bolton, J. J. & Beukes, D. R. Antiplasmodial and antimicrobial activities of South African marine algal extracts. *Pharmaceutical Biology* **47**, 408–413 (2009).
45. Ravikumar, S., Inbaneson, S. J. & Suganthi, P. Seaweeds as a source of lead compounds for the development of new antiplasmodial drugs from South East coast of India. *Parasitol Res* **109**, 47–52 (2010).
46. Vonthron-Sénécheau, C., Kaiser, M., Devambe, I., Vastel, A., Mussio, I. & Rusig, A.-M. Antiprotozoal activities of organic extracts from French marine seaweeds. *Mar Drugs* **9**, 922–933 (2011).
47. Stout, E. P., Prudhomme, J., Le Roch, K., Fairchild, C. R., Franzblau, S. G., Aalbersberg, W., Hay, M. E. & Kubanek, J. Unusual antimalarial meroditerpenes from tropical red macroalgae. *Bioorg Med Chem Lett* **20**, 5662–5665 (2010).
48. Wright, A. D., König, G. M., Angerhofer, C. K., Greenidge, P., Linden, A. & Desqueyroux-Faúndez, R. Antimalarial activity: the search for marine-derived natural products with selective antimalarial activity. *J. Nat. Prod.* **59**, 710–716 (1996).
49. Mendiola Martínez, J., Hernández, H., Acuña, D., Esquivel, M., Scull Lizama, R. & Abreu Payrol, J. Inhibiting activity of the *in vitro* growth of *Plasmodium falciparum* of extracts from algae of genus *Laurencia*. *Rev Cubana Med Trop* **57**, 192–195 (2005).
50. Stewart, W. D. P. & Stewart, W. D. P. *Algal physiology and biochemistry*. (Oxford: Blackwell, 1974).
51. Rowan, K. S. *Photosynthetic Pigments of Algae*. (CUP Archive, 1989).
52. Manning, W. M. & Strain, H. H. Chlorophyll D, a green pigment of red algae. *J. Biol. Chem.* **151**, 1–19 (1943).
53. Sathasivam, R. & Ki, J.-S. A Review of the biological activities of microalgal carotenoids and their potential use in healthcare and cosmetic industries. *Mar Drugs* **16**, (2018).

54. Li, S.-Y., Shabtai, Y. & Arad, S. Floridoside as a carbon precursor for the synthesis of cell-wall polysaccharide in the red microalga *Porphyridium Sp.* (Rhodophyta). *Journal of Phycology* **38**, 931–938
55. Pereira, M. G., Benevides, N. M. B., Melo, M. R. S., Valente, A. P., Melo, F. R. & Mourão, P. A. S. Structure and anticoagulant activity of a sulfated galactan from the red alga, *Gelidium crinale*. Is there a specific structural requirement for the anticoagulant action? *Carbohydr. Res.* **340**, 2015–2023 (2005).
56. Witvrouw, M. & De Clercq, E. Sulfated polysaccharides extracted from sea algae as potential antiviral drugs. *Gen. Pharmacol.* **29**, 497–511 (1997).
57. Coombe, D. R., Parish, C. R., Ramshaw, I. A. & Snowden, J. M. Analysis of the inhibition of tumor metastasis by sulphated polysaccharides. *Int. J. Cancer* **39**, 82–88 (1987).
58. Cumashi, A. *et al.* A comparative study of the anti-inflammatory, anticoagulant, antiangiogenic, and antiadhesive activities of nine different fucoidans from brown seaweeds. *Glycobiology* **17**, 541–552 (2007).
59. Sanina, N. M., Goncharova, S. N. & Kostetsky, E. Y. Fatty acid composition of individual polar lipid classes from marine macrophytes. *Phytochemistry* **65**, 721–730 (2004).
60. Dembitskii, V. M. & Rozentsvet, O. A. Lipids of macrophyte marine algae. I. Fatty-acid and phospholipid composition of Rhodophyceae. *Chem Nat Compd* **26**, 92–93 (1990).
61. Khotimchenko, S. V., Klochkova, N. G. & Vaskovsky, V. E. Polar lipids of marine macrophytic algae as chemotaxonomic markers. *Biochemical Systematics and Ecology* **18**, 93–101 (1990).
62. Dang, H. T., Lee, H. J., Yoo, E. S., Shinde, P. B., Lee, Y. M., Hong, J., Kim, D. K. & Jung, J. H. Anti-inflammatory constituents of the red alga *Gracilaria verrucosa* and their synthetic analogues. *J. Nat. Prod.* **71**, 232–240 (2008).
63. The innate immunity of a marine red alga involves oxylipins from both the eicosanoid and octadecanoid pathways | Plant Physiology. Disponible sur : <http://www.plantphysiol.org/content/135/3/1838.short>. (Consulté le : 3 Juillet 2018)
64. Kumari, P., Reddy, R. & Jha, B. Quantification of selected endogenous hydroxy-oxylipins from tropical marine macroalgae. *Mar Biotechnol* **16**, 74–87 (2014).
65. Barbosa, M., Valentão, P. & Andrade, P. B. Biologically active oxylipins from enzymatic and nonenzymatic routes in macroalgae. *Mar Drugs* **14**, (2016).
66. Dang, H. T., Lee, H. J., Yoo, E. S., Shinde, P. B., Lee, Y. M., Hong, J., Kim, D. K. & Jung, J. H. Anti-inflammatory constituents of the red alga *Gracilaria verrucosa* and their synthetic analogues. *J. Nat. Prod.* **71**, 232–240 (2008).
67. Plouguerné, E., Kikuchi, H., Oshima, Y., Deslandes, E. & Stiger-Pouvreau, V. Isolation of Cholest-5-en-3-ol formate from the red alga *Grateloupia turuturu* Yamada and its chemotaxonomic significance. *Biochemical Systematics and Ecology* **9**, 714–717 (2006).

68. Fukuzawa, A., Kumagai, Y., Masamune, T., Furusaki, A., Katayama, C. & Matsumoto, T. Acetylpinasterol and pinasterol, ecdysone-like metabolites, from the marine red alga *Laurencia pinnata*. *Tetrahedron Letters* **22**, 4085–4086 (1981).
69. Perumal, P., Sowmiya, R., Prasanna Kumar, S., Ravikumar, S., Deepak, P. & Balasubramani, G. Isolation, structural elucidation and antiplasmodial activity of fucosterol compound from brown seaweed, *Sargassum linearifolium* against malarial parasite *Plasmodium falciparum*. *Nat. Prod. Res.* **32**, 1316–1319 (2018).
70. Lin, A.-S., Engel, S., Smith, B. A., Fairchild, C. R., Aalbersberg, W., Hay, M. E. & Kubanek, J. Structure and biological evaluation of novel cytotoxic sterol glycosides from the marine red alga *Peyssonnelia* sp. *Bioorg Med Chem* **18**, 8264–8269 (2010).
71. de Inés, C., Argandoña, V. H., Rovirosa, J., San-Martín, A., Díaz-Marrero, A. R., Cueto, M. & González-Coloma, A. Cytotoxic activity of halogenated monoterpenes from *Plocamium cartilagineum*. *Z. Naturforsch., C, J. Biosci.* **59**, 339–344 (2004).
72. Sabry, O. M. M., Goeger, D. E., Valeriote, F. A. & Gerwick, W. H. Cytotoxic halogenated monoterpenes from *Plocamium cartilagineum*. *Nat Prod Res* **31**, 261–267 (2017).
73. Barcellos Marini, M., Rodrigues de Freitas, W., Lacerda da Silva Machado, F., Correa Ramos Leal, I., Ribeiro Soares, A., Masahiko Kanashiro, M. & Frazão Muzitano, M. Cytotoxic activity of halogenated sesquiterpenes from *Laurencia dendroidea*. *Phytother Res* **32**, 1119–1125 (2018).
74. Alarif, W. M., Al-Footy, K. O., Zubair, M. S., Halid Ph, M., Ghandourah, M. A., Basaif, S. A., Al-Lihaibi, S. S., Ayyad, S.-E. N. & Badria, F. A. The role of new eudesmane-type sesquiterpenoid and known eudesmane derivatives from the red alga *Laurencia obtusa* as potential antifungal-antitumour agents. *Nat. Prod. Res.* **30**, 1150–1155 (2016).
75. Argandoña, V. H., Rovirosa, J., San-Martín, A., Riquelme, A., Díaz-Marrero, A. R., Cueto, M., Darias, J., Santana, O., Guadaño, A. & González-Coloma, A. Antifeedant effects of marine halogenated monoterpenes. *J. Agric. Food Chem.* **50**, 7029–7033 (2002).
76. Shapumba, C. W., Knott, M. & Kapewangolo, P. Antioxidant activity of a halogenated monoterpene isolated from a Namibian marine algal *Plocamium* species. *J Food Sci Technol* **54**, 3370–3373 (2017).
77. Ferreira, W. J., Amaro, R., Cavalcanti, D. N., de Rezende, C. M., da Silva, V. A. G. G., Barbosa, J. E., Paixão, I. C. N. de P. & Teixeira, V. L. Anti-herpetic activities of chemical components from the Brazilian red alga *Plocamium brasiliense*. *Nat Prod Commun* **5**, 1167–1170 (2010).
78. Kamada, T., Phan, C.-S. & Vairappan, C. S. New anti-bacterial halogenated tricyclic sesquiterpenes from Bornean *Laurencia majuscula* (Harvey) Lucas. *Nat. Prod. Res.* 1–8 (2017).
79. Biá Ventura, T. L., da Silva Machado, F. L., de Araujo, M. H., de Souza Gestinari, L. M., Kaiser, C. R., de Assis Esteves, F., Lasunskaja, E. B., Soares, A. R. & Muzitano, M. F. Nitric Oxide Production Inhibition and Anti-Mycobacterial activity of extracts and

- halogenated sesquiterpenes from the Brazilian Red Alga *Laurencia Dendroidea* J. Agardh. *Pharmacogn Mag* **11**, S611-618 (2015).
80. Perumal, P., Sowmiya, R., kumar, S. P., Ravikumar, S., Deepak, P. & Balasubramani, G. Isolation, structural elucidation and antiplasmodial activity of fucosterol compound from brown seaweed, *Sargassum linearifolium* against malarial parasite *Plasmodium falciparum*. *Natural Product Research* **32**, 1316–1319 (2018).
81. Afolayan, A. F., Mann, M. G. A., Lategan, C. A., Smith, P. J., Bolton, J. J. & Beukes, D. R. Antiplasmodial halogenated monoterpenes from the marine red alga *Plocamium cornutum*. *Phytochemistry* **70**, 597–600 (2009).
82. Lane, A. L., Stout, E. P., Lin, A.-S., Prudhomme, J., Roch, K. L., Fairchild, C. R., Franzblau, S. G., Hay, M. E., Aalbersberg, W. & Kubanek, J. Antimalarial bromophycolides J-Q from the Fijian red alga *Callophycus serratus*. *J Org Chem* **74**, 2736–2742 (2009).
83. Taxonomy Browser : Algaebase. Disponible sur : <http://www.algaebase.org/browse/taxonomy/?id=97240>. (Consulté le : 8 Aout 2018)
84. Cole, K. M. & Sheath, R. G. *Biology of the Red Algae*. (Cambridge University Press, 1990).
85. Courtois, A., Simon-Colin, C., Boisset, C., Berthou, C., Deslandes, E., Guézennec, J. & Bordron, A. Floridoside Extracted from the red alga *Mastocarpus stellatus* is a potent activator of the classical complement pathway. *Mar Drugs* **6**, 407–417 (2008).
86. Barwell, C. J. & Blunden, G. Hordenine from the red alga *Gigartina stellata*. *J. Nat. Prod.* **44**, 500–502 (1981).
87. Hilliou, L., Larontonda, F. D. S., Sereno, A. M. & Gonçalves, M. P. Thermal and viscoelastic properties of κ/ι -Hybrid carrageenan gels obtained from the Portuguese seaweed *Mastocarpus stellatus*. *J. Agric. Food Chem.* **54**, 7870–7878 (2006).
88. Gómez-Ordóñez, E., Jiménez-Escrig, A. & Rupérez, P. Bioactivity of sulfated polysaccharides from the edible red seaweed *Mastocarpus stellatus*. *Bioactive Carbohydrates and Dietary Fibre* **3**, 29–40 (2014).
89. Künzler, K. & Eichenberger, W. Betaine lipids and zwitterionic phospholipids in plants and fungi. *Phytochemistry* **46**, 883–892 (1997).
90. Bjørnland, T. & Aguilar-Martinez, M. Carotenoids in red algae. *Phytochemistry* **15**, 291–296 (1976).
91. Dubber, D. & Harder, T. Extracts of *Ceramium rubrum*, *Mastocarpus stellatus* and *Laminaria digitata* inhibit growth of marine and fish pathogenic bacteria at ecologically realistic concentrations. *Aquaculture* **274**, 196–200 (2008).
92. Allmendinger, A., Spavieri, J., Kaiser, M., Casey, R., Hingley-Wilson, S., Lalvani, A., Guiry, M., Blunden, G. & Tasdemir, D. Antiprotozoal, antimycobacterial and cytotoxic potential of twenty-three British and Irish red algae. *Phytotherapy Research* **24**, 1099–1103

93. Gómez-Ordóñez, E., Jiménez-Escrig, A. & Rupérez, P. Effect of the red seaweed *Mastocarpus stellatus* intake on lipid metabolism and antioxidant status in healthy Wistar rats. *Food Chemistry* **135**, 806–811 (2012).
94. Vonthron-Sénécheau, C., Oliveira Santos, J. S., Mussio, I. & Rusig, A.-M. X-ray structure of floridoside isolated from the marine red algae *Dilsea carnosa*. *Carbohydrate Research* **343**, 2697–2698 (2008).
95. Tariq, V.-N. Antifungal activity in crude extracts of marine red algae. *Mycological Research* **95**, 1433–1435 (1991).
96. Hornsey, I. S. & Hide, D. The production of antimicrobial compounds by British marine algae I. Antibiotic-producing marine algae. *British Phycological Journal* **9**, 353–361 (1974).
97. Rodríguez, M. C., Matulewicz, M. C., Nosedá, M. D., Ducatti, D. R. B. & Leonardi, P. I. Agar from *Gracilaria gracilis* (Gracilariales, Rhodophyta) of the Patagonic coast of Argentina – Content, structure and physical properties. *Bioresource Technology* **100**, 1435–1441 (2009).
98. Valentão, P., Trindade, P., Gomes, D., Guedes de Pinho, P., Mouga, T. & Andrade, P. B. *Codium tomentosum* and *Plocamium cartilagineum*: Chemistry and antioxidant potential. *Food Chemistry* **119**, 1359–1368 (2010).
99. Díaz-Marrero, A. R., Cueto, M., Dorta, E., Roviroso, J., San-Martín, A. & Darias, J. New halogenated monoterpenes from the red alga *Plocamium cartilagineum*. *Tetrahedron* **58**, 8539–8542 (2002).
100. San-Martín, A., Negrete, R. & Roviroso, J. Insecticide and acaricide activities of polyhalogenated monoterpenes from Chilean *Plocamium cartilagineum*. *Phytochemistry* **30**, 2165–2169 (1991).
101. Roviroso, J., Moena, J. & San-Martín, A. Two chemical types of the red alga *Plocamium cartilagineum* from Chile. *Biochemical Systematics and Ecology* **16**, 593–595 (1988).
102. Mynderse, J. S. & Faulkner, D. J. Variations in the halogenated monoterpene metabolites of *Plocamium cartilagineum* and *P. violaceum*. *Phytochemistry* **17**, 237–240 (1978).
103. Barrow, K. D. & Temple, C. A. Biosynthesis of halogenated monoterpenes in *Plocamium cartilagineum*. *Phytochemistry* **24**, 1697–1704 (1985).
104. González, A. G., Manuel Arteaga, J., Martín, J. D., Rodríguez, M. L., Fayos, J. & Martínez-Ripolls, M. Two new polyhalogenated monoterpenes from the red alga *Plocamium cartilagineum*. *Phytochemistry* **17**, 947–948 (1978).
105. Řezanka, T. & Dembitsky, V. M. Polyhalogenated homosesquiterpenic fatty acids from *Plocamium cartilagineum*. *Phytochemistry* **57**, 607–611 (2001).
106. Harden, E. A., Falshaw, R., Carnachan, S. M., Kern, E. R. & Prichard, M. N. Virucidal activity of polysaccharide extracts from four algal species against *herpes simplex* virus. *Antiviral Research* **83**, 282–289 (2009).

107. Banskota, A. H., Stefanova, R., Sperker, S., Lall, S. P., Craigie, J. S., Hafting, J. T. & Critchley, A. T. Polar lipids from the marine macroalga *Palmaria palmata* inhibit lipopolysaccharide-induced nitric oxide production in RAW264.7 macrophage cells. *Phytochemistry* **101**, 101–108 (2014).
108. Dupéron, R., Thiersault, M. & Dupéron, P. Occurrence of steryl glycosides and acylated steryl glycosides in some marine algae. *Phytochemistry* **22**, 535–538 (1983).
109. Jongaramruong, J. & Blackman, A. J. Polyhalogenated monoterpenes from a tasmanian collection of the red seaweed *Plocamium cartilagineum*. *J. Nat. Prod.* **63**, 272–275 (2000).
110. Díaz-Marrero, A. R., Rovirosa, J., Darias, J., San-Martín, A. & Cueto, M. Plocamenols A-C, novel linear polyhalohydroxylated monoterpenes from *Plocamium cartilagineum*. *J. Nat. Prod.* **65**, 585–588 (2002).
111. Darias, J., Rovirosa, J., San Martín, A., Díaz, A. R., Dorta, E. & Cueto, M. Furoplocamioids A-C, novel polyhalogenated furanoid monoterpenes from *Plocamium cartilagineum*. *J. Nat. Prod.* **64**, 1383–1387 (2001).
112. Gao, D., Okuda, R. & Lopez-Avila, V. Supercritical fluid extraction of halogenated monoterpenes from the red alga *Plocamium cartilagineum*. *J AOAC Int* **84**, 1313–1331 (2001).
113. Lee, D., Nishizawa, M., Shimizu, Y. & Saeki, H. Anti-inflammatory effects of dulse (*Palmaria palmata*) resulting from the simultaneous water-extraction of phycobiliproteins and chlorophyll a. *Food Research International* **100**, 514–521 (2017).
114. Yuan, Y. V., Bone, D. E. & Carrington, M. F. Antioxidant activity of dulse (*Palmaria palmata*) extract evaluated *in vitro*. *Food Chemistry* **91**, 485–494 (2005).
115. Yuan, Y. V., Carrington, M. F. & Walsh, N. A. Extracts from dulse (*Palmaria palmata*) are effective antioxidants and inhibitors of cell proliferation *in vitro*. *Food and Chemical Toxicology* **43**, 1073–1081 (2005).
116. Gosch, B. J., Paul, N. A., Nys, R. de & Magnusson, M. Spatial, seasonal, and within-plant variation in total fatty acid content and composition in the brown seaweeds *Dictyota bartayresii* and *Dictyopteris australis* (Dictyotales, Phaeophyceae). *J Appl Phycol* **27**, 1607–1622 (2015).
117. Schmid, M., Guihéneuf, F. & Stengel, D. B. Ecological and commercial implications of temporal and spatial variability in the composition of pigments and fatty acids in five Irish macroalgae. *Mar Biol* **164**, 158 (2017).
118. Wein, S., Maynadier, M., Tran Van Ba, C., Cerdan, R., Peyrottes, S., Fraisse, L. & Vial, H. Reliability of antimalarial sensitivity tests depends on drug mechanisms of action. *J. Clin. Microbiol.* **48**, 1651–1660 (2010).
119. Nardella, F. *Flavones substituées : une nouvelle classe de composés pour le traitement du paludisme : optimisation vers un candidat médicament*. (Strasbourg, 2017).
120. Yaffe, D. Retention of differentiation potentialities during prolonged cultivation of myogenic cells. *PNAS* **61**, 477–483 (1968).

121. Gallé, J.-B., Attioua, B., Kaiser, M., Rusig, A.-M., Lobstein, A. & Vonthron-Sénécheau, C. Eleganolone, a diterpene from the French marine alga *Bifurcaria bifurcata* inhibits growth of the human pathogens *Trypanosoma brucei* and *Plasmodium falciparum*. *Mar Drugs* **11**, 599–610 (2013).
122. Rodríguez-Concepción, M., Boronat (auth.), A., Bach, T. J. & Rohmer (eds.), M. *Isoprenoid Synthesis in Plants and Microorganisms: New Concepts and Experimental Approaches*. (Springer-Verlag New York, 2013).
123. Amsler, C. D. *Algal Chemical Ecology*. (Springer Science & Business Media, 2007).
124. Delgoda, R. & Murray, J. E. Chapter 7 - Evolutionary perspectives on the role of plant secondary metabolites. in *Pharmacognosy* (eds. Badal, S. & Delgoda, R.) 93–100 (Academic Press, 2017).
125. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs from 1981 to 2014. *J. Nat. Prod.* **79**, 629–661 (2016).
126. Wolfender, J.-L., Marti, G., Thomas, A. & Bertrand, S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J Chromatogr A* **1382**, 136–164 (2015).
127. Mendham, J. *Analyse chimique quantitative de Vogel*. (De Boeck Supérieur, 2005).
128. Stéphane, B. *La spectrométrie de masse en couplage avec la chromatographie en phase gazeuse*. (Lavoisier, 2009).
129. Delannay, E., Toribio, A., Boudesocque, L., Nuzillard, J.-M., Zèches-Hanrot, M., Dardennes, E., Le Dour, G., Sapi, J. & Renault, J.-H. Multiple dual-mode centrifugal partition chromatography, a semi-continuous development mode for routine laboratory-scale purifications. *Journal of Chromatography A* **1127**, 45–51 (2006).
130. Günther, H. *La spectroscopie de RMN : principes de base, concepts et applications de la spectroscopie de résonance magnétique nucléaire du proton et du carbone 13 en chimie*. (Masson, 1996).
131. L'expérience DOSY, une puissante méthode RMN pour l'analyse de mélanges complexes et la détection de traces. *ResearchGate* Disponible sur : https://www.researchgate.net/publication/221966846_L%27experience_DOSY_une_puissante_methode_RMN_pour_l%27analyse_de_melanges_complexes_et_la_detection_de_traces. (Consulté le : 31 Juillet 2018)
132. Hubert, J., Nuzillard, J.-M. & Renault, J.-H. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? *Phytochem Rev* **16**, 55–95 (2017).
133. Beutler, J. A., Alvarado, A. B., Schaufelberger, D. E., Andrews, P. & McCloud, T. G. Dereplication of phorbol bioactives: *Lyngbya majuscula* and *Croton cuneatus*. *J. Nat. Prod.* **53**, 867–874 (1990).
134. Wang, M. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* **34**, 828–837 (2016).

135. Nothias-Scaglia, L.-F., Esposito, M., Costa, J., Paolini, J., Touboul, D. & Litaudon, M. Les réseaux moléculaires, une approche bio-informatique globale pour interpréter les données de spectrométrie de masse tandem. *Spectra Analyse* **307**, 73–78 (2015).
136. Pauli, G. F., Kuczkowiak, U. & Nahrstedt, A. Solvent effects in the structure dereplication of caffeoyl quinic acids. *Magnetic Resonance in Chemistry* **37**, 827–836 (1999).
137. Politi, M., Groves, P., Chávez, M. I., Cañada, F. J. & Jiménez-Barbero, J. Useful applications of DOSY experiments for the study of mushroom polysaccharides. *Carbohydrate Research* **341**, 84–89 (2006).
138. Hubert, J., Nuzillard, J.-M., Purson, S., Hamzaoui, M., Borie, N., Reynaud, R. & Renault, J.-H. Identification of natural metabolites in mixture: A pattern recognition strategy based on ^{13}C NMR. *Anal. Chem.* **86**, 2955–2962 (2014).
139. Clendinen, C. S., Lee-McMullen, B., Williams, C. M., Stupp, G. S., Vandeborne, K., Hahn, D. A., Walter, G. A. & Edison, A. S. ^{13}C NMR Metabolomics: Applications at Natural Abundance. *Anal Chem* **86**, 9242–9250 (2014).
140. Bingol, K., Bruschweiler-Li, L., Li, D.-W. & Bruschweiler, R. Customized metabolomics database for the analysis of NMR ^1H - ^1H TOCSY and ^{13}C - ^1H HSQC-TOCSY spectra of complex mixtures. *Anal Chem* **86**, 5494–5501 (2014).
141. Van, Q. N., Issaq, H. J., Jiang, Q., Li, Q., Muschik, G. M., Waybright, T. J., Lou, H., Dean, M., Uitto, J. & Veenstra, T. D. Comparison of 1D and 2D NMR spectroscopy for metabolic profiling. *J. Proteome Res.* **7**, 630–639 (2008).
142. Chikayama, E., Suto, M., Nishihara, T., Shinozaki, K., Hirayama, T. & Kikuchi, J. Systematic NMR analysis of stable isotope labeled metabolite mixtures in plant and animal systems: coarse grained views of metabolic pathways. *PLoS ONE* **3**, e3805 (2008).
143. Robinette, S. L., Ajredini, R., Rasheed, H., Zeinomar, A., Schroeder, F. C., Dossey, A. T. & Edison, A. S. Hierarchical Alignment and Full Resolution Pattern Recognition of 2D NMR Spectra: Application to Nematode Chemical Ecology. *Anal Chem* **83**, 1649–1657 (2011).
144. Bingol, K., Zhang, F., Bruschweiler-Li, L. & Bruschweiler, R. TOCCATA: a customized carbon total correlation spectroscopy NMR metabolomics database. *Anal. Chem.* **84**, 9395–9401 (2012).
145. BMRB - Biological Magnetic Resonance Bank. Disponible sur : <http://www.bmrwisc.edu/>. (Consulté le : 31 Juillet 2018)
146. Wishart, D. S. *et al.* HMDB 3.0 - The Human Metabolome Database in 2013. *Nucleic Acids Res.* **41**, D801-807 (2013).
147. COLMAR ^{13}C - ^1H HSQC Query. Disponible sur : <http://spin.ccic.ohio-state.edu/index.php/hsqc/>. (Consulté le : 31 Juillet 2018)
148. Bingol, K., Bruschweiler-Li, L., Yu, C., Somogyi, A., Zhang, F. & Bruschweiler, R. Metabolomics beyond spectroscopic databases: a combined MS/NMR strategy for the rapid identification of new metabolites in complex mixtures. *Anal. Chem.* **87**, 3864–3870 (2015).

149. Pence, H. E. & Williams, A. ChemSpider: An Online Chemical Information Resource. *J. Chem. Educ.* **87**, 1123–1124 (2010).
150. Watanabe, N. & Niki, E. Direct-Coupling of FT-NMR to high performance liquid chromatography. *Proc. Jpn. Acad., Ser. B* **54**, 194–199 (1978).
151. Lambert, M., Staerk, D., Hansen, S. H., Sairafianpour, M. & Jaroszewski, J. W. Rapid extract dereplication using HPLC-SPE-NMR: analysis of isoflavonoids from *Smirnowia iranica*. *J. Nat. Prod.* **68**, 1500–1509 (2005).
152. Lam, S.-H., Wang, C.-Y., Chen, C.-K. & Lee, S.-S. Chemical investigation of *Phyllanthus reticulatus* by HPLC-SPE-NMR and conventional methods. *Phytochemical Analysis* **18**, 251–255 (2007).
153. Lee, S.-S., Lai, Y.-C., Chen, C.-K., Tseng, L.-H. & Wang, C.-Y. Characterization of isoquinoline alkaloids from *Neolitsea sericea* var. *aurata* by HPLC-SPE-NMR. *J. Nat. Prod.* **70**, 637–642 (2007).
154. Sprogøe, K., Stärk, D., Jäger, A. K., Adsersen, A., Hansen, S. H., Witt, M., Landbo, A.-K. R., Meyer, A. S. & Jaroszewski, J. W. Targeted natural product isolation guided by HPLC-SPE-NMR: Constituents of *Hubertia* Species. *J. Nat. Prod.* **70**, 1472–1477 (2007).
155. Clarkson, C., Hansen, S. H., Smith, P. J. & Jaroszewski, J. W. Discovering new natural products directly from crude extracts by HPLC-SPE-NMR: Chinane Diterpenes in *Harpagophytum procumbens*. *J. Nat. Prod.* **69**, 527–530 (2006).
156. Schroeder, F. C., Gibson, D. M., Churchill, A. C. L., Sojikul, P., Wursthorn, E. J., Krasnoff, S. B. & Clardy, J. Differential analysis of 2D NMR spectra: new natural products from a pilot-scale fungal extract library. *Angew. Chem. Int. Ed. Engl.* **46**, 901–904 (2007).
157. Mahanti, P., Bose, N., Bethke, A., Judkins, J. C., Wollam, J., Dumas, K. J., Zimmerman, A. M., Campbell, S. L., Hu, P. J., Antebi, A. & Schroeder, F. C. Comparative metabolomics reveals endogenous ligands of DAF-12, a nuclear hormone receptor regulating *C. elegans* development and lifespan. *Cell Metab* **19**, 73–83 (2014).
158. Nardella, F., Margueritte, L., Lamure, B., Viéville, J. M. P. & Bourjot, M. Targeted discovery of bioactive natural products using a pharmacophoric deconvolution strategy: Proof of principle with eleanolone from *Bifurcaria bifurcata* R. Ross. *Phytochemistry Letters* **26**, 138–142 (2018).
159. Muñoz, J., Culioli, G. & Köck, M. Linear diterpenes from the marine brown alga *Bifurcaria bifurcata*: a chemical perspective. *Phytochem Rev* **12**, 407–424 (2013).
160. Margueritte, L., Markov, P., Chiron, L., Starck, J.-P., Vonthron-Sénécheau, C., Bourjot, M. & Delsuc, M.-A. Automatic differential analysis of NMR experiments in complex samples. *Magn Reson Chem* **56**, 469–479 (2018).
161. Jeannerat, D. High resolution in heteronuclear ^1H - ^{13}C NMR experiments by optimizing spectral aliasing with one-dimensional carbon data *Magn Reson Chem* **41**, 3-17 (2018).
162. Delsuc M. A. L'expérience DOSY, une puissante méthode RMN pour l'analyse de mélanges complexes et la détection de traces. Disponible sur : https://www.researchgate.net/publication/221966846_L'experience_DOSY_une_puissante

- _methode_RMN_pour_l'analyse_de_melanges_complexes_et_la_detection_de_traces.
(Consulté le : 17 Juillet 2018)
163. Wallaart, T. E., Pras, N., Beekman, A. C. & Quax, W. J. Seasonal variation of artemisinin and its biosynthetic precursors in plants of *Artemisia annua* of different geographical origin: proof for the existence of chemotypes. *Planta Med.* **66**, 57–62 (2000).
 164. Delabays, N., Simonnet, X. & Gaudin, M. The genetics of artemisinin content in *Artemisia annua* L. and the breeding of high yielding cultivars. *Curr. Med. Chem.* **8**, 1795–1801 (2001).
 165. Klayman, D. L., Lin, A. J., Acton, N., Scovill, J. P., Hoch, J. M., Milhous, W. K., Theoharides, A. D. & Dobek, A. S. Isolation of artemisinin (Qinghaosu) from *Artemisia annua* growing in the United States. *J. Nat. Prod.* **47**, 715–717 (1984).
 166. Smyrniotopoulos, V., Merten, C., Kaiser, M. & Tasdemir, D. Bifurcatriol, a new antiprotozoal acyclic aiterpene from the brown alga *Bifurcaria bifurcata*. *Mar Drugs* **15**, (2017).
 167. Dinio, T., Gorka, A. P., McGinniss, A., Roepe, P. D. & Morgan, J. B. Investigating the activity of quinine analogues versus chloroquine resistant *Plasmodium falciparum*. *Bioorganic & Medicinal Chemistry* **20**, 3292–3297 (2012).
 168. Leverrier, A., Bero, J., Frédéricich, M., Quetin-Leclercq, J. & Palermo, J. Antiparasitic hybrids of Cinchona alkaloids and bile acids. *European Journal of Medicinal Chemistry* **66**, 355–363 (2013).
 169. Hornsey, I. S. & Hide, D. The production of antimicrobial compounds by British marine algae. IV. Variation of antimicrobial activity with algal generation. *British Phycological Journal* **20**, 21–25 (1985).
 170. Calvin, N. I. & Ellis, R. J. Water soluble extracts of certain marine red algae and processes for use thereof. (1979).
 171. Kornprobst, J. M. Encyclopedia Of Marine Natural Products. Wiley Blackwell. (2014).
 172. Nasiri, M., Saeidnia, S., Mashinchian-Moradi, A. & Gohari, A. R. Sterols from the red algae, *Gracilaria salicornia* and *Hypnea flagelliformis*, from Persian Gulf. *Pharmacognosy Magazine* **7**, 97 (2011).
 173. Ragasa, C. Structure elucidation of stigmasterol and B-sitosterol from *Sesbania grandiflora* (Linn.) Pers. and B-carotene from *Heliotropium indicum* Linn. by NMR spectroscopy. *Kimika*
 174. Nothias, L.-F., Nothias-Esposito, M., da Silva, R., Wang, M., Protsyuk, I., Zhang, Z., Sarvepalli, A., Leyssen, P., Touboul, D., Costa, J., Paolini, J., Alexandrov, T., Litaudon, M. & Dorrestein, P. C. Bioactivity-based molecular networking for the discovery of drug leads in natural product bioassay-guided fractionation. *J. Nat. Prod.* **81**, 758–767 (2018).
 175. Laurent, D., Jullian, V., Parenty, A., Knibiehler, M., Dorin, D., Schmitt, S., Lozach, O., Lebouvier, N., Frostin, M., Alby, F., Maurel, S., Doerig, C., Meijer, L. & Sauvain, M. Antimalarial potential of xestoquinone, a protein kinase inhibitor isolated from a Vanuatu

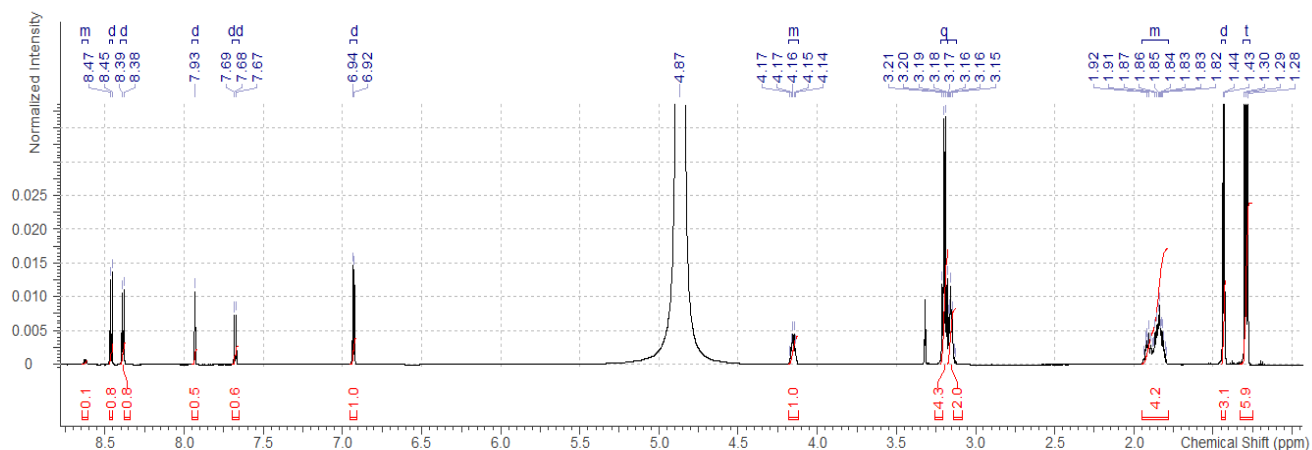
- marine sponge *Xestospongia sp.* *Bioorganic & Medicinal Chemistry* **14**, 4477–4482 (2006).
176. Tasdemir, D., Topaloglu, B., Perozzo, R., Brun, R., O'Neill, R., Carballeira, N. M., Zhang, X., Tonge, P. J., Linden, A. & Rüedi, P. Marine natural products from the Turkish sponge *Agelas oroides* that inhibit the enoyl reductases from *Plasmodium falciparum*, *Mycobacterium tuberculosis* and *Escherichia coli*. *Bioorganic & Medicinal Chemistry* **15**, 6834–6845 (2007).
177. Orhan, I., Şener, B., Kaiser, M., Brun, R. & Tasdemir, D. Inhibitory activity of marine sponge-derived natural products against parasitic protozoa. *Marine Drugs* **8**, 47–58 (2010).
178. Davis, R. A., Duffy, S., Avery, V. M., Camp, D., Hooper, J. N. A. & Quinn, R. J. (+)-7-Bromotrypargine: an antimalarial β -carboline from the Australian marine sponge *Ancorina sp.* *Tetrahedron Letters* **51**, 583–585 (2010).
179. Kirsch, G., König, G. M., Wright, A. D. & Kaminsky, R. A new bioactive sesterterpene and antiplasmodial alkaloids from the marine sponge *Hyrtios cf. erecta*. *J. Nat. Prod.* **63**, 825–829 (2000).
180. Appenzeller, J., Mihci, G., Martin, M.-T., Gallard, J.-F., Menou, J.-L., Boury-Esnault, N., Hooper, J., Petek, S., Chevalley, S., Valentin, A., Zaparucha, A., Al-Mourabit, A. & Debitus, C. Agelasines J, K, and L from the Solomon Islands marine sponge *Agelas cf. mauritiana*. *J. Nat. Prod.* **71**, 1451–1454 (2008).
181. Desoubzdanne, D., Marcourt, L., Raux, R., Chevalley, S., Dorin, D., Doerig, C., Valentin, A., Ausseil, F. & Debitus, C. Alisiaquinones and alisiaquinol, dual inhibitors of *Plasmodium falciparum* enzyme targets from a New Caledonian deep water sponge. *J. Nat. Prod.* **71**, 1189–1192 (2008).
182. Tramesel, D., Catherinot, V. & Delsuc, M.-A. Modeling of NMR processing, toward efficient unattended processing of NMR experiments. *Journal of Magnetic Resonance* **188**, 56–67 (2007).
183. Trager, W. & Jensen, J. B. Human malaria parasites in continuous culture. *Science* **193**, 673–675 (1976).
184. Burkot, T. R., Williams, J. L. & Schneider, I. Infectivity to mosquitoes of *Plasmodium falciparum* clones grown *in vitro* from the same isolate. *Trans. R. Soc. Trop. Med. Hyg.* **78**, 339–341 (1984).
185. Desjardins, R. E., Canfield, C. J., Haynes, J. D. & Chulay, J. D. Quantitative assessment of antimalarial activity *in vitro* by a semiautomated microdilution technique. *Antimicrob. Agents Chemother.* **16**, 710–718 (1979).
186. Wein, S., Maynadier, M., Ba, C. T. V., Cerdan, R., Peyrottes, S., Fraisse, L. & Vial, H. Reliability of antimalarial density tests depends on drug mechanisms of action. *J. Clin. Microbiol.* **48**, 1651–1660 (2010).



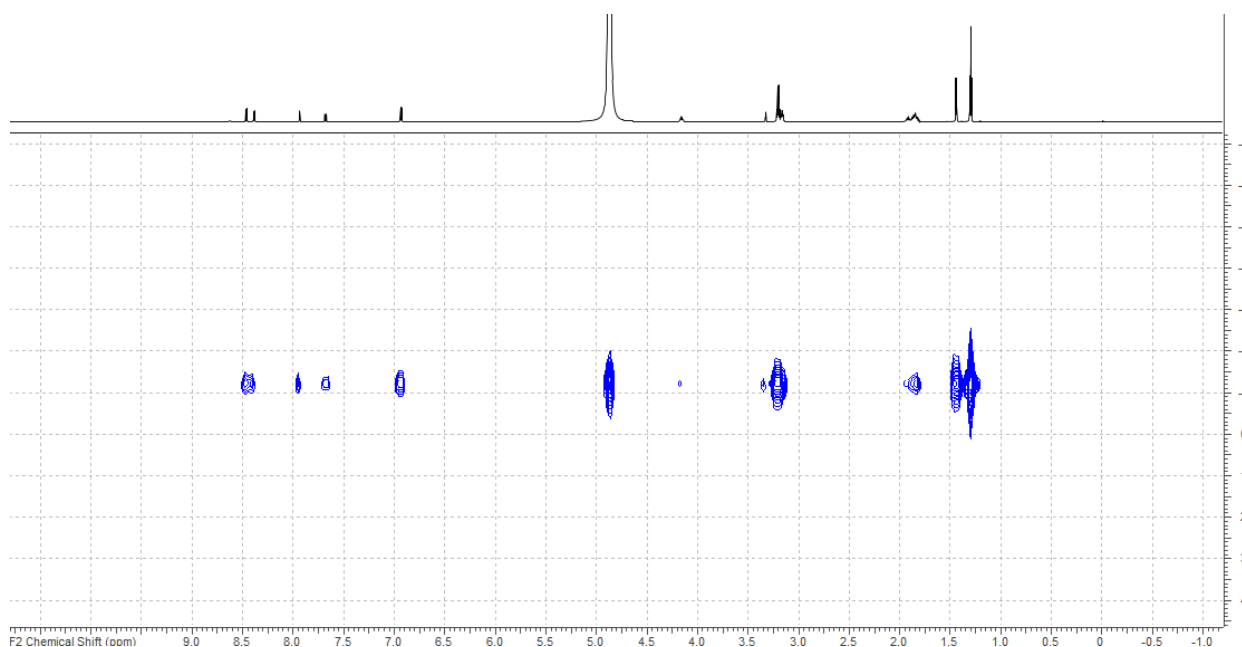
Annexes

Annexes

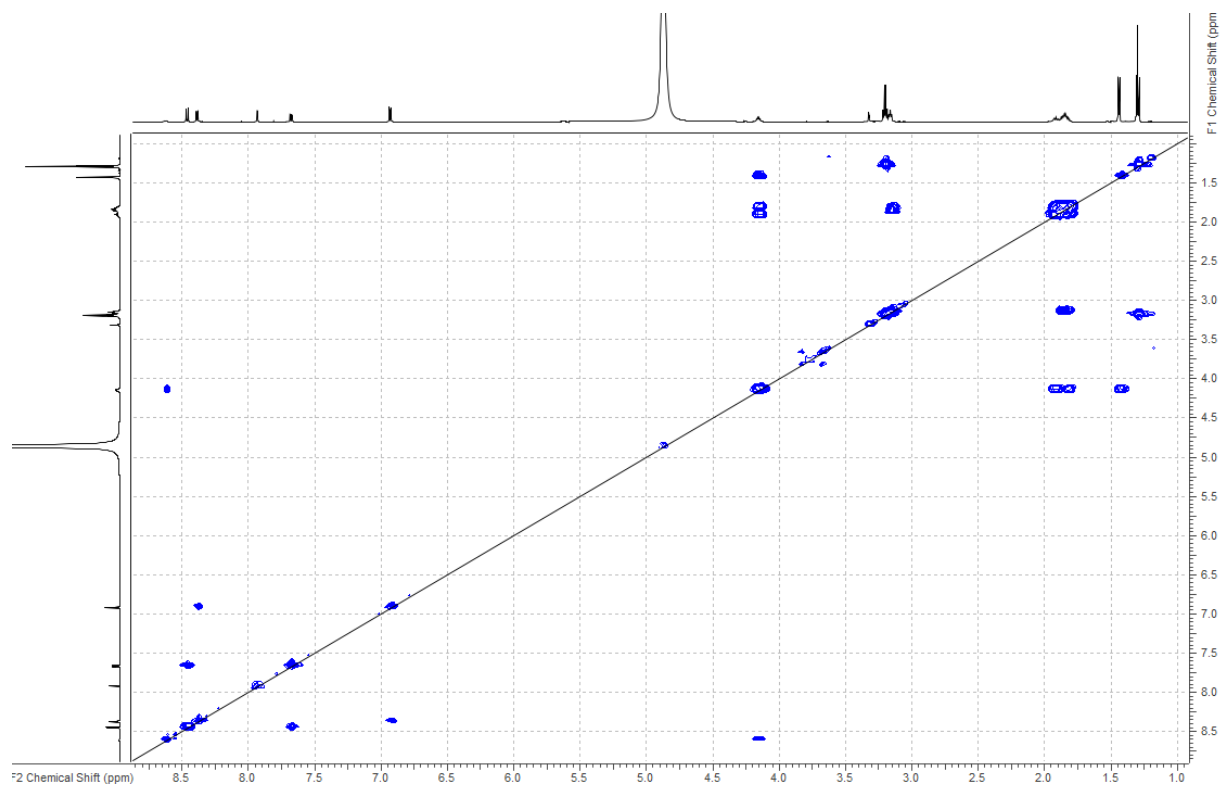
Annexe 1 : Spectres ^1H , ^1H - δ DOSY, ^1H - ^1H COSY, ^1H - ^1H TOCSY, ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC de la chloroquine



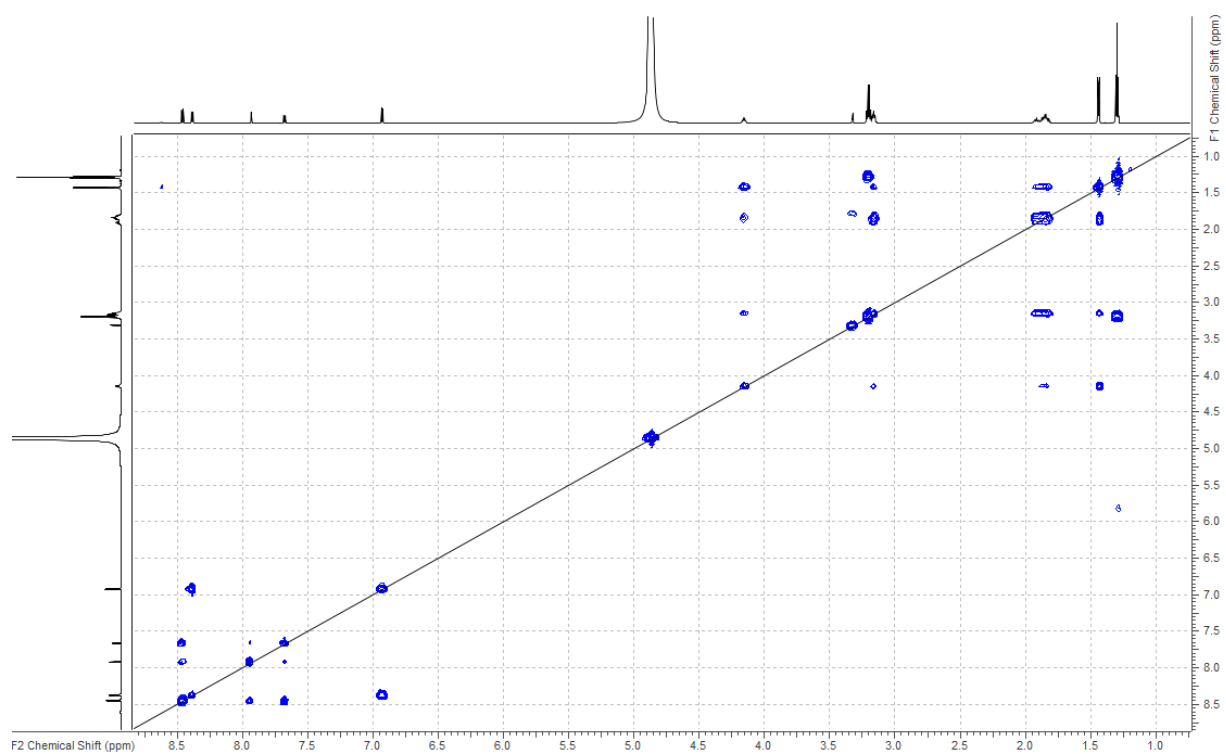
Spectre ^1H de la chloroquine (700 MHz, MeOD)



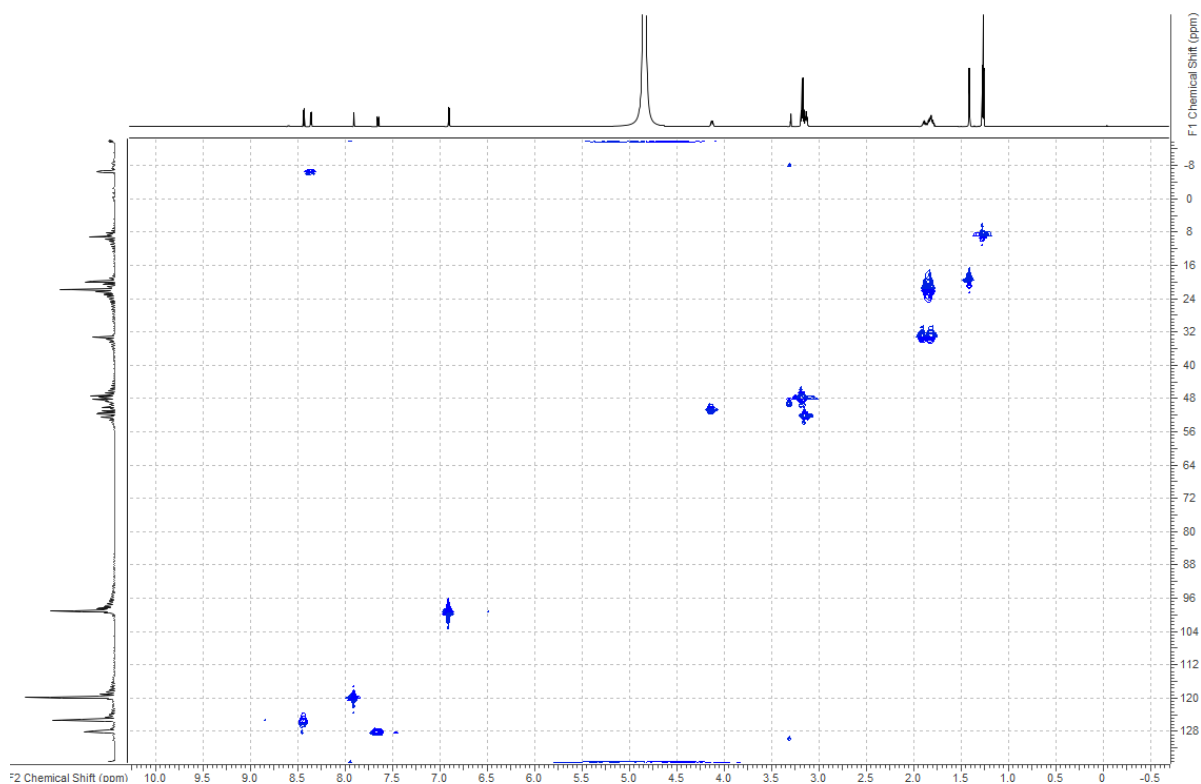
Spectre ^1H - δ DOSY de la chloroquine (700 MHz, MeOD)



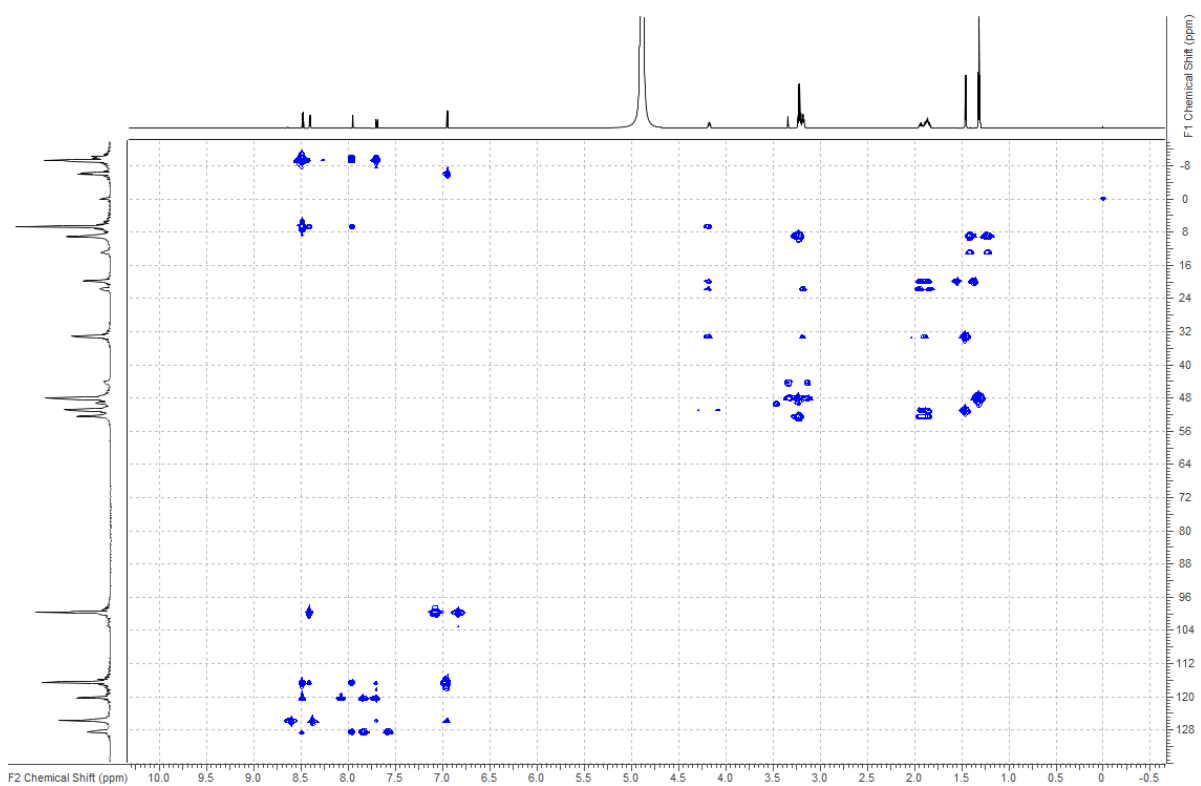
Spectre ^1H - ^1H COSY de la chloroquine (700 MHz, MeOD)



Spectre ^1H - ^1H TOCSY de la chloroquine (700 MHz, MeOD)



Spectre ^1H - ^{13}C HSQC de la chloroquine (700 MHz, MeOD)



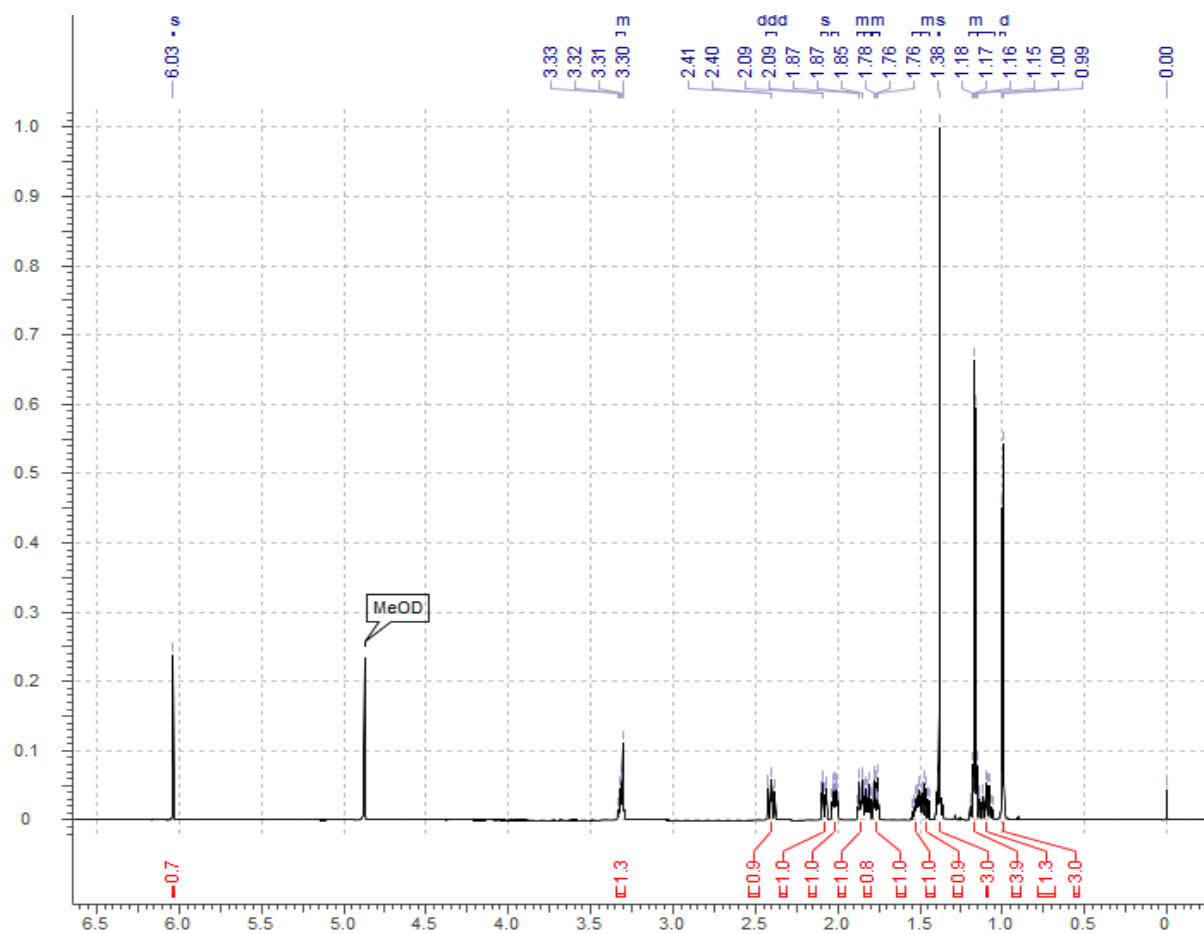
Spectre ^1H - ^{13}C HMBC de la chloroquine (700 MHz, MeOD)

Annexe 2 : Fichiers « Report » et « Analysis » de la série CHLORO

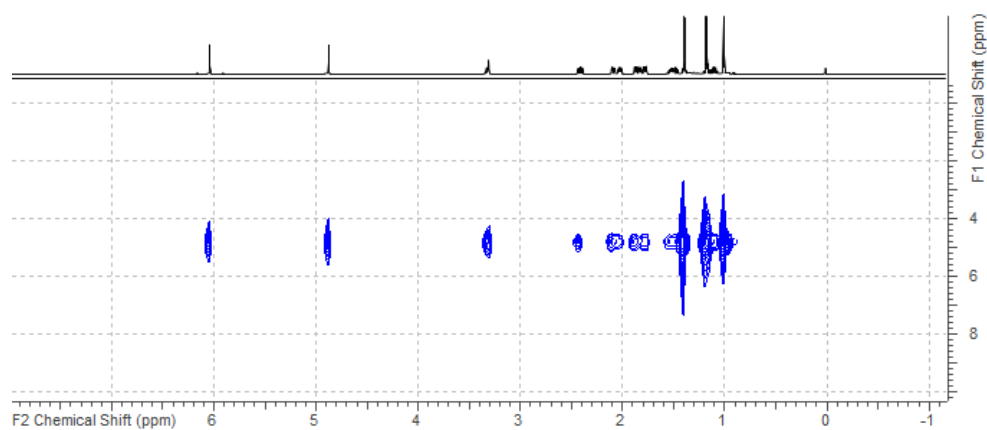
#	# report from ../DECPHA/CHLOROnew													2D					DOSY		
	manip	expho	date	parameters										P1	P2	P3	D9	D20	P30		
				PULPROG	SFO1	NS	TE	TD	RG	SW	O1	DI	P1	SFO1	TD	SW	O1				
Chloro1_161006	1	07/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro1_161006	10	07/10/2016	cosydfp	700.16338	8	295.0593	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		512	12.0222	3380	0.08	-	-	
Chloro1_161006	2	08/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro1_161006	20	07/10/2016	dipsi2p	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		400	12.0222	3380	0.08	-	-	
Chloro1_161006	31	07/10/2016	dsteppg3spr	700.16338	64	294.9623	4096	256	12.0222060421218	3380	1.5	12.5	700.16338		50	12.0222	3380	0.08	0.15	1100	
Chloro1_161006	40	07/10/2016	hsqctgp	700.16338	4	294.9623	2048	2050	10.986446752339	3380	1.5	12.5	176.06593532232		512	150	10563.32232	0.06	-	-	
Chloro1_161006	50	07/10/2016	hmbcgplndqf	700.16338	48	294.9623	4096	2050	10.986446752339	3380	1.5	12.5	176.06593532232		400	150	10563.32232	0.06	-	-	
Chloro2_161006	1	08/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro2_161006	10	08/10/2016	cosydfp	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		512	12.0222	3380	0.08	-	-	
Chloro2_161006	2	08/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro2_161006	20	08/10/2016	dipsi2p	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		400	12.0222	3380	0.08	-	-	
Chloro2_161006	31	08/10/2016	dsteppg3spr	700.16338	64	294.9623	4096	256	12.0222060421218	3380	1.5	12.5	700.16338		50	12.0222	3380	0.08	0.15	1100	
Chloro2_161006	40	08/10/2016	hsqctgp	700.16338	4	294.9623	2048	2050	10.986446752339	3380	1.5	12.5	176.06593532232		512	150	10563.32232	0.06	-	-	
Chloro2_161006	50	08/10/2016	hmbcgplndqf	700.16338	48	294.9623	4096	2050	10.986446752339	3380	1.5	12.5	176.06593532232		400	150	10563.32232	0.06	-	-	
Chloro3_161006	1	06/10/2016	zg	700.16338	64	295.0593	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro3_161006	10	06/10/2016	cosydfp	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		512	12.0222	3380	0.08	-	-	
Chloro3_161006	2	07/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro3_161006	20	06/10/2016	dipsi2p	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		400	12.0222	3380	0.08	-	-	
Chloro3_161006	30	06/10/2016	dsteppg3spr	700.16338	32	294.9623	4096	256	12.0222060421218	3380	1.5	12.5	700.16338		50	12.0222	3380	0.08	0.15	1000	
Chloro3_161006	40	06/10/2016	hsqctgp	700.16338	4	294.9623	2048	2050	10.986446752339	3380	1.5	12.5	176.06593532232		512	150	10563.32232	0.06	-	-	
Chloro3_161006	50	07/10/2016	hmbcgplndqf	700.16338	48	294.9623	4096	2050	10.986446752339	3380	1.5	12.5	176.06593532232		400	150	10563.32232	0.06	-	-	
Chloro4_161006	1	08/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro4_161006	10	08/10/2016	cosydfp	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		512	12.0222	3380	0.08	-	-	
Chloro4_161006	2	09/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro4_161006	20	09/10/2016	dipsi2p	700.16338	8	295.0593	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		400	12.0222	3380	0.08	-	-	
Chloro4_161006	31	09/10/2016	dsteppg3spr	700.16338	64	294.9623	4096	256	12.0222060421218	3380	1.5	12.5	700.16338		50	12.0222	3380	0.08	0.15	1100	
Chloro4_161006	40	09/10/2016	hsqctgp	700.16338	4	295.0593	2048	2050	10.986446752339	3380	1.5	12.5	176.06593532232		512	150	10563.32232	0.06	-	-	
Chloro4_161006	50	09/10/2016	hmbcgplndqf	700.16338	48	294.9623	4096	2050	10.986446752339	3380	1.5	12.5	176.06593532232		400	150	10563.32232	0.06	-	-	
Chloro5_161006	1	09/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro5_161006	10	09/10/2016	cosydfp	700.16338	8	295.0593	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		512	12.0222	3380	0.08	-	-	
Chloro5_161006	2	10/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chloro5_161006	20	09/10/2016	dipsi2p	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		400	12.0222	3380	0.08	-	-	
Chloro5_161006	3	11/10/2016	zg	700.16338	64	294.9623	16384	8	12.0222060421218	3380	1.5	12.5	-								
Chloro5_161006	31	09/10/2016	dsteppg3spr	700.16338	64	294.9623	4096	256	12.0222060421218	3380	1.5	12.5	700.16338		50	12.0222	3380	0.08	0.15	1100	
Chloro5_161006	40	09/10/2016	hsqctgp	700.16338	4	294.9623	2048	2050	10.986446752339	3380	1.5	12.5	176.06593532232		512	150	10563.32232	0.06	-	-	
Chloro5_161006	50	09/10/2016	hmbcgplndqf	700.16338	48	294.9623	4096	2050	10.986446752339	3380	1.5	12.5	176.06593532232		400	150	10563.32232	0.06	-	-	
Chlororef_161006	1	10/10/2016	zg	700.16338	64	294.9623	16384	64	12.0222060421218	3380	1.5	12.5	-								
Chlororef_161006	10	10/10/2016	cosydfp	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		512	12.0222	3380	0.08	-	-	
Chlororef_161006	2	11/10/2016	zg	700.16338	64	294.9623	16384	16	12.0222060421218	3380	1.5	12.5	-								
Chlororef_161006	20	10/10/2016	dipsi2p	700.16338	8	294.9623	4096	18	12.0222060421218	3380	1.5	12.5	700.16338		400	12.0222	3380	0.08	-	-	
Chlororef_161006	31	10/10/2016	dsteppg3spr	700.16338	64	294.9623	4096	256	12.0222060421218	3380	1.5	12.5	700.16338		50	12.0222	3380	0.08	0.15	1100	
Chlororef_161006	40	10/10/2016	hsqctgp	700.16338	4	294.9623	2048	2050	10.986446752339	3380	1.5	12.5	176.06593532232		512	150	10563.32232	0.06	-	-	
Chlororef_161006	50	10/10/2016	hmbcgplndqf	700.16338	48	295.0593	4096	2050	10.986446752339	3380	1.5	12.5	176.06593532232		400	150	10563.32232	0.06	-	-	

# report from ../DECPHA/CHLOROnew/Results				
manip	expno	type	file	content
Chloro1_161006		1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro1_161006		1 1D	1_peaklist.csv	423 in Peak list
Chloro1_161006		2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro1_161006		2 1D	2_peaklist.csv	450 in Peak list
Chloro1_161006		10 cosydfphpr	cosydfphpr_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
Chloro1_161006		10 cosydfphpr	cosydfphpr_10_peaklist.csv	422 in Peak list
Chloro1_161006		20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
Chloro1_161006		20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	190 in Peak list
Chloro1_161006		31 dstebpgp3spr	dstebpgp3spr_31_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
Chloro1_161006		31 dstebpgp3spr	dstebpgp3spr_31_peaklist.csv	150 in Peak list
Chloro1_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
Chloro1_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	305 in Peak list
Chloro1_161006		40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
Chloro1_161006		40 hsqcetgp	hsqcetgp_40_peaklist.csv	92 in Peak list
Chloro2_161006		1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro2_161006		1 1D	1_peaklist.csv	322 in Peak list
Chloro2_161006		2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro2_161006		2 1D	2_peaklist.csv	351 in Peak list
Chloro2_161006		10 cosydfphpr	cosydfphpr_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
Chloro2_161006		10 cosydfphpr	cosydfphpr_10_peaklist.csv	446 in Peak list
Chloro2_161006		20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
Chloro2_161006		20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	169 in Peak list
Chloro2_161006		31 dstebpgp3spr	dstebpgp3spr_31_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
Chloro2_161006		31 dstebpgp3spr	dstebpgp3spr_31_peaklist.csv	160 in Peak list
Chloro2_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
Chloro2_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	256 in Peak list
Chloro2_161006		40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
Chloro2_161006		40 hsqcetgp	hsqcetgp_40_peaklist.csv	77 in Peak list
Chloro3_161006		1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro3_161006		1 1D	1_peaklist.csv	90 in Peak list
Chloro3_161006		2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro3_161006		2 1D	2_peaklist.csv	237 in Peak list
Chloro3_161006		10 cosydfphpr	cosydfphpr_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
Chloro3_161006		10 cosydfphpr	cosydfphpr_10_peaklist.csv	726 in Peak list
Chloro3_161006		20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
Chloro3_161006		20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	191 in Peak list
Chloro3_161006		30 dstebpgp3spr	dstebpgp3spr_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
Chloro3_161006		30 dstebpgp3spr	dstebpgp3spr_30_peaklist.csv	247 in Peak list
Chloro3_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
Chloro3_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	242 in Peak list
Chloro3_161006		40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
Chloro3_161006		40 hsqcetgp	hsqcetgp_40_peaklist.csv	85 in Peak list
Chloro4_161006		1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro4_161006		1 1D	1_peaklist.csv	164 in Peak list
Chloro4_161006		2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro4_161006		2 1D	2_peaklist.csv	223 in Peak list
Chloro4_161006		10 cosydfphpr	cosydfphpr_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
Chloro4_161006		10 cosydfphpr	cosydfphpr_10_peaklist.csv	550 in Peak list
Chloro4_161006		20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
Chloro4_161006		20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	316 in Peak list
Chloro4_161006		31 dstebpgp3spr	dstebpgp3spr_31_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
Chloro4_161006		31 dstebpgp3spr	dstebpgp3spr_31_peaklist.csv	282 in Peak list
Chloro4_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
Chloro4_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	323 in Peak list
Chloro4_161006		40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
Chloro4_161006		40 hsqcetgp	hsqcetgp_40_peaklist.csv	95 in Peak list
Chloro5_161006		1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro5_161006		1 1D	1_peaklist.csv	15 in Peak list
Chloro5_161006		2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro5_161006		2 1D	2_peaklist.csv	16 in Peak list
Chloro5_161006		3 1D	3_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chloro5_161006		3 1D	3_peaklist.csv	41 in Peak list
Chloro5_161006		10 cosydfphpr	cosydfphpr_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
Chloro5_161006		10 cosydfphpr	cosydfphpr_10_peaklist.csv	496 in Peak list
Chloro5_161006		20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
Chloro5_161006		20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	304 in Peak list
Chloro5_161006		31 dstebpgp3spr	dstebpgp3spr_31_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
Chloro5_161006		31 dstebpgp3spr	dstebpgp3spr_31_peaklist.csv	128 in Peak list
Chloro5_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
Chloro5_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	215 in Peak list
Chloro5_161006		40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
Chloro5_161006		40 hsqcetgp	hsqcetgp_40_peaklist.csv	18 in Peak list
Chlororef_161006		1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chlororef_161006		1 1D	1_peaklist.csv	44 in Peak list
Chlororef_161006		2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
Chlororef_161006		2 1D	2_peaklist.csv	162 in Peak list
Chlororef_161006		10 cosydfphpr	cosydfphpr_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
Chlororef_161006		10 cosydfphpr	cosydfphpr_10_peaklist.csv	274 in Peak list
Chlororef_161006		20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
Chlororef_161006		20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	352 in Peak list
Chlororef_161006		31 dstebpgp3spr	dstebpgp3spr_31_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
Chlororef_161006		31 dstebpgp3spr	dstebpgp3spr_31_peaklist.csv	45 in Peak list
Chlororef_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
Chlororef_161006		50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	134 in Peak list
Chlororef_161006		40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
Chlororef_161006		40 hsqcetgp	hsqcetgp_40_peaklist.csv	14 in Peak list

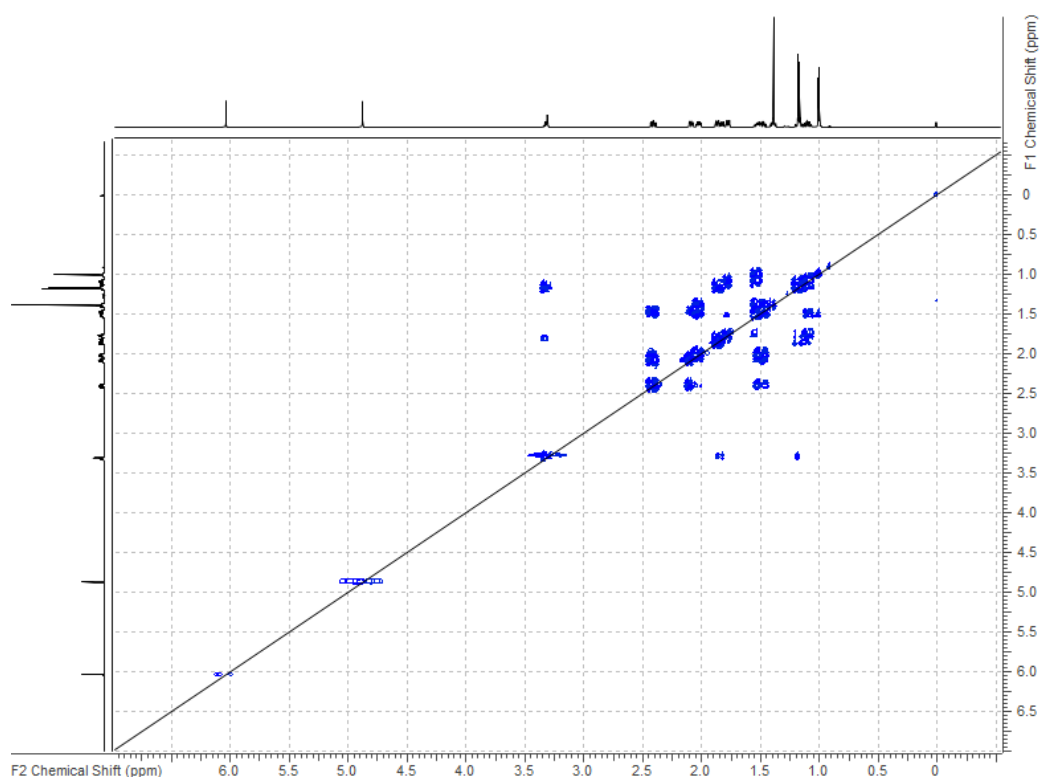
Annexe 3 : Spectres ^1H , ^1H - ^1H COSY, ^1H - δ DOSY, ^1H - ^1H TOCSY, ^1H - ^{13}C HSQC et ^1H - ^{13}C HMBC de l'artémisinine



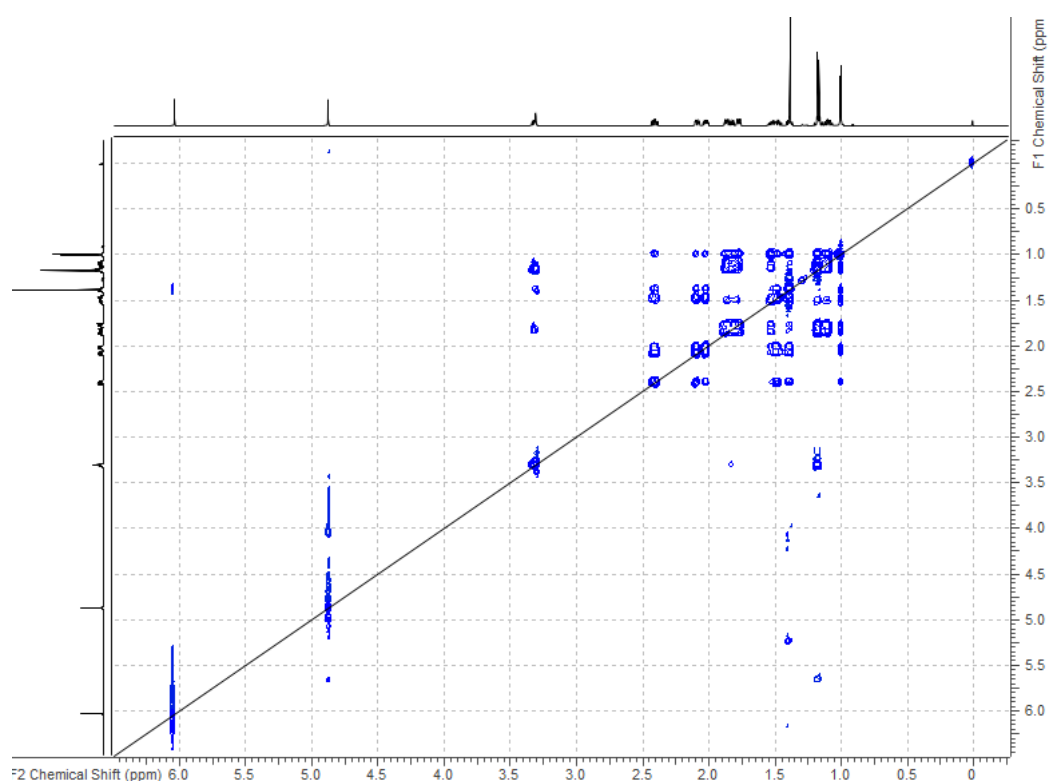
Spectre ^1H de l'artémisinine (700 MHz, MeOD)



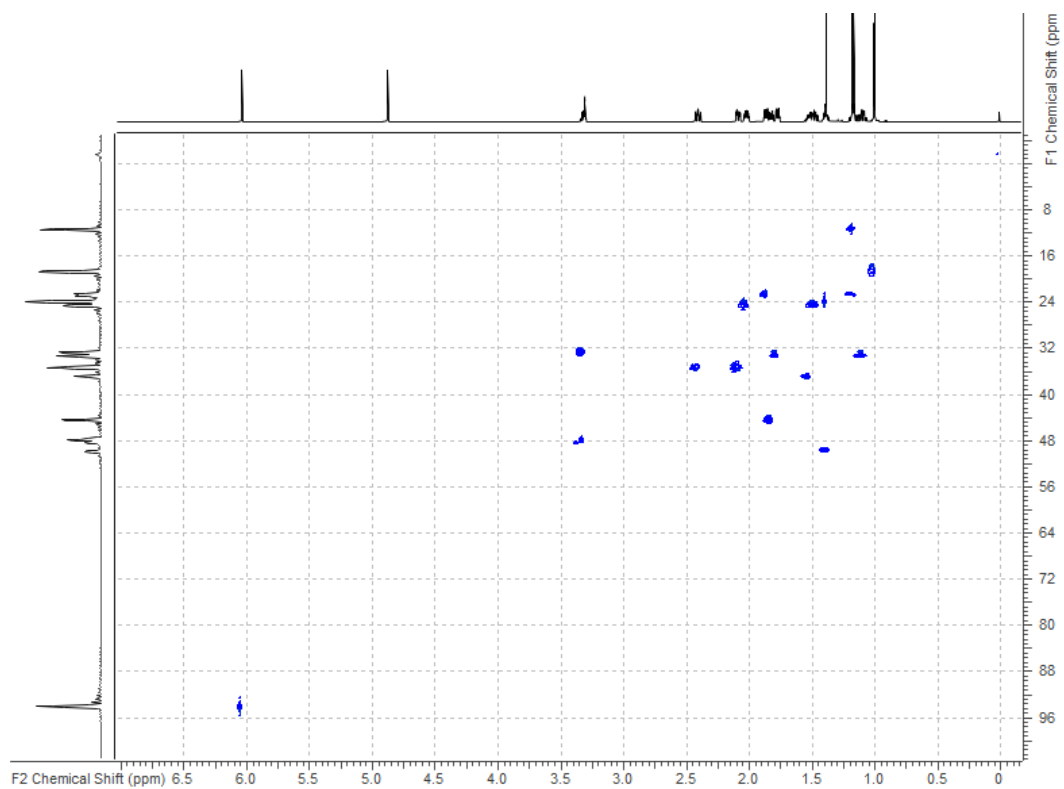
Spectre ^1H - δ DOSY de l'artémisinine (700 MHz, MeOD)



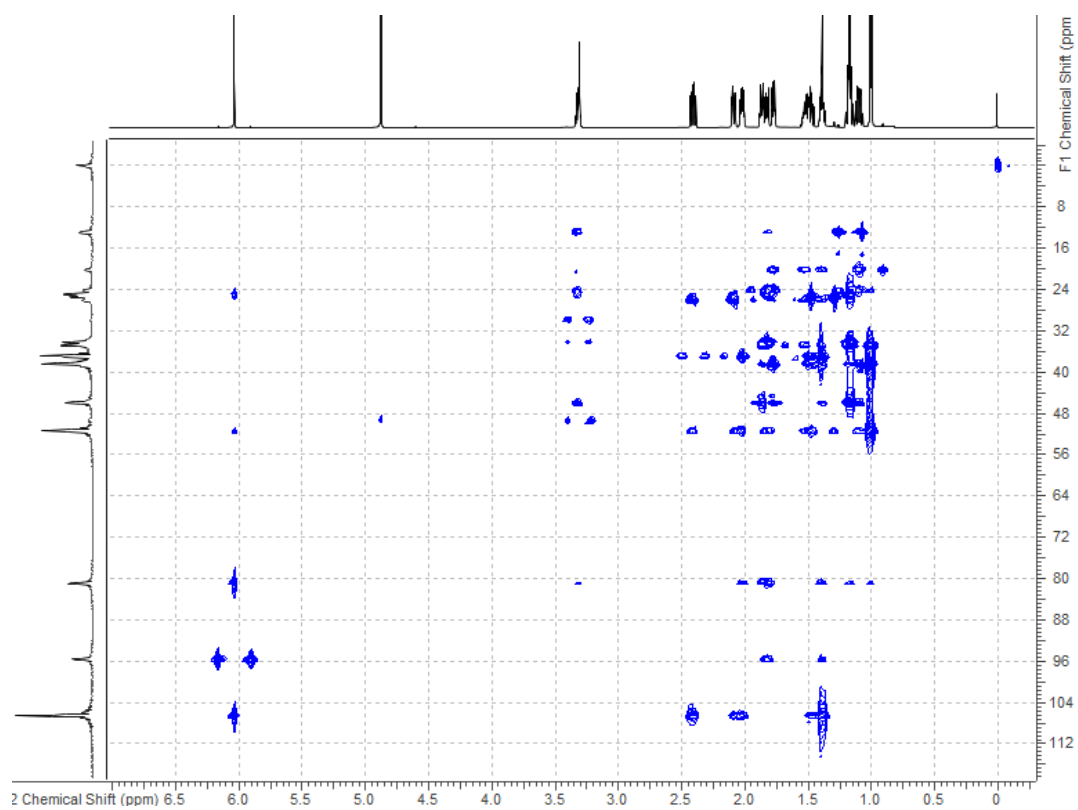
Spectre ^1H - ^1H COSY de l'artémisinine (700 MHz, MeOD)



Spectre ^1H - ^1H TOCSY de l'artémisinine (700 MHz, MeOD)



Spectre ^1H - ^{13}C HSQC de l'artémisinine (700 MHz, MeOD)



Spectre ^1H - ^{13}C HMBC de l'artémisinine (700 MHz, MeOD)

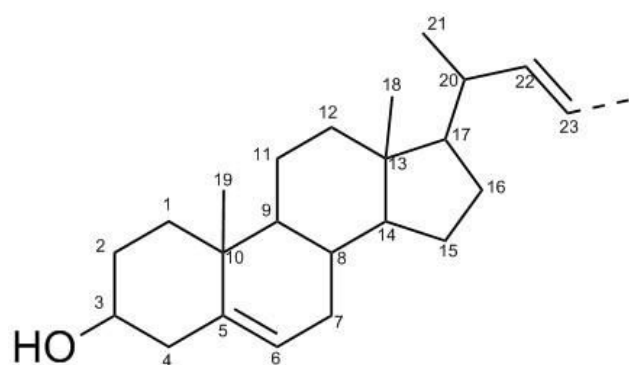
S

# report from ../DECPHA/ARTeNew/Results				
manip	expno	type	file	content
ARTE1_160816	1	1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE1_160816	1	1D	1_peaklist.csv	1355 in Peak list
ARTE1_160816	2	1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE1_160816	2	1D	2_peaklist.csv	1431 in Peak list
ARTE1_160816	10	cosydfph	cosydfph_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
ARTE1_160816	10	cosydfph	cosydfph_10_peaklist.csv	155 in Peak list
ARTE1_160816	20	dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
ARTE1_160816	20	dipsi2phpr	dipsi2phpr_20_peaklist.csv	69 in Peak list
ARTE1_160816	30	dstebpgp3s	dstebpgp3s_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
ARTE1_160816	30	dstebpgp3s	dstebpgp3s_30_peaklist.csv	133 in Peak list
ARTE1_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
ARTE1_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	116 in Peak list
ARTE1_160816	40	hsqctgtp	hsqctgtp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTE1_160816	40	hsqctgtp	hsqctgtp_40_peaklist.csv	52 in Peak list
ARTE2_160816	1	1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE2_160816	1	1D	1_peaklist.csv	605 in Peak list
ARTE2_160816	2	1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE2_160816	2	1D	2_peaklist.csv	431 in Peak list
ARTE2_160816	3	1D	3_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE2_160816	3	1D	3_peaklist.csv	650 in Peak list
ARTE2_160816	10	cosydfph	cosydfph_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
ARTE2_160816	10	cosydfph	cosydfph_10_peaklist.csv	177 in Peak list
ARTE2_160816	20	dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
ARTE2_160816	20	dipsi2phpr	dipsi2phpr_20_peaklist.csv	80 in Peak list
ARTE2_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
ARTE2_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	227 in Peak list
ARTE2_160816	40	hsqctgtp	hsqctgtp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTE2_160816	40	hsqctgtp	hsqctgtp_40_peaklist.csv	61 in Peak list
ARTE2_160816	30	ledbpgp2s	ledbpgp2s_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
ARTE2_160816	30	ledbpgp2s	ledbpgp2s_30_peaklist.csv	250 in Peak list
ARTE3_160816	1	1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE3_160816	1	1D	1_peaklist.csv	658 in Peak list
ARTE3_160816	3	1D	3_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE3_160816	3	1D	3_peaklist.csv	621 in Peak list
ARTE3_160816	10	cosydfph	cosydfph_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
ARTE3_160816	10	cosydfph	cosydfph_10_peaklist.csv	206 in Peak list
ARTE3_160816	20	dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
ARTE3_160816	20	dipsi2phpr	dipsi2phpr_20_peaklist.csv	95 in Peak list
ARTE3_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 2.83 x 11.18 data points
ARTE3_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	64 in Peak list
ARTE3_160816	40	hsqctgtp	hsqctgtp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTE3_160816	40	hsqctgtp	hsqctgtp_40_peaklist.csv	60 in Peak list
ARTE3_160816	30	ledbpgp2s	ledbpgp2s_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
ARTE3_160816	30	ledbpgp2s	ledbpgp2s_30_peaklist.csv	207 in Peak list
ARTE4_160816	1	1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE4_160816	1	1D	1_peaklist.csv	821 in Peak list
ARTE4_160816	2	1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTE4_160816	2	1D	2_peaklist.csv	840 in Peak list
ARTE4_160816	10	cosydfph	cosydfph_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
ARTE4_160816	10	cosydfph	cosydfph_10_peaklist.csv	170 in Peak list
ARTE4_160816	20	dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
ARTE4_160816	20	dipsi2phpr	dipsi2phpr_20_peaklist.csv	84 in Peak list
ARTE4_160816	30	dstebpgp3s	dstebpgp3s_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
ARTE4_160816	30	dstebpgp3s	dstebpgp3s_30_peaklist.csv	207 in Peak list
ARTE4_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
ARTE4_160816	50	hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	356 in Peak list
ARTE4_160816	40	hsqctgtp	hsqctgtp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTE4_160816	40	hsqctgtp	hsqctgtp_40_peaklist.csv	53 in Peak list

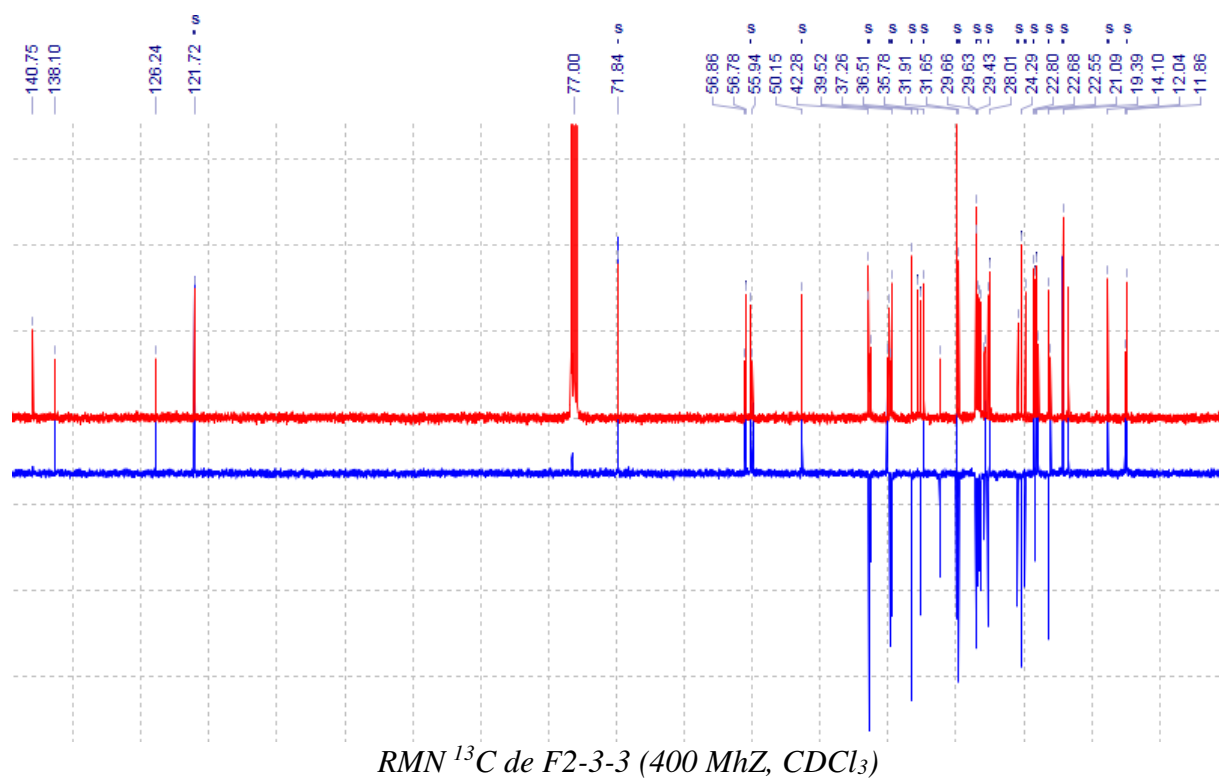
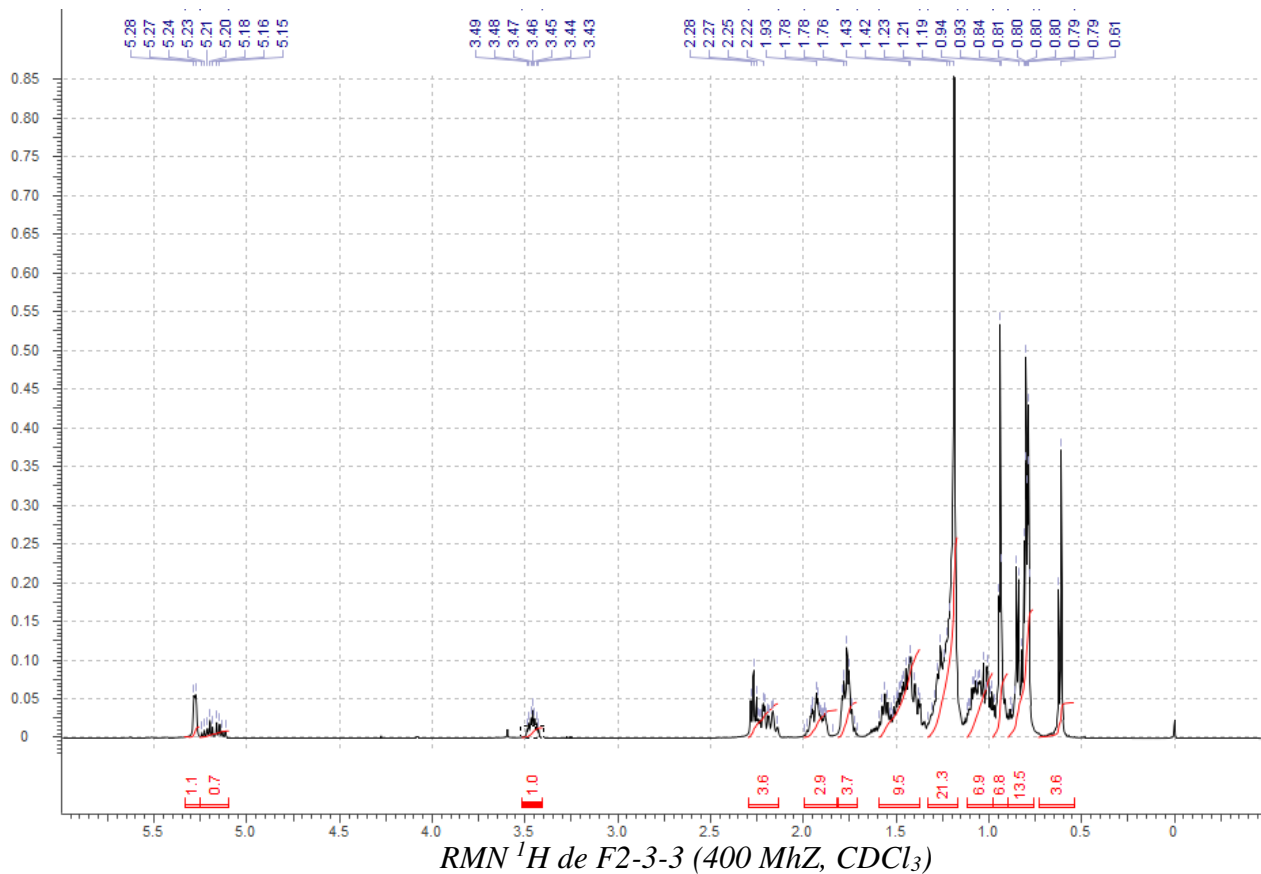
Suite page suivante

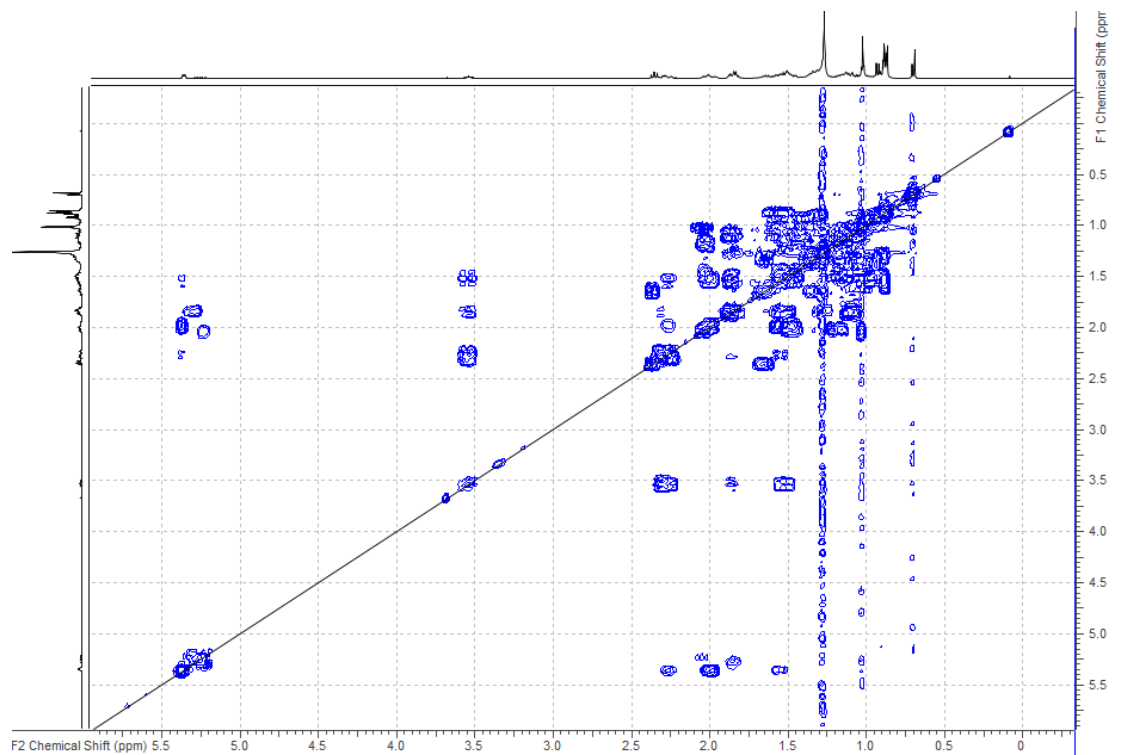
ARTES_160816	1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTES_160816	1 1D	1_peaklist.csv	441 in Peak list
ARTES_160816	2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTES_160816	2 1D	2_peaklist.csv	638 in Peak list
ARTES_160816	10 cosydfph	cosydfph_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
ARTES_160816	10 cosydfph	cosydfph_10_peaklist.csv	235 in Peak list
ARTES_160816	20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
ARTES_160816	20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	145 in Peak list
ARTES_160816	50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
ARTES_160816	50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	281 in Peak list
ARTES_160816	40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTES_160816	40 hsqcetgp	hsqcetgp_40_peaklist.csv	118 in Peak list
ARTES_160816	41 hsqcetgp	hsqcetgp_41_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTES_160816	41 hsqcetgp	hsqcetgp_41_peaklist.csv	79 in Peak list
ARTES_160816	30 ledbpgp2s	ledbpgp2s_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
ARTES_160816	30 ledbpgp2s	ledbpgp2s_30_peaklist.csv	184 in Peak list
ARTEref	1 1D	1_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTEref	1 1D	1_peaklist.csv	546 in Peak list
ARTEref	2 1D	2_bucketlist.csv	901 buckets with a mean size of 13.63 data points
ARTEref	2 1D	2_peaklist.csv	535 in Peak list
ARTEref	10 cosydfph	cosydfph_10_bucketlist.csv	98596 rectangular buckets with a mean size of 2.56 x 10.22 data points
ARTEref	10 cosydfph	cosydfph_10_peaklist.csv	224 in Peak list
ARTEref	20 dipsi2phpr	dipsi2phpr_20_bucketlist.csv	98596 rectangular buckets with a mean size of 2.00 x 10.22 data points
ARTEref	20 dipsi2phpr	dipsi2phpr_20_peaklist.csv	270 in Peak list
ARTEref	30 dstebpgp3s	dstebpgp3s_30_bucketlist.csv	9734 rectangular buckets with a mean size of 2.13 x 5.11 data points
ARTEref	30 dstebpgp3s	dstebpgp3s_30_peaklist.csv	451 in Peak list
ARTEref	50 hmbcgplpndqf	hmbcgplpndqf_50_bucketlist.csv	46786 rectangular buckets with a mean size of 10.67 x 11.18 data points
ARTEref	50 hmbcgplpndqf	hmbcgplpndqf_50_peaklist.csv	519 in Peak list
ARTEref	40 hsqcetgp	hsqcetgp_40_bucketlist.csv	46786 rectangular buckets with a mean size of 6.83 x 5.59 data points
ARTEref	40 hsqcetgp	hsqcetgp_40_peaklist.csv	195 in Peak list

Annexe 5 : Spectres RMN de F2-3-3

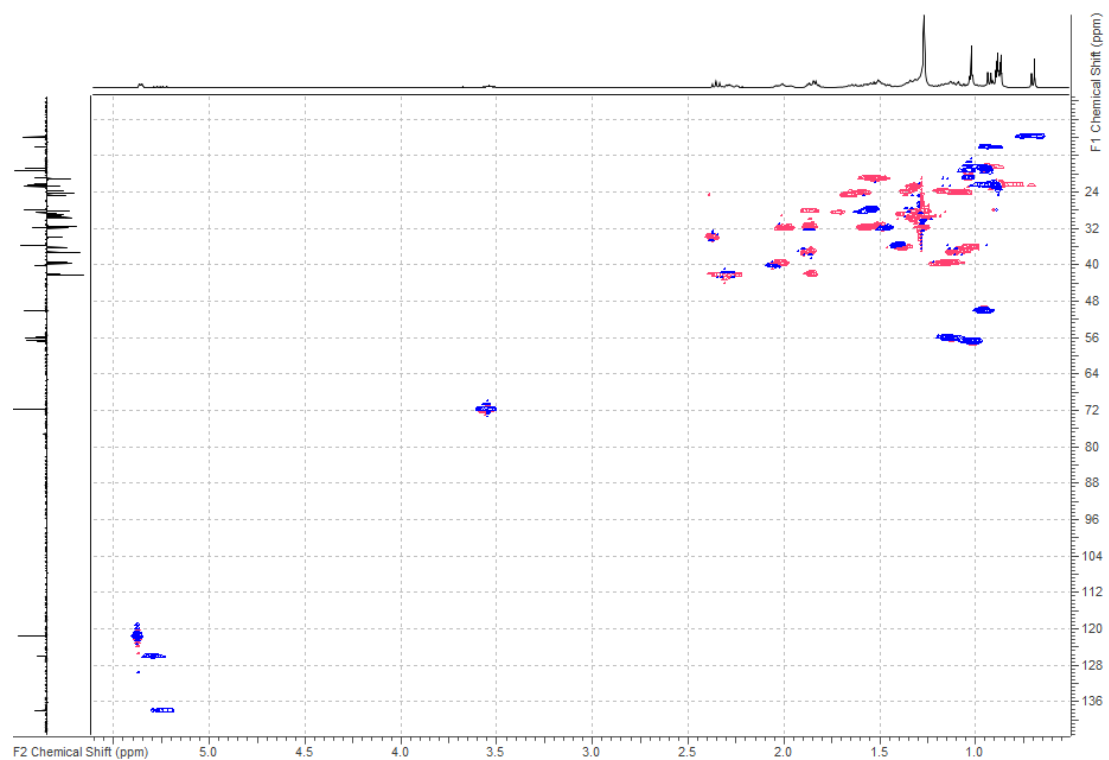


Position	δ_H (ppm)	δ_C (ppm)
1a	1.09	37.2
1b	1.86	''
2a	1.84	31.6
2b	1.53	31.6
3	3.61 (m, 1H)	71.9
4a	2.25	42.1
4b	2.23	''
5	-	140.7
6	5.35	121.7
7a	1.96	32.0
7b	1.54	''
8	1.45	31.9
9	0.94	50.0
10	-	36.5
11a	1.53	21.1
11b	1.47	''
12a	1.14	39.6
12b	2.01	''
13	-	42.3
14	1.00	56.9
15a	1.57	24.2
15b	1.07	''
16a	1.85	28.3
16b	1.25	''
17	1.10	56.2
18	0.7	12.0
19	1.02	19.4
20	2.03	40.0
21	?	?
22	5.22 (m, 1H)	137.8
23	5.29 (m, 1H)	126.0
24	1.82	41.6

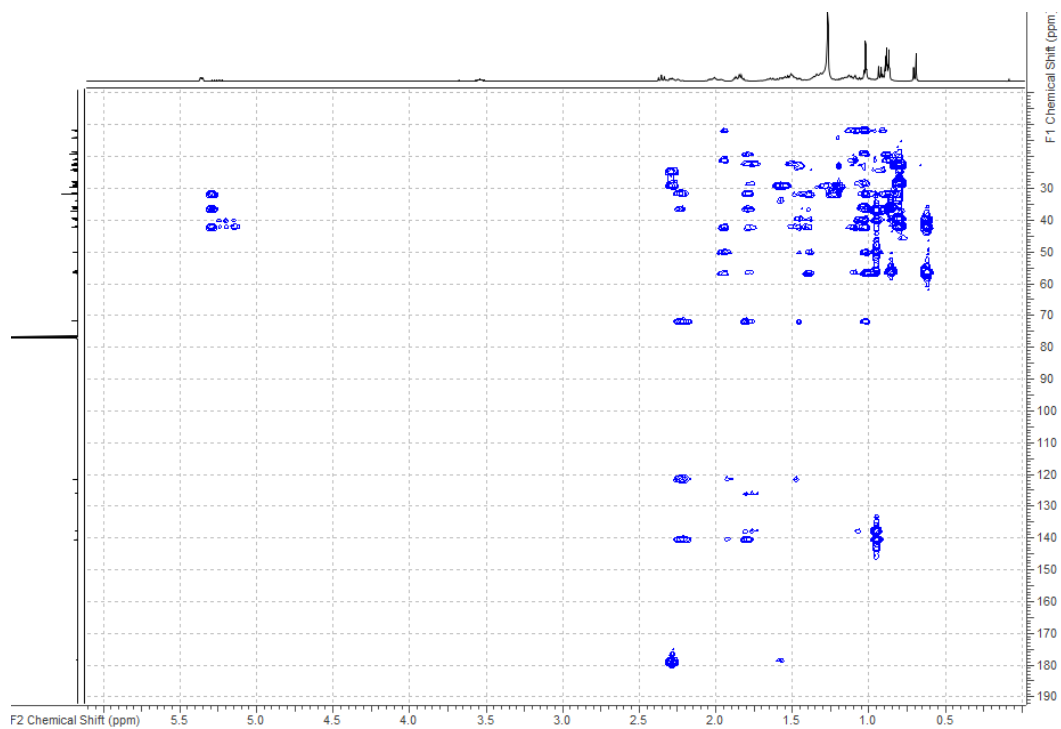




RMN ^1H - ^1H COSY de F2-3-3 (400 MhZ, CDCl_3)



RMN ^1H - ^{13}C HSQC DEPT de F2-3-3 (400 MhZ, CDCl_3)



RMN ^1H - ^{13}C HMBC de F2-3-3 (400 MHz, CDCl_3)

Annexe 6 : BucketUtilities_LM

```

# coding: utf8
"""
The following code contains all the routines used in the analysis of the 2D
bucket lists generated by Spike.
The LM version contains improvements for great and faster homonuclear and
heteronuclear spectra displays
L Margueritte & M-A Delsuc,
"""
from __future__ import print_function

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.cm as cm

NETMODE = 'mieux' # standard / mieux / encore

def loadInt2D(epath, net=False, sym=False):
    """loads intensities from a csv bucket-list file from 2D spectra
    net: determines whether the cleaning method is used
        the method used is defined by NETMODE global
    sym: whether symetrisation is used
    """
    # read the file
    nel = pd.read_csv( epath, header=1, sep = ', ', usecols=[0, 1, 2],
engine='python')
    x1 = np.array(nel['centerF1'])
    xul = np.unique(x1)
    y1 = np.array(nel['centerF2'])
    yul = np.unique(y1)
    z1 = np.array(nel['bucket'])
    # matrix calculation
    Xr1, Yr1 = np.meshgrid(yul, xul)
    Zr1 = np.reshape(z1, (len(xul), len(yul)))
    netmode = NETMODE
    if net:
        if netmode=='standard':
            Zr1 = nettoie(Zr1)
        elif netmode=='mieux':
            Zr1 = nettoie_mieux(Zr1)
        elif netmode=='encore':
            Zr1 = nettoie_encore_mieux(Zr1)
        else:
            raise Exception(netmode + ' : Wrong netmode !')
    if sym:
        Zr1 = symetrise(Zr1)
    return [Xr1, Yr1, np.nan_to_num(Zr1)]

```

```

def loadStd2D(epath, net=False, sym=False):
    """loads std from a csv bucket-list file from 2D spectra
    net: determines whether the cleaning method is used
        the method used is defined by NETMODE global
    sym: whether symetrisation is used
    """
    nel = pd.read_csv( epath, header=1, sep = ', ', usecols=[0, 1, 5],
engine='python')
    x1 = np.array(nel['centerF1'])
    xul = np.unique(x1)
    y1 = np.array(nel['centerF2'])
    yul = np.unique(y1)
    z1 = np.array(nel['std'])

    Xr1, Yr1 = np.meshgrid(yul, xul)
    Zr1 = np.reshape(z1, (len(xul), len(yul)))
    netmode = NETMODE
    if net:
        if netmode=='standard':
            Zr1 = nettoie(Zr1)
        elif netmode=='mieux':
            Zr1 = nettoie_mieux(Zr1)
        elif netmode=='encore':
            Zr1 = nettoie_encore_mieux(Zr1)
        else:
            raise Exception(netmode + ' : Wrong netmode !')
    if sym:
        Zr1 = symetrise(Zr1)
    return [Xr1, Yr1, np.nan_to_num(Zr1)]

def loadInt1D (epath) :
    """loads std from a csv bucket-list file from 1D spectra"""
    nel = pd.read_csv(epath, header=1, sep = ', ', usecols=[0, 1],
engine='python')
    X = np.array(nel['center'])
    Y = np.array(nel['bucket'])
    return [X, Y]

def loadStd1D (epath) :
    """loads std from a csv bucket-list file from 1D spectra"""
    nel = pd.read_csv(epath, header=1, sep = ', ', usecols=[0, 4],
engine='python')
    X = np.array(nel['center'])
    Y = np.array(nel['std'])
    return [X, Y]

def affiche1D(X, Y, scale=1.0) :
    """draw the 1D bucket list of homonuclear experiment"""
    fig = plt.figure()
    ax1 = fig.add_subplot(111)
    ax1.plot(X,Y, c='b')
    ax1.invert_xaxis()
    ax1.set_xlabel(r"ppm")
    ax1.set_ylabel(r"au")
    major_ticks = np.arange(0, 10, 0.5)
    minor_ticks = np.arange(0, 9.5, 0.25)
    ax1.set_xticks(major_ticks)
    ax1.set_xticks(minor_ticks, minor =True)
    return ax1

```

```

def affiche(X, Y, Z, scale=1.0, new=True, cmap=None, reverse=True,
figsize=(10, 8)) :
    """draw the 2D bucket list of homonuclear experiment
    To display two spectra in the same picture, use the parameters : new
    cmap : colormap of matplotlib, example : cm.winter or cm.spring"""
    if new:
        f1, ax = plt.subplots(figsize=(10,8))
    else:
        ax = plt.gca()
# else should be a drawable matplotlib axis.
levelbase = [0.5,1,2,5,10,20,50,100]
m1 = Z.max()
level = [m1*(i/100.0)/scale for i in levelbase ]
ax.contour(X, Y, Z, level, cmap=cmap)
ax.set_xlabel(r"$\delta$ (ppm)$")
ax.set_ylabel(r"$\delta$ (ppm)$")
ax.yaxis.set_label_position('right')
major_ticks = np.arange(0, 9.5, 0.5)
minor_ticks = np.arange(0, 9.5, 0.1)
ax.yaxis.set_ticks_position('right')
ax.set_xticks(major_ticks)
ax.set_xticks(minor_ticks, minor =True)
ax.set_yticks(major_ticks)
ax.set_yticks(minor_ticks, minor=True)
if reverse:
    ax.invert_xaxis()
    ax.invert_yaxis()
return ax

def affiche2(X, Y, Z, scale=1.0, new=True, cmap=None, reverse=True):
    """draw the 2D bucket list of heteronuclear experiment"""
    if new:
        f1, ax1 = plt.subplots(figsize=(10, 8))
    else:
        ax1 = plt.gca()
levelbase = [0.5,1,2,5,10,20,50,100]
m1 = Z.max()
level = [m1*(i/100.0)/scale for i in levelbase ]
ax1.contour(X, Y, Z, level, cmap= cmap)
ax1.set_xlabel(r"$\delta$ (ppm)$")
ax1.set_ylabel(r'$\delta$ (ppm)$')
major_ticksx = np.arange(0, 10, 0.5)
minor_ticksx = np.arange(0, 10, 0.1)
major_ticksy = np.arange(-15, 140, 10)
minor_ticksy = np.arange(-15, 140, 2)
ax1.set_xticks(major_ticksx)
ax1.set_xticks(minor_ticksx, minor = True)
ax1.set_yticks(major_ticksy)
ax1.set_yticks(minor_ticksy, minor = True)
if reverse:
    ax1.invert_xaxis()
    ax1.invert_yaxis()
return ax1

```

```

def affichegrid(X, Y, Z, scale=1.0, new=True, cmap=None, reverse=True):
    """draw the 2D bucket list of homonuclear experiment and a grid (if you
    need ;- ) """
    if new:
        f1, ax1 = plt.subplots(figsize=(10, 8))
    else:
        ax1 = plt.gca()
    levelbase = [0.5,1,2,5,10,20,50,100]
    m1 = Z.max()
    level = [m1*(i/100.0)/scale for i in levelbase ]
    ax1.contour(X, Y, Z, level, cmap= cmap)
    ax1.set_xlabel(r"$\delta$ (ppm)")
    ax1.set_ylabel(r'$\delta$ (ppm)')
    major_ticks = np.arange(0, 10, 0.5)
    minor_ticks = np.arange(0, 10, 0.1)
    ax1.set_xticks(major_ticks)
    ax1.set_xticks(minor_ticks, minor = True)
    ax1.set_yticks(major_ticks)
    ax1.set_yticks(minor_ticks, minor = True)
    #ax1.grid(b = True, which = 'major', axis = 'both')
    #ax1.grid(b = True, which = 'minor', axis = 'both')
    #ax1.set_xlim(xmax=11,xmin=0)
    #ax1.set_ylim(ymax=11,ymin=0)
    ax1.grid(which='major', alpha=1)
    ax1.grid(which='minor', alpha=0.3)
    if reverse:
        ax1.invert_xaxis()
        ax1.invert_yaxis()
    return ax1

def affratio(I1, I2, new=True, scale=1.0, cmap=None, reverse=True):
    "display the ratio of two pictures from homonuclear experiments"
    affiche(I1[0], I1[1], I1[2]/(I2[2]+1e-5), scale=scale, new=new,
    cmap=cmap, reverse=reverse)

def affratio2(I1, I2, new=True, scale=1.0, cmap=None, reverse=True):
    "display the ratio of two pictures from heteronuclear experiments"
    # il faudrait tester que les X Y sont les même !!!
    affiche2(I1[0], I1[1], I1[2]/(I2[2]+1e-5), scale=scale, new=new,
    cmap=cmap, reverse=reverse)

def affsub(I1, I2, new=True, scale=1.0, cmap=None, reverse=True):
    "display the substraction of two pictures from homonuclear experiments"
    # il faudrait tester que les X Y sont les même !!!
    affiche(I1[0], I1[1], I1[2]-I2[2], scale=scale, new=new, cmap=cmap,
    reverse=reverse)

def affsub2(I1, I2, new=True, scale=1.0, cmap=None, reverse=True):
    "display the substraction of two pictures from homonuclear experiments"
    # il faudrait tester que les X Y sont les même !!!
    affiche2(I1[0], I1[1], I1[2]-I2[2], scale=scale, new=new, cmap=cmap,
    reverse=reverse)

def symetrise(ZZ):
    "symetrisation os spectra - simple minimum operation"
    return np.minimum(ZZ, ZZ.T)

```

```
def nettoie(ZZ, factor=2.0):
    " clean noise in matrix - hard thresholding"
    ZZr = ZZ.copy()
    thresh = factor*np.median(ZZ)
    print( thresh)
    ZZr[ZZ<thresh] = 1.0    # 1.0 allows to display log(Z) !
    return ZZr

def nettoie_mieux(ZZ, factor=2.0):
    " clean noise in matrix - hard thresholding columnwise"
    ZZr = ZZ.copy()
    for i in range(ZZ.shape[1]):
        iZZ = ZZ[:,i]
        thresh = factor*np.median(iZZ)
        ZZr[iZZ<thresh,i] = 1.0
    return ZZr

def nettoie_encore_mieux(ZZ, factor=2.0):
    " clean noise in matrix - soft thresholding columnwise"
    ZZr = ZZ.copy()
    for i in range(ZZ.shape[1]):
        iZZ = ZZ[:,i]
        thresh = factor*np.median(iZZ)
        ZZr[:,i] = np.where(iZZ<thresh, 1, iZZ-thresh+1)
    return ZZr

def normalize(Z):
    "normalise les histogrammes"
    ZZ = np.log(Z)
    mu = ZZ.mean()
    sigma = ZZ.std()
```

Annexe 7 : Publication


Received: 24 July 2017 | Revised: 20 October 2017 | Accepted: 3 November 2017

DOI: 10.1002/mrc.4683

SPECIAL ISSUE RESEARCH ARTICLE

WILEY

Automatic differential analysis of NMR experiments in complex samples

Laure Margueritte¹ | Petar Markov² | Lionel Chiron³ | Jean-Philippe Starck³ |
Catherine Vonthron-Sénécheau¹ | Mélanie Bourjot¹ | Marc-André Delsuc⁴ ¹Laboratoire d'Innovation Thérapeutique (LIT) UMR CNRS 7200, LabEx Medalis, Faculté de Pharmacie, Université de Strasbourg, Strasbourg, France²Structural Biophysics Group, School of Optometry and Vision Sciences, Cardiff University, Cardiff, UK³CASC4DE Le Lodge, 20, Avenue du Neuhoef, Strasbourg 67100, France⁴Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U596, CNRS UMR 7104, Université de Strasbourg, Strasbourg, France**Correspondence**Marc-André Delsuc, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U596, CNRS UMR 7104, Université de Strasbourg, Strasbourg, France.
Email: delsuc@igbmc.fr

Liquid state nuclear magnetic resonance (NMR) is a powerful tool for the analysis of complex mixtures of unknown molecules. This capacity has been used in many analytical approaches: metabolomics, identification of active compounds in natural extracts, and characterization of species, and such studies require the acquisition of many diverse NMR measurements on series of samples.

Although acquisition can easily be performed automatically, the number of NMR experiments involved in these studies increases very rapidly, and this data avalanche requires to resort to automatic processing and analysis.

We present here a program that allows the autonomous, unsupervised processing of a large corpus of 1D, 2D, and diffusion-ordered spectroscopy experiments from a series of samples acquired in different conditions. The program provides all the signal processing steps, as well as peak-picking and bucketing of 1D and 2D spectra, the program and its components are fully available. In an experiment mimicking the search of a bioactive species in a natural extract, we use it for the automatic detection of small amounts of artemisinin added to a series of plant extracts and for the generation of the spectral fingerprint of this molecule.

This program called Plasmodesma is a novel tool that should be useful to decipher complex mixtures, particularly in the discovery of biologically active natural products from plants extracts but can also in drug discovery or metabolomics studies.

KEYWORDS

automatic processing, mixture analysis, recursive feature elimination, spectral fingerprint

1 | INTRODUCTION

Liquid state nuclear magnetic resonance (NMR) is a powerful tool for the analysis of mixtures containing unknown molecules. All species in the solution display their NMR spectra, with a signal intensity proportional to their relative concentrations, provided that slow tumbling rates or relaxation agents do not hide the lines by fast relaxation processes. This capacity has been used in many analytical approaches: metabolomics, identification of

bioactive compounds in natural extracts, and characterization of species.^[1-5] See references Larive et al,^[6] Mahrous et Farag^[7] and Wolfender et al^[8] for recent reviews.

Such studies require the acquisition of many diverse NMR measurements on series of samples. Modern NMR spectrometers allow sequential actions (introduction of the sample, probe tuning, and acquisition) in order to produce automatically the corresponding 1D and 2D data-sets.

Unfortunately, if the acquisition is quite easily performed, the access to the final informations is less straightforward: Signal processing (Fourier transform, phasing, baseline correction, and peak detection) and finally spectrum interpretation are not trivial tasks. Moreover, the use of 2D spectra implies more complex steps and additional tasks such as reduction of t_1 -noise and t_1 -ridges or the determination of contour levels for display.

In the case of metabolomics studies or natural extracts screening, the number of NMR experiments increases very rapidly, and this data avalanche calls to resort to automatic processing. Although metabolomics is aimed at measuring precisely the amount of well-known compounds and to quantify precisely their variations from sample to sample, the identification of an active molecule in a natural extract starts with its detection and then its characterization as an unknown compound or as a family of related species.

In this article, we present the specific development of a computer program allowing the autonomous, unsupervised processing of a large corpus of 1D and 2D experiments from a series of samples acquired in different conditions. Results obtained using this program on series of complex natural extracts highlight the time saving and the efficiency increase regarding classical “handmade” processing of raw data.

2 | SOFTWARE DEVELOPMENTS

2.1 | Plasmodesma

The program Plasmodesma^[9] developed for this project relies on the Spectrometry Processing Innovative KERNel (SPIKE) library for most of its operation.^[10] It is intended to process autonomously a large series of different spectra originated from different samples, obtained in varying conditions. This analytical process involves the handling of a complex set of NMR experiments (1D and 2D homonuclear or heteronuclear spectra), at the end, a spectral report summarizing the analysis is expected, containing all figures, peak, and bucket lists for each sample.

The current work is based on SPIKE, a comprehensive software library aimed at the processing and analysis of Fourier transform spectroscopies. It provides basic functionalities such as apodization, Fourier transform, phasing, peak-picking, line-fitting, baseline correction, and more advanced tools. It is easily extensible through a plug-in mechanism. SPIKE combines the use of a parallel multiprocessor approach to a low memory footprint, thus insuring rapid processing with an optimal use of the computer hardware. Moreover, SPIKE allows the efficient handling and visualization of very large data-sets limited only by disk space. SPIKE is a continuation of the

previous Gifa and NPK^[11,12] NMR processing software and was developed to include other Fourier transform spectroscopies, in particular FT Mass spectrometry (Orbitrap and FT-ICR) and 2D-FT-ICR.^[13–15] For portability reasons and ease of development, the program is written in Python and relies on external libraries such as `numpy`, `scipy`, `pandas`, `scikit-learn`, and `hdf5`.^[16–21] SPIKE is freely available at bitbucket.org/delsuc/SPIKE.

2.2 | Principle of operations

The program Plasmodesma operates without any human interaction. When applied to a folder, all NMR files are imported and processed, and a global report is generated for the totality of the analysis. All the processing and analysis steps are optimized depending on the acquisition parameters found in the data-sets, either 1D or 2D data. No other input is required. The 1D and 2D experiments are processed sequentially. The data are apodized, Fourier transformed, and the baseline corrected. 1D spectra are also automatically phased. Additionally, an efficient denoising step^[22] is performed on the 2D experiments, in order to reduce the t_1 -noise. The F1 Fourier transform step of the 2D data-sets is performed depending on the spectral type and on the acquisition protocol. The calibration is then determined precisely from the reference signal (in this case, 0 ppm for the TMS). A peak-picking and bucket analysis are then performed (see below). Peak lists and bucket lists are generated as csv files for each experiment. Finally, figures of each spectrum are created, with and without peaks displayed. A final report that contains all acquisition and processing parameters is generated (see Supporting Information).

2.3 | Specific developments

Some functions used by Plasmodesma have been developed specifically for this analysis and were implemented as SPIKE plug-ins.

2.3.1 | Autophasing

In the context of metabolomics and screening studies, the possibility to detect and quantify precisely the intensity of vanishing small peaks is paramount. The phase of a 1D spectrum, if set slightly off, may have a strong impact on the possibility to detect small signals, in particular close to a large one. Errors of only a few degrees introduce bias resulting to too low or too high quantization, as well as shifts of the maximum. Automatically acquired natural extract spectra are usually difficult to phase because of strong solvent lines and other artifacts

present in the spectra. The improvement of the simple but robust automatic phasing procedure developed in NPK^[12] contributes efficiently to resolve this problem. The principle is to minimize the negative wing of the 1D spectrum, by performing a grid search first on 0th order (frequency independent) alone, then on both 0th and 1st order (frequency dependent) corrections, the larger peak being used as the 1st order pivot. An automatic baseline correction (see below) is performed at each correction step, and an optional *inwater* mode allows to ignore the central spectral zone.

2.3.2 | Autobaseline

A flat baseline is also a requisite for correct analysis, and a specific plug-in has been developed in this respect. We developed a new approach, which relies on an iterative statistical treatment on the signal split into pieces of constant lengths and fitting the baseline by piecewise linear segments. The fit is based on the use of a linear regression minimizing the $\ell_p(x) = (\sum(|x|^p))^{1/p}$ norm of the difference. A rough estimate of the spectral baseline is first generated using $p = 1$ on each piece. Then, the estimate is iteratively improved by removing that part of the signal above the current baseline approximation and using $p = 3$ for fitting. This method guarantees baselines that stick well to the signal avoiding spurious oscillations that higher-order polynomials or splines may produce.

2.3.3 | Bucketing

Bucketing is an important operation in the processing pipeline. It consists in computing the area under the spectrum over small spectral segments that cover the whole spectral width. The segments should be large enough to blur the small discrepancies that appear from one sample to another, while preserving the resolutive power of the spectra. A bucket size of 0.01 ppm was used for 1D ¹H spectra.

Bucketing also reduces the size of the data that will be submitted to statistical analysis. This is of foremost importance in the analysis of 2D spectra, which routinely contains millions of points. The reduction of 2D data-sets to tractable sizes in statistical tools requires nevertheless bucket sizes on the order of 0.03 to 0.05 ppm in ¹H spectroscopy and to 1.0 ppm in ¹³C. Such sizes are certainly too large to capture all the details contained in the 2D spectra. One solution to this difficulty could be to use segments of varying sizes; however, we rather chose to enrich the information by adding to the area of each bucket, additional information. For each bucket, computed over 1D or 2D spectra, the coordinates of the bucket center and its size in pixel were stored, along with the area information

computed as the mean over the bucket and enriched with the values of the min and max points and the standard deviation of data over the bucket.

2.3.4 | Processing of DOSY experiment

Diffusion-ordered spectroscopy (DOSY) spectra are extremely efficient in deciphering complex mixtures and have been used in many different work (see Mahrous et al.^[7] and reference therein). They require a specific processing for the analysis of the exponential decays observed along the indirect dimension of the 2D spectrum. In this work, this specific processing was performed by using the recently introduced PALMA algorithm^[23] that implements a rapid inverse Laplace transform analysis, using a hybrid constraint, maximizing the entropy while minimizing the ℓ_1 norm of the reconstructed spectrum. This algorithm was developed using the SPIKE library, so it was particularly easy to insert it into the processing pipe-line. As a consequence, all DOSY are systematically processed, and a peak list and an adapted bucket list are also generated.

2.3.5 | Report

Finally, a concise report is produced as a csv file (see Supporting Information). The report contains all the important parameters related to data acquisition and processing. They are finally displayed, using the pandas python library.^[18]

2.4 | Analysis

Given a set of 1D and 2D NMR raw experiments, the approach described above is able to produce, in full automation and without any human interaction, a set of correctly processed spectra, along with complete peak lists and enriched bucket lists.

The artifacts observed in the spectra, such as antidiagonals and t_1 noise and ridges were corrected on the bucket list. On modern spectrometers, these artifacts are usually at a low intensity, however their presence is detrimental as the purpose here is to detect species at low concentration.

The 2D bucket list is corrected for remaining t_1 noise and t_1 ridges for each column in the matrix, by setting to null all buckets below twice the median value of the considered column. The bucket list originated from symmetric spectra, such correlation spectroscopy (COSY) and total correlation spectroscopy (TOCSY), were further corrected for departure of this symmetry by setting symmetrical buckets to the minimum value of the pair. These two operations have the effect of preserving the most significant buckets, without losing the weak spectral areas.

3 | MATERIAL AND METHODS

3.1 | Chemicals

Artemisinin 98% was purchased from Sigma-Aldrich and deuterated methanol (10 × 0.75 ml) from Eurisotop (Saint Aubin, France).

3.2 | Algae collection and identification

The alga *Sargassum muticum* was collected in June 2006 in Cap Lévy (Manche), France. Taxonomic determination was performed by Dr. A.-M. Rusig and a voucher specimen was deposited in the Herbarium of the University of Caen. Extraction was realized as in Vonthron-Sénécheau et al.^[24]

3.3 | Samples preparation

ARTE sample: The crude algal extract of *S. muticum* was differently supplemented with artemisinin, an antiplasmodial active principle. Five samples containing 10 mg of *S. muticum* hydroalcoholic extract were prepared. An artemisinin DMSO solution at 3 mg/ml was added to the samples in order to obtain final concentrations of artemisinin of 0.2, 0.3, 0.4, and 2.7 mg/ml, as summarized in Table 1. All samples were lyophilized and dissolved in 750 µl of methanol d4 and put in 5-mm NMR tubes.

SMARTE sample: A *S. muticum* hydroalcoholic extract was supplemented with artemisinin according to natural concentrations observed in *Artemisia annua* and its extract.^[25] The 9.4 mg of artemisinin was added to 350 mg of extract (2.7% m/m). This supplemented extract was fractionated by flash chromatography using a normal phase column (PuriFlash PF-30SiHP/40G, Interchim). The sample was eluted with a heptane/dichloromethane mixture of increasing polarity (50/50 to 0/100 in 60 min), then with a dichloromethane/methanol mixture (100/0 to 70/30 in 60 min) and a flowrate of 16 ml/min. The column was washed with CH₂Cl₂/acetone/MeOH/formic acid mixture (5/3/0.5/0.5). Five fractions were performed, each 30 min (F1 to F4), and F5 is the washing. All samples were lyophilized and dissolved in 750 µl of methanol d4 and put in 5-mm NMR tubes.

A sample of pure artemisinin was prepared in methanol d4 and studied by NMR for reference.

TABLE 1 ARTE sample preparation

Sample n°	1	2	3	4	5
<i>Sargassum muticum</i> extract (mg)	10	10	10	10	10
added artemisinin (mg)	0	0.15	0.24	0.32	2

3.4 | In vitro antimalarial activity—pLDH assay.

A bioactivity assay was performed on the ARTE and SMARTE samples by measuring their ability to inhibit the growth of *Plasmodium falciparum* in vitro. pLDH immunodetection assays were performed against *P. falciparum* 7G8 chloroquine-resistant strain, with a commercially available sandwich enzyme immunosorbent assay (ApDia) as reported previously.^[26] The parasite culture (1% parasitaemia and 2% hematocrit) was incubated with two concentrations of the extract (100 and 10 µg/mL) for 96 hr at 37 °C under reduced oxygen conditions (candle jar). The tests were performed twice, each time in triplicate, and chloroquine was used as a positive control. IC₅₀ were calculated with the Graphpad Prism 6 software.

3.5 | NMR spectroscopy

Acquisitions were performed on a Bruker Avance-III spectrometer operating at 700 MHz and equipped with a TCI cryo probe and a standard Bac60 sample changer. The prepared NMR tubes were spun with a small bench centrifuge to help sedimentation of insoluble parts and placed in the NMR sample changer. Each sample was automatically inserted into the spectrometer, tuned, and shimmed after a stabilization delay of 120 s. All experiments were automatically run on each sample, the whole sequence being programmed using a TopSpin macro (see Supporting Information). Spectral parameters ($\pi/2$ pulses, receiver gain, ...) were optimized on one sample and used for the whole series without further check.

Spectral widths were set to 12 ppm in ¹H and to 150 ppm in ¹³C. 1D spectra were acquired on 64 scans, 16,384 points, and a relaxation delay of 1.5 s, for a total time of 3 min. COSY experiments were performed with eight scans, with 512 increments of 4,096 points each, for a total acquisition time of 2 hr. TOCSY experiments were performed with eight scans, with 400 increments of 4,096 points each, and using a DISPSI-2 mixing sequence of 80 ms duration. TOCSY acquisition time was 1 hr and 40 min. DOSY experiments were performed with 32 scans, with 50 increments of 4,096 points each, for a total acquisition time of 50 min. Heteronuclear single quantum correlation (HSQC) experiments were performed with four scans, with 512 increments of 2,048 points each, for a total acquisition time of 1 hr. HMBC experiments were performed with 48 scans, with 400 increments of 4,096 points each, for a total acquisition time of 10 hr.

The complete acquisition time for one sample, including sample injection and tuning, took about 16 hr. The five samples were acquired in one continuous run.

The artemisinin sample was studied by NMR: Artemisinin: ¹H NMR (CD₃OD, MHz) δ 0.99 (3H, d, J = 6.2 Hz, 6-CH₃), 1.16 (3H, d, J = 7.2 Hz, 9-CH₃), 1.38 (3H,

s, 3-CH₃), 2.08 (1H, ddd, H4), 2.40 (1H, ddd, H4), 2.01 (1H, m, H5), 1.47 (1H, m, H5), 1.38 (1H, m, H5a), 1.52 (1H, m, H6), 1.09 (1H, m, H7), 1.77 (1H, m, H7), 1.17 (1H, m, H8), 1.86 (1H, m, H8), 1.82 (1H, m, H8a), 3.31 (1H, dq, H9), 6.03 (1H, dq, H12), ¹³C NMR (CD₃OD, 700 MHz) δ 106.7 (C, C3), 25.2 (CH₃, C3), 36.6 (CH₂, C4), 25.7 (CH₂, C5), 51.2 (CH, C5a), 38.1 (CH, C6), 19.9 (CH₃, C6), 34.6 (CH₂, C7), 24.0 (CH₂, C8), 45.6 (CH, C8a), 34.0 (CH, C9), 12.7 (CH₃, C9), 81.0 (CH, C12a), 95.5 (CH, C12).

3.6 | Data processing

Spectral processing was integrally performed using the Plasmodesma program presented here. The program is written in python and is compatible both with Pythons 2 and 3. It is based on the SPIKE library^[10], and the DOSY processing was performed using the PALMA approach^[23] embedded in SPIKE as a plug-in.

Complete processing took less than an hour on a MacOs machine with eight cores, running the python anaconda distribution 4.2 from Continuum Analytics (Austin, TX).

3.7 | Statistical analysis

The bucket lists and peak lists produced by the Plasmodesma run were analysed with a python script based on the pandas and scikit-learn libraries,^[18,19] using the Jupyter notebook environment.

The Plasmodesma program, along with examples, experimental data related to the artemisinin series, Jupyter notebooks presenting the data analysis, and the specific SPIKE plug-ins are freely available at github.com/delsuc/plasmodesma repository.

4 | RESULTS

To mimic the presence of a bioactive molecule at different concentrations in complex mixtures, several samples were prepared. ARTE series was prepared from a crude algal extract supplemented at different concentrations with artemisinin, a naturally occurring and structurally known sesquiterpene lactone, and five different samples were prepared. Another series equivalent to ARTE

but where chloroquine was used instead of artemisinin was also prepared and analysed. Results from this sample are not presented here and were equivalent to the one observed with ARTE. Typical 1D and DOSY spectra were presented in a previous work.^[23]

Starting from the same crude algal extract, the SMARTE series was prepared by adding artemisinin at levels equivalent to its natural concentration. The complemented extract was then fractionated by flash chromatography to yield five fractions, and the antimalarial activity was evaluated for each fraction using the pLDH bioassay. This second series more closely resembles the case of a bioguided isolation task, usually used to identify bioactive compounds from a crude plant extract.

All five samples from both ARTE and SMARTE extracts were placed in the sample changer, and NMR data were acquired in an automatic manner, after an initial tuning of the first sample.

4.1 | Data processing

The raw data-sets were processed as described above, and the peak lists and bucket lists generated. Figures 1 and 2 show an example of the result of such a processing.

The bucketing procedure is used to summarize the spectral content and by reducing the data size, allows further statistical analysis. However, it can be seen in Figure 2 that t_1 -noise and other spectral artifacts are present, in particular in the standard deviation analysis, which enhances the local signal variations. It appears that corruption of buckets from spectral artifacts appears more deleterious in 2D spectroscopy than in classical 1D. For this reason, the areas and standard deviation values of the bucket list were subjected to the following simple procedure. The first step consists in setting to a null value all values below a certain threshold computed from the median over the vertical column of the considered bucket. This procedure allows to remove a large part of the noise and to retain only the peaks well above the threshold. The threshold level adapted for each column permit to efficiently clean the strong t_1 -noise stripes, while preserving weak peaks located in less crowded regions. In a second step, a symmetrization procedure is applied to homonuclear

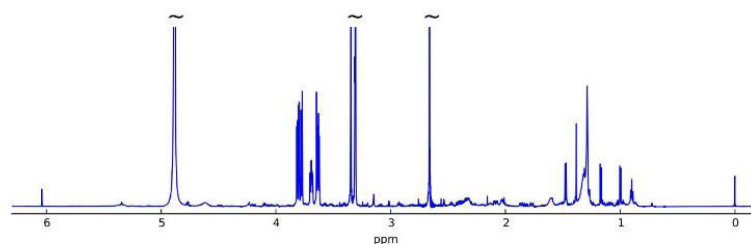


FIGURE 1 1D ¹H spectrum of sample ARTE n°4

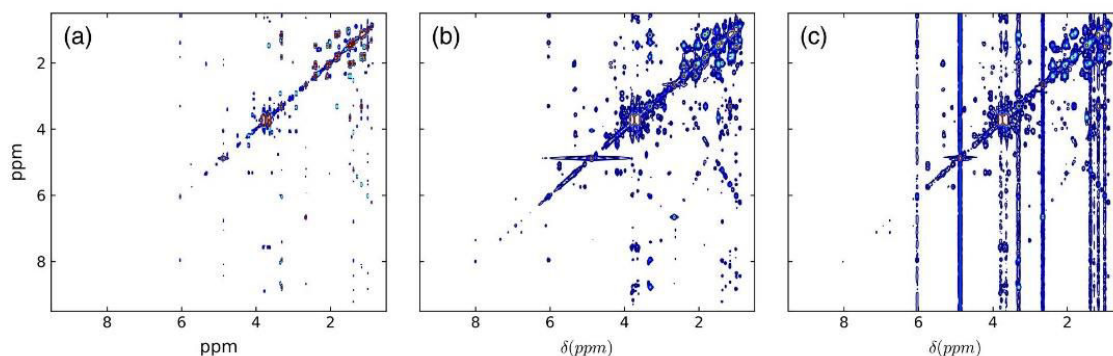


FIGURE 2 Correlation spectroscopy spectrum of sample ARTE n°5. (a) automatically generated contour plot, (b) graphic representation of the raw bucket list, and (c) same as (b) but showing the standard deviation of each bucket

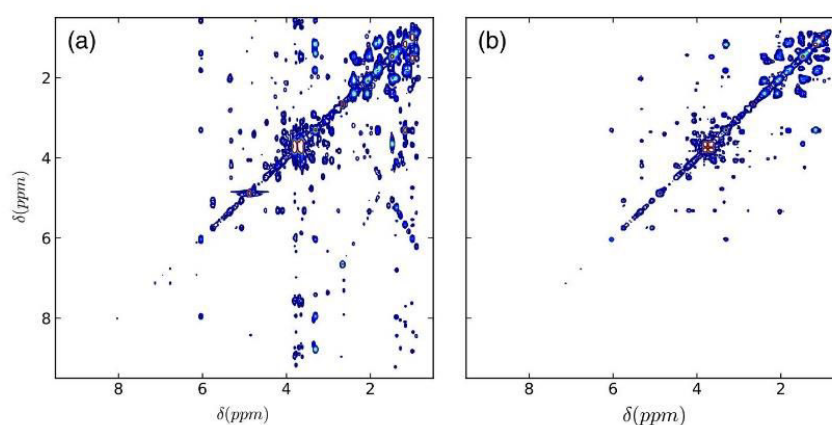


FIGURE 3 (a) Standard deviation the bucket list from Figure 2c after t_1 -noise removal, (b) same data as in (a) after symmetrization

experiments; it was done here by simply taking the smaller of the two values related by symmetry. This procedure is much simpler and more robust on bucket lists than on real spectra, as the bucketing has already homogenized the spectral axes and produced squared buckets. This procedure was applied on the area and the standard deviation values of the bucket list.

Figure 3 shows the result of each cleaning step. It can clearly be seen that this procedure allows an improvement of the quality of data and a better compatibility with automatic analysis.

4.2 | Data analysis

The cleaned bucket lists can be efficiently used for detection of the spectral features varying from spectrum to another. This can be done on any 1D or 2D spectra: Figure 4 shows the result on the analysis of the COSY spectrum.

Obviously, the positions of the signals of the artemisinin spectrum are detected and separated from the constant background, even though the background is of much larger intensity. Here, the original spectrum is not

genuinely recovered, not only because some signals are missing but also principally because of the loss of the intensities. However, the generated spectral pattern can be used to extract chemical shifts and topologies and to recognize a molecular pattern, which can be used as a fingerprint. The same result could not easily be obtained directly from the spectrum. The bucketing standardizes the spectra, and the standard deviation measures the fluctuations rather than the intensities. Finally, the cleaning operation smooths out the random fluctuations that would otherwise hamper the direct comparison.

The procedure above is not very sensitive, and the samples with lower level of artemisinin could not be processed efficiently. A second procedure was tested by taking this time the ratio of the bucket standard deviation values. The results are shown in Figure 5. Despite the low concentration level (few hundred micrograms of artemisinin in 10 mg of crude material), the spectral fingerprint is recovered. In this case, the diagonal of the homonuclear spectrum is not recovered; this does not have a strong impact, as it can be fully inferred from the off-diagonal fingerprint dots.

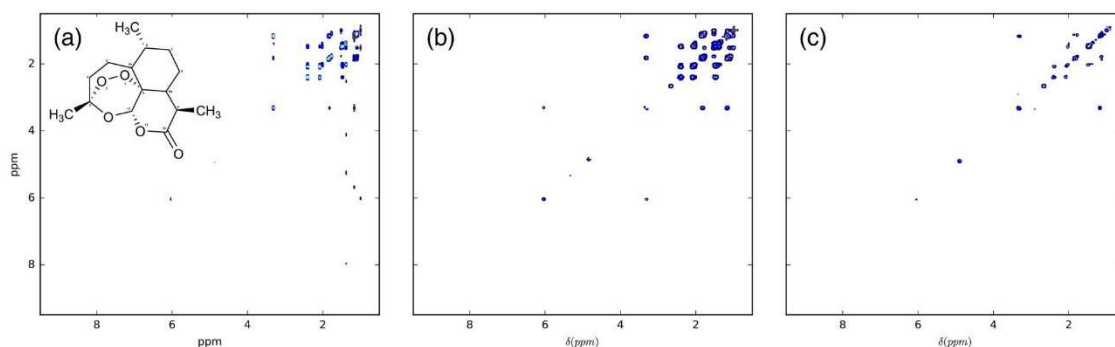


FIGURE 4 Automatic extraction of spectral fingerprint from automatic spectral analysis; (a) the standard correlation spectroscopy spectrum of artemisinin, (b) fingerprint obtained from the subtraction of the standard deviation channels from the bucket lists of sample ARTE-5 and sample ARTE-1. (c) The same as (b) for sample ARTE-4 and sample ARTE-1

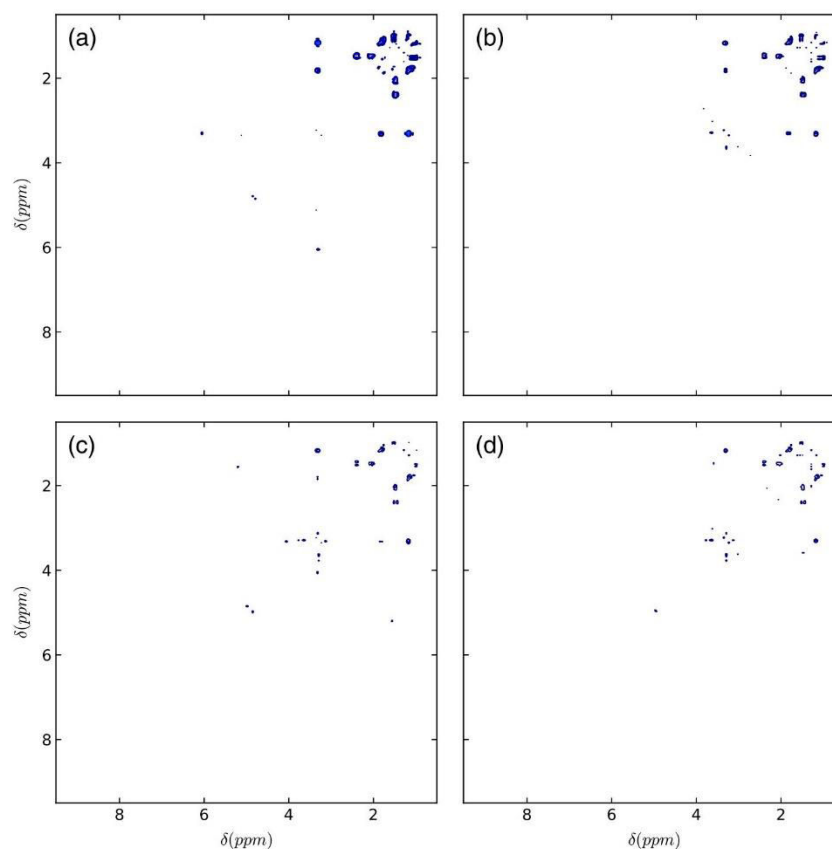


FIGURE 5 Spectral fingerprint obtained by the ratio of bucket standard deviation, (a) comparison of sample ARTE-5 and sample ARTE-1, (b-d) same as (a) for sample ARTE-4, ARTE-3, and ARTE-2, respectively

4.2.1 | Linear regression

This first approach applies a pairwise comparison and can be used on homonuclear spectra, as shown here, as well as on heteronuclear ones. Using the whole set of spectra at once requires to have an additional information,

eventually imprecise, on the amount of active material in each sample. In this case, signals coming from the studied molecule are expected to be proportional to its concentration, and this property can be exploited to separate those signals, varying along with the concentration value, to the other uncorrelated signals. This was per-

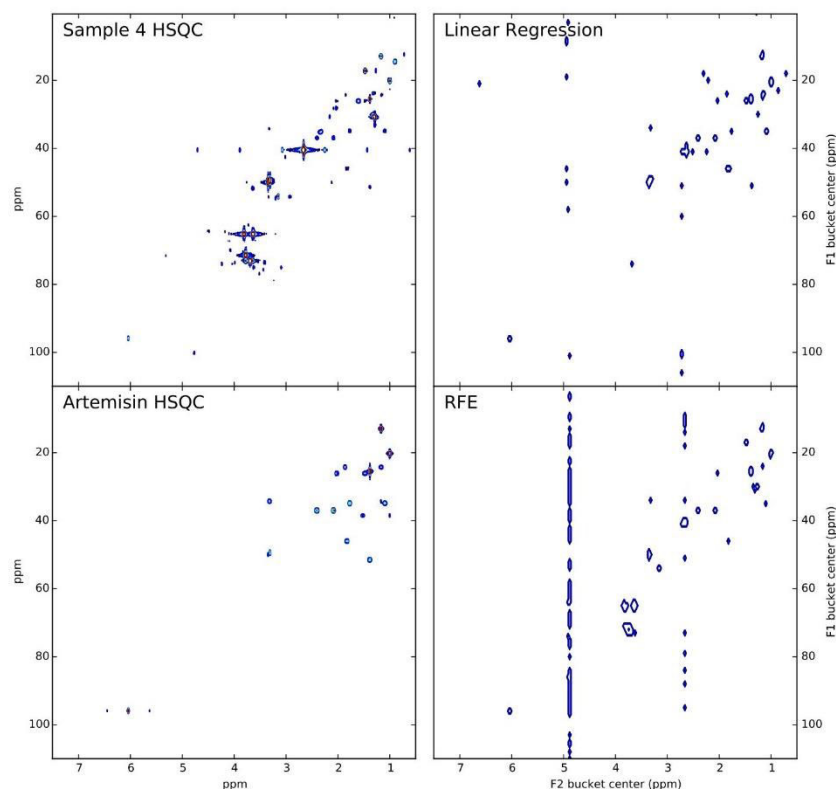


FIGURE 6 Heteronuclear single quantum correlation (HSQC) spectral fingerprint determined from the linear regression and recursive feature elimination (RFE) analyses over the four ARTE samples with the lowest concentrations in artemisinin

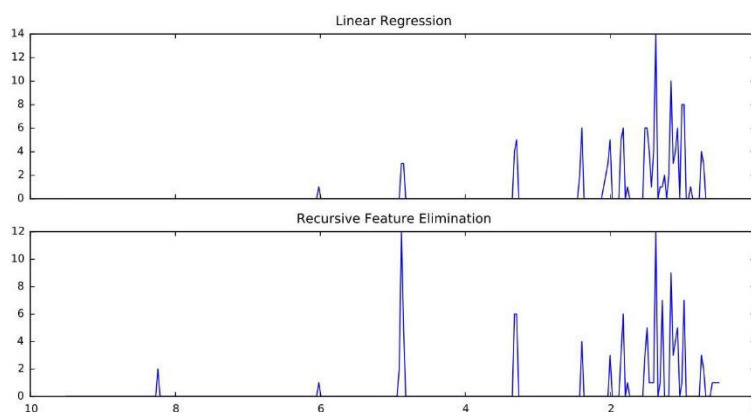


FIGURE 7 1D spectral fingerprints obtained by analysing the SMARTe diffusion-ordered spectroscopy spectra with linear regression and by recursive feature elimination from the bioactivity assay IC_{50} values, to be compared with spectrum of artemisinin in Figure 4a

formed by using the `scikit-learn` library,^[19] a generic tool for machine learning, written in python with full interoperability with python, Jupyter, and SPIKE. The `linear_model.LinearRegression()` function and the recursive feature elimination tool were used and applied on the the bucket lists area values (see Supporting Information for detailed operation).

These tools allow to select a small subset of parameters, which best correlate with the estimated concentration. The selected features are then supposed to define a spectral fingerprint in a manner equivalent to the previous approach but with a quantitative aspect this time. As the whole spectrum series is used, it is expected to produce better results. Results are shown in Figure 6 for the HSQC spectra

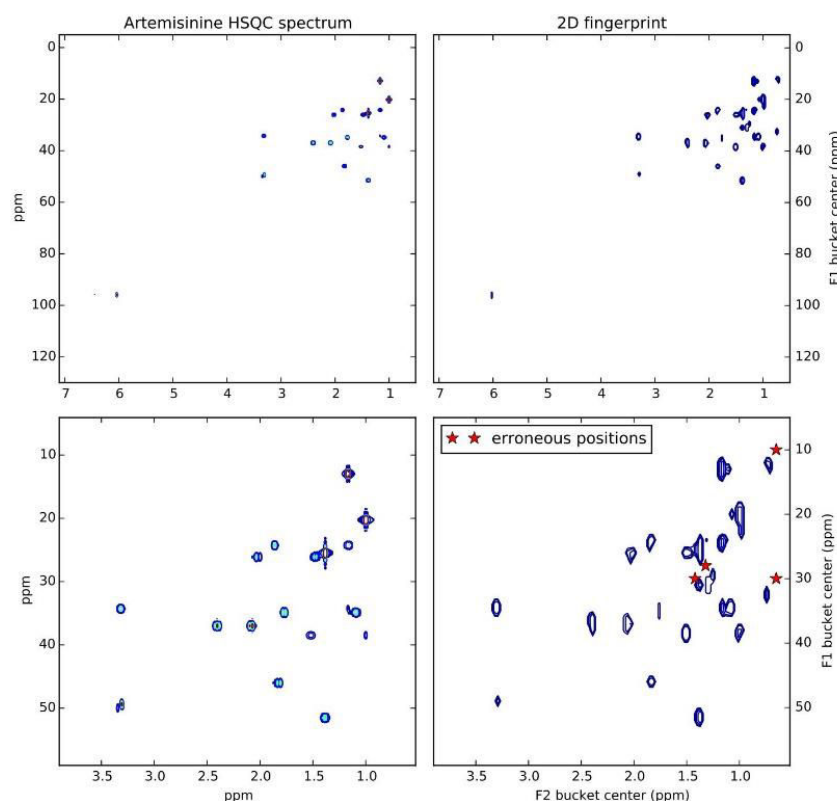


FIGURE 8 2D heteronuclear single quantum correlation (HSQC) fingerprint obtained from the SMARTE HSQC spectra by recursive feature elimination (RFE) and combined with the 1D RFE profile presented in Figure 7. The real NMR spectrum (left) is compared to the automatically determined fingerprint (right), for the whole spectrum (top) and for a zoom on the aliphatic region (bottom). Artifacts are marked with a red star

obtained on the four samples presenting the lowest concentrations. It can be seen that the HSQC spectrum of artemisinin is extracted from the complex spectrum of the mixture. The main artifacts observed in the fingerprint are associated with the solvent lines (here water, methanol, and DMSO).

4.3 | Improving the analysis with the help of DOSY experiments

The DOSY experiment is an efficient tool for the analysis of complex mixtures,^[7] and DOSY spectra were acquired in this study for all samples. However, the format of this 2D experiment is different from a regular 2D NMR experiment, as the vertical axis does not correspond to chemical shift but to translational diffusion coefficients, resulting in a spectral format difficult to integrate into an analysis pipeline. We chose to apply a bucketing on the DOSY spectra using the logarithm of the diffusion coefficient as a pseudo chemical shift axis. The bucket lists were thus processed as presented above, and the results used as 1D fingerprint by projecting the matrix onto the 1D ¹H chem-

ical shift axis (see Supporting Information for details). Figure 7 presents the fingerprint obtained from the DOSY spectra. The 1D ¹H spectrum can be somehow recovered in the profile (compare with Figure 4a).

In a second step, the HSQC spectra of the SMARTE series were analysed as in the ARTE series, and the resulting HSQC fingerprint was further filtered, keeping only the patterns compatible with the ¹H profile determined from the DOSY experiments. Figure 8 compares the obtained pattern with the regular HSQC spectrum of artemisinin. No signal is missing, and only a few minor additional signals can be seen.

5 | DISCUSSION

The measurement of NMR spectra over series of complex samples and their analysis is a common procedure in screening or bioactivity studies. The acquisition part is usually well covered through the use of sample changers and associated software, with the eventual help of companion programs, allowing for an optimized set-up.^[27]

Here, we extend the set of tools to the possibility of automatic processing and statistical analysis of the set of 1D and 2D spectra. There are already basic tools that allow to perform the first processing steps of the data; however, they usually rely either on preset parameters values (phase corrections and window function) or crude estimate of the optimum parameters (baseline correction). In contrast, the program *Plasmodesma* presented here works in an autonomous manner, without any user interaction, relying on a small set of preset global parameters. It is able to autonomously process 1D, 2D, and DOSY experiments. Processing parameters is optimized either from the experiment type (window function) or optimized automatically on the data (phase correction and baseline). In addition, advanced methods are used for the denoising of 2D spectra or the analysis of DOSY experiments. Finally, the program generates spectral plots, peak lists, and bucket lists for all spectra, as well as reports on the data and the analysis.

The use of the SPIKE library,^[10] a general processing library for NMR and other Fourier spectroscopies, allowed a rapid development of the program, as well as the use of advanced tools. Not relying on parameters previously set by an operator allows to process directly after the measure and use the result of the processing as a token for the quality of the acquisition. The processing step being faster than the acquisition, the processing was repeated during the acquisition in order to monitor the progress and quality of the current experiment.

The series of spectra are further algorithmically analysed using machine-learning inspired approaches. However, each 2D spectrum is typically composed of several million of data points, and this size hampers the possibility to algorithmically compare efficiently several spectra. For instance, the series of spectra generated in this study represent more than 20 million points overall, and some preconditioning of the data is required. For this reason, the automatic analysis of the spectra is here principally performed on the bucket lists, which provide a reduced but faithful representation of the spectra. Moreover, the bucketing operation smooths out slight spectral variations from sample to sample, unavoidable for spectra acquired automatically on samples of various origin.

Many artifacts, such as antidiagonals and t_1 noise and ridges, are also present. These artifacts are of rather low intensities, however, as the purpose here is the detection of compounds at low concentration, their presence is detrimental, and we chose to correct them on the bucket lists rather than on the original spectra. The smoothing and spectral normalization afforded by the bucketing operation allow optimal spectral corrections and make spectra comparison easier. Finally, quantities, such as standard

deviation of the signal and minimum and maximum values, allow a finer description of the spectra.

From this material, a spectral fingerprint of the molecule of interest can be determined from a pairwise spectra comparison with a presence/absence of the compound. In this case, the ratio of the bucket standard deviations is able to detect and extract the spectral features of the compound from the COSY spectra, even at low concentrations. Linear regression over the whole series of spectra is also used to generate a faithful fingerprint. One step regression and recursive feature selection were used, and both proved to be efficient in extracting a spectral fingerprint for both homonuclear and heteronuclear experiments (see Figures 6 and 8 and Supporting Information).

The ARTE experiment, where the same plant extract is spiked with various quantities of the small molecule shows that concentrations as low as 1.5% w/w can be detected and analysed. The SMARTE experiment, with a set of different samples obtained from the fractionation of the plant extract shows that the same procedure can be used to extract the spectral signature of an active molecule. This can certainly be of great help in the search for new active compounds, for assignment and dereplication.

Each experiment type can be used for the determination of the fingerprint. COSY, DOSY, and HSQC experiments were explored here. It is possible to perform the same analysis on a concatenation of all experiments; however, such an approach did not provide a correct result, probably because of the heterogeneity of the different spectra types.

6 | CONCLUSION

The program developed in this work represents an efficient alternative to the autonomous processing of a series of NMR data (1D and 2D). It fully exploits the NMR technique as a fingerprinting technique: Complete 2D NMR fingerprint of the compound is recovered through differential analysis performed either by the detection of local variations in the spectra or by linear regression between signal intensity and the concentrations of the natural product in the sample, either known or estimated from its activity.

The extended bucketing procedure allows a strong reduction of the size of the data, while preserving a large part of the molecular information present in the original spectra. Basic machine learning approaches were used to analyse this compressed but rich information and proved sufficient to readily extract the spectral fingerprint of the unknown molecule, either from spectral comparison or by handling whole spectral series at once.

This approach can contribute efficiently to the discovery of structurally unknown molecules present in natural

extracts, with no or limited chromatographic separation. This program, called Plasmodesma, is a novel tool which should be useful to decipher complex mixtures, particularly in the discovery of biologically active natural products from plants extracts but also in drug discovery or metabolomics studies. It has been developed with the aim of providing this capability to users, in an open-science manner. It is based on freely available tools, and the code and examples are provided under a liberal licence at github.com/delsuc/plasmodesma.

ACKNOWLEDGEMENTS

The authors are very grateful to Labex Medalis and Région Alsace for a fellowship (L. M.) and Europe for an Erasmus fellowship (P. M.). We are also grateful to A.-M. Rusig for the collection and the identification of the algal material, to J. Viéville for help in the NMR set-up, F. Nardella in bioassays, and to G. Bret for help in the statistical analysis. We acknowledge Wikimedia^[28] for the *Sargassum muticum* picture used in the Graphical Abstract.

ORCID

Marc-André Delsuc  <http://orcid.org/0000-0002-1400-5326>

REFERENCES

- [1] A. Bakiri, J. Hubert, R. Reynaud, S. Lanthony, D. Harakat, J.-H. Renault, J.-M. Nuzillard, *J. Nat. Prod.* **2017**, *80*(5), 1387.
- [2] B. D'Abrosca, M. Lavorgna, M. Scognamiglio, C. Russo, V. Graziani, C. Piscitelli, A. Fiorentino, M. Isidori, *Food Chem. Toxicol.* **2017**, *105*, 52.
- [3] A. Abdelsalam, E. Mahran, K. Chowdhury, A. Boroujerdi, A. El-Bakry, *Physiol. Mol. Biol. Plants* **2017**, *23*(2), 369.
- [4] J. Hubert, J.-M. Nuzillard, S. Purson, M. Hamzaoui, N. Borie, R. Reynaud, J.-H. Renault, *Anal. Chem.* **2014**, *86*(6), 2955.
- [5] S.-K. Oettl, J. Hubert, J.-M. Nuzillard, H. Stuppner, J.-H. Renault, J.-M. Rollinger, *Anal. Chim. Acta.* **2014**, *846*, 60.
- [6] C. K. Larive, G. A. Barding Jr, M. M. Dinges, *Anal. Chem.* **2015**, *87*(1), 133.
- [7] E. Mahrous, M. Farag, *J. Adv. Research* **2015**, *6*(315).
- [8] J.-L. Wolfender, G. Marti, A. Thomas, S. Bertrand, *J. Chromatogr. A.* **2015**, *1382*, 136.
- [9] Wikipedia, Plasmodesma are microscopic channels which traverse the cell walls of plant cells and some algal cells, enabling transport and communication between them, <https://en.wikipedia.org/wiki/Plasmodesma>, [Online; accessed 16-May-2017].
- [10] L. Chiron, M.-A. Coutouly, J.-P. Starck, C. Rolando, M.-A. Delsuc, *arXiv* **2016**, 1608.06777.
- [11] J. L. Pons, T. E. Malliavin, M.-A. Delsuc, *J. Biomol. NMR* **1996**, *8*, 445.
- [12] D. Tramesel, V. Catherinot, M.-A. Delsuc, *J. Magn. Reson.* **2007**, *188*(1), 56.
- [13] P. Pfändler, G. Bodenhausen, J. Rapin, R. Houriet, T. Gäumann, *Chem. Phys. Lett.* **1987**, *138*(2), 195.
- [14] P. Pfändler, G. Bodenhausen, J. Rapin, M. Walser, T. Gäumann, *J. Am. Chem. Soc.* **1988**, *110*(17), 5625.
- [15] M. A. van Agthoven, M.-A. Delsuc, G. Bodenhausen, C. Rolando, *Anal. Bioanal. Chem.* **2013**, *405*, 51.
- [16] S. van der Walt, S. Colbert, G. Varoquaux, *Comput. Sci. Eng.* **2011**, *13*, 22.
- [17] E. Jones, T. Oliphant, P. Peterson, Others, SciPy: Open source scientific tools for Python, **2001**.
- [18] W. McKinney, Pandas, Python Data Analysis Library. 2015, Reference Source, **2014**.
- [19] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, *J. Machine Learning Res.* **2011**, *12*, 2825.
- [20] The HDF Group, Hierarchical Data Format, version 5.
- [21] F. Alted, I. Vilata, et al., PyTables: Hierarchical datasets in Python, **2002**.
- [22] L. Chiron, M. A. van Agthoven, B. Kieffer, C. Rolando, M.-A. Delsuc, *Proc. Natl. Acad. Sci. (USA)* **2014**, *111*(4), 1385.
- [23] A. Cherni, E. Chouzenoux, M.-A. Delsuc, *Analyst* **2017**, *142*(5), 772.
- [24] C. Vonthron-Sénécheau, M. Kaiser, I. Devambe, A. Vastel, I. Mussio, A.-M. Rusig, *Mar. Drugs* **2011**, *9*, 922.
- [25] D. L. Klayman, A. J. Lin, N. Acton, J. P. Scovill, J. M. Hoch, W. K. Milhous, A. D. Theoharides, A. S. Dobek, *J. Nat. Products* **1984**, *47*(4), 715.
- [26] P. S. Atchade, C. Doderer-Lang, N. Chabi, S. Perrotey, T. Abdelrahman, C. D. Akpovi, L. Anani, A. Bigot, A. Sanni, E. Candolfi, *Malaria Journal* **2013**, *12*(1), 279.
- [27] L. J. Clos, M. F. Jofre, J. J. Ellinger, W. M. Westler, J. L. Markley, *Metabolomics* **2013**, *9*(3), 558.
- [28] W. Commons, File:Sargassum muticum Yendo Fensholt 1955 Lamiot WimmereuxHautsDeFrance Estran Juillet 2016a9.jpg — Wikimedia Commons, the free media repository, **2016**.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Margueritte L, Markov P, Chiron L, et al. Automatic differential analysis of NMR experiments in complex samples. *Magn Reson Chem.* 2018;1–11. <https://doi.org/10.1002/mrc.4683>

Développement d'une méthode de déconvolution pharmacophorique pour la découverte accélérée d'antipaludiques chez les Rhodophytes

Résumé

Le paludisme est responsable de plus de 445 000 morts en 2016. L'apparition et la propagation de *P. falciparum* résistants aux associations à base d'artémisinine est un problème majeur en Asie du Sud-Est. Il est nécessaire de rechercher des molécules présentant des mécanismes d'actions nouveaux qui pourraient entrer dans de nouvelles associations efficaces. Une cible d'intérêt est l'apicoplaste, un organe de *Plasmodium* issu d'une endosymbiose avec une algue rouge. Les algues rouges pourraient présenter une source privilégiée de nouvelles molécules antipaludiques ciblant notamment une voie de biosynthèse apicoplastique, celle de la biosynthèse des isoprénoïdes. Cette thèse porte sur l'identification de ces métabolites secondaires bioactifs via le développement d'une nouvelle stratégie analytique. Un programme informatique nommé Plasmodesma a été créé et permet l'analyse différentielle de manière automatique de spectres RMN 2D ^1H - ^1H et ^1H - ^{13}C . Les molécules actives peuvent ensuite être isolées et identifiées par HPLC-SPE-RMN. L'activité antipaludique mise en évidence dans les algues rouges étudiées serait due à la présence de molécules de classes chimiques différentes dont des stérols.

Mots-clés : Paludisme, déconvolution, analyse différentielle, RMN, déréplication, mélange complexe, algue rouge

Resume

Malaria is responsible for 445 000 deaths in 2016. The emergence and the spread of resistant *P. falciparum* to artemisinin-based combinations is a major health problem in South-East Asia. Research must continue to find compounds with a novel mechanism of action. The apicoplast is an interesting target. It is a *Plasmodium* organelle derived from a secondary endosymbiosis with a red alga. Red algae could be a special source of new antiplasmodial compounds targeting the isoprenoid biosynthesis pathway in the apicoplast. This thesis focuses on bioactive secondary metabolites identification via a new analytical strategy. A computer program called Plasmodesma was created and achieves an automatic differential analysis of ^1H - ^1H and ^1H - ^{13}C 2D NMR spectra. Bioactive compounds are isolated and identified by hyphenated HPLC-SPE-NMR. The presence of different classes of compounds including sterols could explain the antiplasmodial activity in studied the red algae species.

Keyword: Malaria, deconvolution, differential analysis, NMR, dereplication, complex mixture, red algae