



HAL
open science

Développement de stratégies originales d'extraction des protéines du tourteau de colza pour la production d'isolats par une démarche de modélisation / optimisation multicritères

Claire Defaix

► To cite this version:

Claire Defaix. Développement de stratégies originales d'extraction des protéines du tourteau de colza pour la production d'isolats par une démarche de modélisation / optimisation multicritères. Génie des procédés. Université de Lorraine, 2019. Français. NNT : 2019LORR0176 . tel-02517398

HAL Id: tel-02517398

<https://hal.univ-lorraine.fr/tel-02517398v1>

Submitted on 12 Jun 2024

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



Université de Lorraine (UL)

Ecole Doctorale SIMPPÉ (Science et Ingénierie des Molécules, des Produits, des Procédés et de l'Énergie)

Laboratoire Réactions et Génie des Procédés (LRGP)

Thèse

Présentée et soutenue publiquement le 2 juillet 2019 pour l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ DE LORRAINE

Mention « Procédés Biotechnologiques »

Par Claire DEFAIX

Développement de stratégies originales d'extraction des protéines du tourteau de colza pour la production d'isolats par une démarche de modélisation / optimisation multicritères

Membres du jury

Rapporteurs	Pierre-Yves PONTALIER	Maître de conférences, ENSIACET, Toulouse
	Françoise NAU	Professeur, Agrocampus Ouest, Rennes
Examineurs	Isabelle CHEVALOT	Professeur des Universités, ENSAIA, Nancy
	Catherine AZZARO-PANTEL	Professeur des Universités, LGC Toulouse
	Valérie LULLIEN-PELLERIN	Directeur de recherches, INRA, Montpellier
	Frantz FOURNIER	Maître de conférences, ENSAIA, Nancy
	Romain KAPEL	Maître de conférences, IUT Nancy Brabois
Invité	Olivier GALET	Responsable R&I, groupe Avril

LRGP – UMR CNRS 7274

Plateforme SVS, 13 rue du Bois de la Champelle, 54500 Vandœuvre-lès-Nancy

Remerciements

Je tiens à remercier toutes les personnes qui ont contribué au travail réalisé au cours de ces trois dernières années. Tout d'abord, je souhaite adresser ma reconnaissance aux membres du jury, Françoise Nau, Pierre-Yves Pontalier, Isabelle Chevalot, Catherine Azzaro-Pantel, et Valérie Lullien-Pellerin d'avoir accepté de participer à mon jury de thèse.

J'adresse ma plus sincère gratitude à mes encadrants de thèse pour leur confiance et leur précieuse aide au cours de ces travaux de thèse. Romain, merci pour ta complicité et ton soutien dans les moments les plus critiques. Tu as su cultiver ma volonté de relever des défis afin de développer mes compétences et de m'apprendre à faire de la recherche. Merci Frantz pour tes enseignements et ta pédagogie, pour avoir toujours trouvé le temps de me recevoir pour discuter modélisation et optimisation.

Je remercie également le Groupe Avril représenté au cours de la thèse par Olivier Galet, qui a apporté une dimension industrielle à ma thèse. Merci également à mes collègues, Marine et Mballo pour leur sympathie et pour leur disponibilité pour répondre à mes interrogations sur l'échelle industrielle.

Je souhaite remercier les personnels permanents de l'équipe BioProMo pour leur accueil. Merci Xavier et Arnaud de m'avoir formée aux techniques du laboratoire, pour votre disponibilité et votre assistance dans tous les petits soucis du quotidien au laboratoire. Merci Christelle pour ton investissement dans l'équipe. Je tiens également à remercier toutes les personnes que j'ai côtoyées à l'IUT. Merci Dominique de m'avoir fait confiance pour réaliser des vacances à l'IUT. Je garderai un excellent souvenir des TP de 11 h avec Gigi et Augustin. Et puis merci à tous pour votre bonne humeur et vos conseils.

Un grand merci à tous mes compagnons de laboratoire qui m'ont accompagnée sur la route du doctorat. Merci pour les soirées passées ensemble, pour la bonne humeur au laboratoire qui ont embelli mon quotidien, et surtout, merci à tous pour votre soutien dans les moments difficiles. A Sara, merci pour ta générosité inconditionnelle, pour ton soutien et ta patience avec moi. A Sophie, merci ta confiance et pour tes conseils design dans les présentations et posters. J'ai beaucoup apprécié de partager avec toi les cours à l'IUT et la thèse en 180 secondes. Mélody, tu m'as appris à faire mes premières extractions et tu as toujours été présente pour m'aider au laboratoire, merci d'avoir été mon binôme, jusqu'au bout. Odile, j'ai beaucoup apprécié ta personnalité et le travail avec toi. Mélanie, merci pour toutes les « pépites » qui me donnent encore le sourire aujourd'hui. Mais aussi merci à tous ceux que j'ai côtoyé au cours de ces années Alice, Jean-Luc, Léna, Aline, Mengyao, Felipe, Julien, Elise, Louis-Alexis, Thi Tuong, Audrey, Damien, Armel.

Je ne serai pas arrivée à ce point sans le soutien des amis qui m'ont soutenue pendant toutes ces années d'études : Annie, Aurélie, Amélie, Emilie, merci pour votre présence et les bons moments passés ensemble. Et puis, le déroulement de la thèse aurait été bien différent sans le soutien de mes amis du Nippon Kempo de Nancy. Je voudrais adresser ma gratitude à Mesut et David pour leur confiance, ainsi qu'à tous les membres du dojo : Sarah, Joao, Yacine, Aude, Flora & Pi, Stéphane, Julien et tous les autres. Merci d'avoir été ma bouffée d'air pendant ces années, l'ambiance du dojo va me manquer terriblement.

Et bien sûr, merci à ma famille, en particulier à mes parents et à ma grand-mère, pour leur soutien inconditionnel, vous m'avez inculqué les valeurs de travail et de persévérance, « acier rouge et main d'or ». Je pense aussi à mes grands-parents qui sont partis trop tôt et qui ont laissé en moi une marque indélébile, je ne vous oublie pas. Enfin, Thomas, aucun mot ne peut exprimer ma reconnaissance pour ton soutien inébranlable et pour tout ce que tu m'apportes au quotidien. Tu es ma force et ma joie de vivre.

Abréviations

AOAC	Association of Official Agricultural Chemists
AP	Acide Phytique
BCA	BiCinchonique Acid
CE	Capacité Emulsifiante
CIE	Commission Internationale de l'Eclairage
CLHP	Chromatographie Liquide Haute Performance
CLHP-ES	Chromatographie Liquide Haute Performance d'Exclusion Stérique
DC	Dichroïsme Circulaire
DLS	Dynamic Light Scattering
DoE	Design of Experiments
DV	Diavolume
EDTA	EthyleneDiamineTetraacetic Acid
EA	Emulsifying Activity
FAO	Food and Agriculture Organization of the United Nations
FRV	Facteur de Réduction Volumique
ICH	International Council for Harmonisation
MS	Matière Sèche
MM	Masse Molaire
MW	Molecular Weight
OMS	Organisation Mondiale de la Santé
PA	Phytic acid
PNNS	Plan National Nutrition Santé
SE	Stabilité Emulsifiante
SE-HPLC	Size-Exclusion High Performance Liquid Chromatography
TCA	TriChloroacetic Acid
TFA	TriFluoroacetic Acid
UF	Ultrafiltration
UV	Ultraviolet

Nomenclature

A	Absorbance
c	Concentration massique (g.L^{-1})
l	Longueur de cuve (cm)
ϵ	Coefficient d'extinction molaire ($\text{L.mol}^{-1}.\text{cm}^{-1}$)
θ	Ellipticité molaire ($\text{deg.cm}^2.\text{dmol}^{-1}$)
V	Volume (L)
V_{eq}	Volume équivalent (L)

Table des figures

Figure 1 : Schéma des problématiques identifiées dans la littérature et les travaux mis en œuvre au cours de la thèse pour y répondre.....	5
Figure 2 : Structure de la graine de colza et stockage des protéines. a) Représentation schématique d'une graine de colza (réalisée à partir de (Hu, et al., 2013) ; b) Coupe transversale d'une graine de colza au cours de son développement (Borisjuk, et al., 2013) ; c) Schéma de l'évolution du stockage d'un type de protéine du colza, les globulines, au cours de la maturation de la graine de colza (Borisjuk, et al., 2013).	11
Figure 3 : Composition des graines de colza (représentation de (Terres Univia, 2019) selon Feedbase).	12
Figure 4: Bioraffinerie du colza et positionnement de la problématique des protéines végétales pour l'alimentation humaine (modifié à partir d'un document interne Groupe Avril).....	12
Figure 5: Etapes de la trituration industrielle des graines de colza en huile et tourteau (schéma adapté de (Fine, et al., 2013)).	13
Figure 6: Composition de l'huile de colza (représentation de (Terres Univia, 2019)).	14
Figure 7: Exemple de composition du tourteau de colza (réalisé à partir de (INRA, CIRAD, & Association Française de Zootechnie, 2019)).....	15
Figure 8: Images de microscopie de tourteau de colza, partiellement décortiqué (<i>B. napus</i>), montrant en rouge les protéines, et en bleu les polymères osidiques composant la paroi des cellules végétales (Rommi, et al., 2014). Les images 1 et 2 sont prises à des grossissements différents.	16
Figure 9 : Changements de conformation des globulines en fonction de la force ionique et du pH du milieu (schéma adapté de (Schwenke, Raab , Plietz , & Damaschun, 1983).	17
Figure 10: a) Structure de l'acide phytique (Cheryan & Rackis, 1980); b) Graphes de solubilité de l'azote (cercles) et des phytates (triangles) (De Rham & Jost, 1979).	20
Figure 11 : Schéma des paramètres connus pour leur influence sur l'extraction des protéines.	22
Figure 12 : Schéma de la diafiltration par ultrafiltration appliquée à un extrait de protéines.....	29
Figure 13 : Schéma bilan des procédés breveté pour la production d'isolats de colza	32
Figure 14 : Déviation d'une lumière polarisée par passage à travers une solution optiquement active. Schéma de (Baillet, Consulté en septembre 2019).....	40
Figure 15 : Structures secondaires des protéines observables par dichroïsme circulaire. a) représente le spectre d'une protéine riche en hélices α , b) en feuillet β , c) en coudes β , d) en pelote (figure annotée de (Kelly & Price, 2000)).	41
Figure 16 : Spectres de DC, UV proches, de globulines du colza à différents pH. Phe = phénylalanine, Tyr = tyrosine et Trp = tryptophane. Figure modifiée de (Perera, McIntosh, & Wanasundara, 2016) 41	

Figure 17 : Modes de représentation en DLS. Schéma traduit de (Malvern Instruments, Consulté en février 2019)..... 42

Figure 18 : Profils de solubilité des concentrats de protéines de colza dans une solution à 0.1 M de NaCl : en rouge un concentrats de protéines totales (albumines et globulines), en vert un concentrat d'albumines et en bleu, un concentrat de globulines. Figure modifiée à partir de (Fetzer, Herfellner, & Eisner, 2019) 45

Figure 19 : couleurs de différents concentrats protéiques issus de tourteaux gras de colza à pH 5,7-5,8 (nat) ou à pH 7 (neut). UFP correspond à des protéines totales (albumines et globulines) purifiées par ultrafiltration, APP correspond à des protéines isolées par précipitation isoélectrique (globulines) et SNP correspond aux protéines solubles après précipitation isoélectrique (albumines). Figure modifiée à partir de (Fetzer, Herfellner, & Eisner, 2019)..... 51

Figure 20 : Schéma résumant les étapes du procédé mis en œuvre pour l'extraction des protéines de colza dans le cadre d'un procédé séquentiel..... 58

Figure 21 : Schéma résumant les étapes du procédé mis en oeuvre pour l'extraction des protéines de colza dans le cadre d'un procédé d'extraction sélective des albumines de colza 60

Figure 22 : Matrice expérimentale utilisée pour l'étude de l'extraction des protéines dans le cadre d'un procédé séquentiel et représentation des expériences en fonction du pH et de la concentration en NaCl dans le milieu..... 62

Figure 23 : Graphical abstract showing that the method can be implemented to follow-up a process of production of rapeseed proteins 65

Figure 24 :Total rapeseed proteins isolate analysis, A. Chromatograms of the SE-HPLC method at 214 and 325 nm ; B. Fractionation of the main peaks observed in Size-Exclusion HPLC and SDS-PAGE analysis of the obtained fractions..... 72

Figure 25 : Globulins and albumins isolates analysis by Size-Exclusion at 214 nm and SDS-PAGE 73

Figure 26 : Methodology for the validation of the SE-HPLC method. Light blue boxes represent analytical steps and grey boxes show the calculated values. Below are the values obtained by both methods represented as functions of the isolate concentration. The K_a is the ratio of both slopes. . 75

Figure 27 : Stacked chromatograms of extracts at 214 nm and for three extraction conditions. The blue line corresponds to pH 2 0 M NaCl, the orange to pH 7 0 M NaCl and the green to pH 9 0.3 M NaCl. On the right side of the figure is a zoom of the protein peaks. 77

Figure 28 : Stability study of several rapeseed protein extracts under different storage conditions (room temperature, 4 °C and -20 °C) after 24 h (A) or 48 h (B). Values on top of the bars represent the absolute relative difference of area with the initial measure. 79

Figure 29 : L1 and S1 are the chromatograms at 214 nm ; L2 and S2 are the stacked chromatograms at 280 nm and 325 nm ; L3 and S3 are zooms from 214 nm chromatograms and show the different fractions shown on SDS-PAGE gels in reducing conditions, L4 and S4.	82
Figure 30 : Schéma de la démarche d'optimisation des procédés	91
Figure 31 : Procédés mis en œuvre au cours de la thèse et cahier des charges associé.....	92
Figure 32 : Démarche du procédé séquentiel en deux étapes avec leurs conditions opératoires et les critères de performances cibles.....	95
Figure 33: Preliminary experiments to assess the working domain of pH for the sequential extraction of globulin and albumin from rapeseed meal. A) Screening of pH on the extraction of rapeseed proteins by SDS-PAGE and, B) histogram of the protein yield (Kjeldahl N x5.35) and proportion of globulin (SE-HPLC) in rapeseed extracts.	101
Figure 34 : Contour plots of the yields in raw protein, globulin, and albumin in the extracts.....	106
Figure 35 : Contour plots of the phenolic complexation in the extract and B) of the solubility of the purified proteins at pH 5.....	107
Figure 36 : Area of pH and quantity of NaCl where the constraints were all respected: overall yield higher than 45%, solubility at pH 5 higher than 95% and phenolic complexation lower than 13%...	109
Figure 37 : Results of the two extraction steps process and characterization of the final product A) Solubility of the purified protein, B) photography of the light powder obtained after freeze-drying; C) chromatograms at 280 (blue) and 325 nm (red) by SE-HPLC.	110
Figure 38 : Surface de réponse de la proportion en globulines en fonction du pH et de la quantité de NaCl ajoutée.....	113
Figure 39 : Bilan des conditions optimales du procédé d'extraction séquentiel par rapport à la qualité des produits	113
Figure 40 : Solubilité du concentrat de globulines en fonction du pH	114
Figure 41 : Schéma de synthèse de l'étude de l'extraction sélective des albumines de colza.....	118
Figure 42 : Response Surface of the predicted performances and plots of the observed values as functions of predicted values for the performance criteria regarding a) and b) the extraction of albumin and c) and d) the remaining meal.....	129
Figure 43 : Response Surface of the predicted performance and plot of the observed values as a function of predicted values for the protein losses during the microfiltration step (determined by the Kjeldahl method).....	131
Figure 44 : Pareto front and Pareto domain resulting from the multicriteria optimization of albumin extraction according to three criteria: the albumin yield, the albumin proportion and the phytic acid content in the meal after extraction. The colors represent the ranking of the solution after applying	

the decision tool: in red are the first 10%, in green the solutions ranked between 10 and 30%, in black between 30 and 70% and in blue between 70 and 100%. 132

Figure 45 : Solubility of the albumin isolate as a function of the pH. The rectangle below the point at pH 7 shows the color of the isolate in a 1% solution diluted in a phosphate buffer pH 7. 134

Figure 46 : Structural analyses by Circular Dichroism (CD) of rapeseed albumin A) shows the spectra observed by CD far-UV at room temperature and at pH 7. B) is the spectra in CD near-UV at room temperature at pH 4,7 and 9. C) and D) show the DLS repartition respectively in volume and in intensity as functions of the diameter of the particles..... 135

Figure 47 : Contour plots of the performance criteria regarding the extraction of albumin. a) and b) show the albumin proportions, respectively at 20 and 55°C; c) and d) show the albumin yields respectively at 20 and 55°C. 147

Figure 48 : Contour plots of the phytic acid content in the meal after extraction of albumin: a) from the Design of Experiments (DOE) at 20°C. and b) from the DoE at 55°C. 149

Figure 49 : Graph of the desirable operating conditions to produce rapeseed albumin. The area in blue corresponds to an albumin yield higher than 55%, an albumin proportion higher than 95% and a phytic acid content in the meal lesser than 4%. The orange area corresponds to an albumin yield higher than 70%, an albumin proportion higher than 95% and a phytic acid content in the meal lesser than 4%. 150

Figure 50 : Structural analyses by Circular Dichroism for both albumin products from 20 and 55°C processes. a) shows the spectra observed by CD far-UV at room temperature at pH 7. b) shows spectra in CD near-UV at room temperature at pH 7. c) and d) show the DLS repartition for the albumin at 20°C and at 55°C respectively in volume and in intensity as functions of the diameter of the particles at pH 7. 152

Figure 51 : Functional properties of rapeseed albumin isolates. Albumin produced according to the process at 20°C are represented in blue, the albumin produced at 55°C in red. a) shows the solubility of albumin at various pH; b) represents the emulsifying activity; c) is the foaming capacity and d) is the foaming stability; e) and f) show the chromatograms of both proteins at 20 and 55°C at 214 nm and 280 and 325 nm; g) and h) show photographs, respectively of the albumin at 20 and 55°C. 155

Table des tableaux

Tableau 1 : Recommandations de l'OMS pour la consommation d'acides aminés essentiels (World Health Organization, 2007) en comparaison à la teneur en acides aminés essentiels contenus dans le tourteau de colza (Canola council of Canada, 2015)	9
Tableau 2 : Comparaison des propriétés moussantes de protéines de différentes origines. Ces données sont à considérer étude par étude car les protocoles ne sont pas strictement identiques d'une étude à l'autre.	47
Tableau 3 : Comparaison des propriétés émulsifiantes de protéines de différentes origines. Ces données sont à considérer étude par étude car les protocoles ne sont pas strictement identiques d'une étude à l'autre.....	49
Tableau 4 : Comparison of the SE-HPLC method with the Kjeldahl method and with the SDS-PAGE densitometry.....	76
Tableau 5 : Validation of the method according to the International Conference on Harmonisation (ICH) guideline.....	80
Tableau 6 : Matrix for the design of experiments for the study of the sequential extraction of globulin and then albumin and related results.....	103
Tableau 7 : Coefficients for the studied models	104
Tableau 8 : Matrix for the design of experiments at 20 °C for the study of albumin extraction and ANOVA analysis of the main indicators used to validate the models.....	127
Tableau 9 : Coefficients for the performance models and their p-values	128
Tableau 10 : Matrix for the design of experiments at 20°C for the study of albumin extraction.	145
Tableau 11 : Coefficients for the studied models of both DOE at 20 and 55°C.....	146
Tableau 12 : Aminograms of albumin produced at 20°C and at 55°C in comparison with the World Health Organization (WHO) recommendations.....	153

Sommaire

INTRODUCTION GENERALE.....	1
CHAPITRE I : Etude Bibliographique	6
Introduction	7
I Les protéines dans l'alimentation.....	7
I.1 Description des protéines et de leurs sources.....	7
I.2 Les protéines en tant qu'ingrédient.....	8
I.3 Les besoins nutritionnels en protéines	8
I.4 Potentiel du colza comme source de protéines	9
I.4.1 Histoire et développement des variétés de colza.....	9
I.4.2 Culture du colza	10
II La bioraffinerie du colza.....	11
II.1 Les graines de colza.....	11
II.2 Description de la bioraffinerie	12
II.3 Composition de l'huile de colza	14
II.4 Composition des tourteaux gras de colza.....	15
II.4.1 Les glucides	15
II.4.2 Les protéines	16
II.4.3 Les molécules antinutritionnelles	19
III Les procédés de production d'isolats protéiques à partir de tourteau de colza	21
III.1 L'extraction des protéines du tourteau de colza	21
III.1.1 Effet du procédé de production du tourteau	22
III.1.2 Effet du pH	23
III.1.3 Effet du ratio solide / liquide	23
III.1.4 Effet de la quantité de sel	24
III.1.5 Effet de la température.....	25
III.1.6 Effet de la durée d'extraction	25
III.2 La purification des protéines de colza	26
III.2.1 La précipitation isoélectrique	26
III.2.2 La précipitation à froid ou la micellisation.....	27
III.2.3 La purification par des méthodes chromatographiques.....	28
III.2.4 La purification par procédés membranaires.....	28
III.3 Les procédés brevetés	30
IV Les outils analytiques pour le dosage et la caractérisation des protéines.....	33

IV.1	Les techniques de dosage des protéines	33
IV.1.1	La méthode de dosage de référence, le Kjeldahl.....	33
IV.1.2	Le dosage par analyse élémentaire, la méthode de Dumas	34
IV.1.3	Les dosages colorimétriques.....	34
IV.1.4	Caractérisation des protéines par électrophorèse SDS-PAGE	35
IV.1.5	Les méthodes chromatographiques	36
IV.2	L'étude structurale des protéines.....	39
IV.2.1	Le dichroïsme circulaire (DC)	39
IV.2.2	La diffusion dynamique de la lumière (DLS)	42
IV.3	L'étude des propriétés des isolats protéiques.....	43
IV.3.1	La solubilité	43
IV.3.2	Propriétés moussantes	45
IV.3.3	Propriétés émulsifiantes	48
IV.3.4	Propriétés gélifiantes	49
IV.3.5	La couleur.....	50
IV.3.6	Une autre propriétés organoleptique, le goût.....	51
V	Démarche de modélisation et d'optimisation des étapes de la production d'isolats de protéines végétales	52
V.1	La méthodologie des plans d'expériences.....	53
V.2	L'optimisation multicritères.....	54
V.3	Les outils d'aide à la décision.....	55
	CHAPITRE II : Matériel et méthodes.....	57
I	Le tourteau étudié	58
II	Extraction des protéines dans le cadre d'un procédé séquentiel.....	58
II.1	Première étape d'extraction des protéines	59
II.2	Clarification des extraits par microfiltration.....	59
II.3	Purification des protéines et chromatographie liquide basse pression	59
II.4	Protocole de la seconde étape d'extraction des protéines	60
III	Extraction sélective des albumines.....	60
III.1	Protocole d'extraction des protéines	61
III.2	Purification par procédés membranaires	61
IV	Les outils de modélisation et d'optimisation.....	61
IV.1	La méthodologie des plans d'expériences.....	61
IV.2	Les algorithmes d'optimisation multicritères	62
IV.3	Les algorithmes d'aide à la décision	63
V	Méthodes analytiques pour le suivi du procédé.....	63
V.1	Détermination de la teneur en protéines.....	63
V.1.1	La méthode de Kjeldahl	63

V.1.2	La densitométrie de gels SDS-PAGE.....	64
V.1.3	Développement d'une nouvelle méthode de quantification des protéines par Chromatographie Liquide Haute Performance d'Exclusion Stérique	64
V.1.4	Article publié dans le journal Food Chemistry (2019).....	64
V.1.5	Contribution de l'article	83
V.2	Détermination de la teneur en acide phytique.....	83
V.2.1	Extraction	83
V.2.2	Formation d'un complexe.....	83
V.2.3	Dosage des ions Fe ³⁺ libres	84
V.2.4	Calculs pour la détermination de la teneur en acide phytique.....	84
V.3	Détermination de la complexation phénolique	85
VI	Les méthodes de caractérisation des isolats	85
VI.1	Caractérisation générale.....	85
VI.1.1	Matière sèche	85
VI.1.2	Pureté et solubilité.....	86
VI.2	Les méthodes d'analyse structurale	86
VI.2.1	Le dichroïsme circulaire	86
VI.2.2	Détermination de la taille des particules par Diffusion Dynamique de la Lumière	86
VI.3	Les méthodes d'analyse des propriétés fonctionnelles.....	87
VI.3.1	Qualification de la couleur	87
VI.3.2	Capacité et stabilité moussante	87
VI.3.3	Capacité et stabilité émulsifiante	87
	Introduction des résultats et discussion	89
	CHAPITRE III : Etude d'un procédé d'extraction séquentiel des globulines et albumines de colza ..	93
I	Introduction	94
II	Manuscript : impact of pH and NaCl on protein extractability, irreversible phenolic complexation and protein solubility from a cold-press rapeseed meal.....	95
II.1	Introduction	95
II.2	Material and methods	97
II.2.1	Material.....	97
II.2.2	Rapeseed meal.....	97
II.2.3	Protein extraction	97
II.2.4	Analytics	98
II.2.5	Design of Experiments methodology.....	99
II.3	Results and discussion	100
II.3.1	Protein extractability	100
II.3.2	Impact of the pH and NaCl concentration on albumin, globulin and total protein yield	102

II.3.3	Impact of the pH and NaCl content on the phenolic complexation	107
II.3.4	Impact of the pH and NaCl content on the solubility of purified proteins	108
II.3.5	Determination of optimal operating conditions	108
II.3.6	Application of the two extraction steps process	109
II.3.7	Conclusion	110
II.4	Contribution de l'article	111
III	Un procédé séquentiel pour la production séparément des globulines et des albumines	111
III.1	Détermination de conditions opératoires optimales	112
III.2	Mise en œuvre du procédé	113
IV	Conclusion du Chapitre III	114
CHAPITRE IV. Modélisation de l'effet des conditions opératoires sur l'extraction des albumines		116
I	Introduction	117
II	Manuscript: Multicriteria optimization of albumin extraction from rapeseed meal for both the food and feed industry	118
II.1	Introduction	118
II.2	Material and methods	120
II.2.1	Raw material	120
II.2.2	Protein extraction	120
II.2.3	Analytics	121
II.2.4	Modelling and optimization	122
II.2.5	Purification of albumin	124
II.2.6	Albumin isolate characterization	125
II.3	Results and discussion	126
II.3.1	Effect of pH and NaCl on albumin extractions and remaining meal composition	126
II.3.2	Multicriteria optimization	132
II.4	Conclusion	136
II.5	Contribution de l'article	137
III	Manuscript: Influence of the extraction temperature on the quality of albumins from rapeseed meal	138
III.1	Introduction	138
III.2	Material and methods	139
III.2.1	Raw material	139
III.2.2	Protein extraction	139
III.2.3	Analytics	140
III.2.4	Modelling and optimization	141
III.2.5	Albumin isolate production	142
III.2.6	Functional properties assessment	143
III.2.7	Structural analyses	144

III.3	Results and discussion	145
III.3.1	Impact of temperature, pH and NaCl concentration on rapeseed albumin extraction 145	
III.3.2	Establishement of the optimal pH and NaCl concentration at 20°C and 55°C	148
III.3.3	Characterization of albumin obtained at 20°C and 55°C	150
III.4	Conclusion.....	156
III.5	Contribution de l'article.....	157
III.6	Extrapolation du procédé	157
IV	Conclusion du Chapitre IV	157
	CONCLUSION GENERALE ET PERSPECTIVES	159
I	Conclusion.....	160
II	Perspectives	161

INTRODUCTION GENERALE

Les protéines sont l'un des trois nutriments dans l'alimentation avec les lipides et les glucides. Alors qu'elles sont présentes naturellement en grandes quantités dans certains aliments (viande, œufs ou certains végétaux comme les lentilles), elles sont aussi utilisées en tant qu'ingrédient dans les aliments manufacturés, car elles ont des propriétés fonctionnelles leur permettant de structurer les matrices alimentaires sous forme de mousses, d'émulsions ou de gels.

Une étude de 2016 propose le terme de Matières Protéiques Végétales (MPV) pour désigner des ingrédients protéiques issus d'une transformation à partir de produits végétaux (Guéguen, Walrand, & Bourgeois, 2016). Un tel ingrédient est appelé un isolat lorsqu'il est composé à plus de 90 % de protéines ou un concentrat lorsqu'il contient entre 65 et 90 % de protéines. La même étude affirme que les protéines végétales les plus répandues en 2016, en tant que MPV, proviennent du soja et du blé. En parallèle, de nombreuses études s'intéressent à l'exploitation d'autres ressources végétales, telles que le colza, pour ce type d'applications. La ressource doit être disponible du point de vue des quantités produites et de la qualité nutritionnelle et fonctionnelle des protéines qu'elle contient.

En France, alors qu'en 1962, moins de 100000 ha étaient dédiés à la culture du colza avec une production de 160000 tonnes de graines, en 2016, cette surface a atteint 1,5 millions d'hectares pour produire 4,7 millions de tonnes de graines (soit 30 fois plus). Cette évolution s'est déroulée en plusieurs étapes. Dans les années 1970 et 1980, des travaux importants de sélection variétale, visant à limiter les teneurs en acide érucique et en glucosinolates des graines, ont rendu l'huile de colza propre à l'alimentation humaine. Parallèlement, ceci a fait du tourteau issu de l'extraction d'huile une ressource intéressante pour l'alimentation animale. Ces nouvelles variétés sont appelées « double zéro » ou « 00 ». Ensuite, dans les années 1990, le biodiesel s'est développé et a généré une nouvelle vague d'augmentation des cultures de colza. Ainsi, depuis les années 2000, un accroissement des surfaces de culture est observé, en corrélation avec une diminution de l'import de tourteaux de soja, une plante largement cultivée en Asie et en Amérique du Nord.

Actuellement, le marché de l'huile de colza est tendu au niveau mondial pour les applications non alimentaires. Par ailleurs, les développements scientifiques et industriels sur l'huile (oléochimie) atteignent une limite : de 2012 à 2017, les quantités de biodiesel produites stagnent à environ 2200 tonnes par an en France (UNdata, 2019) et peu d'évolutions technologiques significatives ont été réalisées dans le domaine. C'est pourquoi, l'exploitation des tourteaux pour des applications à plus haute valeur ajoutée que l'alimentation animale constitue un enjeu industriel majeur. En 2014, le projet de recherche porté par le Groupe Avril et le Laboratoire Réactions et Génie des Procédés (LRGP) est lauréat du concours mondial de l'innovation et reçoit un financement de la Banque Publique

d'Investissement (BPI). Il repose sur le développement de production d'isolats protéiques à partir des tourteaux de colza et de tournesol pour une utilisation en alimentation humaine.

La production d'isolats de protéines se base sur une extraction des protéines suivie d'étapes de purification. De nombreux paramètres opératoires sont impliqués : la durée de l'extraction, la quantité de tourteau par rapport à la quantité de milieu d'extraction, le pH, l'ajout de sel et la température. Toutes ces conditions opératoires ont un impact sur un grand nombre de critères de performance : les rendements, la composition en protéines, la complexation phénolique, la solubilité et la fonctionnalité des protéines après purification. Malgré de nombreux articles scientifiques et brevets, aucun procédé de production d'isolat de protéines de colza n'est aujourd'hui industrialisé. Compte tenu de la complexité du système d'extraction, ceci peut être dû à la difficulté de mettre en évidence des conditions opératoires appropriées avec les démarches expérimentales classiques qui considèrent le rendement avant tout. Le manque de valorisation du résidu solide (jusqu'à 50% de la matière sèche initiale) appauvri en protéines peut également être une seconde explication.

L'objectif de la thèse a été de mettre en œuvre une démarche originale de modélisation et de simulation afin de déterminer les conditions opératoires permettant de satisfaire un cahier des charges quant aux performances du procédé et du produit : un rendement d'extraction d'au moins 60 %, une solubilité de l'isolat d'au moins 95 % à pH 5 ainsi qu'une couleur de poudre la plus claire possible, afin de rendre le produit acceptable pour une grande variété d'applications en alimentation humaine. Pour atteindre cet objectif, une démarche en quatre étapes a été adoptée.

1. La première a consisté à identifier les facteurs les plus influents sur le procédé et à déterminer les domaines d'étude pour chacun.
2. La deuxième étape s'est reposée sur la méthodologie des plans d'expériences afin de générer des équations décrivant l'influence des facteurs sur les performances du procédé.
3. La troisième étape a été l'application d'algorithmes génético-évolutionnaires sur les équations précédentes, afin de sélectionner les meilleurs compromis entre les conditions opératoires répondant au cahier des charges.
4. Enfin, la quatrième étape a consisté à l'utilisation d'un outil d'aide à la décision afin de faire un choix rationnel entre les compromis.

Cette démarche a été appliquée pour explorer différentes voies de productions possibles dans des zones de conditions opératoires variées. Puis un procédé de purification des protéines par ultrafiltration membranaire, adapté au scénario d'extraction, a été proposé afin de permettre une caractérisation physico-chimique des isolats produits.

La Figure 1 résume la structure du manuscrit de thèse. Le Chapitre I fait l'état de l'art antérieur sur les méthodes de caractérisation des protéines de colza, les procédés d'extraction existants et les verrous scientifiques liés. Le côté droit de la Figure 1 montre les solutions qui ont été proposées dans les travaux de thèse et leur répartition dans les chapitres afin de lever ces verrous.

La plupart des études sur l'extraction des protéines de colza considèrent les rendements d'extraction en protéines totales. Or ces protéines se composent de deux fractions aux propriétés bien distinctes, les albumines et les globulines. L'étude bibliographique ne montre pas de technique analytique rapide qui permette à la fois la quantification individualisée de ces protéines ainsi que d'apporter des informations quant aux phénomènes complexation irréversibles des phénols. Ainsi, **la première étape des travaux de thèse a consisté au développement d'une nouvelle méthode analytique palliant ce problème, par Chromatographie Liquide Haute Performance d'Exclusion Stérique (CLHP-ES). Grâce à cette méthode, plusieurs critères de performance originaux ont été rendus accessibles : la proportion en albumines / globulines, les rendements d'extraction et la complexation phénolique. L'intégration de ces critères dans le projet de thèse a contribué à améliorer la compréhension du système d'extraction et d'interaction complexe entre les protéines et les autres molécules présentes dans le tourteau.**

L'étude de l'état de l'art a permis d'identifier des problématiques relatives au procédé de production d'isolats de colza. Dans la grande majorité des études, un tourteau industriel est utilisé après avoir subi des traitements thermiques et mécaniques. Une hypothèse des travaux de thèse a été de considérer que ce mode de production des tourteaux est probablement néfaste à la qualité des protéines qu'ils contiennent. C'est pourquoi, les travaux de thèse ont été réalisés sur des tourteaux gras pressés à froid.

Le Chapitre III de la thèse porte sur un procédé de plusieurs extractions séquentielles afin de maximiser l'extraction des protéines tout en considérant la qualité du produit fini. Pour cette étude, la démarche d'étude décrite précédemment a été appliquée à un domaine de conditions expérimentales classiques, c'est-à-dire, à pH neutre voire faiblement basique et avec une concentration modérée en sel (0 à 0.2 M). Les travaux d'optimisation ont permis de proposer deux voies d'extractions distinctes répondant à un cahier des charges préalablement établi. Ces conditions ont ensuite été mises en œuvre et les protéines ont été purifiées par ultrafiltration afin de produire des isolats et de les caractériser. Le premier scénario a consisté en la production d'un unique isolat résultant de deux extractions séquentielles dans l'objectif d'avoir un produit final avec une solubilité élevée, une couleur claire et cela, avec un rendement d'extraction équivalent aux modes opératoires classiques. Le second scénario s'est reposé sur l'utilisation du procédé séquentiel pour la production

séparément d'un isolat de globulines et d'un isolat d'albumines. Cependant, un tel procédé avec des extractions séquentielles a appauvri considérablement la teneur en protéines du tourteau, et donc sa valeur pour une utilisation en alimentation animale en a été réduite.

Dans une troisième partie la démarche de modélisation et de simulation a été appliquée à une zone de pH acide dans la même gamme de concentration en NaCl. Dans cette zone, les albumines sont extraites sélectivement mais les rendements observés dans la littérature restent très faibles. Ainsi, pour la thèse, une première sous-partie de l'étude a porté sur l'optimisation de l'extraction des albumines de colza à 20 °C. La seconde sous-partie concernait l'impact de la température (20 ou 55 °C) sur le procédé d'extraction : pour ce faire, des conditions optimales ont été déterminées par application de contraintes, et les protéines ont été purifiées afin de les comparer du point de vue de leurs propriétés fonctionnelles et structurales. Une originalité est que cette étude prend en compte à la fois des performances du procédé mais aussi de la qualité des albumines extraites et de la qualité du tourteau résiduel, alors riche en globulines.

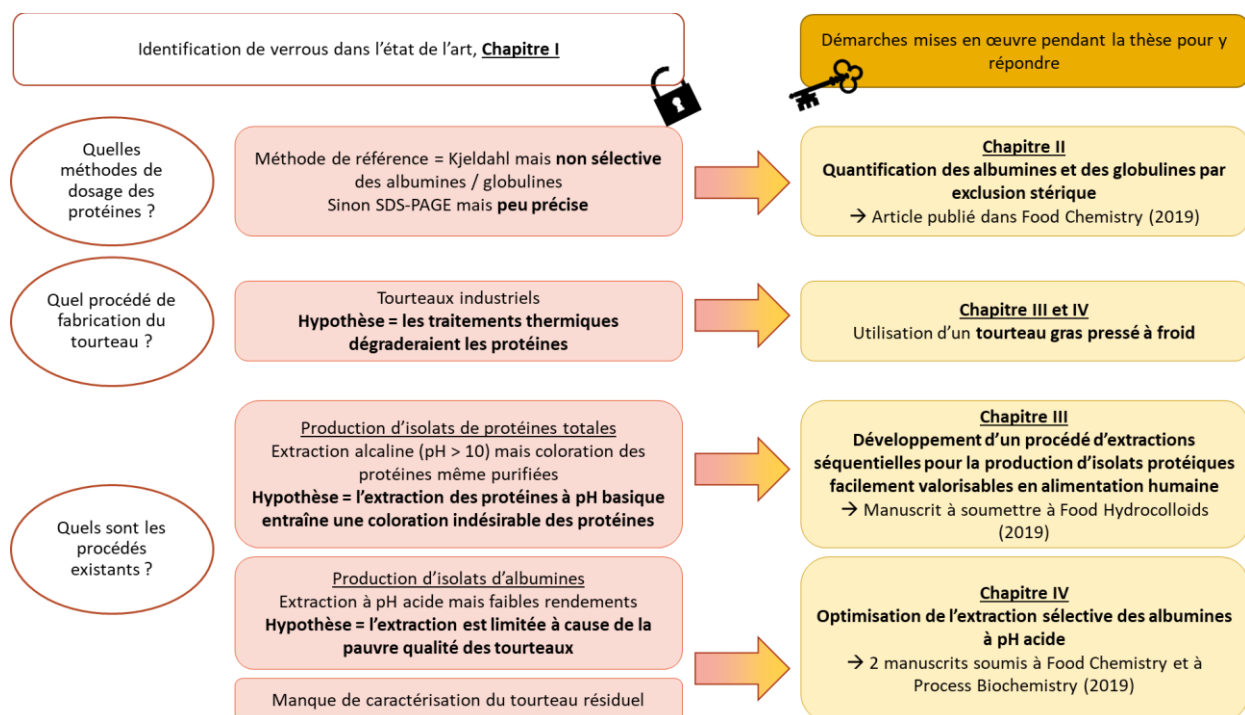


Figure 1 : Schéma des problématiques identifiées dans la littérature et les travaux mis en œuvre au cours de la thèse pour y répondre

CHAPITRE I : ETUDE BIBLIOGRAPHIQUE

Introduction

Afin d'introduire les problématiques et les méthodes utilisées dans le projet de thèse, l'étude bibliographique est divisée en 5 parties. Dans un premier temps, la place des protéines dans l'alimentation humaine est brièvement décrite. Dans une deuxième partie, les enjeux liés à l'exploitation de ressources végétales pour notre alimentation en France sont décrits et mis en perspective vis-à-vis de la bioraffinerie du colza. Une troisième partie explore les procédés d'extraction et de purification des protéines, issues de cette ressource. La quatrième partie est une revue des méthodes analytiques, couramment utilisées pour l'analyse des protéines et des molécules co-extraites, comme l'acide phytique ou les composés phénoliques, c'est-à-dire pour leur identification ou leur quantification. Leurs avantages et inconvénients sont décrits afin de justifier les choix des outils utilisés au cours des travaux de thèse. Enfin, une dernière partie présente la démarche d'optimisation du procédé d'extraction et les outils utilisés au cours de ces travaux.

I Les protéines dans l'alimentation

I.1 Description des protéines et de leurs sources

Avec les glucides et les lipides, les protéines constituent l'une des trois familles majeures de macromolécules nécessaires au fonctionnement de l'organisme et doivent être renouvelées régulièrement par l'alimentation. Ce sont des constituants essentiels des muscles, du système immunitaire et elles assurent la biocatalyse de l'ensemble des réactions biochimiques et physiologiques (Gibney, Lanham-New, Cassidy, & Vorster, 2009). Ces macromolécules sont constituées d'acides aminés qui sont caractérisés par deux fonctions chimiques : une fonction acide dont le pKa est de 2 – 3 et une fonction amine avec un pKa proche de 9. Naturellement, il existe 20 acides aminés dont 9 sont dits essentiels, du fait qu'ils soient nécessaires au bon fonctionnement de l'organisme, mais ils ne sont pas synthétisés par l'être humain. Ainsi, ils doivent être apportés par l'alimentation via les protéines.

Les protéines sont présentes dans de nombreux aliments, qui peuvent être d'origine animale, comme la viande, le poisson ou les œufs, ou végétale, comme le soja, les lentilles ou le blé. Ainsi, en France, entre 2000 et 2013, plus de 54 % des protéines consommées proviennent de sources végétales (Food and Agriculture Organization of the United Nation (FAO), 2019).

De nombreux aliments d'origine végétale contiennent naturellement des protéines en proportions significatives. Par exemple, les noix et graines peuvent contenir 20 % de protéines. Les légumineuses,

sont une famille de plantes produisant des graines riches en protéines et très utilisées dans notre alimentation comme les lentilles, les pois, les fèves ou les haricots et contiennent 20 à 25 % de protéines. Les céréales sont également une source importante de protéines dans l'alimentation. En Europe occidentale, à la fin des années 1980, environ 60 % des protéines consommées étaient d'origine végétale dont 24 % étaient issues des céréales (Gibney, Lanham-New, Cassidy, & Vorster, 2009). Notamment, la graine de blé contient environ 14 % de protéines sur la matière sèche (ANSES Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail, 2019). En particulier, en France, une étude a mis en évidence que le pain est l'aliment qui contribue le plus aux apports en protéines avec 7 à 10 % (Agence Française de Sécurité Sanitaire des Aliments (AFSSA), 2007).

I.2 Les protéines en tant qu'ingrédient

En parallèle de cet apport direct, des protéines sont également consommées en tant qu'ingrédient dans des produits alimentaires formulés. Elles sont utilisées ici pour leurs propriétés nutritionnelles mais aussi pour leurs propriétés fonctionnelles qui permettent de donner de la structure aux aliments (mousses, émulsions, gels). Quand elles sont d'origine végétale, ces protéines sont obtenues par extraction solide/liquide à partir d'une agro-ressource ou d'un coproduit de transformation d'agro-ressource riche en protéines et par purification. En fonction du niveau de pureté en protéines, il s'agira d'isolat (>90 % en protéines) ou de concentrats (>65 % en protéines). Des concentrats peuvent également être obtenus par fractionnement d'une ressource solide en voie sèche. Plusieurs marchés sont porteurs : dans les produits traiteurs, les protéines végétales contenues dans les produits commercialisés en 2019 ont pour rôle de se substituer à la viande ou au poisson. Un second marché, en essor est celui des en-cas ou snack. Dans ce cas, ce sont les propriétés fonctionnelles des protéines qui sont ciblées par les industries alimentaires, afin de proposer des produits avec des textures variées. Un troisième segment de marché est la nutrition-santé : il s'agit alors de formuler des ingrédients répondant aux besoins de populations spécifiques, les enfants en bas âge, les seniors ou les sportifs (Guéguen, Walrand, & Bourgeois, 2016).

I.3 Les besoins nutritionnels en protéines

En 2007, l'OMS (Organisation Mondiale de la Santé) a publié un rapport d'expertise concernant les besoins nutritionnels humains en protéines et en acides aminés. Ainsi, pour un adulte sain, un apport de 0,83 g de protéines par kilogramme par jour est recommandé pour satisfaire ses besoins journaliers. Le Tableau 1 recense les apports recommandés en acides aminés essentiels (World Health Organization, 2007).

Cependant, les protéines végétales sont souvent limitées quant à leur profil en acides aminés essentiels : il manque généralement un ou plusieurs de ces éléments, ce qui nécessite de diversifier les sources de protéines végétales pour satisfaire ses besoins journaliers. Par exemple, les légumineuses manquent souvent d'acides aminés comme la méthionine, le tryptophane et la cystéine (Singh N. , 2017). De plus, des études ont comparé la digestibilité des protéines végétales avec les protéines d'origine animale. Les protéines contenues dans les grains de maïs, de blé complet ou d'avoine présentent une digestibilité de l'ordre de 86 % contre plus de 94 % pour la viande, le lait et les œufs. Ces scores de digestibilité sont ainsi légèrement plus bas que pour les produits d'origine animale mais demeurent élevés (Young & Pellett, 1994).

Tableau 1 : Recommandations de l'OMS pour la consommation d'acides aminés essentiels (World Health Organization, 2007) en comparaison à la teneur en acides aminés essentiels contenus dans le tourteau de colza (Canola council of Canada, 2015)

Acide aminé	Quantité recommandée par l'OMS		Profil en acides aminés du tourteau de colza
	en mg / kg par jour	en mg / g de protéines	en mg / g de protéines
Histidine	10	15	33,9
Isoleucine	20	30	34,7
Leucine	39	59	61,9
Lysine	30	45	59,2
Méthionine	10	16	19,4
Cystéine	4	6	11,6*
Phénylalanine + tyrosine	25	38	65.6*
Thréonine	15	23	42,7
Tryptophane	4	6	13,3
Valine	26	39	49,7

*Valeur calculée à partir des données de (Canola council of Canada, 2015).

Néanmoins, comme le montre le Tableau 1, le profil en acides aminés essentiels des protéines contenues dans le colza est équilibré par rapport aux recommandations de l'OMS et par rapport à d'autres protéines végétales. En particulier, la richesse en lysine et en acides aminés soufrés (cystéine et méthionine) est notable. Ainsi, la production de protéines végétales pour l'alimentation humaine, pourrait être un nouveau débouché d'une filière existante, la bioraffinerie du colza.

I.4 Potentiel du colza comme source de protéines

I.4.1 Histoire et développement des variétés de colza

Le colza, *Brassica napus*, appartient à la même famille que le chou et la navette, dont il est un croisement. La culture du colza remonte à l'Antiquité et aux civilisations grecques, romaines et chinoises pour son huile en tant que carburant pour les éclairages. Ce n'est qu'à partir du XVIIIe siècle

que l'huile de colza est utilisée en alimentation en Europe. La culture est exportée au Canada au milieu du XXe siècle.

Des travaux de sélection variétale ont permis de développer des plantes, avec des taux en acide érucique et en glucosinolates réduits, pour la consommation humaine. L'acide érucique est un acide gras dominant dans l'huile de colza où il représente environ 45 % des acides gras totaux (Shahidi, Canola and Rapeseed, 1990). Une exposition prolongée à cette molécule provoquerait une accumulation anormale de lipides dans le cœur (EFSA (European Food Safety Authority) Panel Contaminants Food, 2016). Les glucosinolates sont des métabolites secondaires de défense de la plante : leur hydrolyse enzymatique produit des composés toxiques. Ainsi, une diminution de la teneur en glucosinolates a été recherchée afin d'augmenter la digestibilité des produits issus de la transformation du colza (Shahidi, Canola and Rapeseed, 1990).

Ces nouvelles variétés, appelées « double zéro », ou « 00 », ont permis de généraliser l'utilisation de l'huile de colza. D'autant que dans les années 1970, la France était dépendante des Etats-Unis pour s'approvisionner en tourteau de soja, qui était alors une ressource majeure en alimentation animale. En 1974, un embargo a été décrété et a imposé à la France de développer ses filières d'oléoprotéagineuses (Terres OléoPro, 2019). Par ailleurs, dans les années 1990, le biodiesel a été inventé et industrialisé. Issu de la transestérification des triglycérides de l'huile avec un alcool (méthanol), sa production repose essentiellement sur l'utilisation d'huile de colza. Ces événements successifs ont transformé la bioraffinerie du colza et ont permis son essor en France. Cela s'est traduit par une augmentation des surfaces cultivées de 70000 ha, au début des années 1960, à plus de 1,5 millions d'hectares en 2016. Ainsi, la production annuelle de colza est passée de 107000 tonnes à plus de 4,7 millions de tonnes générant à la fois de l'huile mais aussi des tourteaux (Food and Agriculture Organization of the United Nation (FAO), 2019).

I.4.2 Culture du colza

Les graines de colza sont généralement semées à la fin de l'été, pour une floraison en avril et une récolte en juin voire juillet. Le colza est une tête de rotation dans les cultures, c'est-à-dire que, sur un cycle complet, le colza est la plante cultivée en premier. En effet, il a été observé de nombreux effets bénéfiques de sa culture en précéssion à celles de céréales, notamment en termes de rendements, mais aussi pour la diminution des maladies affectant les céréales (Terres Univia, 2019).

II La bioraffinerie du colza

II.1 Les graines de colza

La graine de colza est une structure d'environ 2 mm qui contient un embryon végétal. La Figure 2 a) montre la structure schématique d'une graine. Elle est constituée d'un tégument, appelé « coque », qui permet de protéger l'embryon. Le colza étant une plante dicotylédone, deux cotylédons sont présents dans la graine. Si la graine germe, ces cotylédons formeront les premières structures aériennes de la plantule et la radicule, la première structure racinaire. Les réserves de la graine, sous forme de lipides et de protéines, sont situées dans ses cotylédons.

La Figure 2 b) est une coupe transversale de la graine de colza en cours de développement, avec une coloration, en violet, de l'une des deux protéines majoritaires du colza, les globulines. La Figure 2 c) montre schématiquement la chronologie de stockage des globulines dans la graine. Ces deux figures montrent que les protéines sont situées dans les vacuoles de stockage, elles-mêmes situées dans les cellules composant les deux cotylédons et la radicule. Dans cette étude, il a été montré que les autres protéines du colza, les albumines et les oléosines, ont une répartition identique aux globulines (Borisjuk, et al., 2013). Les oléosines permettent la formation d'oléosomes, qui sont des globules contenant de l'huile. Ils sont situés dans le cytoplasme des cellules.

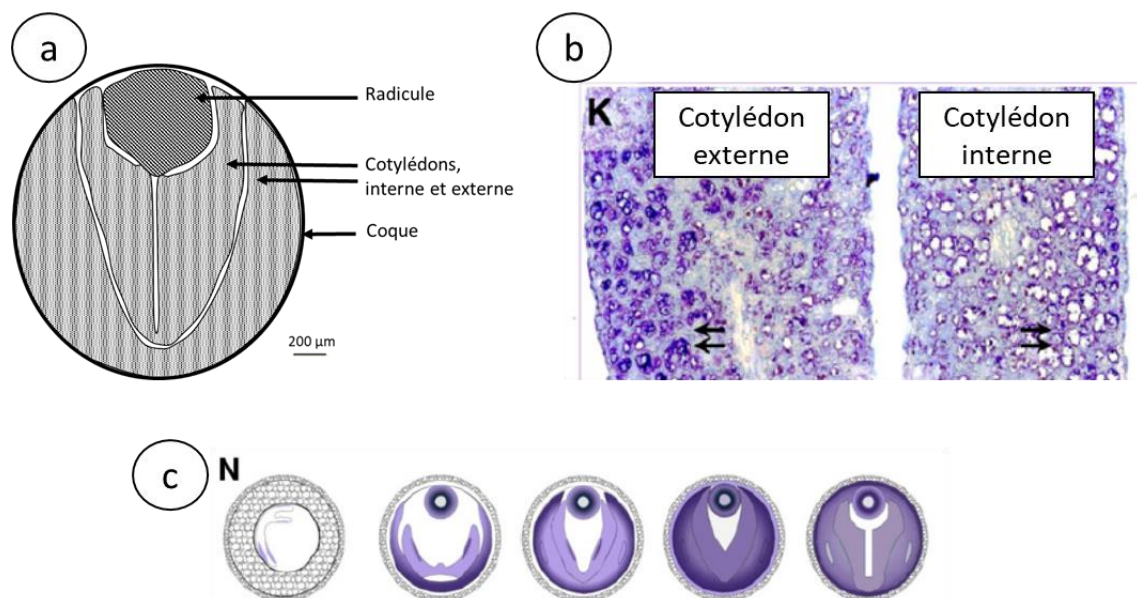


Figure 2 : Structure de la graine de colza et stockage des protéines. a) Représentation schématique d'une graine de colza (réalisée à partir de (Hu, et al., 2013) ; b) Coupe transversale d'une graine de colza au cours de son développement (Borisjuk, et al., 2013) ; c) Schéma de l'évolution du stockage d'un type de protéine du colza, les globulines, au cours de la maturation de la graine de colza (Borisjuk, et al., 2013).

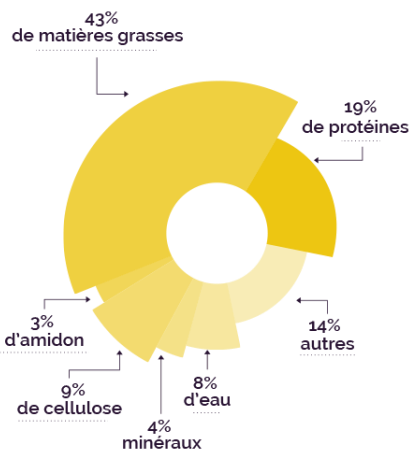


Figure 3 : Composition des graines de colza (représentation de (Terres Univia, 2019) selon Feedbase).

Comme présenté en Figure 3, les graines de colza contiennent 43% de matières grasses, sous forme d'huile, et 19% de protéines. Les autres composés sont majoritairement des sucres, notamment sous forme de réserves d'amidon et de fibres, comme la cellulose qui est un composant essentiel des parois végétales.

II.2 Description de la bioraffinerie

L'objectif de la bioraffinerie du colza est de valoriser les différents produits de sa culture, tel que représenté en Figure 4. Les différentes étapes, de la sélection des variétés jusqu'à l'utilisation de l'huile et des tourteaux impliquent de nombreux acteurs de la filière.

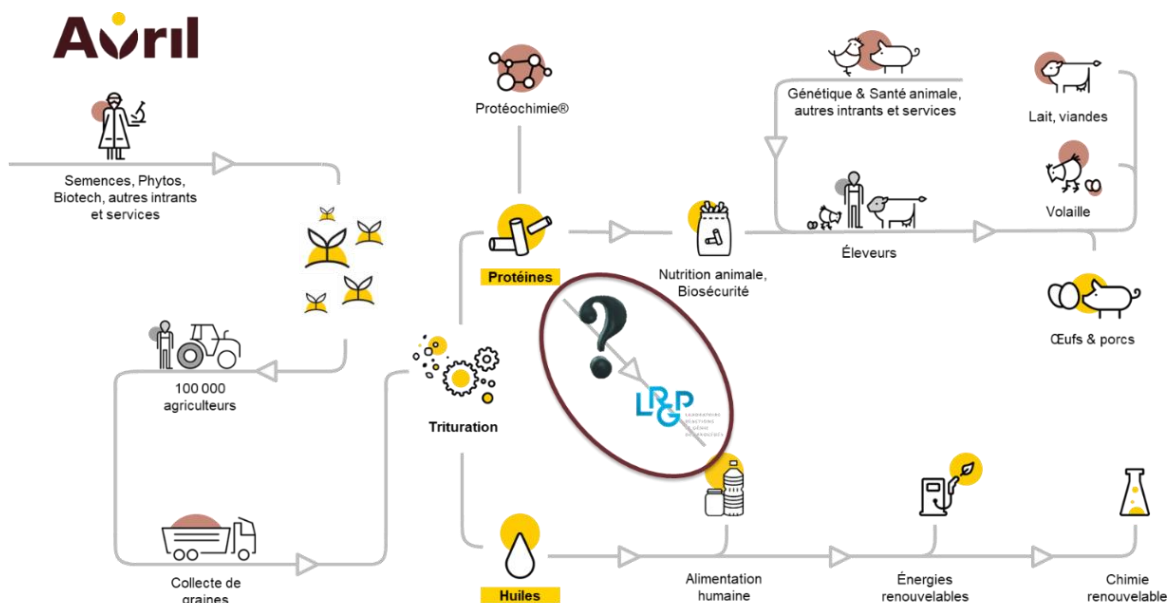


Figure 4: Bioraffinerie du colza et positionnement de la problématique des protéines végétales pour l'alimentation humaine (modifié à partir d'un document interne Groupe Avril).

Après récolte, les graines sont traitées dans des usines de trituration pour en récupérer l'huile. Les étapes de la trituration sont présentées en Figure 5. Les graines sont d'abord nettoyées, séchées, éventuellement décortiquées, avant d'être conditionnées : pour ce faire, les graines sont chauffées avec pour objectif premier de limiter l'éclatement des graines lors de l'étape suivante. Le floconnage consiste à aplatir les graines de colza et à rompre les parois cellulaires. Vient ensuite la cuisson des flocons entre 80 et 105 °C, qui permet de réduire la viscosité de l'huile et ainsi de rendre son extraction plus aisée. Enfin, les flocons cuits sont pressés pour obtenir de l'huile brute et un **tourteau gras**, contenant encore 10 à 20 % d'huile (Rous & Groupe Avril, 2016) (Canola Council, 2009) . Ensuite, ce dernier subit une étape de déshuilage, c'est-à-dire une extraction assistée par un solvant, le plus courant étant l'hexane. Cette étape permet de récupérer une fraction supplémentaire d'huile mélangée à de l'hexane. Une étape de distillation permet de séparer ces deux éléments (Fine, et al., 2013). Enfin, le raffinage de l'huile consiste en plusieurs étapes permettant d'éliminer certains composés indésirables de l'huile, comme par exemple la cire et les pigments qui affectent la couleur et les propriétés de l'huile. Le raffinage rend l'huile utilisable en alimentation, oléochimie ou pour la production de biodiesel (diester).

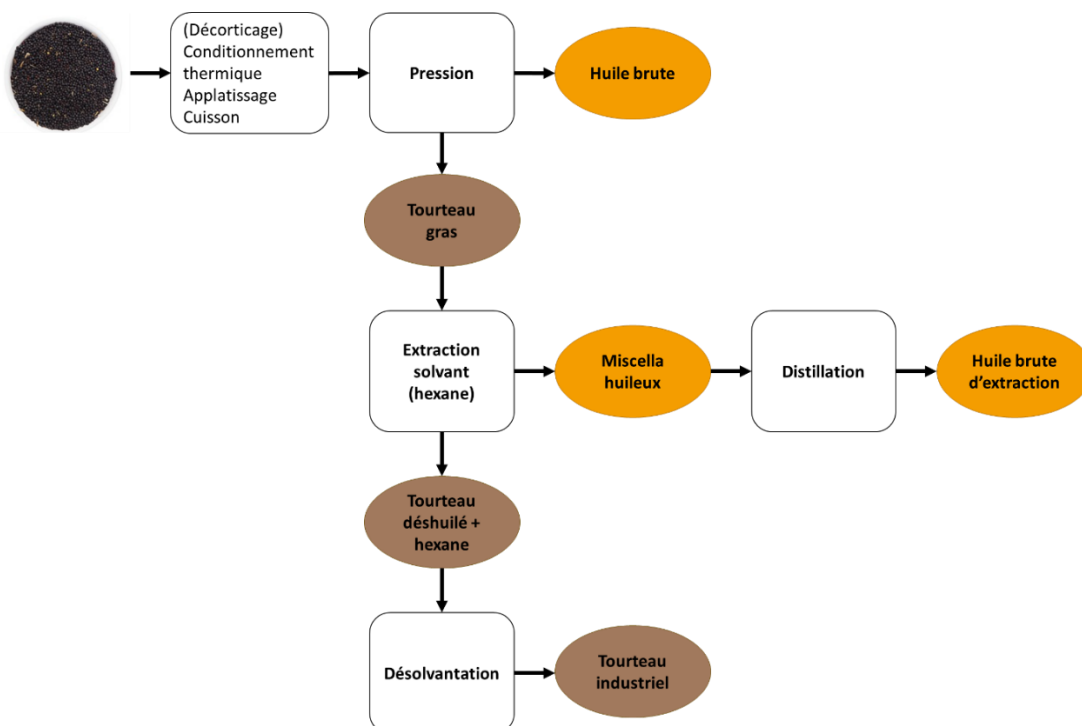


Figure 5: Etapes de la trituration industrielle des graines de colza en huile et tourteau (schéma adapté de (Fine, et al., 2013)).

Des études sont actuellement menées pour trouver des substituts à l'hexane comme solvant pour extraire l'huile du tourteau gras. Une étude de 2015 a montré qu'il était possible de remplacer l'hexane par du 2-méthyltetrahydrofuran, un solvant organique produit à partir d'agroressources

comme la canne à sucre ou le maïs. Comparé à l'hexane, ce solvant serait moins nocif et en plus, il est produit à base de biomasse et non pas à partir d'une ressource fossile (Sicaire, et al., 2015).

Pour les procédé classique utilisant de l'hexane, le tourteau est ensuite désolvanté. Pour cela, le tourteau est chauffé, à une température pouvant atteindre jusqu'à 100 °C, afin d'évaporer l'hexane. Le solide résiduel, le **tourteau industriel**, ne contient plus que 2 % d'huile environ.

Cependant, dans le cadre d'une utilisation du tourteau pour des applications en alimentation humaine, l'utilisation de l'hexane est remise en cause. En effet, il s'agit d'un solvant dangereux : il est facilement inflammable et est un neurotoxique reconnu (Fine, et al., 2013). C'est pourquoi, les tourteaux utilisés, au cours des travaux de thèse, sont des tourteaux gras, issus de la pression à froid des graines de colza.

II.3 Composition de l'huile de colza

Actuellement, l'huile de colza a deux usages majeurs. La production de biodiesel utilise environ 2/3 de l'huile produite alors que le tiers restant est réservé à l'alimentation humaine (Terres Univia, 2017). Comme le présente la Figure 6, le colza contient majoritairement des acides gras insaturés. Une particularité de l'huile de colza est son profil équilibré en oméga 3 et 6, qui sont dits essentiels car ils ne sont pas synthétisés par l'être humain. D'après le Plan National Nutrition Santé (PNNS) français, un ratio de 1 acide gras oméga-3 pour 4 acides gras oméga-6 serait l'idéal. Dans l'huile de colza, ce ratio est de 1 : 2,5.

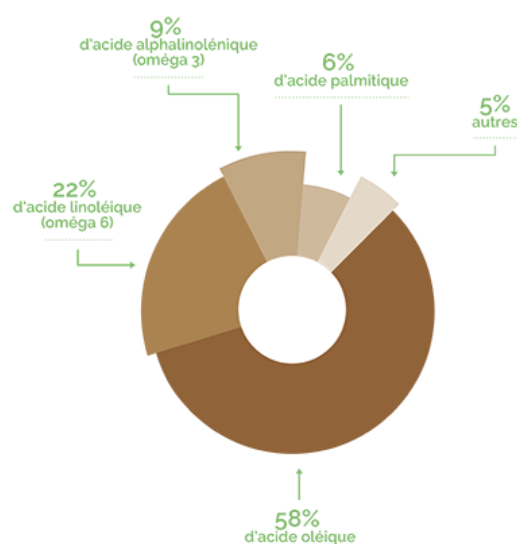


Figure 6: Composition de l'huile de colza (représentation de (Terres Univia, 2019)).

II.4 Composition des tourteaux gras de colza

Comme le présente la Figure 7, le tourteau de colza contient encore entre 10 et 20 % d'huile résiduelle après pressage. Il contient majoritairement des protéines (30 à 35 %) et des fibres. Cette composition est directement liée aux éléments composant les cellules végétales et, particulièrement, la paroi riche en polysaccharides.

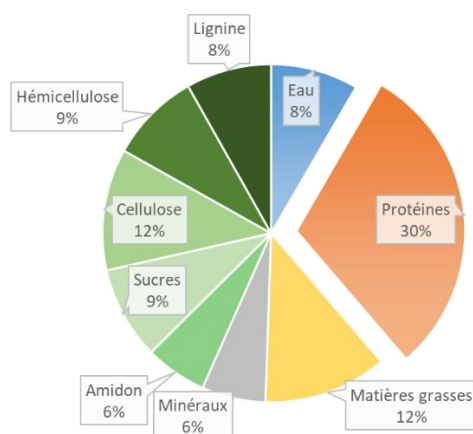


Figure 7: Exemple de composition du tourteau de colza (réalisé à partir de (INRA, CIRAD, & Association Française de Zootechnie, 2019))

Les tourteaux de colza actuellement produits et utilisés en alimentation animale sont déshuilés avec de l'hexane. Toute origine confondue (soja, colza ou tournesol), les tourteaux représentent environ 15 % de l'alimentation du bétail. En France, en 2012 et 2013, plus de 2,5 millions de tonnes de tourteau de colza ont été produites et consommées pour cet usage (Terres Univia, 2019).

II.4.1 Les glucides

Les glucides du tourteau de colza sont majoritairement des fibres qui composent les cellules végétales: la cellulose, l'hémicellulose, la lignine ou encore les pectines. La Figure 8 montre des images obtenues par microscopie sur un tourteau décortiqué de colza. Il y apparaît des composants de parois cellulaires végétales, prouvant qu'ils sont encore présents après pressage de l'huile. L'étape de pression à froid des graines de colza entraîne la rupture, au moins partielle, des polysaccharides composant la paroi végétale et des corps protéiques par action mécanique (Rommi, et al., 2014).

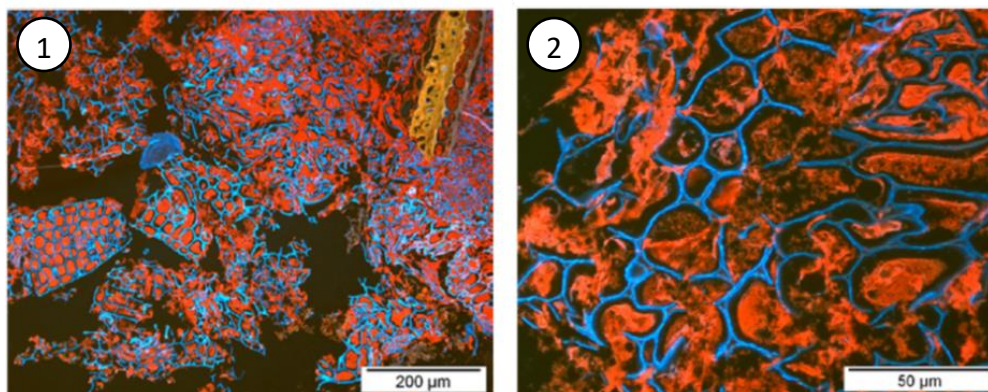


Figure 8: Images de microscopie de tourteau de colza, partiellement décortiqué (*B. napus*), montrant en rouge les protéines, et en bleu les polymères osidiques composant la paroi des cellules végétales (Rommi, et al., 2014). Les images 1 et 2 sont prises à des grossissements différents.

II.4.2 Les protéines

Le colza contient majoritairement deux familles de protéines de stockage, des albumines et des globulines, ainsi que des oléosines qui sont des protéines de structure. Au total, il y a une proportion massique de 30 à 35 % de protéines dans le tourteau de colza. Comme détaillé dans le Tableau 1, les protéines du tourteau de colza ont un profil équilibré en acides aminés essentiels avec des teneurs élevées en acides aminés soufrés et en lysine.

Une classification usuelle des protéines selon leur solubilité dans différents solvants a été définie en 1897 (Osborne, 1897) : les albumines sont définies comme étant des protéines solubles dans l'eau ; les globulines dans des solutions salées ; les prolamines dans des mélanges eau / éthanol ; et les glutélines dans des solutions alcalines.

Les protéines sont aussi généralement caractérisées par un coefficient de sédimentation. Son unité est le Svedberg, tirant son nom de son inventeur Theodor Svedberg, lauréat du prix Nobel de chimie en 1926. Le coefficient de sédimentation d'une protéine dépend à la fois de son volume, de sa forme et de sa masse moléculaire. Il est obtenu par mesure de la vitesse de sédimentation par ultracentrifugation.

Enfin, de par leur structure et composition en acides aminés, chaque protéine possède un point isoélectrique, noté pI , qui lui est propre. Il est défini comme étant le pH auquel la molécule est dans un état de charge neutre. A ce pH, la protéine atteint généralement son minimum de solubilité.

II.4.2.1 Les cruciférines ou globulines

Les cruciférines sont des globulines, majoritaires dans le colza, avec une proportion de 50 à 60 % des protéines totales. Ce sont des protéines 11 – 12 S et leur masse molaire est comprise entre 300 et 350 kDa. Elles sont composées de 6 sous unités, d'environ 50 kDa, liés par des interactions électrostatiques (Schwenke, Raab , Plietz , & Damaschun, 1983). Chaque sous-unité est composée de deux chaînes

polypeptidiques, α et β , correspondant respectivement à des chaînes d'environ 30 et 20 kDa. Elles sont liées par un pont disulfure à l'état natif. Le point isoélectrique des globulines a été trouvé entre 5 et 7 (Cheung, Wanasundara, & Nickerson, The Effect of pH and NaCl Levels on the Physicochemical and Emulsifying Properties of a Cruciferin Protein Isolate, 2014).

Par ailleurs, les globulines sont sujettes à des changements de conformation selon les conditions du milieu. Dans un milieu de faible force ionique, les globulines seraient majoritairement présentes sous formes de trimères 7 S, alors qu'en présence d'une force ionique plus importante, la forme 12 S serait dominante. Il a été montré que ce changement de conformation entre 7 S et 12 S est réversible. En revanche, à pH acide, la protéine serait dissociée irréversiblement en monomères 2 S (Schwenke, Raab, Plietz, & Damaschun, 1983) (Perera, McIntosh, & Wanasundara, 2016). Dans cette dernière étude, il a également été montré que les globulines sont moins stables que les albumines lorsqu'elles sont exposées à une température élevée : leur température de dénaturation a été mesurée entre 83 et 85 °C, contre plus de 100 °C pour les albumines (Perera, McIntosh, & Wanasundara, 2016).

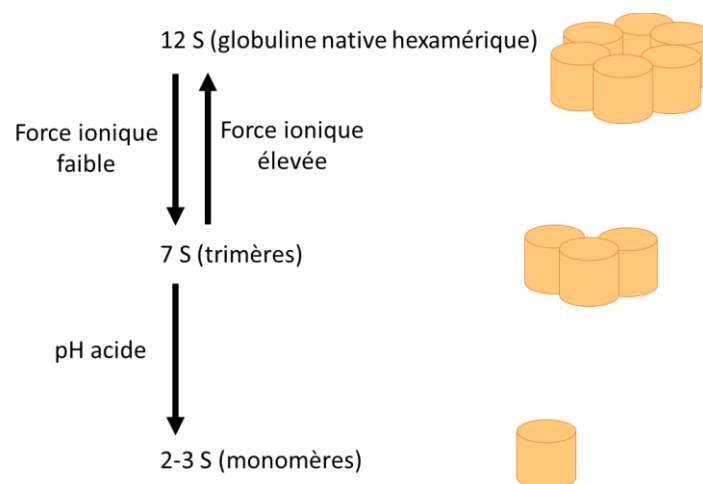


Figure 9 : Changements de conformation des globulines en fonction de la force ionique et du pH du milieu (schéma adapté de (Schwenke, Raab, Plietz, & Damaschun, 1983).

Cette structure confère aux globulines des propriétés fonctionnelles particulières et distinctes des albumines. A pH 3, il a été trouvé une solubilité élevée, de l'ordre de 90 %, alors qu'à pH 5 et 7 en l'absence de sel, la solubilité est de 10 %. Dans cette étude, l'ajout de sel NaCl a permis d'augmenter significativement la solubilité des globulines (Cheung, Wanasundara, & Nickerson, 2014).

Les globulines sont capables de former des gels par chauffage. Une étude de 1998 a comparé les gels formés par des globulines ou des albumines de colza. Il a été montré que la formation des gels de globulines est induite à une température plus basse, environ 70 °C contre 90 °C pour les albumines. Dans cette étude, le pH de formation du gel influence peu la température de gélification, mais elle

change la force du gel. Pour les napines, les gels formés sont les plus faibles comparés aux cruciférines et au mélange des deux protéines (Schwenke, Dahme, & Wolter, 1998).

Les globulines peuvent également former des émulsions. Cependant, le pH et la quantité de sel ont une forte influence sur la capacité à les former. Les émulsions les plus stables ont été formées sans sel et aux pH 3 et 5 (Cheung, Wanasundara, & Nickerson, 2014). Une équipe a observé une meilleure stabilité des émulsions formées par des globulines que par des albumines. De plus, la présence d'albumines dans ces émulsions réduirait leur stabilité (Wu & Muir, 2008).

II.4.2.2 Les napines ou albumines

Les napines sont des albumines, qui représentent 35 à 40 % des protéines totales. Leur poids moléculaire est entre 12 et 15 kDa. Elles sont composées de deux chaînes polypeptidiques d'environ 4 et 9 kDa, liées par deux ponts disulfures. Leur point isoélectrique est proche de pH 10-11 (Lonnerdal & Janson, 1972), leur conférant une solubilité élevée sur une large gamme de pH. De manière générale, les napines sont des protéines particulièrement résistantes aux changements de pH, de force ionique et de température. Ce dernier point est particulièrement intéressant par rapport à d'autres protéines végétales, puisque les albumines du colza ont des températures de dénaturation très élevées, supérieures à 100 °C (Perera, McIntosh, & Wanasundara, 2016).

Les albumines sont capables de former des émulsions mais avec une stabilité inférieure aux émulsions formées par des globulines. Les gouttes d'huile dans les émulsions formées par les napines ont été observées de taille plus importante qu'avec les globulines, induisant ainsi une plus faible stabilité (Wu & Muir, 2008).

Les albumines sont également capables de former des mousses de volume important. Il a été aussi montré que la stabilité de ces mousses est meilleure que celles d'isolats protéiques d'autres sources, comme le soja (Wanasundara J. , McIntosh, Perera, Withana-Gamage, & Mitra, 2016).

En ce qui concerne les propriétés de gélification, les albumines nécessitent des températures plus élevées que les globulines pour former des gels qui sont dans tous les cas très instables. Ceci serait dû aux propriétés des napines à résister à des températures élevées. Ainsi, après chauffage, les napines retrouveraient leur conformation initiale ce qui déstabiliserait le gel (Schwenke, Dahme, & Wolter, 1998).

Par ailleurs, le colza appartient à la même famille que la moutarde qui est reconnue comme un allergène majeur. Ainsi, dans la législation française, la moutarde fait partie des 14 allergènes qui doivent être déclarés dans tout aliment ou ingrédient en contenant (Direction de l'information légale et administrative (Premier ministre), Consulté le 16 janvier 2019). C'est pourquoi, les napines font

l'objet d'études sur leur potentiel allergénique (Puumalainen, et al., 2005). Une explication proposée serait que les albumines, ayant une structure très compacte, sont peu attaquées par les enzymes digestives et conserveraient intact leur épitope : il s'agit d'une structure permettant leur reconnaissance par les immunoglobulines du système immunitaire gastro-intestinal (Mills, Madsen, Shewry, & Wichers, 2003) (Wanasundara J. , McIntosh, Perera, Withana-Gamage, & Mitra, 2016).

II.4.3 Les molécules antinutritionnelles

II.4.3.1 L'acide phytique

L'acide phytique (myo-inositol hexaphosphate) est présent dans les céréales et les graines des plantes oléoprotéagineuses. Il constitue la forme de stockage principale de phosphore dans les graines, Figure 10 a), avec environ 80 % du phosphore total (Yang, Huang, Kuo, & Chiou, 2017). Ces phytates sont présents dans les vacuoles des cellules végétales, c'est-à-dire dans un compartiment séparé des protéines. Cependant, au cours de l'étape de pressage des graines, les structures cellulaires sont rompues, ce qui permet le contact entre les molécules contenues dans différents compartiments.

L'acide phytique est chargé négativement sur une large gamme de pH (Martin & Evans, 1986) ce qui le rend réactif vis-à-vis de molécules chargées positivement, comme certains minéraux, ou les protéines (Cheryan & Rackis, 1980). En effet, lorsque le pH est inférieur au point isoélectrique des protéines, ces dernières ont une charge nette positive. Par ailleurs, ces interactions semblent particulièrement affecter les albumines (Mothes, Schwenke, Zirwer, & Gast, 1990).

Une étude a comparé la solubilité de l'acide phytique et de l'azote constituant les protéines du soja en fonction du pH (De Rham & Jost, 1979). La Figure 10 b) montre leurs courbes de solubilité respectives. Entre les pH 6 et 11, les deux composés ont des solubilités élevées. Au contraire, à pH supérieur à 11 ou inférieur à 6 l'acide phytique est insoluble. Cette dernière zone est particulièrement intéressante puisqu'il y a une différence notable de solubilité entre les deux composés. Ainsi, les auteurs ont utilisé cette caractéristique pour générer des isolats avec une teneur réduite en acide phytique (De Rham & Jost, 1979).

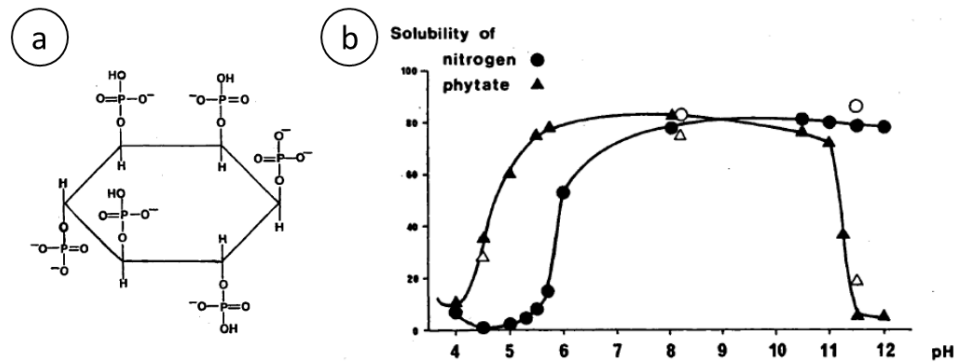


Figure 10: a) Structure de l'acide phytique (Cheryan & Rackis, 1980); b) Graphes de solubilité de l'azote (cercles) et des phytates (triangles) (De Rham & Jost, 1979).

II.4.3.2 Les composés phénoliques

Les composés phénoliques sont présents dans les plantes oléoprotéagineuses et dans leurs graines. Ces molécules, issues du métabolisme secondaire des plantes, sont aujourd'hui recherchées car elles possèdent souvent des propriétés intéressantes pour le traitement de maladies, par des actions antioxydantes. Dans le tourteau de colza, les composés phénoliques sont présents en proportion massique de 1,5 à 2 %. Ils peuvent être libres, sous forme d'esters ou liés à la fraction insoluble du tourteau.

Les composés phénoliques présents dans le colza, sont à plus de 90 % sous formes d'esters. Sinon, ils sont sous forme libre ou sous forme de composés insolubles liés. Les polyphénols estérifiés majoritaires sont des dérivés de l'acide sinapique (70 à 85 % des composés phénoliques libres) (Shahidi & Naczk, An Overview of the Phenolics of Canola and Rapeseed: Chemical, Sensory and Nutritional Significance, 1992). Dans le cadre de l'extraction des protéines du tourteau de colza, la coextraction de composés phénoliques peut avoir des conséquences négatives sur la qualité des protéines. En effet, de par leur structure, les composés phénoliques sont susceptibles de se complexer avec les protéines par interactions faibles mais aussi de manière covalente. En particulier, à pH basique, les composés phénoliques peuvent s'oxyder et s'associer irréversiblement aux protéines (Naczk, Amarowicz, Sullivan, & Shahidi, 1998). La présence de composés phénoliques dans des isolats de protéines peut affecter significativement la couleur et le goût des isolats de protéines (Sosulski, 1979) (Shahidi & Naczk, An Overview of the Phenolics of Canola and Rapeseed: Chemical, Sensory and Nutritional Significance, 1992).

En résumé :

Le tourteau de colza est une source de protéines d'intérêt pour l'alimentation humaine. Cependant, c'est une ressource complexe contenant de nombreux composés susceptibles d'affecter l'extraction des protéines, voire leur qualité après purification comme c'est le cas des phénols ou de l'acide phytique qui sont capables de se complexer aux protéines de manière irréversible. La connaissance de cette composition est essentielle pour le choix des conditions opératoires appliquées aux étapes du procédé de production d'isolats de protéines.

III Les procédés de production d'isolats protéiques à partir de tourteau de colza

Les propriétés nutritionnelles et fonctionnelles des protéines font du colza une ressource de choix pour la production d'isolats de protéines. Cette troisième partie fait l'état de l'art des procédés, proposés jusqu'à aujourd'hui, avec leurs avantages et inconvénients.

Le procédé de production d'isolats de protéines peut se décomposer en deux étapes principales. D'abord, une extraction solide-liquide permet d'extraire les protéines contenues dans le tourteau. Puisque l'objectif du projet de thèse est de produire des protéines pour l'alimentation humaine, seuls les procédés utilisant des solvants aqueux sont considérés. Ensuite, une seconde étape est nécessaire dans le but de purifier les protéines tout en éliminant les composés co-extraits qui pourraient nuire à leur qualité.

III.1 L'extraction des protéines du tourteau de colza

L'extraction des protéines est réalisée par la mise en suspension de tourteau de colza dans de l'eau. Plusieurs caractéristiques, du tourteau ou du milieu d'extraction, sont susceptibles d'affecter l'extraction des protéines : le procédé de fabrication du tourteau, le pH, l'ajout de sel, la température, la durée de l'extraction ou la proportion de solide / liquide. La Figure 11 résume les paramètres reconnus dans la littérature scientifique pour leur influence sur l'extraction des protéines.

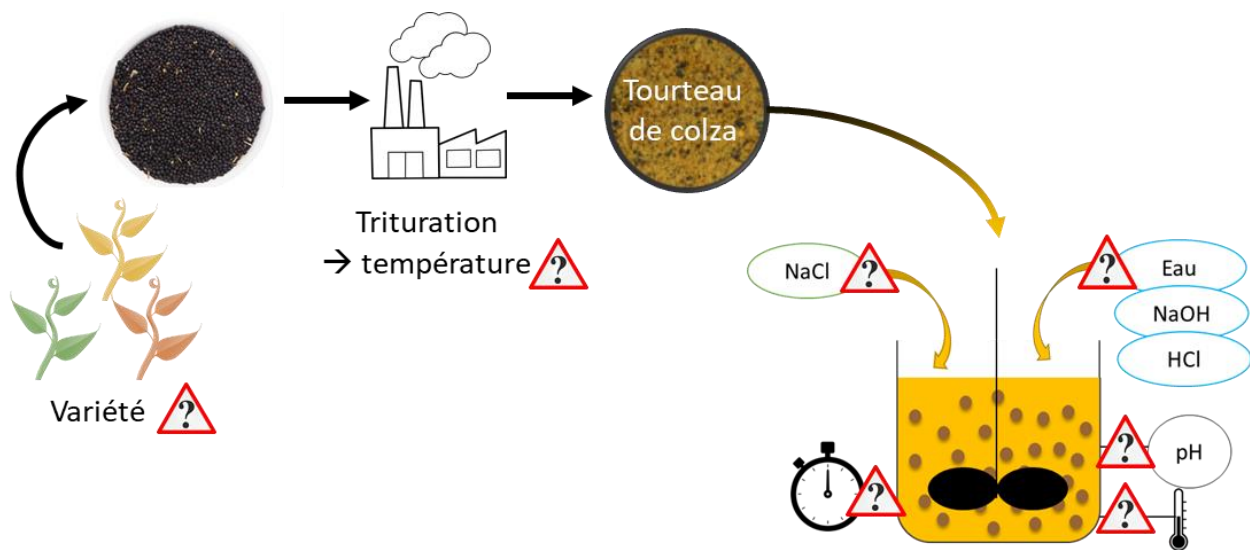


Figure 11 : Schéma des paramètres connus pour leur influence sur l'extraction des protéines.

III.1.1 Effet du procédé de production du tourteau

L'intérêt de l'extraction des protéines de colza pour une utilisation en alimentation humaine a commencé à se développer dans les années 1970 – 1980. La bioraffinerie du colza visait alors à la production de grands volumes d'huile, sans considération pour la qualité des protéines contenues dans le tourteau. Afin d'augmenter les rendements en huile, une étape de déshuilage, avec de l'hexane, était réalisée, et nécessitait ensuite un traitement thermique à haute température (voir partie II. La bioraffinerie du colza). Cependant, les protéines sont des molécules hautement structurées qui sont altérées par exposition à des températures élevées. Actuellement, les usines fonctionnent encore selon ce procédé de pressage puis de déshuilage du tourteau avec de l'hexane. Néanmoins, avec la prise de conscience du potentiel de valorisation du tourteau de colza, des études ont été menées pour quantifier l'impact de ces procédés sur la qualité des protéines.

Ainsi, le but d'une étude récente était d'enquêter sur la qualité des protéines à différentes étapes du procédé de la production du tourteau. Deux étapes sont susceptibles d'exposer les tourteaux à de fortes températures : le pressage qui peut être réalisé à froid (maximum 40 °C) ou à chaud (supérieur à 100 °C) ; et la désolvantation subséquente du déshuilage à l'hexane, réalisée à 60 °C dans l'étude. Il a été montré que la température de traitement des tourteaux contribue à réduire l'extractibilité des protéines : elle est comprise entre 10 % et 40 % sur la gamme de pH entre 3 et 12, pour un tourteau industriel déshuilé et chauffé jusqu'à 110 °C (toasté), contre 35 à plus de 80 % pour un tourteau pressé à froid et chauffé à 60 °C lors du déshuilage (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018).

Ainsi, l'utilisation de tourteaux exposés à de hautes températures n'est pas désirable pour la production d'isolats protéiques puisqu'ils ne permettraient pas d'atteindre des rendements aussi élevés qu'avec un tourteau pressé à froid et déshuilé dans des conditions douces de température.

III.1.2 Effet du pH

Afin d'extraire les protéines du tourteau de colza, les premiers procédés développés ont considéré l'extraction d'azote dans sa globalité, sans distinction entre les globulines et les albumines. Ainsi, il a été montré que l'extractibilité de l'azote, contenu dans le tourteau de colza est la plus élevée lorsque le pH est basique (Radwan & Lu, 1976) (Tzeng, Diosady, & Rubin, 1988) (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018). C'est pourquoi, de nombreuses études ont été réalisées à pH élevé, souvent entre 10 et 12 pour extraire un maximum de protéines. Par la suite, il a été montré que ces conditions d'extraction présentent des inconvénients. D'abord, les protéines sont susceptibles d'être hydrolysées, ce qui entraîne une altération de leurs propriétés fonctionnelles. Dans une étude de 2004, les protéines extraites à pH 10 ont de meilleures propriétés moussantes et émulsifiantes que des protéines extraites à pH 11 et 12 (Pedroche, et al., 2004). De plus, dans un environnement à pH basique, les composés phénoliques co-extraits avec les protéines sont sujets à des phénomènes d'oxydation, les rendant capables de former des liaisons covalentes avec les protéines (Naczka, Amarowicz, Sullivan, & Shahidi, 1998). Ceci a pour effet de donner une couleur indésirable aux isolats protéiques.

Plus récemment, il a été montré qu'à pH fortement acide, entre 2 et 4, les albumines sont extraites sélectivement. Cependant, les concentrations en protéines dans ces extraits sont faibles par rapport aux autres points de pH (Nioi, Kapel, Rondags, & Marc, 2012). Un criblage de pH d'extraction réalisé par Fetzer *et al.* a confirmé ces résultats (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018). La faible extractibilité des protéines à pH acide explique le peu d'études dans ces conditions. Par ailleurs, il a été montré que cette extraction sélective des albumines peut s'expliquer par une dénaturation irréversible des globulines à pH acide qui restent ainsi dans la phase solide, alors que les albumines seraient solubles en phase aqueuse (Perera, McIntosh, & Wanasundara, 2016).

Entre les pH 5 et 9, il semble qu'un mélange de protéines est extrait avec une majorité de globulines. Sur cette gamme de pH, l'extractibilité des protéines est plus élevée (50 – 60 %) qu'à pH 2 – 4 (35-40 %) (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018).

III.1.3 Effet du ratio solide / liquide

La quantité de solide mis en suspension dans le milieu d'extraction est un paramètre susceptible d'avoir une influence sur les rendements d'extraction. En 2012, Nioi *et al.* ont montré, pour une

extraction à pH 2, une diminution des rendements protéiques au-delà d'un ratio S/L de 12 % (m/m). De plus, la viscosité de la phase liquide dans ces conditions était problématique quant à l'agitation du mélange (Nioi, Kapel, Rondags, & Marc, 2012). Fetzer *et al.* ont réalisé des expériences similaires à pH 7 et ont montré qu'un ratio S/L de 1 : 9 donne le même rendement qu'un ratio de 1 : 19 (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018). A pH basique, la dilution du mélange n'a pas permis d'extraire plus de protéines (Gerzhova, Mondor, Benali, & Aider, 2016). Ainsi, dans le cas de l'extraction des protéines de colza, l'augmentation du ratio S/L ne permet pas d'augmenter les rendements d'extraction. De plus, une quantité trop importante de solide entraîne une réduction des rendements avec une augmentation de la viscosité de la phase liquide. C'est pourquoi, afin de limiter la consommation d'eau du procédé, tout en facilitant les étapes de purification, de nombreux auteurs ont utilisé une proportion de solide proche de 10 % (m/m) (Aluko & McIntosh, 2001; Pedroche, et al., 2004).

III.1.4 Effet de la quantité de sel

L'ajout de sel, le plus souvent du NaCl, dans un milieu aqueux, peut contribuer à une meilleure solubilité des protéines en phase aqueuse. Cet effet, appelé « salting-in », consiste en une interaction entre les charges à la surface des protéines et les sels qui améliore l'état de solvation des protéines, c'est-à-dire, leur dissolution dans le milieu aqueux. Cela entraîne une diminution des interactions protéine-protéine.

Des études ont été menées sur l'extraction des protéines de colza en présence de sel. Notamment, en 2016, une étude a comparé la quantité de protéines extraites, dans la gamme de pH entre 10 et 12, lorsque du NaCl est ajouté au milieu d'extraction en concentrations de 0 à 1 M (Gerzhova, Mondor, Benali, & Aider, 2016). L'ajout de sel à pH 10 augmente la quantité de protéines extraites, de façon plus marquée qu'à pH 11 ou 12. De plus, ceci a conduit à une diminution de la proportion d'albumines. Par conséquent, l'ajout de sel est susceptible de changer les proportions des deux fractions de protéines dans les extraits.

Fetzer *et al.* ont étudié l'influence de l'ajout de NaCl à pH 7 avec des concentrations entre 0 et 1 M. Entre 0 et 0,25 M, le rendement d'extraction en protéine augmente significativement. Cependant, au-delà de cette concentration, les rendements observés sont statistiquement identiques (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018).

Par ailleurs, dans l'industrie, les cuves en inox sont généralement résistantes à des concentrations en ions chlorure d'environ 0,2 M. Ceci est un paramètre à prendre en compte pour l'extrapolation du procédé ainsi que pour les problématiques de traitement des eaux.

III.1.5 Effet de la température

La température est un paramètre connu pour augmenter les constantes de partage entre deux phases, liquide-liquide ou bien solide-liquide. Elle favorise la solubilisation des molécules dans un milieu. Cependant, dans le cas du procédé d'extraction des protéines de colza, son influence a été peu étudiée par rapport aux autres paramètres. Cela s'explique par le fait qu'en augmentant la température d'un milieu aqueux, le risque de prolifération bactérienne s'accroît, particulièrement entre 30 et 40 °C. Les températures supérieures à 60 °C semblent également peu désirables puisque les protéines risqueraient de se dénaturer, dégradant ainsi la qualité du produit final.

Une étude de 1976 a montré les profils d'extractibilité de l'azote, en fonction du pH et à différentes températures entre 25 et 55 °C (Radwan & Lu, 1976). Deux points sont notables :

- les rendements sont globalement plus élevés avec des valeurs comprises entre 10 et 25 %, pour une température de 55 °C. Cet effet est modulé par le pH car il est remarquable que les plus forts gains de rendements soient observés soit à pH inférieur à 3, soit à pH supérieur à 10.
- le minimum d'extractibilité de l'azote est déplacé : à 25 °C il est proche de pH 4 – 5, alors qu'à 55 °C il est situé à pH 7.

Ainsi, cette étude démontre un fort effet de la température sur l'extractibilité des protéines. Afin de compléter et d'interpréter ces données, il serait pertinent de disposer des proportions en chacune des deux familles de protéines. Cela permettrait d'évaluer une éventuelle différence d'extractibilité des albumines et des globulines du colza sur toute la gamme de pH puisque leurs points isoélectriques diffèrent.

III.1.6 Effet de la durée d'extraction

La durée de l'extraction dépend de la cinétique de transfert des protéines du tourteau vers la phase liquide. Plusieurs articles montrent que des durées de 15 à 30 minutes sont suffisantes pour atteindre l'équilibre. Notamment, Nioi *et al.* ont montré que, pour une extraction réalisée à pH 2, l'équilibre est atteint à partir de 15 minutes (Nioi, Kapel, Rondags, & Marc, 2012). Dans l'étude de Fetzer *et al.*, le rendement en protéines est statistiquement identique entre 30 minutes et 240 minutes, lorsque l'extraction est réalisée à pH 7 (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018).

En résumé :

De nombreux paramètres ont été identifiés comme ayant une influence sur l'extractibilité des protéines. Cependant, un grand nombre d'études est focalisé sur l'obtention de forts rendements sans considérer les qualités requises pour un isolat protéique utilisable en alimentation humaine, comme la couleur ou encore les propriétés fonctionnelles. Ainsi, une gamme réduite de conditions a été décrite. De plus, le solide résiduel est souvent négligé alors qu'il pourrait / devrait être valorisé.

III.2 La purification des protéines de colza

Suite à l'extraction, la fraction liquide, appelée extrait, est séparée de la phase solide par centrifugation ou décantation. L'extrait clarifié est un mélange complexe de molécules en solution dont la matière sèche contient environ 30 % de protéines. Le reste est constitué d'autres composés solubles : des sucres, des minéraux, des acides aminés libres, de l'acide phytique et des composés phénoliques. Afin d'obtenir un isolat de protéines utilisable en alimentation humaine, il est essentiel de purifier les protéines. A l'issue de la purification, plusieurs qualités de produits sont distinguées : d'une part les concentrats, correspondant à des produits avec une proportion d'au moins 65 % de protéines, et, d'autre part, les isolats, avec plus de 90 %.

Les paragraphes suivants décrivent l'état de l'art des méthodes de purification appliquées aux protéines issues du tourteau de colza ainsi que leurs avantages et inconvénients, pour une extrapolation à l'échelle industrielle.

III.2.1 La précipitation isoélectrique

Le procédé de purification par précipitation isoélectrique est utilisé industriellement pour la production d'isolats de protéines de soja. Il repose sur le fait que les protéines présentent un minimum de solubilité lorsque le pH du milieu correspond à leur pH isoélectrique. La première étape consiste en une extraction à un pH neutre, voire basique, des protéines. Une séparation solide-liquide permet ensuite d'obtenir l'extrait. Dans une deuxième étape, le pH de l'extrait est ajusté à un pH proche du point isoélectrique des protéines recherchées, afin d'entraîner leur précipitation. Le culot de précipitation est ensuite lavé et lyophilisé. Dans le cas du soja, les protéines majoritaires sont les globulines (7 S et 11 S qui représentent environ 80 % des protéines). Le passage à un pH entre 3 et 5 entraîne une précipitation sélective de ces protéines.

Ce procédé a été mis en œuvre en 1997 sur un tourteau de colza par Klockerman *et al.* L'extraction est réalisée à pH basique et la précipitation isoélectrique à pH 3,5. Ainsi, les globulines sont précipitées et les albumines, dont le point isoélectrique est proche de 10-11, restent en solution. Dans cet article, seul le précipité est lavé et lyophilisé. Le produit final est ainsi composé de globulines et les albumines,

restées en solution contribuent à une perte du rendement en protéines (Klockeman, Toledo, & Sims, 1997). Cependant, comme le montre Aluko *et al.*, ce type de procédé ne permet pas de produire des isolats : sur quatre variétés de colza, ce procédé a généré des protéines, avec une pureté de 77 à 82 % (Aluko & McIntosh, 2001).

Plus récemment, un procédé a été mis en œuvre en incluant une étape supplémentaire de purification des albumines, présentes dans le surnageant (Akbari & Wu, 2015). Cependant, les rendements de ce procédé sont faibles avec environ 50 % des protéines totales récupérées sous la forme d'un concentrat d'albumines et d'un isolat de globulines.

III.2.2 La précipitation à froid ou la micellisation

La micellisation est un procédé de purification, proposé pour la première fois dans un brevet de 1979, utilisable pour la production de produits protéiques à partir de diverses sources végétales telles que les graines de plantes amidonnées ou les graines de plantes oléoprotéagineuses (Etats-Unis Patent No. 4,169,090, 1979). Le brevet précise que les protéines, de par leur caractère amphiphile, s'arrangent spontanément en structures micellaires lorsque la force ionique du solvant aqueux diminue subitement. Cette structure s'explique par une diminution de solvatation des protéines, entraînant une augmentation des interactions protéines – protéines de type Van der Waals, pour former des agrégats globulaires appelés « micelles ». Ces dernières sont essentiellement formées de protéines de haut poids moléculaires et peuvent être séparées du liquide, contenant les solutés hydrosolubles, par une séparation solide / liquide. Ce phénomène est potentialisé à faible température car les interactions de Van der Waals sont alors favorisées.

Un brevet de 2010 propose un procédé de production d'isolats de protéines de colza composé majoritairement de globulines 7 S (United States Patent No. US 2010/0063255 A1, 2010) (Aider & Barbana, 2011) en plusieurs étapes. D'abord, une extraction des protéines par une solution saline (environ 0,2 M de NaCl) est réalisée. L'extrait est ensuite concentré par procédé membranaire jusqu'à une concentration très élevée comprise entre 200 et 300 g/L. La force ionique dans le rétentat est maintenue constante en choisissant un seuil de coupure de membrane suffisamment élevé pour retenir les protéines tout en laissant passer le NaCl. Dix à quinze volumes d'eau à 4°C sont ensuite ajoutés par volume d'extrait concentré dans le but de former les micelles qui sont enfin récupérées par centrifugation.

L'avantage de ce type de procédé est que, contrairement à la précipitation acide ou isoélectrique, les globulines ne sont pas dénaturées. En revanche, des faibles rendements en protéines sont observés. Une explication probable réside dans les proportions en globulines et en albumines dans le tourteau

de colza (50 à 55 % et 30 à 35 % respectivement) qui sont relativement équilibrées par rapport à d'autres agro-ressources comme le soja où les globulines représentent environ 90 % des protéines de la graine.

III.2.3 La purification par des méthodes chromatographiques

La chromatographie est une technique permettant de séparer des molécules contenues dans un échantillon, en fonction de leurs propriétés physico-chimiques telles que l'hydrophobie, la charge ou encore la taille. La chromatographie liquide repose sur l'entraînement des molécules de l'échantillon par une phase mobile, l'éluant, à travers une phase stationnaire. Les phases mobiles et stationnaires sont sélectionnées pour créer des affinités différentielles entre les molécules du mélange et les deux phases. Ainsi, en début de la colonne de séparation, les molécules ayant le plus d'affinité pour la phase stationnaire sont retenues alors que les molécules qui le sont moins sont entraînées par la phase mobile. Les molécules sont alors caractérisées par un temps de rétention et l'intensité d'un signal (souvent UV à 214 nm ou 280 nm), dépendant de sa concentration.

Des procédés de purification chromatographique des protéines d'extraits de tourteau de colza ont été proposés (2016). Ces procédés sont souvent mis en œuvre en plusieurs étapes. Par exemple, une séquence proposée est la suivante :

1. d'abord la chromatographie d'exclusion stérique permet la séparation entre les protéines de taille élevée et les sels ;
2. ensuite, une étape utilisant la chromatographie d'échange ionique (cation), permet la séparation des albumines et des globulines. Les napines interagissent avec la phase stationnaire de la colonne alors que les globulines sont éluées rapidement ;
3. enfin, une troisième étape, de chromatographie d'exclusion stérique, permet la purification des napines (Perera, McIntosh, & Wanasundara, 2016).

Cette méthode présente plusieurs inconvénients : d'abord, l'absence de données concernant les rendements laisse supposer que le nombre d'étapes entraîne probablement une réduction du taux de recouvrement des protéines ; ensuite, la chromatographie, bien qu'extrapolable, n'est pas compatible, en termes de coût et de productivité, avec le traitement annuel de milliers de tonnes de protéines.

III.2.4 La purification par procédés membranaires

Les procédés membranaires reposent sur l'utilisation de membranes dont les tailles des pores permettent de séparer des molécules selon leur taille. La microfiltration retient les éléments de plus de 0,1 μm ; l'ultrafiltration, les molécules de plus de 0,01 μm environ ; la nanofiltration est utilisée

pour les éléments de taille encore inférieure. La fraction retenue par la membrane est appelée le rétentat, et la fraction passant à travers les pores correspond au perméat. Ce faisant, le volume de rétentat diminue, entraînant donc une concentration des molécules dans ce compartiment. Afin de purifier les molécules du rétentat de façon plus poussée, il est possible d'ajouter une alimentation en eau pour maintenir le volume du rétentat constant. Ce mode de filtration, appelé diafiltration, est largement utilisée pour la purification de protéines. La Figure 12 montre les différents compartiments ainsi que les flux impliqués dans la diafiltration.

Les procédés membranaires reposent sur le principe de la filtration tangentielle c'est-à-dire avec une circulation du flux tangentiellement à la membrane. Une pression appliquée dans le système permet le passage des plus petites molécules à travers les pores de la membrane. Ces procédés sont particulièrement utilisés pour la filtration de petites molécules.

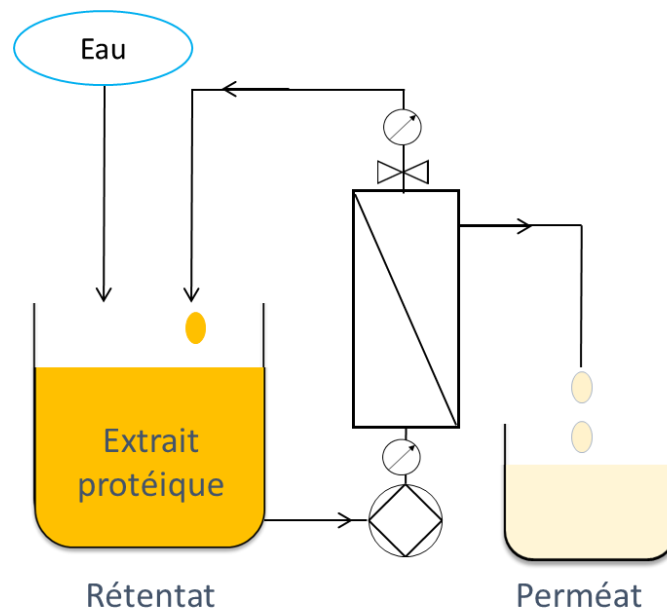


Figure 12 : Schéma de la diafiltration par ultrafiltration appliquée à un extrait de protéines

C'est pourquoi, tôt dans le développement des méthodes de purification des protéines, la filtration à flux tangentiel a été utilisée sur des membranes d'ultrafiltration. Dès 1988, un procédé combinant procédés membranaires et précipitation isoélectrique des globulines a abouti à deux produits : un isolat de globulines et un concentrat d'albumines. Pour cela, une première étape de diafiltration a été réalisée avec une membrane de 30 kDa pour atteindre un facteur de concentration de 10 afin de concentrer les protéines avant précipitation isoélectrique. Ensuite, le surnageant a été purifié sur une membrane de 10 kDa avec 5 diavolumes d'eau acidifiée. Un recouvrement d'environ 60 à 70 % des protéines a été reporté par un tel procédé. Mais le « meilleur produit » d'après les auteurs est celui qui était riche en albumines, et qui a été obtenu avec un rendement de 18 % (Tzeng, Diosady, & Rubin,

1988). Une stratégie similaire a été mise en œuvre en 2015 avec, en plus, une étape préalable de lavage acide du tourteau afin d'éliminer les composés phénoliques et l'acide phytique. Ensuite, une extraction en conditions basiques, suivie d'une précipitation isoélectrique, a permis de séparer les albumines et les globulines. Comme précédemment, le surnageant de la précipitation isoélectrique a été purifié par diafiltration sur une membrane de 10 kDa. La pureté la plus élevée a été obtenue avec un facteur de concentration de 20 et une diafiltration avec 10 diavolumes d'eau. Cependant, le grade d'isolat n'a pas été atteint, puisque le produit final riche en albumines contient 82 % de protéines pour un rendement de 12 %. Il est également à noter que les deux fractions protéiques ont été obtenues avec des teneurs faibles en phénols et en acide phytique (< 2%) (Akbari & Wu, 2015).

Ainsi, les procédés membranaires ont prouvé leur efficacité dans la purification des protéines extraites du tourteau de colza. De plus, un avantage majeur est qu'ils sont facilement extrapolables à une échelle industrielle.

En résumé

Dans la littérature scientifique, de nombreuses techniques de purification des protéines ont été mises en œuvre pour l'étude des protéines de colza. Les procédés membranaires se démarquent des autres par leur facilité à être extrapolés industriellement, ainsi que pour leurs rendements. De plus, c'est une méthode qui n'est pas dénaturante pour les protéines. C'est pourquoi, cette méthode a été principalement utilisée dans les travaux de thèse.

III.3 Les procédés brevetés

Dans les années 1980, plusieurs brevets ont été déposés concernant l'extraction de protéines à partir d'agro-ressources et particulièrement, plusieurs le sont par Edward D. Murray, à travers les sociétés General Foods Limited et B.M.W. Canola Inc. Les procédés qui y sont décrits reposent principalement sur la précipitation micellaire dans le but de purifier les globulines de ces agro-ressources.

En 1996, un brevet présente un procédé partant d'un tourteau gras de colza (Etats-Unis Patent No. 08/594909, 1996). L'extraction est réalisée dans un milieu à un pH de 5 à 6,8, avec une force ionique d'au moins 0,2 M et à une température de 15 à 75 °C permettant d'extraire un mélange d'albumines et de globulines. Après séparation solide-liquide, l'extrait protéique et la phase aqueuse sont séparées. Une étape de concentration par ultrafiltration (30kDa) est réalisée en maintenant la force ionique. De l'eau ensuite est ajoutée et la force ionique devient alors inférieure à 0,2 M ce qui a pour effet d'entraîner la formation de masses micellaires composées des globulines. Ces protéines sont ensuite séparées du liquide et séchées pour donner un isolat. Ce procédé, réalisé à un pH modérément acide, permet de limiter les phénomènes d'oxydation des polyphénols, typique des pH basiques.

Cependant, comme détaillé précédemment (Partie III.1.2. Effet du pH), les pH proches de la neutralité ne permettent pas d'atteindre des rendements élevés. De plus, ce mode opératoire ne permet que la récupération des globulines, présentes en plus faible proportion que dans d'autres graines telles que le soja. Cela signifie donc la perte des albumines, restées solubles, et finalement un faible rendement global.

Afin de valoriser les deux fractions protéiques, albumines et globulines du colza, la société Burcon Nutrascience Corp. propose à partir de 2002 une série de brevets concernant un procédé à partir de tourteaux, issus d'un déshuilage à l'hexane ou d'un pressage à froid. Les étapes d'extraction et de purification par précipitation micellaire sont identiques à celles présentées dans les brevets de Murray. Néanmoins, une purification par ultrafiltration des albumines du surnageant est proposée, permettant ainsi d'aboutir à deux isolats distincts de globulines et d'albumines (Etats-Unis Patent No. US 7,611,735 B2, 2009).

La société Burcon dépose d'autres brevets décrivant différentes technologies pour la production d'isolats, dont un proposant une extraction des albumines à chaud (entre 70 et 100°C) et en milieu salin. Dans ces conditions, une extraction préférentielle des protéines 2S du colza est observée et elles peuvent ensuite être purifiées par procédés membranaires. Dans le même brevet, un procédé d'extractions séquentielles (globulines puis albumines) en deux étapes est décrit. La première étape permettant d'extraire majoritairement des globulines, est réalisée à un pH entre 5 et 6,8 sans sel et entre 55 et 65 °C. La seconde étape, pour l'extraction des albumines, est réalisée à un pH entre 5,3 et 6,2 avec 0,05 à 1M de NaCl et entre 20 à 30 °C. Comme précédemment, la purification de chaque fraction peut être réalisée par ultrafiltration ou par précipitation micellaire pour les globulines. De nouveau, un tel procédé permet d'obtenir deux produits (Etats-Unis Patent No. US 7,955,625 B2, 2011). En 2010, la FDA (Food and Drug Administration) autorise le statut GRAS (Generally Recognized as Safe) pour la Puratein (riche en globulines) ainsi que pour la Supertein (riche en albumines) de Burcon (Archer Daniels Midlans Co. (ADM), 2010). Néanmoins, aucune exploitation industrielle de ces brevets n'est faite à ce jour.

En Europe, une demande similaire a été déposée auprès de la Commission Européenne afin d'autoriser la mise sur le marché des protéines issues du colza. Cette demande a été déposée par la société Helm AG, et reprise par la société allemande Siebte. En 2014, la Commission Européenne a approuvé cette demande qui a été publiée au journal officiel de l'Union européenne (Commission européenne, 2014).

La société Siebte, qui deviendra par la suite la société Teutexx puis Bioexx Speciality Proteins LTD, est également détentrice de plusieurs brevets concernant l'extraction des protéines du colza. Le procédé proposé consiste en une extraction à un pH plus basique que précédemment avec une gamme de pH entre 7 et 10. Plusieurs étapes d'extraction sont réalisées dans les mêmes conditions. Les protéines sont ensuite purifiées par procédés membranaires (Europe Patent No. EP 2 293 685, 2009). Ce procédé permet d'aboutir à un produit nommé Isolexx, qui a obtenu une notification GRAS également. Le même procédé peut également être utilisé pour obtenir un autre produit, Vitalex (également GRAS), obtenu dans les mêmes conditions mais avec l'ajout d'une étape de protéolyse enzymatique (Keller and Heckman LLP, 2011).

Plus récemment, la société DSM a déposé plusieurs brevets pour l'extraction de protéines de plantes oléoprotéagineuses. L'extraction est aussi réalisée à un pH neutre voire basique sans sel. Les protéines sont ensuite concentrées et diafiltrées. Puis, les protéines sont précipitées avec de l'éthanol et séchées (Etats-Unis Patent No. US 2015/0073127 A1, 2015).

Globalement, de nombreux brevets ont été déposés pour extraire les protéines du tourteau de colza depuis les années 1980. Les procédés décrits dans cette partie sont résumés en Figure 13.

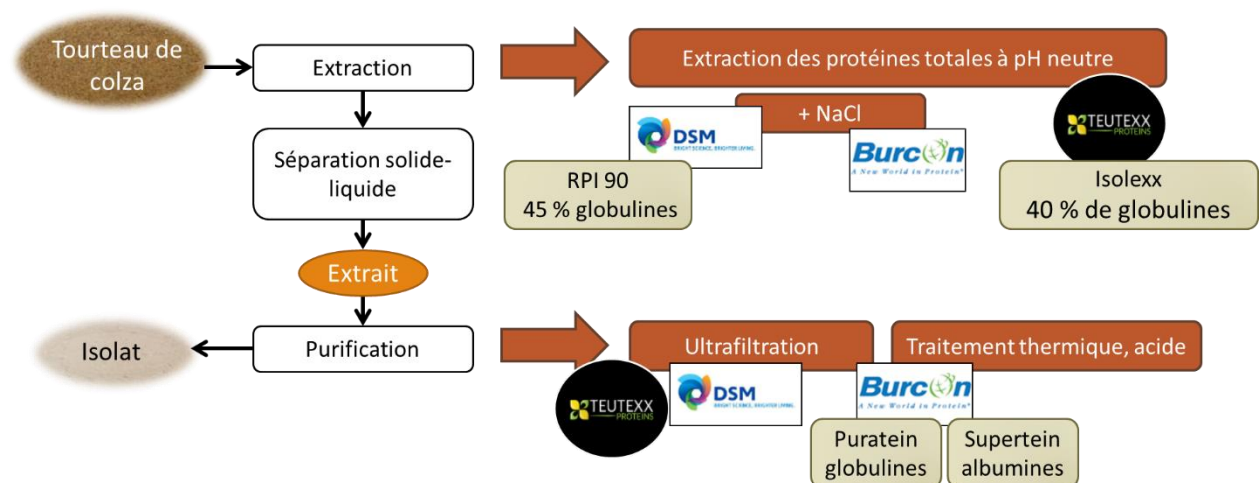


Figure 13 : Schéma bilan des procédés brevetés pour la production d'isolats de colza

Contrairement au soja, la stratégie reposant sur la purification des protéines par précipitation isoélectrique des globulines est absente des brevets, probablement à cause des faibles rendements. En effet, contrairement au soja, le colza contient une grande proportion d'albumines. C'est pourquoi, les brevets décrits se sont d'abord orientés vers l'extraction de mélanges de protéines à pH neutre/basique. Par la suite, les brevets ont reflété la volonté de séparer les deux familles de protéines afin d'en exploiter leurs fonctionnalités propres. Cependant, malgré ces avancées, aucun procédé

n'est industrialisé aujourd'hui pour produire un isolat protéique commercialisable. Une hypothèse expliquant cela serait que les procédés proposés ne permettent pas d'atteindre des rendements suffisamment élevés pour être rentables : soit à cause d'une limitation d'extractibilité des protéines, soit à cause de limitations dues à la purification. En effet, les procédés brevetés présentent souvent plusieurs étapes de purification : dans certains cas, la première étape sert à séparer les fractions globulines et albumines qui sont ensuite purifiées séparément par procédés membranaires. La présence de multiples étapes de séparation et de purification entraîne d'autant plus de pertes en protéines, ce qui se reflète par de faibles rendements.

Il est également à noter que les tourteaux utilisés, à l'exception de ceux des brevets de Murray, étaient pour la plupart déshuilés à l'hexane avec une étape de désolvantation à chaud. Il est possible que ce traitement thermique soit à l'origine des faibles rendements. C'est pourquoi, une hypothèse de départ des travaux de thèse a été de considérer qu'un tourteau gras, issu de pressage à froid, permettrait d'atteindre de plus hauts rendements d'extraction qu'un tourteau déshuilé.

IV Les outils analytiques pour le dosage et la caractérisation des protéines

Afin d'améliorer la compréhension des phénomènes mis en jeu au cours de l'extraction des protéines de colza, il est nécessaire de disposer de méthodes analytiques pour caractériser et quantifier les molécules hydrosolubles présentes dans les extraits. Les paragraphes suivants dressent un état de l'art des outils de dosage et de caractérisation des protéines afin de définir et de justifier l'utilisation des outils utilisés au cours du travail de thèse.

IV.1 Les techniques de dosage des protéines

IV.1.1 La méthode de dosage de référence, le Kjeldahl

La méthode de Kjeldahl permet le dosage de l'azote dans un échantillon liquide ou solide, ce qui n'est pas le cas des méthodes colorimétriques ou chromatographiques. Ensuite, l'application d'un facteur de conversion permet de déterminer la concentration en protéines. D'après l'Association of Official Agricultural Chemists (AOAC 979.09), il s'agit de la méthode de référence pour la quantification des protéines.

L'azote est un constituant majeur des protéines, mais il peut être présent dans d'autres molécules. Pour distinguer ces deux types d'azote, deux analyses sont nécessaires : l'une sur l'échantillon brut et l'autre, en réalisant une précipitation des protéines avec de l'acide trichloroacétique (TCA). Dans ce second cas, le surnageant, après précipitation, est analysé et donne accès à la teneur en azote non

protéique. La différence entre l'azote total et l'azote non protéique donne la teneur en azote protéique.

Pour l'analyse de graines et de céréales, le facteur de conversion protéique le plus utilisé est égal à 6,25. En ce qui concerne le colza, d'autres valeurs ont été proposées à 5,29 et 5,41 à partir d'aminogrammes (Mariotti, Tomé, & Patureau Mirand, 2008) (Sarwar, et al., 1983) (Tkachuk, 1969).

La méthode de Kjeldahl étant la seule reconnue officiellement comme référence pour le dosage des protéines, elle a été utilisée au cours des travaux de thèse. Cependant, elle présente des inconvénients : certaines substances utilisées sont concentrées et peuvent être dangereuses (corrosives, cancérigènes, mutagènes et reprotoxiques). Par ailleurs, dans le cadre du projet de thèse, cette méthode présente une limite majeure puisqu'elle ne permet pas la détermination des teneurs spécifiques en chacune des deux protéines majoritaires.

IV.1.2 Le dosage par analyse élémentaire, la méthode de Dumas

Une autre méthode couramment utilisée est la méthode de Dumas. Elle consiste en une combustion de l'échantillon à très haute température (950 – 1050 °C) produisant des molécules de NO_x et N₂. Les oxydes d'azote sont ensuite réduits en diazote. La mesure de vapeurs de N₂ par un détecteur à conductivité thermique donne la quantité d'azote dans l'échantillon.

L'avantage de cette méthode est qu'elle est facile et rapide à mettre en œuvre et n'implique pas l'utilisation de produits dangereux. Cependant, sa précision est comparable au Kjeldahl et ne permet pas non plus la différenciation entre les différents types de protéines (Moore, DeVries, Lipp, Griffiths, & Abernethy, 2010) (Jung, et al., 2003).

IV.1.3 Les dosages colorimétriques

De nombreuses méthodes de dosage colorimétriques ont été développées pour le dosage des protéines. La plupart reposent sur le principe de la formation de complexes colorés par réaction avec des ions cuivre Cu²⁺. En particulier, la méthode de Lowry a été longtemps utilisée. La première étape consiste à faire réagir les protéines de l'échantillon avec des ions cupriques et du tartrate de potassium ou de sodium. L'ajout du réactif de Folin-Ciocalteu permet la formation d'un complexe absorbant à 750 nm (Lowry, Rosebrough, Farr, & Randall, 1951). Comparée au Kjeldahl, cette méthode surestime la teneur en protéines des extraits de tourteau de colza à cause de la forte teneur en composés phénoliques co-extraits avec les protéines (Lindeboom & Wanasundara, 2007). Elle présente aussi un inconvénient majeur qu'elle nécessite un étalonnage préalable puisque la réponse observée dépend de la protéine analysée.

Bien que cette méthode soit facile à mettre en œuvre, peu dangereuse et peu coûteuse, la présence de polyphénols dans les extraits biaise significativement le dosage des protéines. De plus, elle ne donne pas d'information quant à la nature des protéines présentes dans l'échantillon.

IV.1.4 Caractérisation des protéines par électrophorèse SDS-PAGE

Une méthode usuelle est l'utilisation d'un traitement d'image suite à une méthode d'électrophorèse, le SDS-PAGE. Elle consiste à faire migrer les protéines selon leur poids moléculaire : les molécules les plus petites se déplaceront plus rapidement.

Les échantillons sont préparés dans le tampon de dépôt ou tampon de Laemmli, qui contient notamment du dodécylsulfate de sodium (SDS) (Laemmli, 1970). Le SDS est un surfactant anionique qui se fixe aux protéines : les structures secondaires et tertiaires sont supprimées et l'association du SDS avec les portions hydrophobes de la protéine fait qu'elle adopte une structure linéaire. La migration peut être réalisée en conditions réductrices ou non. En conditions réductrices, les ponts disulfures formés par les résidus de cystéines sont réduits ce qui a pour effet de dissocier les sous-unités des protéines liées par ce type de liaison. Pour ce faire, du β -mercaptoéthanol est ajouté au tampon de dépôt, et l'échantillon est chauffé durant 5 minutes à 95 °C, afin de dénaturer les protéines.

Après une migration de plusieurs heures, le gel est coloré avec du « bleu brillant de Coomassie ». Ce dernier se fixe aux protéines par des liaisons hydrophobes et hétéropolaires (Georgiou, Grintzalis, Zervoudakis, & Papapostolou, 2008). Les gels subissent ensuite plusieurs étapes de décoloration afin d'éliminer l'excès de colorant pour que, finalement, les protéines deviennent apparentes sous forme de bandes bleues. A l'aide d'un marqueur de taille, il est possible d'associer une masse moléculaire à une bande.

La densitométrie peut ensuite être utilisée pour quantifier les protéines contenues dans un échantillon. Pour cela, une analyse de l'image du gel est réalisée à l'aide d'un logiciel informatique. L'intensité de chaque bande peut être déterminée et comparée à une calibration préalable. Cette méthode a été appliquée sur des protéines de blé en 1992 pour sa capacité à quantifier et identifier les protéines en une seule analyse. Cependant, de fortes variations ont été observées entre les mêmes échantillons, jusqu'à 12 % sur un même gel (Kolster, Krechting, & Van Gelder, 1992).

Ainsi, l'électrophorèse SDS-PAGE permet l'analyse qualitative, voire quantitative, sélectivement des protéines dont les poids moléculaires diffèrent. Néanmoins, cette méthode est longue, elle nécessite de nombreuses étapes et elle est peu précise.

IV.1.5 Les méthodes chromatographiques

Des méthodes chromatographiques sont aussi utilisées pour analyser les protéines végétales. Dans la plupart des cas, il s'agit de Chromatographie Liquide Haute Performance (CLHP) avec une détection dans l'UV aux longueurs d'onde de 214 ou 280 nm.

La longueur d'onde de 214 nm correspond au domaine d'absorbance des chaînes latérales des acides aminés et de la liaison peptidique. Dans ce cas, le signal est particulièrement sensible mais peu spécifique : les coefficients d'extinction molaire des résidus d'acides aminés ont des valeurs comprises entre 21 L.mol⁻¹.cm⁻¹ pour la glycine, 29060 L.mol⁻¹.cm⁻¹ pour le tryptophane et 923 L.mol⁻¹.cm⁻¹ pour la liaison peptidique (Kuipers & Gruppen, 2007). Néanmoins, à cette longueur d'onde, des acides organiques absorbent également. Il y a donc des risques importants d'interférence du signal dans le cadre d'analyse de protéines dans des matrices végétales complexes.

A 280 nm, ce sont les chaînes latérales du tryptophane et de la tyrosine qui contribuent essentiellement à l'absorbance. Le signal est moins sensible mais plus spécifique qu'à 214 nm : les coefficients d'extinction molaire du tryptophane et de la tyrosine sont respectivement de 5500 et 1490 L.mol⁻¹.cm⁻¹. Cependant, dans le cas de mélanges complexes issus de végétaux, d'autres molécules absorbent, comme les composés phénoliques et des fragments protéiques ce qui est susceptible de biaiser l'analyse.

Deux phases stationnaires de chromatographie sont principalement utilisées pour l'analyse de protéines végétales : la phase inverse et l'exclusion stérique. Les paragraphes suivants décrivent ces deux types accompagnés d'exemples concrets d'utilisation sur des protéines issues d'agro-ressources.

IV.1.5.1 La chromatographie en phase inverse

Dans ce cas, la phase stationnaire est apolaire : elle est constituée de silice greffée avec des chaînes carbonées de longueur allant de C4 à C18. Au contraire, la phase mobile est polaire d'où le nom de phase inverse. La séparation des molécules en phase inverse est basée sur un critère d'hydrophobie des molécules.

En ce qui concerne les protéines, c'est le nombre de chaînes latérales apolaires exposées (tryptophane, phénylalanine, leucine, isoleucine, valine, etc.) qui conditionne leur affinité avec la phase stationnaire. Dans la plupart des cas, l'élution est faite par un gradient croissant d'acétonitrile dans l'eau ce qui permet de réduire la polarité (GraceVydac, 2002). Dans un premier temps, les molécules les plus apolaires se fixent sur la phase stationnaire alors que les autres sont éluées avec la phase mobile. Dans un second temps, il y a désorption des molécules adsorbées du fait de la réduction de la polarité du solvant.

En 2007, un article a été publié concernant une méthode d'analyse des protéines du maïs par CLHP en phase inverse. Les principales familles de protéines des graines ont ainsi été séparées : les albumines, les globulines, les prolamines et les glutélines. Cette méthode a été utilisée à 280 nm pour comparer les profils protéiques entre des plantes transgéniques et sauvages (Rodríguez-Nogales, Cifuentes, García, & Marina, 2007). Cependant, cette méthode n'a pas permis la quantification des teneurs en protéines.

La chromatographie par phase inverse a également été utilisée en tandem avec la spectrométrie de masse à ionisation par un électro-nébuliseur (ESI). Ceci afin de détecter et quantifier des traces de certaines protéines dans un cadre de détection de fraudes. Il s'agit par exemple de détecter des protéines végétales incorporées dans des poudres de lait ou dans des produits carnés (Luykx, et al., 2007) (Hoffmann, Münch, Schwägele, Neusüss, & Jira, 2017). Toutefois, ces méthodes couplées à la spectrométrie de masse nécessitent des équipements très coûteux.

IV.1.5.2 La chromatographie d'exclusion stérique

Les phases stationnaires d'exclusion stérique (ES) sont aussi utilisées pour l'analyse de protéines végétales. La séparation se fait sur un critère de taille moléculaire : elle repose sur la capacité de diffusion des molécules à l'intérieur des pores de billes de phase stationnaire de diamètre calibré. Les diamètres de ces pores définissent alors le domaine de séparation, qui est une échelle de masse molaire. Dans la plupart des cas, l'élution est isocratique, c'est-à-dire qu'il n'y a pas de modification de la composition de la phase mobile au cours de l'élution.

Dans le cas des protéines, il est particulièrement important de bien sélectionner la phase mobile. En effet, dans un cadre analytique, il est préférable d'utiliser un éluant qui va limiter l'agrégation des protéines et permettre une bonne séparation.

Ainsi, l'ES a été utilisée pour l'évaluation de la proportion et de la taille des protéines contenues dans des farines de blé. Dans une étude, la phase stationnaire utilisée était une Biosep-SEC-S4000 et la phase mobile était composée à 50 % d'acétonitrile et à 50 % d'eau avec 0,1 % de TFA. La méthode a permis de séparer les gliadines des albumines et des globulines par une élution de 22 minutes (Bangur, Batey, McKenzie, & MacRitchie, 1997). Une autre méthode, également sur le blé, a été développée pour séparer les albumines et les globulines. Cependant, l'ES était utilisée afin de comparer des procédés de cuisson du pain mais pas pour quantifier précisément les teneurs des différentes protéines (Singh H. , 2005). Enfin, plus récemment, une étude a porté sur les protéines de colza par CLHP-ES avec une élution par un tampon Tris à pH 8,5 et en présence de sel. Dans ce cas, les proportions de chaque protéine ont été déterminées à 280 nm (Boukroufa, et al., 2017). Cependant,

cet article repose sur l'hypothèse d'un coefficient d'extinction molaire identique entre les albumines et les globulines, ce qui n'est pas évident puisque les compositions en acides aminés, et plus particulièrement les aromatiques, de ces protéines sont bien distinctes. De plus, comme évoqué précédemment, à cette longueur d'onde d'autres molécules comme les phénols absorbent également, et sont très probablement présentes dans les extraits, ce qui risque d'introduire un biais dans l'analyse.

IV.1.5.3 Perspectives de développement pour l'analyse des protéines de colza

Ainsi, ces méthodes chromatographiques permettent une séparation efficace des protéines, ce qui permet d'envisager plusieurs applications :

- La détection des protéines végétales sur la base d'un temps de rétention grâce à une comparaison à des standards
- La détermination d'une composition en protéines, par la comparaison de l'abondance relative de protéines cibles
- La quantification de la concentration protéique grâce à un étalonnage préalable

Par ailleurs, Kuipers et Gruppen ont montré qu'il était possible d'évaluer le coefficient d'extinction molaire d'une protéine, à 214 nm, à partir de sa séquence en acides aminés (Kuipers & Gruppen, 2007). Pour ce faire, la première étape du travail a consisté à déterminer les coefficients d'extinction molaires de chaque résidu d'acide aminé à cette longueur d'onde. Ensuite, la seconde étape a consisté à déterminer les coefficients de petits peptides afin de quantifier l'absorbance de la liaison peptidique, mais aussi pour évaluer l'influence de la position d'un acide aminé dans la protéine : notamment, la proline n'a pas la même absorbance à 214 nm si elle est située en milieu de chaîne ou en position N-terminale. Les données ainsi obtenues ont permis d'obtenir des valeurs utilisables pour prédire l'absorbance de peptides et protéines à cette longueur d'onde.

Cette propriété a été utilisée en CLHP-ES pour quantifier des concentrations en fractions peptidiques dans un mélange complexe à partir du signal UV et de la composition en acides aminés de la protéine de départ (Bodin, Framboisier, Alonso, Marc, & Kapel, 2015). Cette méthode présente l'avantage qu'elle ne nécessite pas d'étalonnage préalable mais elle n'a pas été utilisée à ce jour pour la quantification de protéines.

En résumé

De nombreuses méthodes analytiques pour le dosage de l'azote protéique ont été développées reposant sur des techniques distinctes et courantes. Cependant, le fait que les protéines étudiées soient d'origine végétale implique la présence d'autres molécules dans les échantillons à analyser, comme les polyphénols qui sont susceptibles d'introduire des interférences et donc de biaiser les résultats. Une seule méthode a été reconnue officiellement comme référence pour la quantification des protéines, la méthode de Kjeldahl. L'électrophorèse SDS-PAGE est une seconde technique largement utilisée pour caractériser la nature des protéines dans un mélange, voire estimer une teneur en protéines. Cet état de l'art montre qu'il existe une lacune dans le cadre de l'étude d'un mélange de protéines végétales, puisqu'à ce jour, aucune méthode analytique ne permet la quantification précise des globulines et des albumines. Cependant, des méthodes chromatographiques ont permis cette distinction et il serait envisageable de les utiliser pour développer une méthode de dosage.

IV.2 L'étude structurale des protéines

Les procédés d'extraction et de purification sont susceptibles d'affecter la structure des protéines et donc leur qualité. C'est pourquoi l'analyse structurale est essentielle pour l'évaluation de la qualité des protéines, notamment dans le cadre de comparaison des procédés de production.

IV.2.1 Le dichroïsme circulaire (DC)

Le dichroïsme circulaire est une technique analytique utilisée dans l'étude des protéines pour caractériser leurs structures secondaires et tertiaires. Elle repose sur des phénomènes complexes d'optique. Les premiers travaux mêlant optique et biochimie ont été décrits au XIX^e siècle par Jean-Baptiste Biot, inspiré des théories d'Augustin Fresnel en optique.

Avant de décrire le dichroïsme circulaire, il est nécessaire de rappeler ce qu'est une lumière polarisée et son utilisation par Biot. Contrairement à la lumière du soleil qui se diffuse dans toutes les directions, la lumière polarisée se propage dans un seul plan. Elle peut ensuite être vue comme composée de deux champs avec des intensités et des longueurs d'onde identiques, mais avec des directions de rotation opposées (vers la droite ou vers la gauche) (Snatzke, 1968). Cette lumière polarisée peut être déviée d'un angle α lorsqu'elle traverse une solution dite optiquement active, c'est-à-dire contenant des molécules chirales, comme illustré en Figure 14. La loi de Biot a permis de démontrer qu'il existe un lien direct entre la concentration de la molécule chirale considérée, notée C , l'angle de déviation α , la longueur traversée l et le pouvoir rotatoire spécifique de la molécule $[\alpha]$:

$$\alpha = [\alpha] \times C \times l \quad (\text{Eq. 1})$$

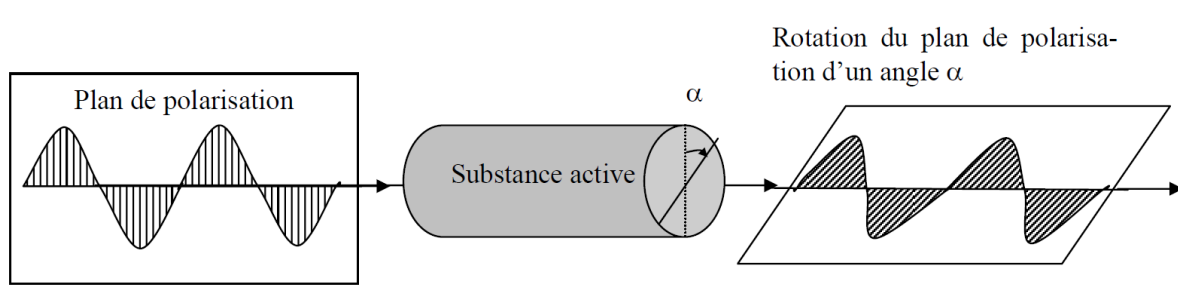


Figure 14 : Déviation d'une lumière polarisée par passage à travers une solution optiquement active. Schéma de (Baillet, Consulté en septembre 2019).

Dans le cas du dichroïsme circulaire, la lumière est polarisée de façon circulaire et non plus linéaire. Il consiste alors à mesurer la différence d'absorption $\Delta\varepsilon$ d'une solution optiquement active entre des rayons polarisés circulaires vers la droite ou vers la gauche. Ainsi, il est caractérisé, à une longueur d'onde donnée, par deux coefficients d'absorptivité molaire $\varepsilon_{\text{droite}}$ et $\varepsilon_{\text{gauche}}$. Pour exprimer un résultat en DC, l'ellipticité θ est la plus souvent utilisée : il s'agit de l'angle formé par le polariseur en entrée du système et le polariseur en sortie de l'appareil (aussi appelé l'analyseur). L'ellipticité (en $\text{deg.cm}^2.\text{dmol}^{-1}$) et la différence d'absorptivité molaire $\Delta\varepsilon$ (en $\text{L.mol}^{-1}.\text{cm}^{-1}$) sont reliés par la formule empirique suivante (Kelly & Price, 2000) :

$$\theta = \Delta\varepsilon \times 3298 \quad (\text{Eq. 2})$$

En pratique, les mesures en dichroïsme circulaire sont des spectres d'absorption sur un domaine de longueurs d'ondes donné : soit les UV proches (250 à 330 nm), soit les UV lointains (180 à 260 nm).

Le DC est une technique qui peut être appliquée aux protéines, qui sont des molécules chirales du fait que tous les acides aminés naturels, à l'exception de la glycine, le sont. Ainsi, une protéine en solution sera optiquement active et aura une réponse spécifique à la lumière polarisée.

Dans l'UV lointain, les spectres d'absorbance permettent de déterminer les proportions de chacune des structures secondaires des protéines (hélices α , feuillets β , etc.). La Figure 15 montre les spectres caractéristiques de ces structures secondaires. Un traitement des données par un logiciel informatique permet de déterminer précisément ces proportions.

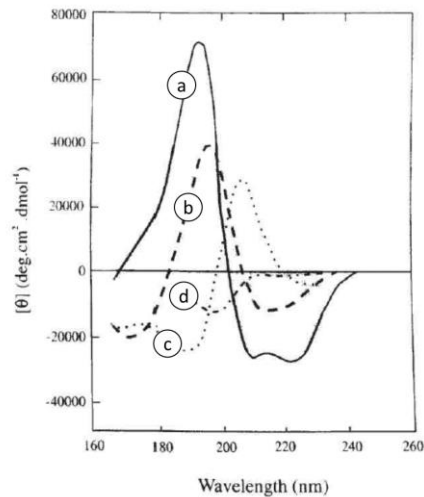


Figure 15 : Structures secondaires des protéines observables par dichroïsme circulaire. a) représente le spectre d'une protéine riche en hélices α , b) en feuillet β , c) en coudes β , d) en pelote (figure annotée de (Kelly & Price, 2000)).

L'analyse en UV proche repose sur l'absorbance des chaînes latérales des acides aminés aromatiques (tyrosine, phénylalanine et tryptophane) visibles par trois pics distincts comme sur la Figure 16. Cette figure est issue d'une publication de 2016 et correspond à l'analyse de globulines de colza dans des solutions à différents pH (Perera, McIntosh, & Wanasundara, 2016). Chaque pic correspond à un acide aminé : le tryptophane est visible entre 290 et 305 nm, le pic de la tyrosine s'observe entre 275 et 282 nm et le celui de la phénylalanine se situe entre 255 et 270 nm (Kelly & Price, 2000).

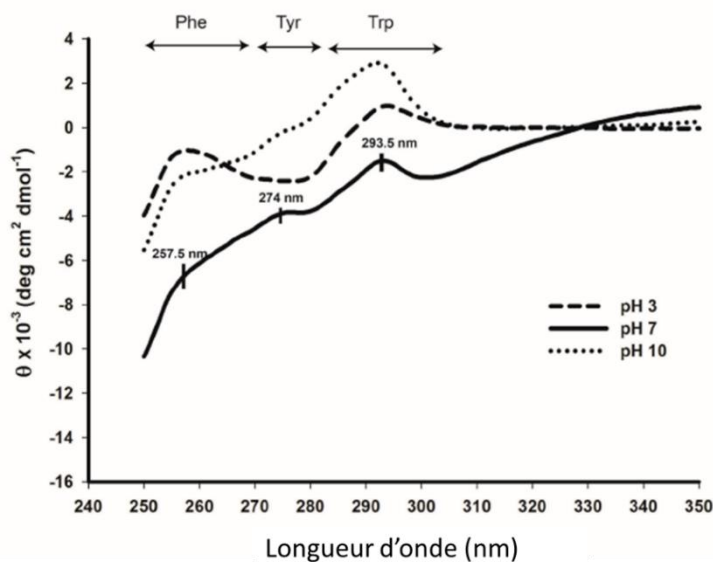


Figure 16 : Spectres de DC, UV proches, de globulines du colza à différents pH. Phe = phénylalanine, Tyr = tyrosine et Trp = tryptophane. Figure modifiée de (Perera, McIntosh, & Wanasundara, 2016)

Cette analyse peut être utilisée afin de détecter finement des changements de conformation des molécules placées dans différents milieux (pH, force ionique, température) ou encore pour étudier des effets d'interactions.

IV.2.2 La diffusion dynamique de la lumière (DLS)

La DLS est une technique d'analyse répandue afin de déterminer le diamètre dynamique de particules en suspension dans une solution. Elle peut être mise en œuvre afin de caractériser des agrégats protéique qu'il est intéressant de quantifier puisque leur présence peut avoir un impact sur les structures (gels, émulsions) formées par ces protéines (Nicolai, Britten, & Schmitt, 2011).

Son principe est basé sur la diffraction de la lumière et plus précisément sur la diffusion de Rayleigh : lorsqu'un rayon lumineux passe à travers une solution et rencontre des petites molécules, son trajet est dévié dans toutes les directions. Cette théorie est valide à condition que la longueur d'onde utilisée (de l'ordre de la centaine de nanomètres) soit beaucoup plus grande que la taille des particules étudiées, ce qui est le cas avec les protéines dont la taille est de l'ordre de la dizaine de nanomètres. La diffusion de Rayleigh dépend de la taille, de la forme des molécules et de la température. Puisque les molécules observées sont dans un liquide, elles se déplacent dans le milieu selon un mouvement brownien. C'est pourquoi, la DLS est réalisée sur plusieurs minutes afin de mesurer les variations d'intensité de la lumière diffusée dues au déplacement des molécules (Pecora, 2000).

La distribution des particules en DLS peut être représentés en fonction de l'intensité relative du rayon diffracté ou bien en fonction du volume des particules, Figure 17.

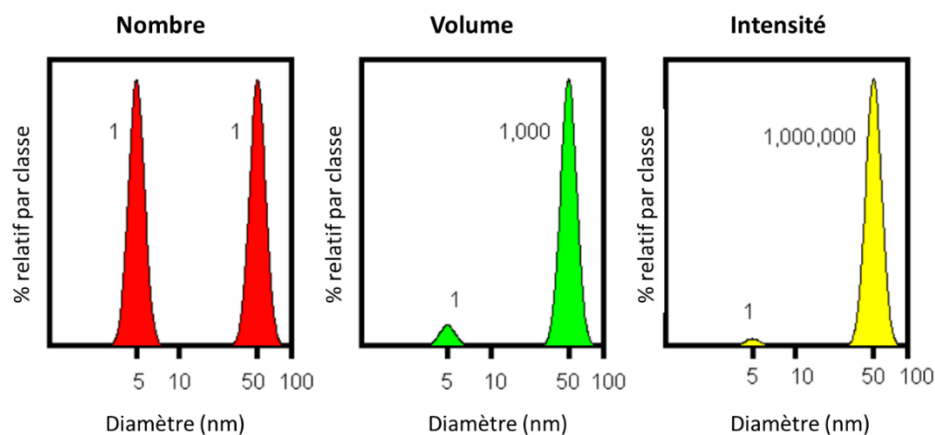


Figure 17 : Modes de représentation en DLS. Schéma traduit de (Malvern Instruments, Consulté en février 2019)

Ainsi, pour une solution contenant un nombre théorique identique de deux types de particules aux diamètres distincts (Malvern Instruments, Consulté en février 2019) :

- Le signal en intensité permet d'observer deux pics correspondant à chaque type de particule. Le pic correspondant aux particules les plus grosses est plus grand que l'autre. Cela s'explique par l'approximation de Rayleigh : l'intensité de diffusion est proportionnelle à la puissance 6 du diamètre de la particule.
- De même, pour le signal en volume il y a un facteur de 1000 entre les 2 pics malgré que chaque population ait le même nombre d'individus.

Appliquée aux protéines, la DLS permet de caractériser les éventuels agrégats formés avec leur distribution de taille. Dans le cas d'une solution contenant un seul type de protéine, la représentation en volume permet de déterminer le diamètre des protéines composant les agrégats.

IV.3 L'étude des propriétés des isolats protéiques

La valeur d'un isolat protéique est conditionnée par certaines de ses caractéristiques : ses propriétés fonctionnelles (solubilité, propriétés moussantes, émulsifiantes et gélifiantes), sa couleur et ses propriétés organoleptiques. D'une part, les propriétés fonctionnelles vont permettre d'ajuster la rhéologie ou la structure de formulations alimentaires. D'autre part, la couleur de l'isolat (en poudre ou en solution) ainsi que ses propriétés organoleptiques sont essentielles car son ajout ne doit pas modifier le goût ou la couleur du produit après sa formulation.

IV.3.1 La solubilité

La solubilité est un paramètre thermodynamique défini comme étant la concentration en molécule cible dans une phase liquide, que ce soit un milieu simple ou biphasique (solide-liquide ou liquide-liquide) (Pelegri & Gasparetto, 2005). En pratique, la solubilité des protéines en milieu aqueux est définie comme le titre de la fraction protéique et est alors exprimée comme un pourcentage. Elle dépend des propriétés de la phase aqueuse telles que le pH, la température, la force ionique ou la présence d'adjuvants.

C'est une propriété majeure puisqu'elle conditionne les autres fonctionnalités des protéines que sont les propriétés moussantes, émulsifiantes ou gélifiantes. Ces dernières sont des propriétés de systèmes colloïdaux dispersé en système polyphasique. Par exemple, une mousse est composée d'une phase continue aqueuse dans laquelle sont dispersées des bulles d'air. Dans ces systèmes, les protéines jouent un rôle majeur de stabilisation des structures formées. Le principe général consiste en l'adsorption des protéines à l'interface entre les deux phases ce qui requiert que les protéines soient solubles dans au moins l'une des deux phases.

Par conséquent, plus un isolat aura une solubilité élevée sur une large gamme de pH, plus les matrices d'utilisation seront variées. Par exemple, les boissons de type soda tendent à avoir un pH plutôt acide alors que les aliments transformés ont des pH proches de 5, c'est notamment le cas des produits de boulangerie utilisant des levures.

Dans la littérature scientifique, peu de travaux présentent des courbes de solubilités des isolats de protéines, probablement à cause du grand nombre de paramètres à prendre en compte. Chacune des étapes d'extraction et de purification des protéines a un impact sur leur solubilité. Par exemple, des traitements thermiques au cours d'un procédé risquent de limiter cette dernière. De plus, les

proportions en albumines et globulines sont peu évaluées par les auteurs, ce qui ne permet pas de rendre compte de la solubilité de mélanges de protéines de colza. Par ailleurs, d'autres composés de l'isolat sont susceptibles d'interagir avec les protéines : c'est notamment le cas de l'acide phytique dont la complexation aux protéines entraîne une diminution significative de leur solubilité, et ce, même à faible concentration (Cheryan & Rackis, 1980) (cf. Chapitre 1, II.4.3. Les molécules antinutritionnelles).

En 2014 et 2015, une équipe canadienne a publié deux articles sur les fonctionnalités des albumines et des globulines. Ils mettent en évidence que les albumines de colza sont solubles à plus de 95 % aux pH 3, 5 et 7 alors que les globulines ne le sont qu'à pH 3. En revanche, l'ajout de sel permet d'augmenter significativement leur solubilité aux mêmes pH que les albumines. Ceci est en accord avec la définition même de ces protéines (Cheung, Wanasundara, & Nickerson, 2014) (Cheung, Wanasundara, & Nickerson, 2015).

Un article très récent présente des profils de solubilité de concentrats de protéines de colza à partir de tourteau gras (Fetzer, Herfellner, & Eisner, 2019). Trois procédés sont mis en œuvre pour générer des produits différents :

- un concentrat d'une teneur de 78 % de protéines, contenant un mélange d'albumines et de globulines, produit par extraction des protéines à pH natif (5,7 – 5,8) suivie d'une purification par ultrafiltration ;
- un concentrat avec une pureté protéique de 84 % et composé de globulines, qui ont été isolées par précipitation isoélectrique ;
- un concentrat d'albumines, qui étaient présentes dans la phase liquide suite à la précipitation isoélectrique des globulines et qui ont été purifiées par ultrafiltration, pour atteindre une pureté de 78 %.

Les teneurs en acide phytique ont aussi été évaluées dans cette étude et étaient similaires entre les produits cités avec des valeurs entre 0,8 et 1,6 % d'acide phytique par rapport à la matière sèche. Cela permet d'exclure toute action de l'acide phytique sur les profils de solubilité observés. Les profils de solubilité de ces concentrats sont représentés en Figure 18 : les courbes correspondant aux produits obtenus à partir d'un tourteau gras pressé à froid sont colorées.

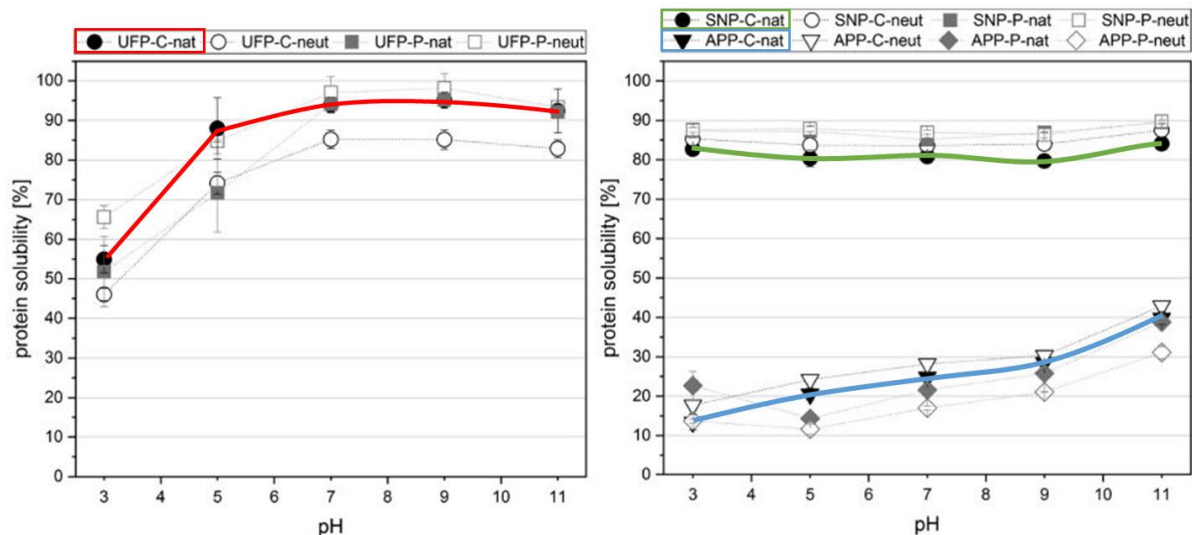


Figure 18 : Profils de solubilité des concentrats de protéines de colza dans une solution à 0.1 M de NaCl : en rouge un concentrats de protéines totales (albumines et globulines), en vert un concentrat d'albumines et en bleu, un concentrat de globulines. Figure modifiée à partir de (Fetzer, Herfellner, & Eisner, 2019)

Les profils de solubilité des albumines (courbe verte) et des globulines (courbe bleue) sont très différentes : les albumines sont solubles à plus de 80 % sur l'ensemble de la gamme de pH alors que les globulines le sont à moins de 40 % avec une augmentation progressive vers les pH les plus basiques. Il est à noter que la solubilité dans cette étude est réalisée en présence de sel. En comparaison aux études de Cheung *et al.*, les solubilités restent basses. Les protéines, directement extraites en milieu aqueux, présentent des solubilités de plus de 85 % à partir de pH 5, avec un minimum de solubilité à pH 3, qui s'explique probablement par la faible solubilité des globulines aux pH acides (Perera, McIntosh, & Wanasundara, 2016).

IV.3.2 Propriétés moussantes

Les mousses sont formées par l'incorporation d'air dans une solution aqueuse formant ainsi un système colloïdal où des bulles d'air sont réparties de manière homogène (phase discontinue) dans une phase aqueuse continue. Grâce à leur caractère amphiphile, les protéines de la phase aqueuse jouent le rôle de tensio-actif et stabilisent la structure. Elles sont adsorbées à l'interface air-eau ce qui induit une modification de leur structure tertiaire par un déroulement de la protéine, permettant éventuellement de constituer plusieurs couches protéiques stabilisant les bulles d'air (Mezzenga, Schurtenberger, Burbidge, & Michel, 2005).

Les mousses sont évaluées selon leur capacité et leur stabilité. La première correspond au volume de mousse formé à partir d'un volume de liquide alors que la seconde représente la proportion de mousse après une durée finie par rapport au volume initialement formé.

Deux caractéristiques des mousses sont généralement évaluées : la capacité moussante, qui correspond au volume de mousse formé à partir d'un volume de liquide, et la stabilité moussante, qui représente la proportion de mousse après une durée finie par rapport au volume initialement formé.

La stabilité d'une mousse s'explique par deux phénomènes physiques (Mezzenga, Schurtenberger, Burbidge, & Michel, 2005). Le premier est le drainage : il est le résultat de la différence de densité entre le liquide et le gaz formant la mousse et qui tend à séparer les deux phases. Le second, est la conséquence de la pression de Laplace : à l'interface entre le milieu gazeux et le milieu liquide, les bulles de gaz sont soumises à une pression qui dépend de leur rayon R. La différence de pression ΔP entre l'intérieur de la bulle de gaz et le milieu liquide s'exprime alors par :

$$\Delta P = \frac{2 \times \gamma}{R} \quad (\text{Eq. 3})$$

Dans cette équation γ correspond à la tension superficielle à l'interface gaz-liquide. Ainsi, plus le rayon de la bulle de gaz est important, plus sa pression interne sera faible. Si des bulles de différents rayons sont dans la mousse, les plus petites seront attirées par les plus grosses à cause de la différence de pression entre elles, il s'agit du phénomène de coalescence. La stabilité de la mousse est donc un paramètre important et qui peut être modulé, notamment par la nature des protéines présentes (taille, conformation spatiale, balance lipophile/hydrophile).

Les propriétés moussantes sont recherchées pour de nombreux applicatifs alimentaires (boissons de type soda ou bière, les gâteaux, les desserts ou encore certaines confiseries) afin de conférer une texture légère et crémeuse au produit, ce qui est recherché par le consommateur. Aujourd'hui, les protéines les plus utilisées dans l'alimentation pour former des mousses sont les protéines du lait et des œufs.

En ce qui concerne les protéines végétales, des études ont comparé leur capacité moussante avec celle des protéines du colza. Le Tableau 2 recense les résultats de quelques publications sur le sujet. Il est à noter que les études sont à considérer séparément : en effet, aucun protocole standard n'existe à ce jour pour la mesure de la capacité et de la stabilité moussante. Ainsi, il est très difficile de trouver dans la littérature des études parfaitement comparables. De plus, de nombreux facteurs ont une influence sur la capacité des protéines à former une mousse comme le pH, qui est directement lié à l'état structural de la protéine, la température, le tampon de dilution, la concentration en protéines ou encore le type d'agitation mis en œuvre pour former la mousse.

Tableau 2 : Comparaison des propriétés moussantes de protéines de différentes origines. Ces données sont à considérer étude par étude car les protocoles ne sont pas strictement identiques d'une étude à l'autre.

Origine	Grade	Capacité moussante	Stabilité moussante	Référence
Colza	Isolat	423 %	/	Paulson <i>et al.</i> , 1984
Soja	Isolat	150 %	/	Paulson <i>et al.</i> , 1984
Tournesol	Concentrat	423 %	/	Paulson <i>et al.</i> , 1984
Colza	Isolat, 99 % de protéines (globulines)	275 %	80 % après 1 h	Xu & Diosady, 1994
Colza	Isolat, 90 % de protéines (albumines)	281 %	73 % après 1 h	Xu & Diosady, 1994
Soja	Isolat, 93 % de protéines	252 %	85 % après 1 h	Xu & Diosady, 1994
Colza	Concentrat (85 % de protéines)	120 %	35 %	Nosenko, 2017
Soja	Concentrat (84 % de protéines)	60 %	55 %	Nosenko, 2017
Tournesol	Concentrat (71 % de protéines)	20 %	45 %	Nosenko, 2017
Colza	Isolat (pH 10)	280 %	73 % après 1 h	Pedroche <i>et al.</i> 2004
Colza	Isolat (pH 11)	163 %	51 % après 1 h	Pedroche <i>et al.</i> 2004
Colza	Isolat (pH 12)	181 %	51 % après 1 h	Pedroche <i>et al.</i> 2004
Colza	Isolat de globulines (98 % de protéines)	130 %	0 % après 15 minutes	Vioque <i>et al.</i> , 2000
Soja	Isolat (91 % de protéines)	130 %	/	Wong & Kitts, 2003
Blanc d'œuf	Isolat (94 % de protéines)	202 %	/	Wong & Kitts, 2003
Babeurre	35 % protéines	380 %	/	Wong & Kitts, 2003
Lait écrémé en poudre	35 % de protéines	920 %	/	Wong & Kitts, 2003
Protéines de pois	80 %	81 %	/	Stone <i>et al.</i> , 2015
Soja	80 %	171 %	/	Stone <i>et al.</i> , 2015
Blé	83 %	182 %	/	Stone <i>et al.</i> , 2015
Lactosérum	81 %	277 %	/	Stone <i>et al.</i> , 2015
Œuf	78 %	116 %	/	Stone <i>et al.</i> , 2015

Dans les études comparant les protéines de sources variées, le colza apparaît parmi les valeurs hautes des capacités moussantes. Ce constat reste néanmoins à prendre avec précaution puisque peu

d'auteurs détaillent la composition en albumines et globulines dont les proportions affectent significativement la capacité mais aussi la stabilité moussante (Chabanon, 2005).

L'étude de Fetzer de 2019 compare les capacités moussantes de différents concentrats de protéines de colza. Dans cet article, le concentrat riche en globulines a la plus faible capacité moussante (940 %) alors que les concentrats d'albumines ou de protéines totales sont comparables avec respectivement 2700 et 2800 %. Les stabilités moussantes obtenues pour tous les produits sont également très élevées avec des valeurs supérieures à 90 % après 60 minutes (Fetzer, Herfellner, & Eisner, 2019).

IV.3.3 Propriétés émulsifiantes

Une émulsion est obtenue en mélangeant deux phases liquides non miscibles. De même que pour les mousses, il s'agit d'un système colloïdal où une phase discontinue est dispersée dans une phase continue. L'émulsion la plus courante est obtenue par mélange d'huile, qui forme des gouttes dans une solution aqueuse. Les protéines sont susceptibles de stabiliser cette structure grâce à leur caractère amphiphile. Cette caractéristique dépend de la séquence en acides aminés des protéines, selon que les résidus exposés en surface soient hydrophiles ou hydrophobes. Les mêmes phénomènes physiques que pour les mousses sont en jeu.

Deux paramètres sont mesurés : la capacité et la stabilité émulsifiante. Pour ces mesures, la première étape consiste à former une émulsion par agitation d'une solution de protéines dans laquelle de l'huile est ajoutée. La capacité émulsifiante s'exprime alors comme le rapport entre le volume d'émulsion après centrifugation sur le volume avant centrifugation. La seconde étape, consiste à chauffer l'émulsion à 85 °C et à mesurer ensuite son volume. Le rapport du volume de l'émulsion avant et après chauffage représente la stabilité de l'émulsion.

D'autres mesures ont été utilisées dans la littérature pour quantifier et comparer les émulsions de protéines. Ainsi, certains chercheurs utilisent l'indice d'activité émulsifiante (abrégié EAI) qui correspond à :

$$EAI = \frac{2 \times T}{\varnothing C} \quad (\text{Eq. 4})$$

Dans cette équation, T correspond à la turbidimétrie, \varnothing est le volume de phase dispersée et C est la concentration massique en protéines (Uruakpa & Arntfield, 2005).

Une autre méthode trouvée dans la littérature consiste à déterminer le volume de phase dispersée à ajouter dans l'émulsion pour observer une inversion de phase : par exemple, de l'huile est incorporée dans une phase aqueuse et forme une émulsion huile dans l'eau, jusqu'à un certain volume où l'émulsion deviendra inverse, c'est-à-dire eau dans l'huile (Fetzer, Herfellner, & Eisner, 2019).

Ainsi, de nombreuses méthodes ont été mises au point pour quantifier les émulsions. Et, de même que pour les propriétés moussantes, il n'existe pas de protocoles standardisés de l'étude des propriétés émulsifiantes des protéines. C'est pourquoi, il est d'autant plus difficile dans ce cas de trouver des études bibliographiques parfaitement comparables. Néanmoins, le Tableau 3 montre les résultats d'études comparant des émulsions formées avec des protéines du colza et un isolat de soja, selon la méthode qui sera appliquée au cours des travaux de thèse.

Des activités émulsifiantes de 18 et 35 % ont été observées respectivement pour des isolats de globulines et d'albumines de colza et leur stabilité était de 30 et 48 % (Chabanon, 2005).

Tableau 3 : Comparaison des propriétés émulsifiantes de protéines de différentes origines. Ces données sont à considérer étude par étude car les protocoles ne sont pas strictement identiques d'une étude à l'autre.

Origine	Grade	Capacité émulsifiante	Stabilité de l'émulsion	Référence
Colza	Isolat, 99 % de protéines (globulines)	56 %	99 %	Xu & Diosady, 1994
Colza	Isolat, 90 % de protéines (albumines)	11 %	111 %	Xu & Diosady, 1994
Soja	Isolat, 93 % de protéines	72 %	99 %	Xu & Diosady, 1994
Colza	Isolat, 91 % de protéines (globulines)	18 %	30 %	Chabanon, 2005
Colza	Isolat, 91 % de protéines (albumines)	35 %	48 %	Chabanon, 2005

Ainsi, dans la littérature, les protéines de colza sont peu reconnues pour leurs propriétés émulsifiantes (Vioque, Sánchez-Vioque, Clemente, Pedroche, & Millán, 2000). Des études se sont penchées sur l'amélioration des propriétés fonctionnelles des protéines de colza par des modifications chimiques ou enzymatiques. Cependant, peu de données sont disponibles concernant l'impact du procédé de production des isolats de protéines de colza sur leurs propriétés fonctionnelles.

IV.3.4 Propriétés gélifiantes

Les gels sont des structures très utilisées dans l'alimentation avec, par exemple l'utilisation des caséines du lait dans la fabrication des yaourts ou de fromages. Les gels sont majoritairement constitués d'eau et d'un réseau de polymères qui forment une maille solide ordonnée. Les interactions entre les molécules peuvent être covalentes, comme la formation de ponts disulfures, ou non covalentes comme la formation de liaisons hydrogènes ou des interactions électrostatiques (Wanasundara, McIntosh, Perera, Withana-Gamage, & Mitra, 2016).

Le principe général de la formation des gels repose sur la dénaturation de la protéine qui se déplie sous l'effet d'une perturbation de son environnement (pH, la température, la force ionique, ajout d'enzymes) (Mezzenga, Schurtenberger, Burbidge, & Michel, 2005). Ce dépliement entraîne la formation de liens entre les polymères de protéines qui s'agglomèrent. Par exemple, dans le cas du

fromage, l'ajout de présure dans le lait entraîne l'agrégation et la précipitation des caséines : ce précipitât est un gel formé par l'action d'une enzyme, la chymosine, qui hydrolyse les caséines et entraîne leur coagulation. Un autre moyen répandu de former des gels consiste à chauffer progressivement la solution aqueuse contenant la protéine.

Les paramètres mesurés fréquemment pour caractériser un gel sont la concentration minimale de gélation, la température de gélation ou encore la force du gel. Dans une étude de 1998, une équipe a comparé des gels formés par des albumines ou des globulines de colza. Il a été montré que les albumines du colza ont une température de gélation plus élevée (85 °C) que les globulines (72 °C) (Schwenke, Dahme, & Wolter, 1998). Dans cette étude, la force des gels a également été mesurée grâce à un équipement dédié : il a ainsi été montré que les gels formés par les globulines étaient plus résistants que ceux formés avec les albumines. Ceci peut s'expliquer par le fait que les protéines de haut poids moléculaire sont susceptibles de former des gels plus solides que des petites protéines par la formation de réseaux en 3 dimensions (Tan, Mailer, Blanchard, & Agboola, 2011).

Afin d'améliorer les propriétés de gélation des protéines, il est possible d'ajouter d'autres composés tels que des polysaccharides, ou encore une enzyme comme la transglutaminase. Dans ce dernier cas, les gels avec les protéines de colza ont été formés à 40 °C et l'action de l'enzyme a permis de former des liaisons entre les sous-unités des protéines rendant le gel plus solide. Finalement, l'élasticité de ces gels a été comparée à des produits commerciaux et les modèles obtenus par l'équipe ont montré qu'il était possible de former un gel similaire à la bolognaise avec les protéines de colza (Pinterits & Arntfield, 2008).

IV.3.5 La couleur

Les isolats de protéines végétales sont souvent colorés à cause de la présence de polyphénols. Ces derniers sont capables de se complexer aux protéines et donnent une couleur prononcée au produit final (Sosulski, 1979). Ces phénomènes de complexation sont discutés en *partie II.4.3. Les molécules antinutritionnelles*. Afin d'être utilisées comme ingrédients, il est essentiel de disposer d'outils permettant d'évaluer la couleur des isolats et concentrats dans le but de contrôler ce paramètre au cours du procédé et pour définir les applications acceptables en fonction de la couleur du produit.

La couleur d'isolats de protéines est encore rarement évaluée dans la littérature. Néanmoins, des outils ont été développés pour caractériser la couleur de produits solides ou liquides. Ces outils mesurent la différence d'intensité de rayons lumineux émis et réfléchis par un échantillon solide ou avant et après passage à travers une solution.

Le plus souvent, les résultats sont exprimés dans un espace formé par trois axes, définis par la Commission Internationale de l'Eclairage (CIE). Chaque couleur correspond ainsi à un ensemble de trois coordonnées L^*a^*b . Le terme L représente la variation du noir au blanc avec des valeurs allant de 0 à 100. Le terme a est associé à l'axe vert – rouge avec des valeurs comprises entre -128 et + 127. Enfin, le terme b correspond à la teinte entre le bleu et le jaune, également entre -128 et + 127. Ce mode de représentation des couleurs fait l'objet d'une norme, NF EN ISO 11664-4.

Ce référentiel de couleur a été récemment utilisé pour comparer des produits issus de différents procédés d'extraction / purification des protéines de colza. Dans cet article, des photographies ont été prises et analysés par un logiciel. Ainsi, Fetzer *et al.* ont montré des différences significatives entre leurs produits concernant les termes a et b mais pas par rapport au terme L, Figure 19(Fetzer, Herfellner, & Eisner, 2019) .

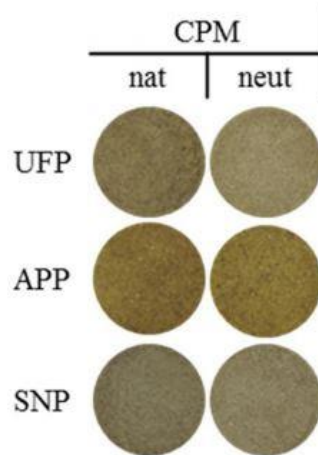


Figure 19 : couleurs de différents concentrats protéiques issus de tourteaux gras de colza à pH 5,7-5,8 (nat) ou à pH 7 (neut). UFP correspond à des protéines totales (albumines et globulines) purifiées par ultrafiltration, APP correspond à des protéines isolées par précipitation isoélectrique (globulines) et SNP correspond aux protéines solubles après précipitation isoélectrique (albumines). Figure modifiée à partir de (Fetzer, Herfellner, & Eisner, 2019)

L'utilisation du système L^*a^*b a ainsi permis de quantifier les couleurs de ces concentrats de protéines et de montrer qu'une différence significative existe entre les protéines issues de la précipitation électrique (APP) et les deux autres procédés. Néanmoins, bien que reconnu comme un système de référence, très peu d'études avec le système L^*a^*b ont été publiées à ce jour.

IV.3.6 Une autre propriétés organoleptique, le goût

Les propriétés organoleptiques d'un ingrédient sont définies comme étant les caractéristiques d'un produit qui affectent les sens : il peut donc s'agir de l'apparence du produit, de sa texture ou encore de son goût. Ce dernier est la seule propriété organoleptique qui n'a pas encore été discutée.

Comme évoqué précédemment, il est très difficile d'obtenir un isolat de protéines absolument pur et exempt de toutes molécules anti-nutritionnelles. Ces dernières, en plus de leur potentiel à colorer la

matière, sont susceptibles d'apporter un goût prononcé aux produits. Dès la fin des années 1970, des études ont été menées pour identifier ces composés et quantifier dans quelle mesure ils contribuent au goût des isolats protéiques. En ce qui concerne le colza, des travaux ont permis d'identifier deux types de molécules contribuant au goût dans les isolats et concentrats :

- les glucosinolates et leurs produits d'hydrolyse qui apporteraient la saveur moutardée au colza ;
- la sinapine et ses dérivés qui contribueraient à l'amertume du produit fini (Sosulski, 1979).

Les polyphénols contribuent grandement à l'amertume des protéines végétales. Un article de 1992 a réalisé des tests sensoriels sur plusieurs composés phénoliques répandus dans les matrices végétales. Il a été montré qu'il existe un effet synergique entre eux, abaissant significativement les seuils de détection de goût des consommateurs observés entre 40 et 240 ppm (Shahidi & Naczki, 1992). Ainsi, de très faibles quantités sont déterminantes pour le goût du produit final.

Un article a récemment été publié concernant des travaux de fractionnement dans le but d'identifier la molécule responsable de l'amertume d'isolats de protéines de colza grâce à des analyses sensorielles. En particulier, une fraction a été isolée par un lavage avec un mélange d'eau et de méthanol (50/50). Ensuite, une extraction en phase solide a permis de raffiner cette fraction de nouveau avec un mélange d'eau et de méthanol (30/70). Enfin, une étape de chromatographie en phase inverse a permis l'isolation d'un composé identifié comme étant le « Kaempferol 3-O-(2'''-O-sinapoyl- β -sophoroside) ». Le goût amer causé par ce composé serait détectable à partir d'une concentration de 3,4 $\mu\text{mol/L}$ (Hald, Dawid, Tressel, & Hofmann, 2019).

V Démarche de modélisation et d'optimisation des étapes de la production d'isolats de protéines végétales

Etant donné le grand nombre de paramètres avec une influence notable sur le procédé d'extraction, des outils d'optimisation ont été utilisés. L'objectif est de définir des conditions opératoires permettant d'atteindre les meilleures performances. A cause de la complexité du mélange obtenu, il n'existe pas aujourd'hui de loi permettant de prévoir les performances de l'extraction des protéines.

Verrou : de nombreux critères de performance définissent le procédé qui est influencé par plusieurs paramètres opératoires. Comment rationaliser la démarche pour l'étude de l'extraction des protéines ?

V.1 La méthodologie des plans d'expériences

Le développement d'un procédé implique son suivi par des indicateurs de performances. Ces derniers varient lorsque les conditions opératoires appliquées au procédé changent. En l'absence de modèles phénoménologiques de l'extraction, la méthodologie des plans d'expériences peut être utilisée afin de pallier ce manque.

Le principe consiste à considérer l'extraction comme une « boîte noire » : des conditions opératoires sont appliquées en entrée du système et les performances sont mesurées en sortie. Le traitement des données est basé sur l'utilisation de modèles linéaires formalisés par une relation matricielle simple :

$$Y^{mod} = A \times X \quad (\text{Eq. 5})$$

Dans cette équation X est une matrice décrivant le modèle polynomial choisi et les conditions opératoires appliquées, A est le vecteur des coefficients du modèle et Y^{mod} est la performance prédite. Il subsiste toujours une erreur de prédiction ε entre la valeur prédite Y^{mod} et la valeur expérimentale :

$$Y^{exp} = A \times X + \varepsilon \quad (\text{Eq. 6})$$

Une méthodologie statistique et matricielle permet de choisir les expériences (X) pour limiter la propagation des erreurs expérimentales sur les coefficients du modèle et sur les prédictions du modèle et de réduire la covariance entre les coefficients. La généralisation des plans d'expériences à l'aide de logiciels informatiques, a rendu simple et accessible cette méthodologie.

La première étape pour mettre en place un plan d'expériences est la définition du problème par l'identification des paramètres (les conditions opératoires) ayant une influence sur la réponse du système (ses performances). Cette étape est cruciale puisque les paramètres non identifiés seront susceptibles de contribuer aux variations observées dans les réponses du système. Au cours de cette étape, les domaines de variation de ces facteurs sont définis et les critères de performances étudiés sont choisis.

La deuxième étape est la conception du plan d'expériences. Un objectif de la méthodologie des plans d'expériences est la réalisation d'un minimum d'expériences pour décrire le plus précisément possible un phénomène. On distingue deux types de plans d'expériences : les plans de screening et les plans de modélisation/optimisation. En screening, les matrices de criblage peuvent être utilisées pour l'étude de nombreux facteurs afin de déterminer s'ils ont une influence sur le système. Ainsi, pour l'étude de 10 facteurs, 11 expériences seront nécessaires, pour ensuite, par calcul matriciel, estimer l'impact de chacun. En modélisation, pour obtenir un modèle plus précis de l'effet des facteurs sur les réponses, on utilise des plans d'expériences dont les matrices reposent sur un modèle polynomial

faisant intervenir des termes quadratiques et des interactions. Par exemple, un plan d'expériences composite centré consiste à faire varier chaque facteur sur 3 niveaux (notés -1, 0 et +1). Ainsi, pour 3 facteurs par exemple, un minimum de 14 expériences est suffisant pour décrire l'espace. Souvent, des expériences supplémentaires, des points dit « au centre », c'est-à-dire dont le niveau est 0, sont répétées pour estimer la variabilité expérimentale ou erreur pure.

La troisième étape est la réalisation des expériences. A l'exception de la variation des facteurs choisis, le protocole doit être strictement identique d'une expérience à l'autre. En effet, toute variation non prévue dans le plan contribuera à la variabilité du système.

La quatrième étape est le traitement des données. Le plus souvent, un logiciel de traitement spécifique est utilisé. A l'aide d'une régression multilinéaire, un modèle polynomial est généré. Dans le cas d'un plan composite centré, l'espace représenté par 3 facteurs F1, F2 et F3 est cubique. Ainsi, les équations des performances du système, noté Y, sont données sous la forme :

$$Y = a_0 + \sum_1^N a_i x_i + \sum_1^N a_{ii} x_i^2 + \sum_{i=1}^N \sum_{j=1}^{j<i} a_{ij} x_i x_j \quad (\text{Eq. 7})$$

Une régression linéaire des données permet de déterminer les coefficients a_i de l'équation. Des analyses statistiques permettent d'affiner les modèles par l'élimination des termes non significatifs.

La dernière étape est l'analyse du modèle grâce à des outils statistiques qui permettent de caractériser les modèles via des indicateurs : le coefficient de détermination R^2 caractérise l'adéquation entre les données expérimentales et le modèle ; Q^2 correspond à la qualité de prédiction du modèle ; la validité du modèle représente la cohésion des données expérimentales par rapport au modèle par la comparaison de l'erreur pure et de l'erreur du modèle ; et la reproductibilité est calculée comme étant le complément de la variation de la réponse lorsque le système est soumis à des conditions opératoires identiques.

V.2 L'optimisation multicritères

Puisque de nombreux critères de performances caractérisent le procédé d'extraction, il est fort possible qu'il n'existe pas une seule et unique solution optimale pour l'ensemble des critères considérés. Les critères de performances peuvent en effet être antagonistes : c'est-à-dire que leur optimum est obtenu pour des conditions opératoires différentes. Dans ce cas, il n'existe pas de solution optimale au sens strict et il faut rechercher la meilleure solution compromis. Un moyen de rationaliser cette démarche s'appuie sur le concept de domination au sens de Pareto. Ce concept s'inspire de la notion de bien-être collectif dans une société où on ne peut améliorer le bien-être (critère) d'une personne sans dégrader celui d'une autre (un autre critère). Pour appliquer ce concept,

une comparaison est réalisée de l'ensemble des performances générées par une solution (un jeu de conditions opératoires) à l'ensemble des performances générées par une autre. La solution est dite *non dominée* si toutes ses performances sont au moins aussi bonnes que celles de la solution à laquelle elle est comparée et si au moins une de ses performances est strictement meilleure. Les solutions non dominées par les autres représentent les compromis acceptables du problème. Ce concept a été intégré dans un algorithme génético-évolutionnaire développé au laboratoire qui détermine l'ensemble des solutions non dominées d'un problème. Les solutions sont généralement représentées dans deux espaces : la Zone et le Front de Pareto. Le front de Pareto est la représentation des compromis dans l'espace des critères de performances. La zone de Pareto est la représentation des compromis dans l'espace des conditions opératoires.

L'algorithme génético-évolutionnaire du laboratoire utilise les principes de la génétique et de la théorie de l'évolution. Dans une métaphore, un ensemble de conditions opératoires est représenté par un individu qui est caractérisé par un phénotype, lui-même résultat de l'expression d'un génotype diploïde associé à un indicateur de dominance/récessivité. Les performances de chaque individu sont déterminées par un modèle polynomial et la capacité de survie des individus est liée au nombre de dominations subies. Des opérations de croisements homozygotes et/ou hétérozygotes et de mutations sont combinées à une sélection élitiste des individus de la population. Au fil des générations, l'algorithme converge vers une population ne contenant que des individus non dominés marquant la fin de l'optimisation. Ainsi, toutes les solutions obtenues par cet algorithme sont des compromis acceptables au sens de Pareto.

Cet algorithme permet ainsi de déterminer tous les compromis sans recourir à l'intégration de préférences (agrégation) ou à une hiérarchisation des critères (ε -constrained) qui transforment les problèmes multicritères en problèmes monocritères mais où la subjectivité introduite au début du processus en biaise la solution (Ritzel, Eheart, & Ranjithan, 1994) (Mavrotas, 2009). Au contraire, notre démarche consiste à utiliser, en complément de notre algorithme multicritère, un outil d'aide à la décision issu du génie décisionnel pour déterminer quelle solution compromis doit être retenue et mise en œuvre.

V.3 Les outils d'aide à la décision

Les outils d'aide à la décision utilisés au laboratoire visent à classer les solutions compromis en fonction des préférences d'un expert, d'un décideur. Il s'agit d'un algorithme en 2 étapes. D'abord, à partir de quelques solutions échantillonnées dans la population non-dominée, un décideur classe les solutions selon son ordre de préférence. Le classement est analysé via la méthode des Rough Sets pour générer des règles de préférences et de non-préférences (Pawlak, 1982) quand elles existent.

Ces règles sont ensuite appliquées à l'ensemble de la population des solutions non dominées. Ainsi, les différents compromis sont classés de manière rationnelle selon les règles gouvernant les choix du décideur. La solution compromis retenue est sélectionnée parmi les solutions les mieux classées.

CHAPITRE II : MATERIEL ET METHODES

I Le tourteau étudié

Au cours des travaux de thèse, un tourteau gras de colza a été utilisé. Il a été fourni par Oléad (Pessac, France). Les graines ont été décortiquées, floconnées et pressées. Une composition partielle du tourteau a été déterminée : 32 % de protéines (N x 5.35), 3,5 % de fibres, 8,7 % de sucre, 9,6 % d'eau, moins de 1 % de polyphénols et 15,4 % d'huile résiduelle sur la matière sèche.

II Extraction des protéines dans le cadre d'un procédé séquentiel

Deux facteurs principaux ont été considérés concernant la première étape d'extraction : le pH et la quantité de sel ajoutée dans le milieu d'extraction. Les autres conditions opératoires, comme la durée de l'extraction ou le ratio solide : liquide ont été fixées. Les conditions opératoires de la seconde extraction, à pH acide, ont été fixées et déterminées au préalable grâce à une expérience préliminaire. Ce faisant, l'impact de la première étape d'extraction sur la seconde a été évalué. L'utilisation de la méthodologie des plans d'expérience nécessite trois niveaux de variation pour chacun des facteurs. Pour le pH, une expérience préliminaire a mis en évidence l'intérêt de travailler sur une gamme de pH entre 5,5 et 8,5. Concernant l'ajout de sel, son domaine de variation répond à une contrainte industrielle qui préconise une concentration limite d'ions chlorure. Ainsi, le domaine choisi est entre 0 et 0,2 M de NaCl.

Suite à l'extraction, une étape de purification des protéines a été appliquée afin d'en évaluer certaines caractéristiques, comme la solubilité en fonction du pH ou la couleur. Dans cette première étude, une purification en deux étapes a été appliquée : la précipitation des protéines avec du sulfate d'ammonium, suivie d'une étape de dessalage par chromatographie basse pression. Ce mode de purification présente l'avantage d'être non dénaturant pour les protéines, et simple et rapide à mettre en œuvre (Wingfield, 2001). La Figure 20 résume les étapes de ce procédé.

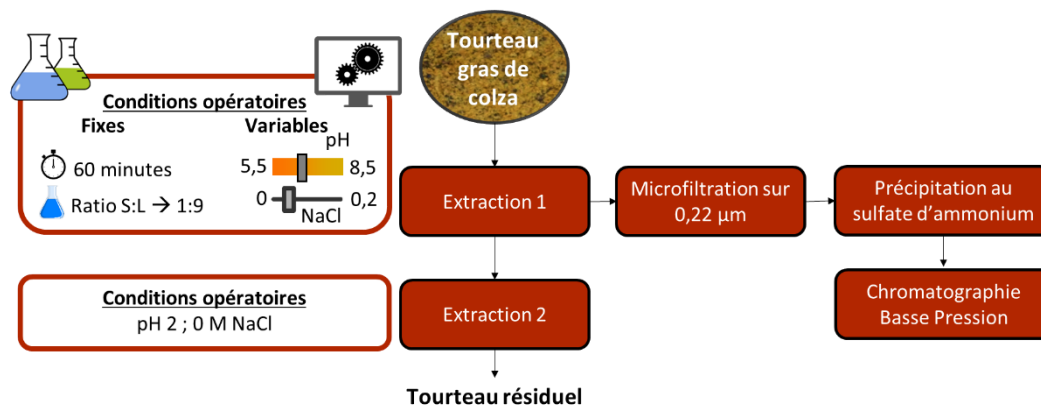


Figure 20 : Schéma résumant les étapes du procédé mis en œuvre pour l'extraction des protéines de colza dans le cadre d'un procédé séquentiel

II.1 Première étape d'extraction des protéines

Les extractions ont été réalisées dans un bioréacteur à double enveloppe de 200 mL, permettant la régulation de la température à 20 °C. Un ratio solide : liquide (eau) de 1 : 9 a été utilisé et l'extraction a duré 60 minutes. Le pH a été ajusté à 5,5, 7 ou 8,5 par ajout d'acide chlorhydrique HCl 1M, ou de soude NaOH 1M.

Du chlorure de sodium NaCl, a été éventuellement ajouté en concentration de 0, 0,1 ou 0,2 M. Ensuite, l'extrait liquide a été séparé du solide par centrifugation pendant 15 minutes à 15000 g. Le liquide est ensuite filtré à travers un filtre en cellulose. Le tourteau résiduel a été également récupéré, lyophilisé et analysé. Il a été utilisé pour la seconde étape d'extraction des protéines.

II.2 Clarification des extraits par microfiltration

Puisqu'un tourteau gras a été utilisé, de la matière grasse sous forme de globules lipidique est présente et est éliminée par un procédé membranaire. La phase liquide a été clarifiée par microfiltration avec une membrane plane Pellicon XL (Millipore, Burlington, Massachusetts, Etats-Unis) avec des pores de 0,22 µm et une surface de 55 cm². Pour ce faire, un système de microfiltration à flux tangentiel, Cogent µScale, de Millipore a été utilisé. Pour chacun des extraits, un Facteur de Réduction Volumique (FRV) de 10 a été appliqué. La Pression TransMembranaire (PTM) a été fixée à 1,5 bars et le débit d'alimentation à 54 mL/min, à température ambiante.

II.3 Purification des protéines et chromatographie liquide basse pression

La purification des protéines a été réalisée en deux étapes. La première étape a été la précipitation des protéines. Elle a été réalisée par ajout de sulfate d'ammonium, (NH₄)₂SO₄, jusqu'à obtenir une saturation de 70 %. Le mélange a été agité pendant 30 minutes. Ensuite, une centrifugation pendant 30 minutes à 15000 g, a permis de récupérer un culot, stocké à 4 °C.

La seconde étape de la purification a été le dessalage par chromatographie basse pression. Tout d'abord, le culot de précipitation contenant les protéines a été solubilisé dans une solution de Tampon Tris pH 8 0,5 M et avec 0,5 M de NaCl. Le volume final était égal au volume d'extrait avant précipitation. Ensuite, l'étape de chromatographie basse pression a été réalisée avec une phase stationnaire qui était une résine de gel filtration Sephadex G-25 (GE Healthcare, Chicago, Illinois, Etats-Unis) et formait un volume de 864 cm³ dans la colonne. La phase mobile était de l'eau déionisée et le débit était de 40 mL.min⁻¹ grâce à un système Äkta Pure (GE Healthcare, Chicago, Illinois, Etats-Unis). Après équilibration, un volume équivalent à 20 % du volume de la colonne a été injecté. L'élution a été réalisée en mode isocratique et la fraction protéique a été recueillie après 104 mL d'élution et jusqu'à 467 mL. Cette fraction a été ensuite lyophilisée pour devenir un isolat de protéines.

II.4 Protocole de la seconde étape d'extraction des protéines

Le tourteau lyophilisé à la fin de la première étape d'extraction a été utilisé pour cette étape. Suite à une expérience préliminaire, les conditions opératoires de la seconde extraction ont été déterminées. Ainsi, le pH a été fixé à une valeur de 2, sans ajout de sel et avec un ratio solide : liquide de 7 % (m/V). Les conditions opératoires ont été maintenues pendant 60 minutes. Ensuite, l'extrait liquide a été séparé du solide par centrifugation pendant 15 minutes à 15000 *g*. Le liquide a été ensuite filtré, d'abord à travers un filtre en cellulose, puis à travers un filtre seringue, dont les pores étaient de 0,22 μm , pour analyses.

III Extraction sélective des albumines

Dans le cadre d'une stratégie d'extraction des albumines à pH acide, trois facteurs ont été étudiés : le pH (2, 3 et 4), l'ajout de sel (0, 0,1 et 0,2 M) et la température du milieu (20 ou 55 °C). Un risque de développement microbologique justifie l'utilisation de ces deux températures. En effet, le domaine de température situé entre 20 et 50 °C serait particulièrement propice au développement de bactéries et de champignons. Alors qu'à 20 ou 55 °C, ce risque est limité. La Figure 21 résume les étapes de ce procédé.

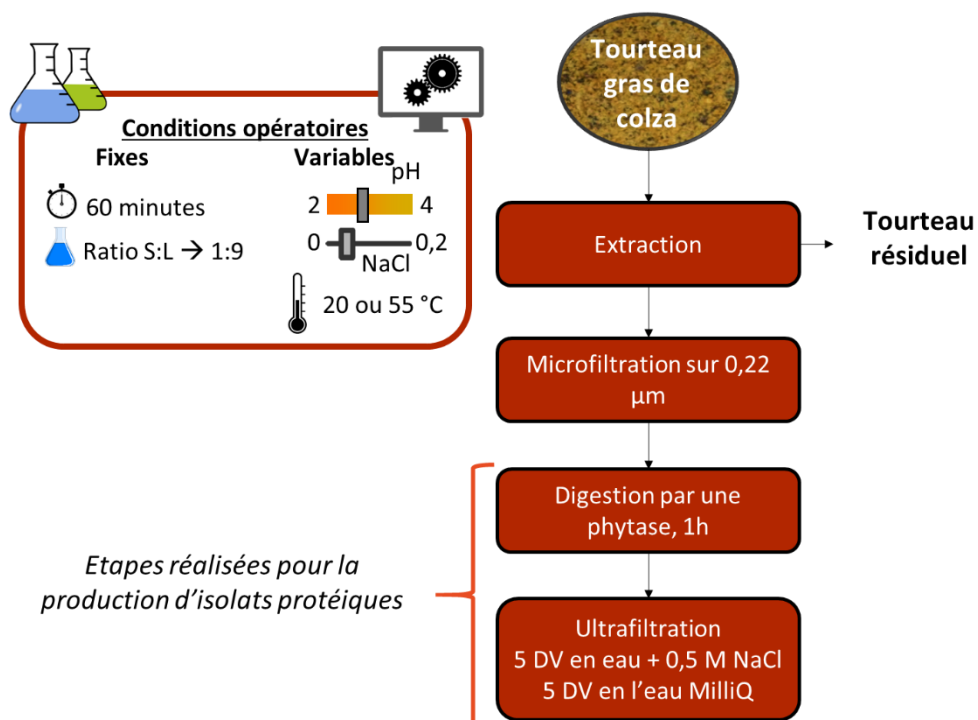


Figure 21 : Schéma résumant les étapes du procédé mis en oeuvre pour l'extraction des protéines de colza dans le cadre d'un procédé d'extraction sélective des albumines de colza

Suite à l'extraction, les protéines ont été purifiées par procédés membranaires. Une étape de microfiltration a permis, comme précédemment, d'éliminer les macromolécules lipidiques présentes dans l'extrait. Ensuite, un procédé d'ultrafiltration a été mis en œuvre pour éliminer les composés de petites tailles, tels que les composés phénoliques ou les acides aminés libres. Un avantage majeur des procédés membranaires réside dans leur facilité d'extrapolation à l'échelle industrielle contrairement à la précipitation au sulfate d'ammonium qui engage une grande quantité de sels.

III.1 Protocole d'extraction des protéines

Le protocole d'extraction des protéines est identique à celui détaillé en section *II.1 Première étape d'extraction des protéines* du Chapitre présent, à l'exception du pH, qui a été ajusté à 2, 3 ou 4. La température a été fixée soit à 20 °C soit à 55 °C.

III.2 Purification par procédés membranaires

Après microfiltration, l'extrait a été traité à 55 °C pendant 1 h à pH 4, afin d'éliminer les polymères osidiques. Pour la production des isolats (hors plans d'expériences), une phytase (Phytase Maxamyl de DSM) a été ajoutée afin de réduire les interactions entre l'acide phytique et les protéines. Après refroidissement, l'extrait est centrifugé 30 minutes à 15000 g.

Ensuite, la phase liquide a été purifiée, à température ambiante, dans un système d'ultrafiltration Äkta Flux 6 (GE Healthcare, Chicago, Illinois, Etats-Unis). Une cartouche à fibres creuses de 3 kDa et d'une surface de 4800 cm² (GE Healthcare, Chicago, Illinois, Etats-Unis) a été utilisée. Un débit d'alimentation de 1 L/min et une pression transmembranaire de 1,5 bars ont été fixées. Cinq diavolumes ont été effectués avec de l'eau salée à une concentration de 0,5 M de NaCl. Le pH était maintenu par l'ajout de HCl à une concentration de 1 M. Ensuite, quatre diavolumes ont été réalisés avec de l'eau ultrapure. Le rétentat final a été lyophilisé (Shanghai Pharmaceutical Machinery Co., LTD type GZL-0.5). Cette procédure de purification des albumines de colza a fait l'objet d'un dépôt de brevet (WO2019096862 - 2019-05-23).

IV Les outils de modélisation et d'optimisation

IV.1 La méthodologie des plans d'expériences

La méthodologie des plans d'expériences a été appliquée à l'étude de l'effet du pH et de l'ajout de NaCl sur l'extraction des protéines. Le logiciel Modde (Umetrics, Umeå, Suède) a été utilisé afin de générer les plans d'expériences et de traiter les résultats. Les plans utilisés étaient des plans de type Central Composite Centré sur les Faces. Ainsi, l'étude de deux facteurs, le pH et l'ajout de sel, é

nécessité un minimum de 11 expériences dont trois répétitions des points centraux. La matrice correspondante ainsi que la représentation des expériences sont présentées en Figure 22.

N°	pH	NaCl (M)
1	5,5	0
2	8,5	0
3	5,5	0,2
4	8,5	0,2
5	5,5	0,1
6	8,5	0,1
7	7	0
8	7	0,2
9	7	0,1
10	7	0,1
10	7	0,1

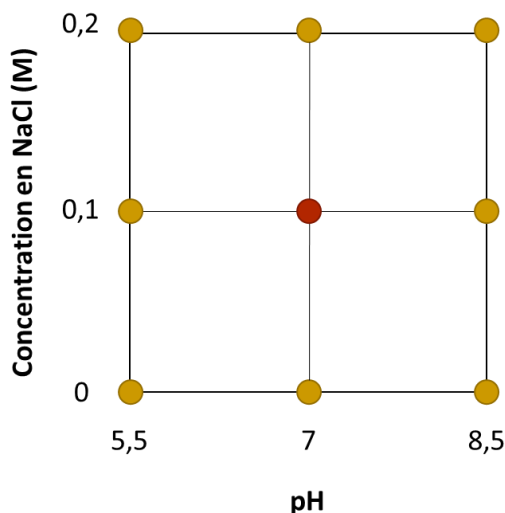


Figure 22 : Matrice expérimentale utilisée pour l'étude de l'extraction des protéines dans le cadre d'un procédé séquentiel et représentation des expériences en fonction du pH et de la concentration en NaCl dans le milieu.

Afin d'interpréter les modèles obtenus, des analyses statistiques ont été réalisées à l'aide du logiciel Modde. Cinq indicateurs sont particulièrement importants afin d'évaluer un modèle : (i) R^2 représente la corrélation entre les valeurs modélisées et expérimentales ; (ii) Q^2 correspond à la précision du modèle lors de la prédiction de nouvelles données, la différence avec R^2 devrait être inférieure à 0,3 pour aboutir à un bon modèle ; (iii) la reproductibilité représente la variation qui peut survenir entre deux expériences identiques ; (iv) un test ANOVA est effectué afin de vérifier si les coefficients du modèle sont significatifs par la détermination de la valeur p ; (v) un autre test ANOVA est utilisé afin de vérifier statistiquement si le modèle manque de cohésion (*lack-of-fit*).

La pertinence des coefficients de chaque modèle a été évaluée statistiquement. Un intervalle de confiance de 90 % a été fixé. Par ailleurs, une contrainte logicielle impose de conserver les termes des effets principaux (pH et ajout de sel) si au moins l'un des coefficients des termes d'interaction est significatif. Par exemple, si le coefficient du terme pH x pH est significatif, alors, le coefficient du terme pH est conservé.

IV.2 Les algorithmes d'optimisation multicritères

L'algorithme d'optimisation multicritère a été développé au laboratoire dans le langage Matlab (MathWorks, Natick, Massachusetts, Etats-Unis). Dans ces travaux, les algorithmes ont été appliqués pour la détermination des compromis selon trois critères de performance antagonistes : le rendement

d'extraction, la proportion en albumines et la teneur en acide phytique du tourteau. Les modèles polynomiaux de chaque critère de performance ont été construits en utilisant le logiciel de plan d'expériences Modde.

D'abord, l'algorithme a généré aléatoirement une population dont les individus étaient caractérisés par une condition opératoire de pH et une de concentration en NaCl et trois critères de performance. Les individus ont été comparés deux à deux et la population non dominée a été conservée. Des croisements entre individus et des mutations ont permis la génération d'une nouvelle population. L'algorithme a été réitéré plusieurs fois, jusqu'à obtenir une convergence de la population. A l'issue de cet algorithme, seules les solutions non dominées, c'est-à-dire dont les compromis sont équivalents, ont été conservés.

Dans le Chapitre IV, cet outil a été utilisé successivement 4 fois, sur des populations de 5000 individus afin de générer une unique population de 20000 individus. Une dernière analyse multicritère a été réalisée sur cette grande population pour éliminer les individus qui se révéleraient de nouveau dominés après la fusion des 4 populations.

IV.3 Les algorithmes d'aide à la décision

La première étape a consisté à générer quatre groupes échantillons constitués de cinq individus distincts. Ensuite, ces individus ont été classés par ordre de préférence selon un expert du procédé, le décideur. Les règles gouvernant les choix de l'expert ont été analysées mathématiquement afin de définir des règles de préférence et de non-préférence. Enfin, ces règles ont été appliquées à l'ensemble des individus afin de classer la population selon les règles de préférence du décideur.

V Méthodes analytiques pour le suivi du procédé

V.1 Détermination de la teneur en protéines

V.1.1 La méthode de Kjeldahl

Les analyses ont été réalisées selon la méthode officielle Kjeldhal AOAC 979.09. Les échantillons ont été préparés avec 4 mL d'acide sulfurique et une spatule du catalyseur Cu-Se. L'étape de minéralisation a été effectuée à 450 °C pendant 90 minutes dans un K-439 Büchi SpeedDigester. L'étape de distillation (Kjelflex K-360, Büchi, Rungis, France) a été réalisée avec de la soude concentrée à 32 % et de l'acide borique à une concentration de 3% m/V. La titration a été effectuée en utilisant une solution de HCl à une concentration de 0,01 M dans un Titrino Plus 877 (Metrohm, Herisau, Suisse).

V.1.2 La densitométrie de gels SDS-PAGE

Les réactifs utilisés pour les analyses SDS-PAGE provenaient de la société Bio-Rad (Hercules, Californie, Etats-Unis). Deux échelles de standards de protéines ont été utilisées : entre 10 et 250 kDa et entre 1,4 et 26,6 kDa. Les échantillons ont été préparés selon la méthode Laemmli (Laemmli, 1970) en conditions réductrices. La migration a été réalisée sur une matrice composée d'un gel de concentration contenant 5 % d'acrylamide et un gel de séparation à 17 %. La coloration a été effectuée par le bleu de Coomassie G 250. Ensuite, une solution de décoloration forte, composée à 40 % d'éthanol et à 10 % d'acide acétique, a été utilisée pendant 30 minutes. Cette étape a été renouvelée une fois. Ensuite, une solution de décoloration faible, avec 5 % d'acide acétique a été appliquée pendant une nuit.

Pour l'analyse densitométrique des bandes le logiciel libre ImageJ (National Institutes of Health, Bethesda, Maryland, Etats-Unis) a été utilisé.

V.1.3 Développement d'une nouvelle méthode de quantification des protéines par Chromatographie Liquide Haute Performance d'Exclusion Stérique

La stratégie mise en œuvre au cours de la thèse nécessitait de disposer des titres en albumines et en globulines dans les extraits, puisque les proportions de chaque fraction protéique ont un impact sur les caractéristiques du produit final. L'état de l'art dans la bibliographie a mis en évidence une lacune dans les méthodes analytiques disponibles pour l'analyse des protéines. La méthode de référence, le Kjeldahl, ne permet pas le dosage spécifique des albumines et des globulines. La densitométrie, suite à une électrophorèse SDS-PAGE, le permet mais elle présente deux inconvénients majeurs. Il est nécessaire de diluer fortement les échantillons avant analyse pour qu'ils aient une concentration comprise dans la gamme de linéarité de la calibration. De plus, les nombreuses étapes de la méthode induisent des incertitudes qui diminuent la précision de l'analyse. Dans le cadre des travaux de thèse, portant sur la production d'isolats d'albumines et de globulines, il a été nécessaire de développer une nouvelle méthode analytique permettant, en une seule analyse, de déterminer les concentrations respectives des deux protéines.

V.1.4 Article publié dans le journal Food Chemistry (2019)

The graphical abstract, Figure 23, shows the main features of the method that was developed and how it can be easily integrated in a process of production of rapeseed albumins.

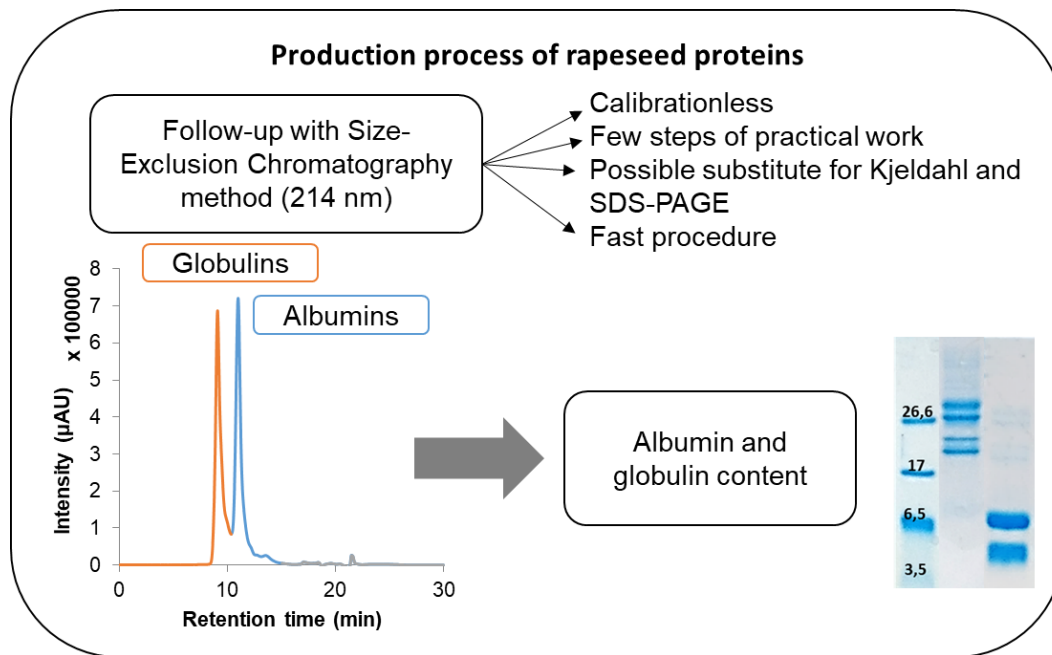


Figure 23 : Graphical abstract showing that the method can be implemented to follow-up a process of production of rapeseed proteins

V.1.4.1 Introduction

According to FAO statistics, in 2016, rapeseed was the second most produced oilseed after soybeans in the world with a respective production of 69 million tons and 335 million tons. Whereas, for sunflower seeds and linseeds, 47 and 3 million tons were produced during the same year (FAOStats, 2019). Because of their richness in fibers and proteins, meals obtained after oil extraction are currently used as a commodity. These proteins are also interesting for human nutrition due to their nutritional properties (World Health Organization, 2007).

In oilseed meals, 2 S and 11-12 S proteins constitute the main protein fractions. In rapeseed, 2 S proteins are called albumins or napins, with a molecular weight (MW) of 10 to 15 kDa. They represent between 30 and 40 % of the total proteins in rapeseed. Among their interesting properties are foaming and emulsifying activities. 12 S protein are called globulins, or cruciferins, and their MW is between 300 and 350 kDa. They represent 50 to 55 % of the total proteins and are suitable to form gels (Crouch & Sussex, 1981) (Aider & Barbana, 2011). In flaxseed and sunflower, similar structure exists with slight variations (Chung, Lei, & Li-Chan, 2005). (Gonzalez-Pérez & Vereijken, 2007).

Currently, protein quantification in aqueous extracts is often done from nitrogen quantification by Kjeldahl method (Sáez-Plaza, Michałowski, Navas, Asuero, & Wybraniec, 2013). The protein amount is deduced from the nitrogen content using a conversion factor (often 6.25 for rapeseed proteins). However, this value has been criticized over the years as it depends on the protein under study (Mariotti, Tomé, & Patureau Mirand, 2008). Furthermore, other biomolecules contain nitrogen. This

is taken into account by using another Kjeldahl analysis after selective protein precipitation by trichloroacetic acid (TCA). Eventually, this method gives an overall protein concentration and does not allow specific albumin or globulin quantification. Spectrophotometric methods like BCA or Lowry's method can alternatively be used. These methods are fast with automated microplate systems. But polyphenols co-extracted bias the quantification (Stepanchenko, Novikova, & Mosh, 2011).

SDS-PAGE separates the polypeptides of a complex mixture according to their size (Chiou & Wu, 1999). By densitometry, it is possible to compare the intensity of protein bands (colored by Coomassie blue) and to get an estimated protein content. But this method is long and the results are likely biased. Alternatively, chromatographic methods have been developed to this day to identify and quantify proteins from plants such as rapeseed (Boukroufa, et al., 2017). Even though these methods discriminate between types of proteins, none allows a clear quantification of the protein content.

The hypothesis of this work is that SE-HPLC can be used to accurately quantify the protein content and the proportion of globulins / albumins in rapeseed extract in order to do a follow-up of protein production processes. The methodology described hereafter is calibrationless: amino acids sequences were used to determine two parameters, the extinction coefficient and the nitrogen-to-protein conversion factor used in the Kjeldahl analysis.

The first step consisted of a proof of concept of the ability of the method to separate both proteins. Secondly, two validation steps were carried out using homemade protein isolates. They were used to compare the proposed method to: (i) the Kjeldahl regarding the quantification of the protein content; (ii) and to SDS-PAGE densitometry for the determination of the proportion of each protein in a sample. Thirdly, the SE-HPLC method was compared to both previously cited reference methods for the analysis of aqueous extracts of proteins. Fourthly, a stability study was conducted on the aqueous extracts. Finally, the methodology was applied to two other oleoproteaginous plants: sunflower and flaxseed.

V.1.4.2 Material and methods

V.1.4.2.a) Materials

- *Meals*

The rapeseed meal used was provided by Olead (Pessac, France): after pressing, a step of oil extraction was carried out with hexane followed by a heating step at 86 °C for 60 minutes to remove the solvent. Its composition was 37.5 % protein, 35.8 % of total dietary fibers, 7.7 % of carbohydrates, 9.4 % of

phytates, 7.3 % of minerals, 3.8% water, 2.0 % of polyphenols and 1 % residual oil on a dry matter basis.

- *Chemical reactants*

Sodium chloride and sodium hydroxide pellets were provided by VWR (Radnor, Pennsylvania, USA). Chlorhydric acid at 37 % was provided by Carlo Erba (Val-de-Reuil, France).

V.1.4.2.b) Production process of rapeseed protein isolates

Albumins and globulins and total proteins were extracted and purified for method calibration and validation. The extraction step was carried out in aqueous medium at room temperature with a solid-liquid ratio of 1:9, 0.1 M NaCl and 0.5 g.L⁻¹ ascorbic acid (Sigma-Aldrich, Saint-Louis, USA) to prevent oxidation. pH was adjusted to 7 with a solution of NaOH 2 M under agitation. The extraction duration was 30 minutes and the pH was maintained at 7. The solid-liquid separation was realized with a Heraeus Megafuge 16 R centrifugation system (Thermo Scientific, Waltham, Massachusetts, USA) at 10,000 g for 30 minutes. The supernatant was recovered and filtered and the solid was discarded. This first step was common to the production of all three proteins. Purification was achieved by an ultrafiltration procedure with an Äkta Flux 6 (GE Healthcare, Chicago, Illinois, USA) using 4800 cm² hollow fibers cartridges (GE Healthcare, Chicago, Illinois, USA) at 2 L/min feed flow and a transmembrane pressure of 1.5 bars. Albumins and globulins were separated using a 500 kDa cut-off membrane. During diafiltration, water or salted water was added to the retentate at the same flow rate as the permeate leaving the system to keep the volume in the retentate constant. Each volume of water equal to the volume of the retentate is called a diavolume (DV). After 7 diavolumes with salted water (0.2 M NaCl), the permeate rich in albumins was recovered, concentrated 5 times and diafiltrated with 8 DV of water using a 3 kDa membrane. The retentate, rich in globulins, was further purified with the 500 kDa membrane using 3 DV of salted water followed by 2 DV of deionized water. Total proteins were directly purified by diafiltration with the 3 kDa membrane with 10 DV of deionized water. Final retentates were freeze dried with an LTD type GZL-0.5 system (Shanghai Pharmaceutical Machinery Co., Shanghai, China).

To determine purity of all three powders, solutions at a concentration of 5 g.L⁻¹ of powder were prepared and homogeneous samples were analyzed according to the official AOAC 979.09 Kjeldahl method (6.25 nitrogen-to-protein conversion factor). Purity was expressed as the ratio of the thus measured protein content divided by the theoretical concentration of the solution.

$$Purity = \frac{\text{Protein content by Kjeldahl method}}{\text{Powder content}} \quad (\text{Eq. 8})$$

After freeze-drying, the powders were characterized as follow: 6 % of water, 102 % of proteins and 4 % of phytic acid for albumins; 2 % water, 72 % proteins and 2 % phytic acid regarding globulins. The total proteins isolate was composed of 5 % water, 94.5 % proteins and 3 % phytic acid. The dry matter was determined according to the method ISO 638:2008. The protein content of the powder corresponds to the purity described previously. The nitrogen-to-protein conversion factor was 6.25 as it is the mostly used value even though it probably overestimates the protein content. And the phytic acid content was determined following the method developed by García-Esteva *et al.* (Garcia-Esteva, Guerra-Hernandez, & Garcia-Villanova, 1999).

V.1.4.2.c) Determination of the protein content by Kjeldahl method

Samples were analyzed according to the official AOAC 979.09 Kjeldahl method. The samples were prepared with 4 mL of sulfuric acid (Sigma-Aldrich, Saint-Louis, USA) and a spatula of catalyst Cu-Se (AppliChem, Gatersleben, Germany). The mineralization step was carried out by heating the flasks at 450 °C for 90 minutes in a Büchi SpeedDigester K-439 (Rungis, France): The distillation step took place in a Kjelflex K-360 (Büchi, Rungis, France) with sodium hydroxide 32 % (VWR, Radnor, Pennsylvania, USA) and boric acid (Sigma-Aldrich, Saint-Louis, USA) prepared in solution at 3 %. The titration was carried out with 0.01 M HCl (Honeywell, Morris Plains, New Jersey, USA) in a Titrino Plus 877 from Metrohm (Herisau, Switzerland).

To determine non protein nitrogen, extracts were diluted twice with a 40 % (w/v) trichloroacetic acid TCA (Sigma-Aldrich, Saint-Louis, USA), stirred 30 minutes and centrifuged at 15000 g during 30

V.1.4.2.d) SDS-PAGE assay and quantification by densitometry

SDS-PAGE reagents were from Bio-Rad (Hercules, California, USA). Two standard protein ladders were used: the first between 10 and 250 kDa and the second between 1.4 and 26.6 kDa. Samples were prepared according to Laemmli method (Laemmli, 1970) in reducing conditions and migrated on a 5 % acrylamide stacking gel and a 17 % separation gel. The coloration was achieved by Coomassie blue G 250 (Sigma-Aldrich, Saint-Louis, USA). For each SDS-PAGE assay, a standard molecular weight kit was used. Band densitometry analysis was achieved using ImageJ software (National Institute of Health, Bethesda, Maryland, USA). From pictures of SDS-PAGE gels, the intensity of the coloration for each band of the gel was determined. A calibration was carried out on globulins and albumins isolate solutions from 0.2 to 4 g.L⁻¹ before determination on the aqueous extracts.

The calibration plots for each protein showed a good correlation, with respective R² of 0.98 and 0.96 (data not shown) for globulins and albumins. However, the linear portion of the calibration is narrow: protein content should be between 0.2 and 0.8 g.L⁻¹ for a reliable determination.

V.1.4.2.e) Quantification of proteins by SE-HPLC

- *Elution method*

SE-HPLC acetonitrile was provided by Biosolve BV (Valkenswaard, the Netherlands) and ultrapure water was generated by a system Purelab Ultra by Elga (High Wycombe, United Kingdom). The chromatographic system was a UFLC machine from Shimadzu (Kyoto, Japan) and the column, Biosep-SEC-s2000 300 x 7.8 mm, 5 μm , was provided by Phenomenex (Torrance, California, USA). Its separation range (300 kDa – 1 kDa) was chosen because of the molecular sizes of albumins and globulins (14 kDa and 300 kDa, respectively).

The mobile phase was 45 % acetonitrile, 55 % ultrapure water and 0.1 % of trifluoroacetic acid (TFA) from ThermoFisher (Waltham, Massachusetts, USA). Elution was done at 0.6 mL / min flow during 30 minutes. The calibration of the chromatographic column was realized using a gel filtration markers kit for protein molecular weights between 12 kDa and 200 kDa from Sigma-Aldrich (Saint-Louis, USA).

- *Quantification method*

Liquid samples from aqueous extraction or solutions prepared from isolates were filtered on 0.22 μm membranes from Pall (Port Washington, New York, USA). The injection volume was 5 μL and the detection was set to 214 nm. For calibration, a 20 $\text{g}\cdot\text{L}^{-1}$ (w/v) protein solution was diluted in a 0.1 M phosphate buffer pH 7 (Sigma-Aldrich, Saint-Louis, USA) to obtain concentrations of 20, 15, 12.5, 10, 7.5, 5 and 1 $\text{g}\cdot\text{L}^{-1}$. Each stock solution was prepared in triplicate and analyzed using the SE-HPLC method. The protein content of each stock solution was also quantified according to the Kjeldahl method.

Data were extracted and processed for integration using an Excel (Microsoft, Redmond, Washington, USA) spreadsheet in order to retrieve the mass of protein injected. First, peaks were integrated:

$$Area = \int_{V_0}^{V_f} A \cdot dV \quad (\text{Eq. 9})$$

In this equation, V_0 and V_f correspond to the elution volume at the beginning and at the end of the protein peak and A is the intensity of the signal.

And according to Beer-Lambert law: $A = \epsilon \cdot l \cdot C$ with ϵ , the extinction coefficient, and l the length of the detection cell.

$$Area = \int_{V_0}^{V_f} \epsilon \cdot l \cdot C \cdot dV = \epsilon \cdot l \cdot \int_{V_0}^{V_f} C \cdot dV = \epsilon \cdot l \cdot m_{protein\ injected} \quad (\text{Eq. 10})$$

Because the detection of the proposed method is UV, a key parameter was ϵ , the extinction coefficient. According to Kuipers and Gruppen in a paper from 2007 (Kuipers & Gruppen, 2007), this value was calculated at 214 nm. Subsequently the mass of protein in each sample can be determined.

$$m_{\text{protein injected}} = \frac{\text{Area}}{\epsilon \cdot l} \quad (\text{Eq. 11})$$

V.1.4.2.f) Validation of the method

- *Aqueous extraction of proteins*

In order to validate the proposed method, aqueous extractions were carried out under several conditions of pH (from 2 to 9) and salt content (from 0 to 0.3 M). To do so, 10 g of meal was used at a ratio of 10 % solid-liquid. pH was adjusted with NaOH or HCl 1 M. Each extraction was carried out during 30 minutes. A centrifugation step allowed the recovery of the extract.

- *Stability of protein extracts*

The SE-HPLC analyses of the extracts previously described were compared after 24 hours of storage under three different conditions: at room temperature, 4 °C or -20 °C. The difference of area under protein peaks between time 0 and after 24 hours of storage are compared to assess extracts stability.

- *Statistical analysis to compare the methods*

In order to compare the SE-HPLC to the reference methods, a statistical analysis was carried out in two steps with Excel (Microsoft, Redmond, Washington, USA). First, a Fisher-test was done in order to compare the variances of the three repetitions of both methods for each validation point. If p-value is superior to 5 %, the variances are deemed equal. The equality or inequality of variances is an input necessary to carry out the next test. Then, a Student T-test was realized in order to compare the expectations taking into account the equality or inequality of variances. If the p-value is superior to 5 %, the expectations of both populations (from the two different methods) were deemed equal and thus, the methods were comparable for the given point of validation.

- *Validation of the method according to the ICH validation guide*

Another step of validation was based on the determination of values advised by the International Conference on Harmonisation (ICH) of technical requirements for registration of pharmaceuticals for human use (Validation of Analytical Procedures: Text and Methodology Q2(R1) , 2005). Hence, linearity was established using 7 concentrations for albumins and for globulins. The correlation coefficient, the y-intercept and the slope of the regression line were determined. The range corresponds to the minimum and maximum concentrations analyzed in the present study, taking into account the purity of the isolates. Accuracy was determined by the comparison with the reference

method for protein content determination, the Kjeldahl method, on the 5 conditions of extractions previously described. The precision was described by the repeatability, or intraday variation, which is calculated as the variation observed when the same sample is analyzed three times the same day, for 3 concentration levels (20, 12.5 and 5 g.L⁻¹). The intermediate precision, or inter-day variation was determined as the variation observed for 3 concentration levels (20, 12.5 and 5 g.L⁻¹) measured on three different days. The detection limit and the quantitation limit were calculated by the comparison of the signal and of the noise of the analysis. The detection limit was determined as the minimal concentration that produced a signal 3 times higher than the noise. The quantitation limit was the minimal concentration with a signal 10 times higher than the noise.

V.1.4.3 Results and discussion

V.1.4.3.a) Separation of proteins by SE-HPLC

The aim of the method was to quantify the two main rapeseed proteins (albumins and globulins) by SE-HPLC. This required to choose a column and an elution condition allowing a separation of these two proteins. Globulins and albumins molar weight is around 300 kDa and 14 kDa respectively. As a consequence, Biosep-SEC-s2000 column was selected because its separation range was given between 1 and 300 kDa. The eluent was a water/acetonitrile mixture (55/45 v/v) containing TFA (0.1% v/v) to avoid protein/protein association or aggregation that would yield multiple peaks for a single protein.

Figure 24 A shows the chromatogram of the rapeseed protein isolate at 214 nm. Two main peaks were observed at respective retention times of 9 and 11 minutes. According to the column calibration (data not shown), these retention times corresponded to molar weights around 50 kDa and 10 kDa respectively. Rapeseed globulins (MW = 300 kDa) are composed of 6 subunits of 50 kDa linked together by electrostatic interactions (Schwenke, Raab, Plietz, & Damaschun, 1983). Albumins (MW = 14 kDa) are composed of 4 and 9 kDa subunits linked by disulfide bridges. Hence, the first peak observed would correspond to globulin subunits dissociated by the elution conditions. Indeed, at the pH of the mobile phase (around 2.5) globulins subunits are positively charged because globulins isoelectric point was reported around 5 (Cheung, Wanasundara, & Nickerson, 2014). This probably results in a subunits dissociation and explains that globulins were eluted in a 50 kDa peak. The second peak at around 10 kDa would correspond to albumins. The elution conditions did not contain any reducing agents, so albumins subunits should remain bonded together. The apparent molar weight given by the calibration was slightly below 13 kDa, probably because of the compactness of albumins due to its reported two intra molecular disulfide bonds (Perera, McIntosh, & Wanasundara, 2016).

To check that out, each peak was fractionated and analyzed by SDS-PAGE in reducing conditions (with β mercaptoethanol). As shown in Figure 24 B. Fractions 2 to 4 (between 8 and 10 minutes of retention time) were composed of polypeptides with molecular weights around 30 and 20 kDa. This corresponds to the globulin α and β -chains respectively. These two chains are known to be bound by disulfide bridges. As a consequence, they were not dissociated by the elution conditions and are well eluted under a 50 kDa globulin subunit. From fraction 6 to 10, two bands are visible corresponding to low molecular weight proteins and are consistent with the two subunits of albumins of approximately 4 and 9 kDa (that remained associated under elution condition). This conformed with the alleged composition of peaks on the chromatogram. The slight peak overlapping, fraction 5 (5% of the total peak area) is composed of a mixture of albumins and globulins.

Previous studies used SE-HPLC for the analysis of proteins. Notably, it was reported that SE-HPLC could be used to follow-up a process of enzymatic cross-linking of milk proteins by the transglutaminase. However, only a comparison of the area under the peaks was carried out and no quantification was done (Faergemand & Qvist, 1997). More specific work was carried out on protein from plants: a study with a Biosep-SEC-s4000 column showed the separation of gliadins and albumins/globulins from wheat after 22 minutes of elution versus 15 minutes in the proposed method. However, no quantification of the protein content was carried out (Bangur, Batey, McKenzie, & MacRitchie, 1997). In 2005, it was shown the separation of wheat proteins between 10 and 22 minutes of elution. This method is the closest to the condition used in the present article with a Biosep-SEC-s4000 column and a mobile phase composed of 50 % of acetonitrile, 50 % water and 0.05 % TFA. However, it showed an elution of proteins between 10 and 22 minutes with a poor resolution even though two peaks were visible (Singh H. , 2005) whereas the present SE-HPLC method elutes proteins between 8 and 15 minutes with a good resolution.

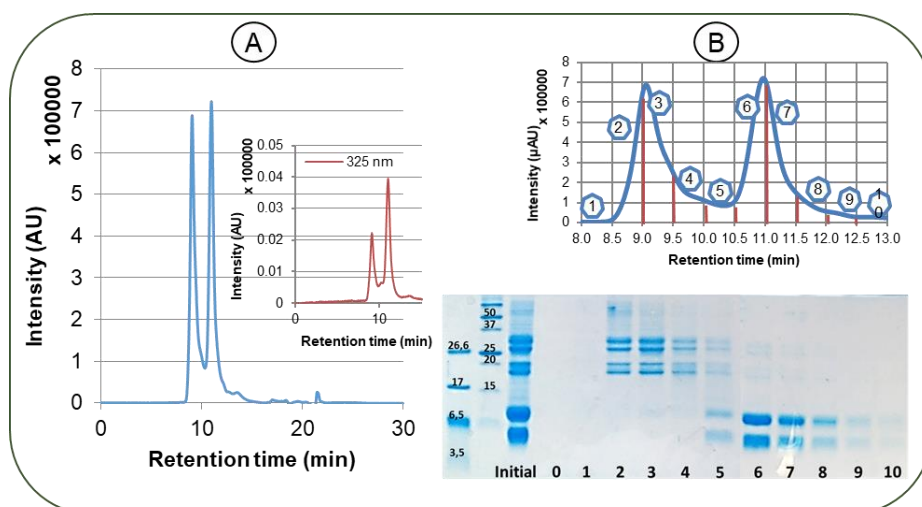


Figure 24 :Total rapeseed proteins isolate analysis, A. Chromatograms of the SE-HPLC method at 214 and 325 nm ; B. Fractionation of the main peaks observed in Size-Exclusion HPLC and SDS-PAGE analysis of the obtained fractions.

V.1.4.3.b) Method validation with protein isolates

The reference assay for food plant protein quantification is the Kjeldahl method. So, the proposed SE-HPLC method was validated by comparison with this method. The protein quantification by Kjeldahl requires the use of a nitrogen-to-protein factor (Ka), since it just quantifies the nitrogen content. The most common Ka used for vegetal proteins is 6.25 (Mariotti, Tomé, & Patureau Mirand, 2008). However, the nitrogen content in proteins depends upon their amino acid sequence. It is possible to calculate a Ka by formula (7) (Mossé, Huet, & Baudet, 1985):

$$Ka = \frac{\sum E_i}{\sum D_i} \quad (\text{Eq. 12})$$

In this equation, i corresponds to each anhydrous amino acid, E_i is its molecular weight and D_i its nitrogen content. For rapeseed, few studies proposed values for Ka but one estimated it as an average 5.53 (Tkachuk, 1969). The calculation by Mariotti gave an average value of 5.35 (Mariotti, Tomé, & Patureau Mirand, 2008). Yet, these values are likely to change with the proportion of globulins and albumins. As a consequence, it was chosen to validate the method with globulins and albumins isolates using Ka values calculated from their amino acids sequences.

Since no standards of globulins and albumins are commercially available, protein isolates were produced according to (US Patent No. 8,541,046 B2, 2013). Figure 25 shows the elution chromatograms at 214 nm of the globulin and albumin isolates prepared. For each isolate, a single peak is observed corresponding to the elution duration observed above. SDS-PAGE analysis confirmed the composition of each isolate. For globulins, a slight contamination by albumins was observed and estimated as 4 % on the SE-HPLC chromatogram. For albumins, the two main bands characteristics of their two subunits were visible by SDS-PAGE with no apparent contamination by globulins. As a consequence, these isolates were considered suitable to carry out the validation of the proposed method.

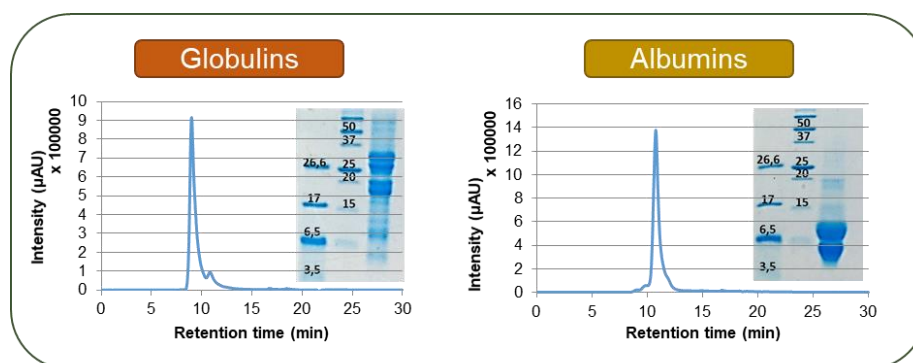


Figure 25 : Globulins and albumins isolates analysis by Size-Exclusion at 214 nm and SDS-PAGE

The validation methodology used is shown on Figure 26. Isolate solutions of globulins and albumins ranging from 1 to 20 g.L⁻¹ were prepared and analyzed by SE-HPLC. On SE-HPLC chromatograms at 214 nm, the absorbance of albumins and globulins peaks were converted into concentrations using the Beer-Lambert law. The obtained peak was integrated to give an amount of proteins in the sample. To do so, the extinction coefficients of albumins and globulins were calculated with amino acids sequences, as described by Kuipers and Gruppen (2007) on the basis of the UniProt database (P80208 for albumins and Q7XB53 for globulins). For globulins, this protein sequence corresponded to a molecular weight of 55687 Da, a theoretical mass extinction coefficient ϵ of 17.22 and a theoretical nitrogen-to-protein conversion factor, K_a , of 5.46. For albumins, it was 14035 Da, the theoretical mass ϵ was 16.44 and the theoretical K_a was 5.24.

The isolate solutions were also analyzed by Kjeldahl method to get the amount of nitrogen. The nitrogen content and the protein concentrations determined by SE-HPLC were plotted against the isolate concentrations. From this plot, an experimental nitrogen-to-protein conversion factor (K_a) was determined as in equation 8.

$$K_a = \frac{\text{Protein content in the sample (SE-HPLC)}}{\text{Nitrogen content (Kjeldahl)}} \quad (\text{Eq. 12 bis})$$

Then the obtained experimental K_a values were compared to theoretical values calculated from the amino acids sequences.

Figure 26 shows the plots of the protein content by SE-HPLC and the nitrogen content by Kjeldahl as functions of the isolate concentration in the solutions for each protein. For both methods, a good linearity is observed with values of R^2 superior to 0.99. The wide range of linearity, between 1 and 20 g.L⁻¹, is particularly convenient as it covers the concentrations of proteins usually found in aqueous extracts. The values of experimental K_a for globulins and albumins deduced from this graph were 5.13 and 4.94 respectively, Tableau 4.

This made relative differences of 6 and 5.7 % from the theoretical values (calculated from protein sequences). This slight discrepancy can be explained by (i) the presence of several isoforms of each protein in a same cultivar of a species leading to an approximation in the values of ϵ and K_a (ii) an approximation in the calculation of the theoretical ϵ as the folding of the protein might change UV absorption; (iii) a contribution to the absorption by remaining polyphenols in the isolates; or (iv) a flawed optical path in the HPLC system. As a consequence, these results show that the use of the SE-HPLC would give very similar results to the reference method, Kjeldahl.

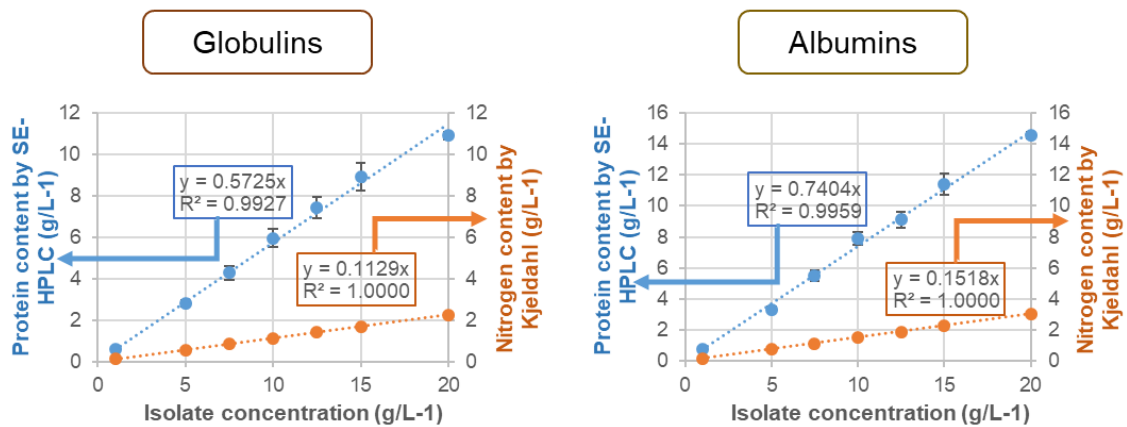
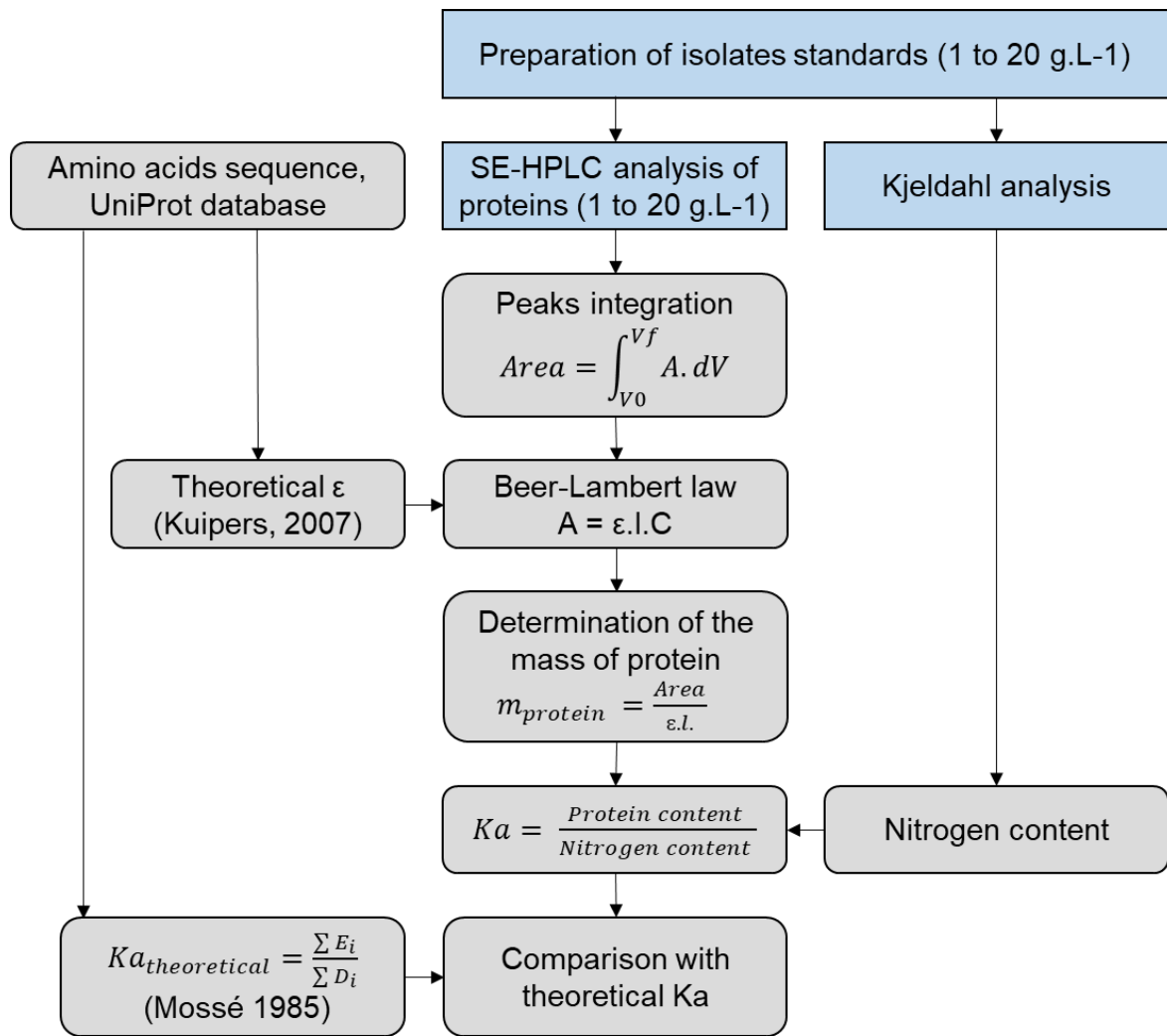


Figure 26 : Methodology for the validation of the SE-HPLC method. Light blue boxes represent analytical steps and grey boxes show the calculated values. Below are the values obtained by both methods represented as functions of the isolate concentration. The Ka is the ratio of both slopes.

Tableau 4 : Comparison of the SE-HPLC method with the Kjeldahl method and with the SDS-PAGE densitometry.

True protein content according to extraction condition	SE-HPLC method (N=3)	Kjeldahl (N=3)	Statistical comparison
			P-value T-test (two tails)
pH 2 0 M NaCl	10.5 ± 1.16	11.8 ± 0.81	0.197
pH 5 0 M NaCl	10.7 ± 0.76	12.3 ± 0.57	0.046 *
pH 7 0 M NaCl	11.0 ± 0.64	10.7 ± 1.04	0.691
pH 7 0.1 M NaCl	16.8 ± 0.79	16.3 ± 0.99	0.464
pH 9 0.3 M NaCl	21.4 ± 2.30	20.1 ± 0.45	0.441
Proportion of albumins	SE-HPLC method (N=3)	SDS-PAGE densitometry (N=3)	Statistical comparison
			P-value T-test (bilateral)
pH 2 0 M NaCl	90.7 ± 0.45	95.8 ± 1.06	0.002 *
pH 5 0 M NaCl	54.5 ± 2.27	55.2 ± 2.29	0.712
pH 7 0 M NaCl	30.7 ± 1.44	37.9 ± 4.33	0.052
pH 7 0.1 M NaCl	50.2 ± 1.20	57.0 ± 1.99	0.007 *
pH 9 0.3 M NaCl	50.3 ± 0.44	61.6 ± 3.02	0.023 *

*The highest part of the Table contains the data used for the characterization of the homemade protein isolates of globulins and albumins from rapeseed meal. The second part consists of the data used to compared the SE-HPLC method with references, Kjeldahl and SDS-PAGE densitometry. N corresponds to the number of repetitions of the experiments. Experimental values are represented as mean ± the standard deviation. In the statistical analysis, values followed by * are below the signification threshold of 5 %.*

V.1.4.3.c) Application of the SE HPLC method for total proteins quantification in rapeseed meal extracts

The main interest is to be able to quantify rapeseed proteins in aqueous extracts from meal solid/liquid extraction. The main parameters that may affect the extraction of one or both proteins are the pH and the ionic strength. As a consequence, aqueous extracts were prepared in several conditions of pH and ionic strength: pH 2 0 M, pH 5 0 M, pH 7 0 M, pH 7 0.1 M and pH 9 0.3 M (Tableau 4). These conditions were chosen in order to get various protein concentrations and globulins/albumins compositions. The extracts were analyzed by SE-HPLC and Kjeldahl. Kjeldahl analysis required an additional step of precipitation with TCA in order to determine the true protein content in samples. Since Ka values of globulins and albumins are very close (respectively 5.46 and 5.24), a Ka value of 5.35 was chosen whatever the globulin / albumin proportion.

Figure 27 represents the stacking of chromatograms at 214 nm for three extraction conditions: pH 2 0 M NaCl, pH 7 0 M NaCl and pH 9 0.3 M NaCl. On the right side is a zoom of the protein peaks. As expected, it shows that the composition of the extract depends very much on the extraction condition: at acidic pH albumin was mainly extracted, at neutral pH globulin represented the majority of the proteins and at alkaline pH, there was a mixture of both.

Other peaks were visible after 15 minutes. Because the separation was based on size-exclusion, these peaks corresponded to low molecular weight molecules. Since phenolic compounds are soluble in water, they are co-extracted with proteins. Moreover, common chromatographic methods to detect and quantify phenolics usually use a UV detection at 325 nm. In the present study, the chromatograms at 325 nm (data not shown) show the presence of several signals at these retention times. They are very likely to correspond to free phenolic compounds. Nevertheless, the presence of these peaks after 15 minutes, did not impact the quantification of rapeseed proteins. In any case, globulin and albumin signals were not biased with other solutes absorbing at 214 nm. Hence, the SE-HPLC method was suitable for protein quantifications from liquid complex extracts.

The results of Kjeldahl and SE-HPLC quantifications are featured in Tableau 4. A statistical investigation was carried out to compare both sets of data. For all sets of data, the Student T-test p-value was higher than 0.05, resulting in the equality of means. except for the pH 5 condition for which there is a small probability of H0 being false ($p=0.046$) for this extract only. As a consequence, the proposed method can be used as a substitute of the Kjeldahl method for quantifying soluble proteins. The SE-HPLC presents many advantages: (i) it is very convenient because it is fully automated; (ii), it is less dangerous for the user (the Kjeldahl method requires the use of strong acids and bases); (iii) it generates less effluents; (iv) it is quick as one run lasts 30 minutes.

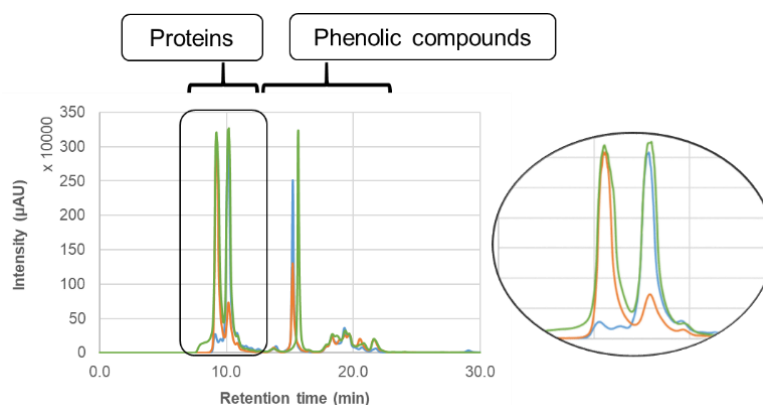


Figure 27 : Stacked chromatograms of extracts at 214 nm and for three extraction conditions. The blue line corresponds to pH 2 0 M NaCl, the orange to pH 7 0 M NaCl and the green to pH 9 0.3 M NaCl. On the right side of the figure is a zoom of the protein peaks.

V.1.4.3.d) Application of the SE HPLC method for albumins/globulins composition determination

The Kjeldahl method does not discriminate between globulins and albumins. That is not the case of the proposed SE-HPLC method. To determine the concentration of one of the proteins in a mixture, the only referenced method is SDS-PAGE densitometry. Hence, the SE-HPLC method was compared to SDS-PAGE densitometry for the quantification of globulins and albumins in aqueous extracts. Tableau 4 shows the values obtained by both methods for the albumin content in the aqueous extracts. The

average values between both methods present up to 19 % difference (for the condition pH 7 and 0.1 M NaCl). As seen before, various compositions of globulins are found between the extraction conditions. These differences can be quantified with both methods.

For three of the conditions (pH 2, pH 7 0.1 M NaCl and pH 9 0.3 M NaCl), there was a significant difference between the methods ($p < 0.05$). The other two conditions (pH 7 0 M and pH 5 0 M) were found statistically equal. However, for the SDS-PAGE densitometry, very large standard deviations were observed which makes the comparison more difficult.

As a consequence, compared to SDS-PAGE densitometry, the SE-HPLC method would be easier to carry out as no particular preparation is required to dose the proteins. Whereas, to prepare SDS samples, it is necessary to dilute the sample quite accurately as the linear parts of the calibrations are between 0.2 and 0.8 g.L⁻¹. Moreover, SDS-PAGE is a method which is longer to apply because: gels must be prepared before the migration step; there are several steps of staining and destaining; and finally, the computational densitometry treatment of the scan or picture of the gel. With a four-gel SDS-PAGE system, up to 36 samples can be loaded in a single run with one marker per gel. The total time including all operations was estimated at 18 hours which would be 30 minutes per sample. Hence, depending on the number of samples to be analyzed, the average time of analysis per sample would be between 30 minutes (for 36 samples) and 18 hours (for one sample). And, contrary to the SE-HPLC analysis, many time-consuming operating steps are required. Indeed, the SE-HPLC method requires only the filtration of the samples. Furthermore, the narrow linearity range will imply to try different sample dilutions thus, even increasing the overall duration. Moreover, the processing of the data with the SDS-PAGE densitometry also involves several steps that were found to induce a bias as shown by the large standard deviations. The Kjeldahl method would require a total time of 4 hours to analyze 5 samples (total nitrogen and non-protein nitrogen analyses), including one blank and one standard analysis. It corresponds to 48 minutes per sample without determination of the globulin and albumin contents. Moreover, it implies a full-time operator exposed to concentrated and dangerous substances. On the contrary, the SE-HPLC method requires only 30 minutes per analysis once the method is set up and after a 30 minutes' equilibration of the system. There are few steps requiring the intervention of an operator, less dangerous substances and a computer-assisted processing of the data.

V.1.4.3.e) Stability study of extracts

Proteins are known to be sensitive to their medium: the presence of other molecules such as polyphenols for example, that are interacting with proteins and can be oxidized under alkaline conditions; pH and ionic strength are directly linked to protein precipitation (Purkayastha, et al., 2014).

All these factors may cause a variation in what is observed by SE-HPLC. Moreover, as shown in , polyphenols are detected at 325 nm. These molecules are very sensitive to oxidation, and hence, their presence in the aqueous medium after extraction, justifies the need to investigate the stability of the samples.

Thus, analyses by SE-HPLC were carried out on extracts produced in a variety of conditions of pH and ionic strength. Figure 28 shows the results of the stability study under several conditions of storage, room temperature, 4 °C and -20 °C, after 24 h or 48 h. The results were compared to the initial analysis of each extract.

As expected, a larger variation was visible on extracts stored at room temperature, all the more after 48 hours. However, pH seemed to affect the stability of certain extracts: the variation was the largest for the samples produced at pH 5 or 7 with 0 or 0.1 M of NaCl. On the contrary, the set of conditions with the most stable extract is pH 9 and 0.3 M NaCl which presented a variation of only 1 % at most whatever the storage condition.

As a consequence, this stability study showed the importance of storage conditions particularly for extracts produced at a neutral pH. These samples must be stored under the coldest conditions as possible to preserve them and constrain alterations.

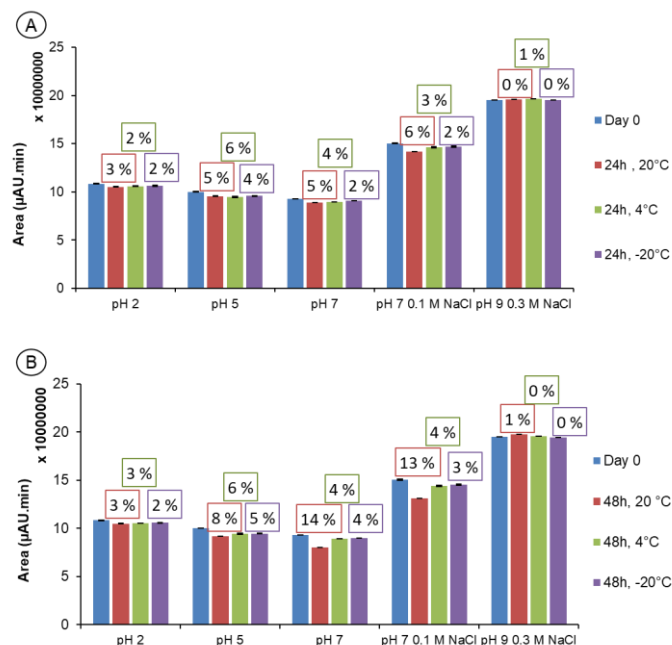


Figure 28 : Stability study of several rapeseed protein extracts under different storage conditions (room temperature, 4 °C and -20 °C) after 24 h (A) or 48 h (B). Values on top of the bars represent the absolute relative difference of area with the initial measure.

V.1.4.3.f) Validation of the method according to the ICH validation guide

The area under the peak (in $\mu\text{AU}\cdot\text{L}$) of both globulin and albumin isolates was plotted against the concentration in the sample (data not shown) between 1 and 20 $\text{g}\cdot\text{L}^{-1}$ of isolate powder. Seven points of concentration were considered with two to three replicates. The concentration in the sample was determined according to the Kjeldahl method using a nitrogen-to-protein coefficient of 5.35. Regarding globulins, it was found a correlation coefficient of 0.9927, the y-intercept was set at 0 and the slope of the regression was 81.5. As for albumins, it was found a regression coefficient of 0.9959 with a y-intercept of 0 and a slope of 75.0. Hence, for both proteins, very good correlations were found. The data regarding the ICH validation is shown in supplementary data, Tableau 5.

Tableau 5 : Validation of the method according to the International Conference on Harmonisation (ICH) guideline.

	Globulins	Albumins
Regression coefficient	0.9927	0.9959
y-intercept	0	0
Slope	81.5	75.0
Concentration range	0.6 to 12.0 $\text{g}\cdot\text{L}^{-1}$	0.4 to 16.2 $\text{g}\cdot\text{L}^{-1}$
Repeatability (intra-day variation)	0.3 %	2.8 %
Intermediate precision (inter-day variation)	0.9 %	3.4 %
Detection limit	0.008 $\text{g}\cdot\text{L}^{-1}$	0.010 $\text{g}\cdot\text{L}^{-1}$
Quantitation limit	0.026 $\text{g}\cdot\text{L}^{-1}$	0.033 $\text{g}\cdot\text{L}^{-1}$

Taking into account the purity of the isolates (with a K_a value of 5.35), the range of concentrations studied was determined. For albumins, it was between 0.4 and 16.2 $\text{g}\cdot\text{L}^{-1}$ and for globulins, it was 0.6 to 12.0 $\text{g}\cdot\text{L}^{-1}$. As a consequence, the method was successfully developed to cover the usual range of concentrations found in rapeseed protein extracts.

For albumins, the repeatability, or intra-day variation was found with a value of 2.8 %. The intermediate precision, or inter-day variation, was 3.4 %. For globulins, repeatability was 0.3 % and intermediate precision was 0.9 %. These values are below 5 % and in consequence, the precision of the method was deemed validated regarding the ICH procedure.

For albumins, the detection limit was found at 0.010 $\text{g}\cdot\text{L}^{-1}$ and the quantitation limit was 0.033 $\text{g}\cdot\text{L}^{-1}$. Similar values were found for globulins with, respectively, 0.008 $\text{g}\cdot\text{L}^{-1}$ and 0.026 $\text{g}\cdot\text{L}^{-1}$.

V.1.4.3.g) Application of the method on other oleoproteaginous resources: examples of linseed and sunflower

Sunflower and linseed are two other major oleoproteaginous produced all over the world. Hence, the same issues regarding the need of a reliable and quick analysis of the proteins are emerging. Indeed, sunflower proteins are also composed of two major types: globulins (11 S), also called helianthinin,

and albumins (2 S) with respective proportions in the meal of about 60 % and 30 %. Their molecular weight is similar to the same proteins in rapeseed (Gonzalez-Pérez & Vereijken, 2007). Linseed, of flaxseed, also have their own 11-12 S globulins, also called linin, and 2 S albumins or conlinin. However, these proteins have been the subject of fewer studies to this day that is why limited data is available regarding this resource (Chung, Lei, & Li-Chan, 2005). As a consequence, the approach used for rapeseed proteins is applied to both linseed and sunflower.

Figure 29 shows the chromatograms of linseed and sunflower with their associated SDS-PAGE. Regarding linseed proteins, two main peaks were observed on the chromatograms. The first peak has a retention time between 8 and 10.5 minutes: it is coherent with a high molecular weight protein, globulins. The second peak, between 10.5 and 14 minutes, would correspond to low molecular weight protein, albumins. These observations were completed by SDS-PAGE analysis after a fractionation every 30 seconds between 7.5 and 15 minutes. In the first fraction, two bands were visible: one thin around 30 kDa and the other one wider between 20 and 25 kDa. These would be the result of the dissociation of disulfide bonds from subunits of globulins. In the second fraction, several bands were present between 6.5 and 15 kDa, consistent with albumins.

In respect to sunflower, two peaks were also present on the chromatograms, the first between 8 and 10.5 minutes and the second between 10.5 and 15 minutes, probably corresponding to respectively globulins and albumins. The SDS-PAGE showed that the first peak, fraction 1, was composed of a high-molecular weight protein, dissociated in 3 bands corresponding to the 3 main subunits of globulins: α and α' chains with 30 to 33 kDa and β chain with 21 kDa. Fraction 2 showed the presence of several bands between 16 to 18 kDa corresponding to sunflower albumins.

As before, thanks to sequences it was possible to determine K_a values and extinction coefficients ϵ for each protein (data not shown). The sequences for linseed were taken from (Nwachukwu & Aluko, 2018) and the calculation gave an average K_a of 5.56. The sequences for sunflower were found in the Uniprot database P19084 (chain) for albumins and P19084 for globulins and resulted in an average K_a of 5.58.

Regarding linseed, the total proteins content in the extract determined by the SE-HPLC method was 1.72 g.L^{-1} versus 1.64 g.L^{-1} with the Kjeldahl method, which represents a relative difference of 4.8 %. As for sunflower, the respective values observed by SE-HPLC and Kjeldahl were 5.70 g.L^{-1} and 6.41 g.L^{-1} meaning a relative difference of 11 %.

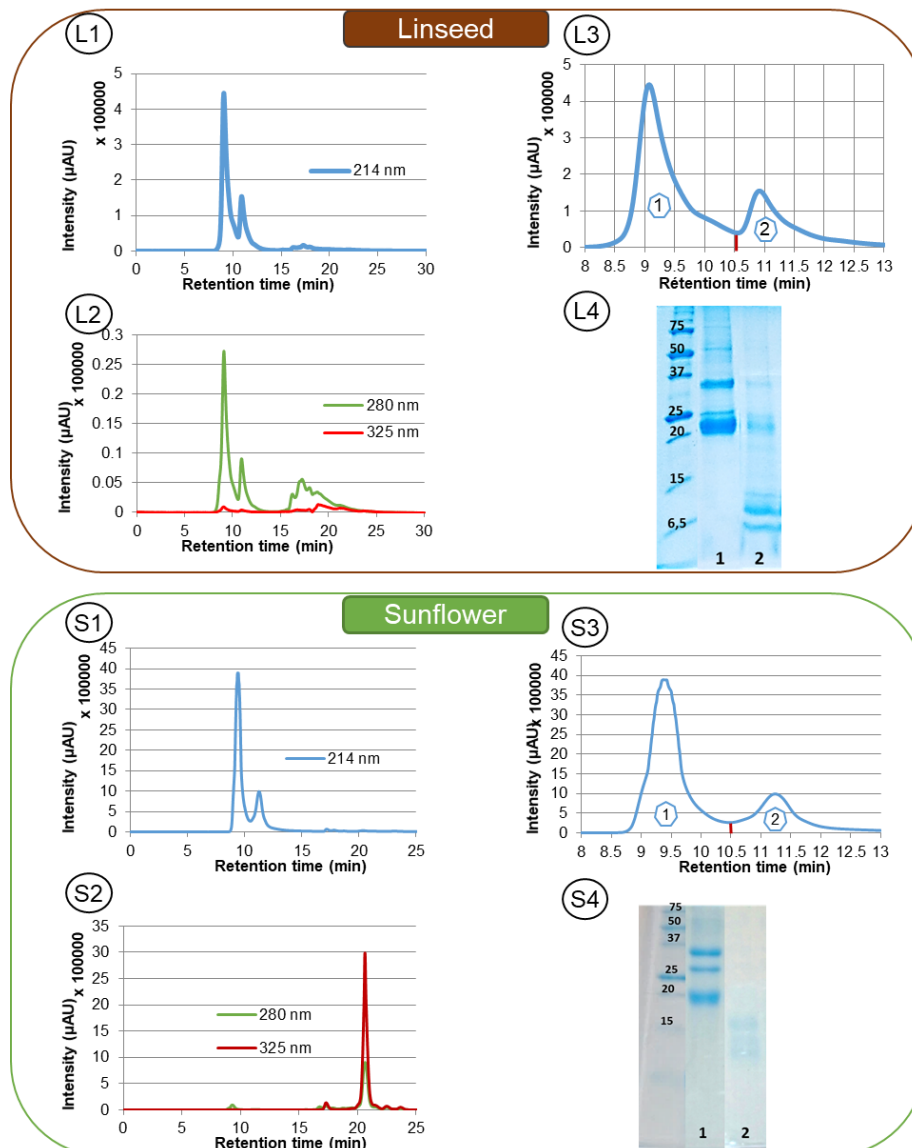


Figure 29 : L1 and S1 are the chromatograms at 214 nm ; L2 and S2 are the stacked chromatograms at 280 nm and 325 nm ; L3 and S3 are zooms from 214 nm chromatograms and show the different fractions shown on SDS-PAGE gels in reducing conditions, L4 and S4.

As a result, these preliminary experiments prove the validity of the SE-HPLC method on these resources as well. Because rapeseed (*Brassica napus*), sunflower (*Helianthus*) and linseed (*Linum usitatissimum*) are different biologically and because the SE-HPLC method works on all three, it is very likely that the proposed method will also work with other oleoproteaginous plant proteins.

V.1.4.4 Conclusion

This study proved the possibility to use SE-HPLC to characterize protein-rich aqueous extracts from rapeseed: indeed, both the soluble protein content and the proportion of each major protein type was determined by this method. It was shown to give statistically equivalent results compared to the Kjeldahl method, for protein isolates but also for proteins in aqueous mediums of extraction. The comparison with SDS-PAGE densitometry was also concluding for the analysis of protein isolates even

though very small concentrations were required. However, it was likely biased by the presence of polyphenols in aqueous extracts resulting in interferences in SDS-PAGE gels.

Furthermore, the proposed method allows not only the quantification of the protein content in a sample but also the determination of the proportion of each major protein. Moreover, this method presents several major advantages: it is quick, with a 30 minutes analysis per sample; no sample preparation is required after filtration; the amount of effluent generated is moderate compared to Kjeldahl; it is accurate and reliable. Finally, it was shown to be usable also on other oleoproteaginous resources.

V.1.5 Contribution de l'article

La méthode développée permet d'obtenir des résultats statistiquement équivalents à ceux obtenus par la méthode de référence, le Kjeldahl. Par ailleurs, elle présente de nombreux avantages pour l'utilisateur : elle est rapide puisqu'un échantillon est analysé en 30 minutes, automatisée, et elle génère moins d'effluents toxiques que les autres méthodes. Ce premier article permet donc d'affirmer que la méthode développée par CLHP-ES peut légitimement être utilisée dans les études suivantes en tant que substitut ou complément de la méthode de Kjeldahl.

V.2 Détermination de la teneur en acide phytique

Le contenu en acide phytique (AP) a été déterminé sur des échantillons solides avec une adaptation de la méthode développée par Garcia-Esteva (Garcia-Esteva, Guerra-Hernandez, & Garcia-Villanova, 1999). Le principe du dosage est détaillé dans les paragraphes suivants.

V.2.1 Extraction

Une masse entre 250 et 500 mg d'échantillon solide a été pesée. L'extraction a été réalisée à un pH proche de 1 dans 20 mL d'une solution de HCl + 10 % Na₂SO₄ (Sigma-Aldrich, Saint-Louis, Etats-Unis) pendant au moins 2 heures. Après une centrifugation de 15 minutes à 15000 g, le surnageant a été récupéré et filtré sur une membrane de 0,22 µm de Pall (Port Washington, New York, Etats-Unis).

V.2.2 Formation d'un complexe

L'extrait filtré a été mélangé à volume égal, 2,5 mL, avec les solutions suivantes : 20 mM de chlorure de fer (Sigma-Aldrich, Saint-Louis, Etats-Unis), 20 % d'acide sulfosalicylique (Sigma-Aldrich, Saint-Louis, Etats-Unis) et une solution de 0,4 M de HCl + 10 % Na₂SO₄ (Sigma-Aldrich, Saint-Louis, Etats-Unis). Le mélange a été agité pour permettre aux ions Fe³⁺ et à l'acide sulfosalicylique de former un complexe de couleur bordeaux capable de réagir avec l'AP. Un traitement thermique a été ensuite réalisé en plongeant les tubes d'échantillons dans un bain à 100 °C pendant 15 à 20 minutes. Au cours de cette étape, le complexe entre les ions Fe³⁺, l'acide sulfosalicylique et l'AP précipite.

Après refroidissement, les échantillons ont été centrifugés 30 minutes à 10000 g.

Ensuite, le précipité a été lavé trois fois avec de l'eau distillée afin de récupérer les ions Fe^{3+} libres. Les surnageants ont été regroupés dans une fiole jaugée de 50 mL.

V.2.3 Dosage des ions Fe^{3+} libres

Un volume de solution de 20 mL a été prélevé. Le pH de cet aliquot est ajusté à $2,5 \pm 0,5$ par ajout de glycine (Sigma-Aldrich, Saint-Louis, Etats-Unis). La solution a été chauffée à une température entre 70 et 80 °C. Enfin, une solution de 2 mM d'EDTA (acide éthylène diamine tétra acétique) a été utilisée en tant que solution titrante. Le volume équivalent a été observé par un changement de couleur du bordeaux au jaune/vert.

V.2.4 Calculs pour la détermination de la teneur en acide phytique.

Le dosage permet de quantifier la quantité d'ions Fe^{3+} libres

$$n(Fe^{3+})_{libre} = V_{eq} \times [EDTA] \times 2,5 \quad (\text{Eq. 13})$$

Avec [EDTA] la concentration en mmol/L, V_{eq} le volume équivalent en L et 2,5 correspondant au volume de 20 mL prélevé de la fiole jaugée de 50 mL.

Un blanc, c'est-à-dire, une solution sans échantillon a permis de déterminer la quantité totale d'ions Fe^{3+} . Ainsi, la formule suivante permet de déduire la quantité d'ions Fe^{3+} contenue dans le précipité.

$$n(Fe^{3+})_{précipité} = n(Fe^{3+})_{total} - n(Fe^{3+})_{libre} \quad (\text{Eq. 14})$$

Dans la littérature, il est généralement admis que six atomes de phosphore se lient à quatre ions Fe^{3+} (Garcia-Esteva, Guerra-Hernandez, & Garcia-Villanova, 1999) (Cheryan & Rackis, 1980).

$$\frac{n(\text{Phosphore})}{n(Fe^{3+})} = \frac{6}{4} \quad (\text{Eq. 15})$$

De plus, il est supposé qu'une molécule d'AP contient six atomes de phosphore, soit :

$$n(\text{Acide Phytique}) = \frac{n(\text{Phosphore})}{6} \quad (\text{Eq. 16})$$

La combinaison des deux formules donne :

$$n(\text{Acide Phytique}) = \frac{n(Fe^{3+})_{précipité}}{4} \quad (\text{Eq. 17})$$

Puisque 2,5 mL d'extrait ont été prélevés à partir de la solution d'extraction de 20 mL, la quantité d'acide phytique dans l'extrait est exprimée par :

$$n(\text{Acide Phytique})_{\text{extract}} = 8 \times n(\text{Acide Phytique})_{\text{précipité}} \quad (\text{Eq. 18})$$

Enfin, la masse d'AP a été déterminée et son contenu a été exprimé relativement à la masse de protéines ou de matière sèche.

$$m(\text{Acide Phytique})_{\text{extrait}} = \frac{n(\text{Acide Phytique})}{M(\text{Acide Phytique})} \quad (\text{Eq. 19})$$

$$\text{Teneur en acide phytique (en \% de matière sèche)} = \frac{m(\text{Acide Phytique})_{\text{extrait}}}{m(\text{Matière sèche})_{\text{extrait}}} \quad (\text{Eq. 20})$$

V.3 Détermination de la complexation phénolique

La méthode CLHP-ES, décrite précédemment, a aussi été utilisée afin d'évaluer l'interaction de composés phénoliques avec les protéines. Les protéines présentent deux maximums d'absorbance, à 214 et à 280 nm. Cette dernière correspond à l'absorption des résidus d'acides aminés aromatiques, le tryptophane, la phénylalanine et la tyrosine, et est spécifique des protéines. La méthode CLHP-ES développée permet l'élution des protéines entre 8 et 14 minutes.

Les composés phénoliques présentent un pic d'absorbance à 325 nm mais n'absorbent pas à 280 nm. Puisque la méthode chromatographique repose sur l'exclusion stérique, les composés phénoliques libres sont élués après les protéines, soit après 15 minutes. Ainsi, la présence de composés phénoliques complexés est avérée si un pic est observé à 325 nm à un temps de rétention correspondant aux protéines, c'est-à-dire, entre 8 et 14 minutes.

C'est pourquoi, la complexation phénolique a été définie comme étant l'aire sous le pic de protéines à 325 nm, divisée par celle à 280 nm.

$$\text{Complexation phénolique} = \frac{\text{Aire sous le pic entre 8 et 14 minutes à 325 nm}}{\text{Aire sous le pic entre 8 et 14 minutes à 280 nm}} \times 100 \quad (\text{Eq. 21})$$

VI Les méthodes de caractérisation des isolats

VI.1 Caractérisation générale

VI.1.1 Matière sèche

La matière sèche (MS) a été obtenue en plaçant une masse donnée d'échantillon (solide ou liquide) dans un four à 110 °C pendant 24 heures. La matière sèche est donnée par :

$$\text{Matière sèche (\%)} = \frac{\text{Masse d'échantillon sec}}{\text{Masse d'échantillon humide}} \quad (\text{Eq. 22})$$

VI.1.2 Pureté et solubilité

Une solution à 5 g.L⁻¹ a été préparée à partir de protéines lyophilisées. L'analyse Kjeldahl (N x 5,35) de cette solution donne la pureté selon la formule suivante :

$$\text{Pureté des protéines (\% MS)} = \frac{\text{Masse de protéines par la méthode Kjeldahl}}{\text{Masse de lyophilisat} \times \text{Matière sèche (\%)}} \times 100 \quad (\text{Eq. 23})$$

Concernant la solubilité, un échantillon de 20 mL a été utilisé pour chaque point de pH étudié. Le pH a été ajusté à 3, 5, 7 ou 9 avec 0,1 M de NaOH ou de HCl. Les volumes de NaOH et HCl ont été pris en compte pour le calcul du volume total de solution. Les protéines insolubles sont séparées, à température ambiante, par une centrifugation de 15 minutes à 10000 g. Le surnageant a été récupéré et filtré sur un filtre de 0,22 µm. Chaque analyse a été dupliquée. La solubilité à un pH donné a été calculée comme le ratio de la quantité de protéines dans le surnageant par la quantité de protéines dans la solution de départ.

$$\text{Solubilité}_{\text{pH } X} = \frac{\text{Masse de protéines dans l'aliquot}}{\text{Masse de protéines dans la solutions initiale}} \quad (\text{Eq. 24})$$

VI.2 Les méthodes d'analyse structurale

VI.2.1 Le dichroïsme circulaire

Les structures secondaires des protéines ont été analysées dans le domaine des UV lointains entre 180 et 260 nm. Des solutions de protéines de 1 mg.mL⁻¹ ont été préparées dans un tampon phosphate à pH 7, à une concentration de 0,5 M. Les analyses spectrales ont été effectuées à température ambiante avec un Chirascan Plus de la société Applied Photophysics (Leatherhead, Grande-Bretagne). Pour l'acquisition du spectre, des échantillons ont été préparés dans des cuves de 0,1 mm de longueur. Pour chaque échantillon, l'analyse a été répétée trois fois, et une moyenne a été tracée. Le logiciel Pro-Data Chirascan de la société Applied Photophysics (Leatherhead, Grande-Bretagne) a été utilisé pour traiter les données.

Le même système a été utilisé pour évaluer les structures tertiaires en utilisant un spectre dans le domaine des UV proches, entre 250 et 320 nm. Dans ce cas, une cuve de quartz de 10 mm de longueur a été utilisée.

VI.2.2 Détermination de la taille des particules par Diffusion Dynamique de la Lumière

La Diffusion Dynamique de la Lumière (DLS), a été utilisée pour déterminer l'état d'agrégation des protéines. Des solutions de protéines à 1 mg.mL⁻¹ ont été préparées dans un tampon phosphate 0,5 M à pH 7. Des cuves en quartz de 10 mm de longueur ont été utilisées. Les analyses ont été effectuées avec un Zetasizer Nano – S de Malvern Instruments (Malvern, Grande-Bretagne).

VI.3 Les méthodes d'analyse des propriétés fonctionnelles

VI.3.1 Qualification de la couleur

La couleur des isolats de protéines a été déterminée en solution aqueuse avec un appareil Lovibond PFX195 (Tintometer, Amesbury, Grande-Bretagne). Des solutions échantillons à 1 % (m/V) de lyophilisat ont été préparées dans un tampon phosphate 0,5 M à pH 7. La couleur des échantillons était mesurée selon l'échelle L.a.b, le terme L étant relatif à la quantité de noir / blanc, a de vert / rouge et b de bleu / jaune.

VI.3.2 Capacité et stabilité moussante

Le protocole a été adapté de Aluko & McIntosh (Aluko & McIntosh, 2001). Une solution de 100 mL à 1 % (m/V) a été préparée dans un tampon phosphate 0,5 M à pH 7 et à partir d'isolats lyophilisés. Un volume de 20 mL de solution a été placé dans un pot en plastique de 100 mL. La solution a ensuite été homogénéisée avec un Ultraturax à 10000 tpm pendant 5 minutes pour permettre le foisonnement de la mousse. La capacité moussante, CM, a été exprimée selon :

$$CM = \frac{\text{Volume de mousse}}{\text{Volume initial de liquide}} \times 100 \quad (\text{Eq. 25})$$

La stabilité de la mousse était calculée comme la proportion de mousse restante après 5, 15, 30, 60 et 120 minutes.

$$\text{Stabilité de la mousse } x \text{ minutes} = \frac{\text{Volume de mousse après } x \text{ minutes}}{\text{Volume initial de mousse}} \quad (\text{Eq. 26})$$

VI.3.3 Capacité et stabilité émulsifiante

Le protocole a été adapté de Naczki *et al.* (Naczki, Diosady, & Rubin, 1985). Une solution de 5 mL à 5 % (m/V) a été préparée à partir de protéines lyophilisées dans un tampon de phosphate 0,5 M à pH 7. La solution a été ensuite homogénéisée avec un Ultraturax à 10000 tpm pendant 30 secondes pour former l'émulsion. 2,5 mL d'huile de colza ont ensuite été ajoutés et le mélange a été émulsifié pendant 30 secondes avec un Ultraturax à 10000 tpm. 2,5 mL d'huile de colza ont été de nouveau ajoutés et homogénéisés pendant 90 secondes. L'émulsion formée a été transférée dans un tube conique de 15 mL et centrifugée pendant 5 minutes à 1100 g à température ambiante. La capacité émulsifiante, CE, a été exprimée telle que :

$$CE = \frac{\text{Volume d'émulsion après centrifugation}}{\text{Volume d'émulsion avant centrifugation}} \times 100 \quad (\text{Eq. 27})$$

La stabilité de l'émulsion, SE, a été évaluée en chauffant le tube conique de 15 mL pendant 15 minutes à 85 °C. Après refroidissement, une étape de centrifugation a été effectuée à température ambiante, pendant 5 minutes à 1100 g :

$$SE = \frac{\text{Volume d'émulsion après chauffage}}{\text{Volume d'émulsion avant chauffage}} \times 100 \quad (\text{Eq. 28})$$

INTRODUCTION DES RESULTATS ET DISCUSSION

Les procédés de production d'isolats de protéines végétales nécessitent deux étapes, l'extraction et la purification. La première étape est particulièrement complexe puisqu'elle nécessite de maîtriser un grand nombre de conditions opératoires (pH, ajout de sel, température, ratio solide/liquide, durée d'extraction, etc.) et que de nombreux critères de performance doivent être considérés (rendements d'extraction, propriétés du produit final en termes de solubilité, de couleur et de fonctionnalités). Les méthodologies mises en œuvre dans l'art antérieur n'ont, à ce jour, pas abouti à un procédé viable économiquement et industrialisé. Pour atteindre des rendements d'extraction élevés, de nombreux auteurs ont travaillé en conditions alcalines. Cependant, de telles conditions entraînent une complexation des polyphénols aux protéines. Il en résulte un produit coloré peu valorisable en alimentation humaine. Le premier verrou réside donc dans l'obtention de hauts rendements d'extraction tout en limitant la complexation phénolique.

A cause de la complexité du système et des démarches empiriques essentiellement expérimentales mises en œuvre, un nombre limité de critères de performance est pris en considération dans l'identification de conditions d'extraction. Ceci occulte probablement des éléments de compréhension des phénomènes en jeu. La principale difficulté d'utiliser un grand nombre de critères de performance est qu'ils sont susceptibles d'être antagonistes : par exemple, dans la littérature, les rendements les plus élevés génèrent aussi une couleur foncée du produit. Avec de nombreux critères, il se pose la problématique de rationaliser les choix. Ainsi, il est particulièrement difficile de choisir une condition opératoire permettant d'obtenir un ingrédient protéique de bonne qualité avec un rendement intéressant.

C'est pourquoi, une méthodologie alternative, basée sur une modélisation des systèmes d'extraction et des approches d'optimisation multicritères, a été mise en œuvre dans le cadre de cette thèse. Les différentes étapes de cette méthodologie sont détaillées en Figure 30. Cette démarche permet d'étudier l'effet de l'ensemble des conditions opératoires sur les critères ciblés dans un domaine de conditions défini.

1. La première étape est l'utilisation des plans d'expériences qui considère l'extraction selon un « système boîte noire ». Elle génère des modèles mathématiques entre les conditions opératoires et les performances de l'extraction.
2. L'étape suivante est l'optimisation multicritères du procédé. Elle repose sur l'utilisation d'algorithmes génético-évolutionnaires et elle permet de faire émerger un ensemble de solutions dites non-dominées : puisque les performances du procédé sont parfois antagonistes, des compromis sont à trouver pour satisfaire le cahier des charges. Ainsi, le front de Pareto illustre les meilleurs compromis entre les performances du procédé.

3. Ensuite, un outil d'aide à la décision est appliqué afin de choisir rationnellement une solution parmi ces compromis.

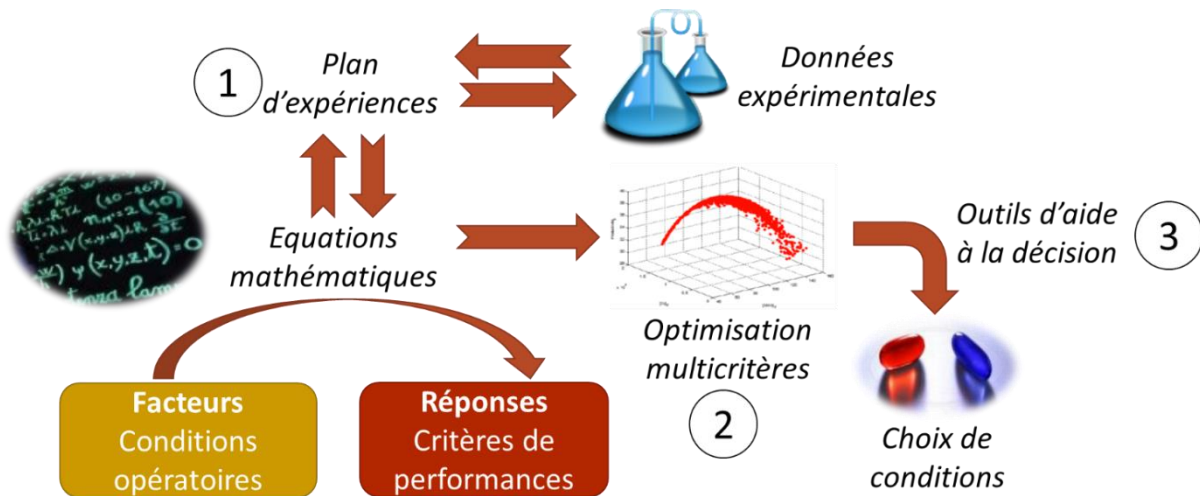


Figure 30 : Schéma de la démarche d'optimisation des procédés

Par ailleurs, les articles publiés à ce jour sur l'extraction des protéines de colza utilisent majoritairement des tourteaux industriellement déshuilés à l'hexane et / ou pressés à chaud (avec une température supérieure à 60 °C). Cependant, ce type de traitement implique généralement une exposition des tourteaux à des températures élevées ce qui détériore la qualité des protéines extraites (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018) et / ou limite considérablement l'extractibilité des protéines. Une des originalités de ce travail de thèse a été d'utiliser un tourteau gras pressé à froid et sans étape de déshuilage. Le parti pris a été de considérer que les 15 – 20 % d'huile restante dans le tourteau (qui constituent une perte de rendement en huile) seraient compensés par le gain en extractibilité et / ou qualité des protéines de l'isolat.

Après avoir évalué l'extractibilité des protéines de ce tourteau gras pressé à froid, la démarche décrite plus haut a été appliquée au cours des travaux de thèse dans des domaines de conditions opératoires différents. La Figure 31 montre ces deux voies vis-à-vis du cahier des charges initial.

Le premier domaine étudié (Chapitre III) correspond aux conditions expérimentales classiques de la littérature : pH d'extraction autour de la neutralité à faiblement basique avec des forces ioniques relativement faibles (0 à 0,2 M). Dans ce cas, deux possibilités sont envisageables : soit réaliser deux extractions successives afin de générer un unique isolat en accordant une attention particulière à ses propriétés fonctionnelles ; soit utiliser les deux extractions pour générer deux produits aux compositions albumines / globulines et donc aux propriétés fonctionnelles différentes. Le second domaine était les pH acides pour une extraction sélective des albumines (Chapitre IV). Pour cette étude, des critères liés à la composition du solide résiduel (après extraction) en protéines et en

antinutritionnel cible (acide phytique) ont été considérés. Ceci afin d'intégrer la problématique de valorisation de ce solide qui représente environ 50% de la matière mise en œuvre.

Les conditions d'extractions intéressantes identifiées grâce aux démarches de modélisation et d'optimisation ont ensuite été mises en œuvre suivies d'une purification des protéines afin de qualifier la qualité des isolats potentiels selon leur couleur, leurs propriétés structurales ou encore leurs propriétés fonctionnelles. Au vu des résultats, certaines conditions ont été extrapolées au niveau d'un démonstrateur par le groupe Avril.

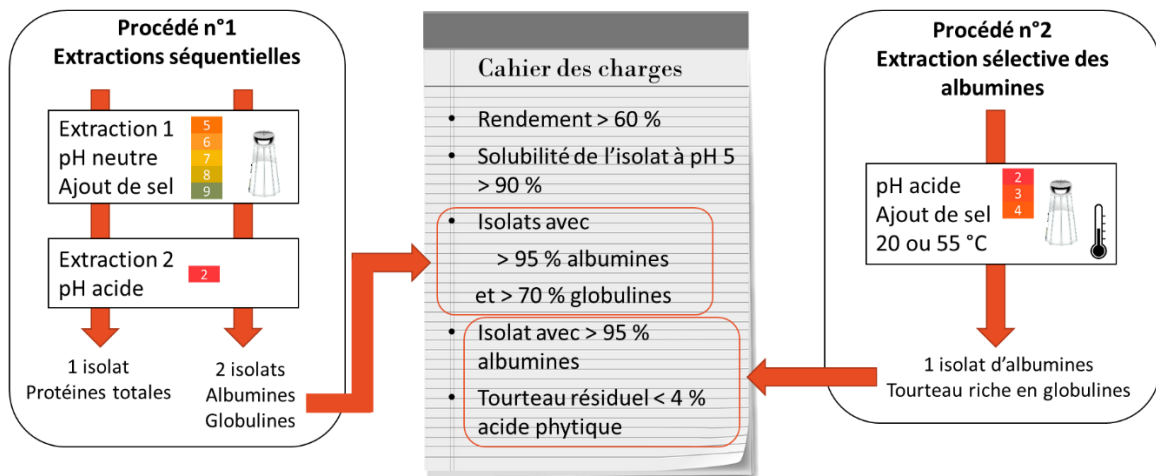


Figure 31 : Procédés mis en œuvre au cours de la thèse et cahier des charges associé

**CHAPITRE III : ETUDE D'UN
PROCEDE D'EXTRACTION
SEQUENTIEL DES
GLOBULINES ET
ALBUMINES DE COLZA**

I Introduction

Les procédés classiques d'extraction des protéines de colza font intervenir au moins une étape d'extraction à pH basique, afin d'atteindre des rendements d'extraction élevés, de l'ordre de 60 %. Cependant, l'extraction en milieu alcalin pose un problème quant à la couleur des protéines : les composés phénoliques, qui sont co-extraits et qui se complexent aux protéines, sont sujets à des phénomènes d'oxydation qui aboutissent à une coloration de l'extrait mais aussi des protéines purifiées. Le premier objectif de la thèse a donc consisté en la recherche de conditions optimales permettant à la fois d'atteindre de hauts rendements d'extraction tout en limitant la complexation des phénols aux protéines.

De plus, afin d'être utilisés dans de nombreuses applications, les isolats de protéines sont recherchés avec une solubilité élevée ainsi qu'une couleur claire. C'est pourquoi une condition optimale a été recherchée sur une gamme de pH autour de la neutralité (entre 5,5 et 8,5). L'influence de l'ajout de sel a également été étudiée. La Figure 32 résume les conditions opératoires étudiées ainsi que les critères de performances cibles. Dans le cadre d'un procédé séquentiels de deux extractions, les objectifs de chacune des extractions sont distincts : la première étant réalisée à pH neutre voire basique, des problématiques de couleurs et de complexation sont à prendre en compte ce qui n'est pas le cas pour la seconde. De plus, l'objectif est de générer un isolat protéique unique aux propriétés intéressantes, c'est pourquoi, des critères de performance sur le procédé dans sa globalité sont intégrés à l'étude : le rendement global, la solubilité et la couleur claire du produit. Ce travail est la première partie de ce chapitre, sous forme d'un manuscrit à soumettre au journal *Food Hydrocolloids*.

La seconde partie de ce travail a consisté à utiliser les données obtenues pour la mise en œuvre d'une seconde stratégie : utiliser chaque extraction du procédé séquentiel pour générer deux produits contenant soit une majorité de globulines soit d'albumines afin d'en comparer les propriétés fonctionnelles.

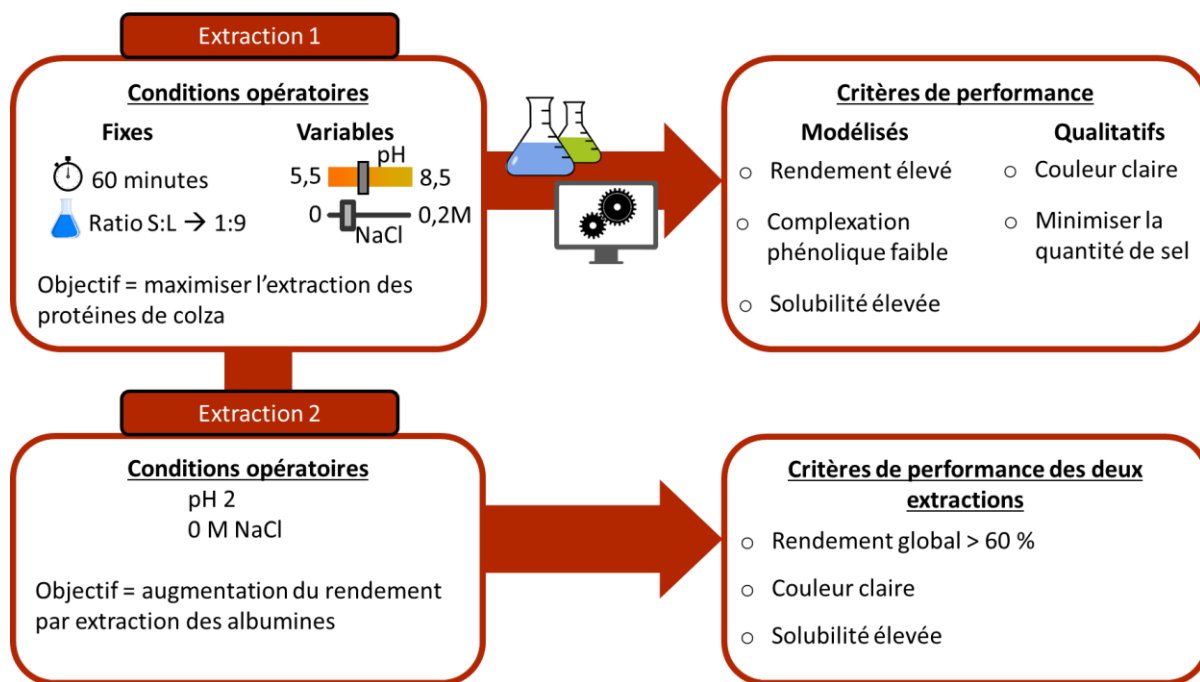


Figure 32 : Démarche du procédé séquentiel en deux étapes avec leurs conditions opératoires et les critères de performances cibles.

II Manuscript : impact of pH and NaCl on protein extractability, irreversible phenolic complexation and protein solubility from a cold-press rapeseed meal

II.1 Introduction

In 2050, nearly 10 billion people are expected on the planet and thus, the issue of the protein world demand is being raised. Sustainable solutions are investigated, and plant proteins constitute promising candidates. There is currently an increasing interest for rapeseed proteins because of their balanced amino acid profile and their richness in lysine and sulfuric residues, cysteine and methionine. Moreover, it is widely cultivated all over the world with more than 325000 km² of cultivated area in 2016 (FAOStats, 2019). Rapeseed bio-refinery produces oil for food and oleochemistry and meal for feed applications. The latter is rich in proteins and thus, a trending idea is to extract the proteins from the meal to use it for human consumption.

Two main storage proteins constitute the nitrogen stock of rapeseed: cruciferin is a globulin, 12 S protein, with a molecular weight of 300 to 350 kDa and an isoelectric point (pI) around 5 (Cheung, Wanasundara, & Nickerson, 2014), corresponding to more than 50% of the total proteins; napin is an albumin, 2 S protein, with 12 to 15 kDa and a pI close to 10-11 and they represent 30 to 40% of the total proteins (Lonnerdal & Janson, 1972). Globulin and albumin were shown to exhibit very distinct

functional properties. Globulin have a low solubility in unsalted water and are able to form stable emulsion and gels (Rehder, Sulewska, Markedal, Sørensen , & Sørensen, 2017) (Cheung, Wanasundara, & Nickerson, 2014) . Albumin was reported as soluble in a wide range of pH values and it was used to form foams and emulsions (Schwenke K. , Rapeseed proteins, 1994).

The classical approach to produce rapeseed protein isolate consists in two main steps, the extraction of proteins in an aqueous medium followed by a purification step. Early studies demonstrated that nitrogen was more extractible when pH was strongly alkaline (Radwan & Lu, 1976) (Tzeng, Diosady, & Rubin, 1988). In the latter study, extractions were carried out at pH 9 or 11 and nearly 30% of proteins were recovered into two fractions after an isoelectric precipitation. It was observed a dark coloration of the protein isolates. Hence, the high pH clearly impacts the quality of the final product: in alkaline medium, phenolic compounds are irreversibly oxidized (Sosulski, 1979) and their complexation with proteins results in a dark coloration of the isolate. To restrict coloration, calcium chloride can be added (Tzeng, Diosady, & Rubin, 1988). The solubility of the extracted and purified protein is also a crucial criterion since it conditions the rest of the functional properties (Li-Chan, 2004). However, there is a lack of knowledge regarding the impact of the operating conditions on rapeseed protein isolate solubility.

Recent studies report interesting way to improve protein extractability. For instance, the extraction assisted by enzymes, that alter the extracellular matrix, showed extraction yields of 74% (Rommi, et al., 2014). Others demonstrated the positive effect of ultrasound (Boukroufa, et al., 2017). However, these strategies are costly or require specialized equipment. Interestingly, extraction yield of proteins at pH 7 were improved by 30% with cold-pressed meal than toasted meal (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018). It was also shown that the addition of salt in the extraction medium was able to increase proteins extractability (Deak, Murphy, & Johnson, 2006). However, to our knowledge, there are no study that reports the impact of extraction conditions on protein extractability, irreversible phenolic complexation and the solubility of the extracted proteins.

In this article, the effect of extraction pH and the NaCl on protein extraction yield, phenolic complexation and protein solubility was studied by Design of Experiments (DoE). Beside extraction yield, the specific extraction yield of albumin and globulin was also considered for the first time to improve the knowledge of this complex extraction system. This work was achieved using a cold-pressed fat meal (without hexane de-oiling) to preserve protein quality. Then the polynomial equations from DoE were used for a multicriteria optimization of the extraction putting constraints on the criterion (extraction yield over 50%, protein solubility at pH 5 over 90% with the lowest irreversible phenolic complexation). The best trade-off extraction condition was applied following by a second

extraction aimed at increasing the overall extraction yield. Eventually, the solubility and the color of the isolate obtained after this original sequential extraction condition were characterized.

II.2 Material and methods

II.2.1 Material

Hydrochloric acid was prepared from a 37% stock solution from Carlo Erba (Val-de-Reuil, France). Sodium hydroxide and sodium chloride were provided by VWR (Radnor, Pennsylvania, USA).

Acetonitrile used for the Size-Exclusion High Performance Liquid Chromatographic (SE-HPLC) method was provided by Biosolve BV (Valkenswaard, the Netherlands)/ ultrapure water was generated by a system Purelab Ultra by Elga (High Wycombe, United Kingdom). Trifluoroacetic acid (TFA) was from ThermoFisher (Waltham, Massachusetts, USA). Ammonium sulfate was from Carlo Erba (Val-de-Reuil, France).

II.2.2 Rapeseed meal

The meal used was a cold-pressed dehulled rapeseed meal provided by Olead (Pessac, France). No hexane de-oiling was carried out. Its proximal composition was: 32% (w/w) (N x 5.35) proteins, 3.5% (w/w) total dietary fibers, 8.7% (w/w) of carbohydrates, 9.6% (w/w) water, < 1% (w/w) polyphenols and 15.4% (w/w) oil on a dry matter basis.

II.2.3 Protein extraction

II.2.3.1 Screening of extraction pH

Extractions were carried out by mixing 5 g of rapeseed meal with water at room temperature. The pH was adjusted to a value between 2 and 10 with HCl or NaOH 1 M. The mixture was agitated for 30 minutes. Then, it was centrifuged 30 minutes at 15,000 *g*. The supernatant was filtered through a filter paper and analyzed: by the Kjeldahl method for the protein content and by SDS-PAGE and SE-HPLC.

II.2.3.2 Extraction of rapeseed proteins

Extractions were carried out in a jacketed bioreactor of 200 mL allowing the regulation of the temperature at 20 °C. 15 g of rapeseed meal were poured in 120 mL water. NaCl was added at a concentration between 0 and 0.2 M. Then, pH was adjusted between 5.5 and 8.5 with HCl and NaOH 1 M. Then the final volume was adjusted with water to reach a solid:liquid ratio of 1:9 (w:w). After 60 min, the liquid extract was separated from the solid by centrifugation (15 minutes at 15000 x *g*). Then, the liquid was filtered through a cellulose filter paper from Fisher Scientific (Pittsburgh, Pennsylvania, USA).

II.2.3.3 Purification of the proteins

The first step of the purification consisted in a microfiltration with a plane Pellicon XL membrane (Millipore, Burlington, Massachusetts, USA) with 0.22 μm pore with a 55 cm^2 surface. The microfiltration system used, Cogent μScale tangential filtration system, was also from Millipore. For each extract, a Volume Reduction Factor (VRF) of 10 was applied.

Ammonium sulfate was added to the microfiltered extract to reach an ammonium sulfate saturation 70%. The mixture was agitated during 30 minutes at 25°C and centrifuged 30 minutes at 15000 g at the same temperature. The solid part was stored at 4 °C before use. Pellets were suspended in a Tris base (Fisher Scientific, Hampton, New Hampshire, USA) buffer pH 8 0.5 M with 0.5 M of NaCl. The volume added corresponded to the volume of the extract before precipitation with ammonium sulfate. Agitation was maintained during 30 minutes. Solubilized proteins were injected on a column packed with a gel filtration resin Sephadex G-25 (GE Healthcare, Chicago, Illinois, USA) with a bed volume of 864 cm^3 . The mobile phase was deionized water and a feed flow of 40 $\text{mL}\cdot\text{min}^{-1}$ was applied by a system Äkta Pure (GE Healthcare, Chicago, Illinois, USA). After equilibration, a volume of extract, equal to 20% of the column bed, was injected. The isocratic elution allowed the separation of proteins from salt. The protein fraction was recovered and freeze-dried on a system LTD type GZL-0.5 (Shanghai Pharmaceutical Machinery Co.).

II.2.4 Analytics

II.2.4.1 Protein quantification by Kjeldahl method

Samples were analyzed according to the official AOAC 979.09 Kjeldahl method. They were prepared with 4 mL of sulfuric acid (Sigma-Aldrich, Saint-Louis, USA) and a spatula of catalyst Cu-Se (AppliChem, Gatersleben, Germany). The mineralization step was carried out by heating the flasks at 450 °C for 90 minutes in a Büchi SpeedDigester K-439 (Rungis, France): The distillation step took place in a Kjelflex K-360 (Büchi, Rungis, France) with sodium hydroxide 32% (VWR, Radnor, Pennsylvania, USA) and boric acid (Sigma-Aldrich, Saint-Louis, USA) prepared in solution at 3%. The titration was carried out with 0.01 M HCl (Honeywell, Morris Plains, New Jersey, USA) in a Titrino Plus 877 from Metrohm (Herisau, Switzerland).

II.2.4.2 SDS-PAGE

SDS-PAGE reagents were from Bio-Rad (Hercules, California, USA). Samples were prepared according to Laemmli method (Laemmli, 1970) in reducing conditions. The migration occurred on a 5% acrylamide stacking gel and a 17% separation gel. The coloration was achieved by Coomassie blue G-250 (Sigma-Aldrich, Saint-Louis, USA). For each SDS-PAGE assay, a standard molecular weight kit was

used (Bio-Rad Precision Plus Protein Standard) allowing the determination of molecular weights between 10 and 250 kDa.

II.2.4.3 Quantification of globulin and albumin

Albumin proportion in liquid phases was assessed by Size-Exclusion High Performance Liquid Chromatography according to (Defaix, et al., 2019). The chromatographic system was a UFLC system from Shimadzu (Kyoto, Japan) and the column was a Biosep-SEC-s2000 300 x 7.8 mm, 5 µm, from Phenomenex (Torrance, California, USA). Eluent was 45% acetonitrile, 55% Ultra-Pure water and 0.1% of TFA (V/V). Elution was done at 0.6 mL / min flow during 30 minutes. Samples were filtered on 0.22 µm membranes from Pall (Port Washington, New York, USA) beforehand. The sample volume was 5 µL and the detection was set to 214 nm. Data were extracted and chromatograms were integrated manually using the spreadsheet software Excel from Microsoft (Redmond, Washington, USA).

II.2.4.4 Assessment of the irreversible phenolic complexation

The SE-HPLC method previously described was also used to assess the irreversible complexation of phenolic compounds to proteins. To do so, the UV-detection was set at both 280 and 325 nm. The phenolic complexation was expressed as the area under the protein peak, between 8 and 14 minutes, at 325 nm divided by the same at 280 nm.

$$\text{Phenolic complexation} = \frac{A_{325 \text{ nm}}}{A_{280 \text{ nm}}} * 100 \quad (\text{Eq. 21})$$

II.2.4.5 Solubility of the purified proteins

250 mg of proteins were suspended in 50 mL of deionized water. 10 mL of this solution was adjusted to pH 5 with HCl 0.1 M. The solution was centrifuged 30 minutes at 15,000 g and the supernatant was filtered on a 0.22 µm syringe filter from Pall (Port Washington, New York, USA). Both this solution and the initial solution (not filtered) were analyzed according to the Kjeldahl method and the solubility at pH 5 was expressed as:

$$\text{Solubility}_{\text{pH } 5} = \frac{\text{mass of proteins in the supernatant at pH } 5}{\text{mass of proteins in 10 mL of the initial solution}} \quad (\text{Eq. 24})$$

II.2.5 Design of Experiments methodology

Two operational conditions were studied: (i) pH with values between 5.5 and 8.5; (ii) the concentration of NaCl added to the extraction medium from 0 to 0.2 M of NaCl. The matrix was a Central Composite Face-centered (CCF) design generated thanks to the software Modde (Umetrics, Umeå, Sweden) and required 12 experiments including 4 repetitions of the center point. The choice to use Central Composite Face-centered (CCF) designs originates from the chosen domain of study of the operating conditions. The protein yield was determined with the Kjeldahl method (N x5.35). The albumin and

globulin yields were determined by the combination of the protein yields and by the globulin / albumin proportion determined by SE-HPLC. The theoretical albumin content in the meal was chosen as 40% of the total proteins in the meal. The solubility at pH 5 was determined as described earlier, with the Kjeldahl method ($N \times 5.35$). The phenolic complexation was determined by SE-HPLC.

II.3 Results and discussion

II.3.1 Protein extractability

The first step consisted in choosing the domain of variation of the factors under study. To do so, a preliminary experiment was carried out to screen extraction pH over a range between 2 and 10.

Figure 33 A shows an SDS-PAGE gel obtained after extraction of proteins at pH ranging from 2 to 10. SDS-PAGE bands at around 30 kDa and 20 kDa corresponded to the globulin α and β -chains. The two bands close to 4 and 9 kDa are albumin small and large subunit respectively. Between pH 2 and 4, a selective extraction of albumin occurred. This result was already observed with rapeseed meals from industrial processes de-oiled with hexane and desolvented at high temperature (Nioi, Kapel, Rondags, & Marc, 2012) (World Patent No. 144939 A1, 2008) The hypothesis for this selective albumin extraction was an irreversible precipitation of globulin under their isoelectric point ($pI = 5$) (Cheung, Wanasundara, & Nickerson, 2014). This was supported by globulin structural analysis at different pH and temperatures (Perera, McIntosh, & Wanasundara, 2016). The heat treatments of the industrial oil extraction process (during seed pressing and hexane removal) might also be responsible for this phenomenon by partly denaturing globulin inside the solid. Interestingly, albumin selective extraction at pH below 5 was also observed with hexane-deoiled cold-pressed and pre-pressed meals, (Fetzer, Herfellner, Stabler, Menner, & Eisner, 2018). In this case, the meal was not exposed to hexane and the pressing step occurred at less than 60°C. So, this selective extraction was only due to globulin denaturation at pH 5.

As shown on the SDS-PAGE, the globulin extraction started at pH 5. Obviously, the extracted proteins composition in globulin rose up to pH 8 and was stable at pH 9 and 10. Figure 33 B. shows the protein extraction yield in the whole pH range (from Kjeldahl analysis). At pH 2, the protein yield was 29%. It decreased slightly at pH 3 and 4 to 23 and 26% respectively. From pH 5 to pH 7 the yield was constant with 30 to 32% before rising from pH 8 up to 60% at pH 10. These results were in good accordance with data from cold-pressed meal (Fetzer, Herfellner, Stabler, Menner, & Eisner, 2018). Extraction yields were shown lower with toasted meals (heat desolventation). This confirms that heat treatments during oil extraction limit the overall protein extractability. In the range of pH corresponding to albumin selective extraction, the protein extraction yield was near 30%. Considering an albumin

proportion of around 35% in the meal (Schwenke K. , Rapeseed proteins, 1994) this would make a high specific albumin extraction yield (close to 85%). Hence, albumin would be far more extractable than globulin.

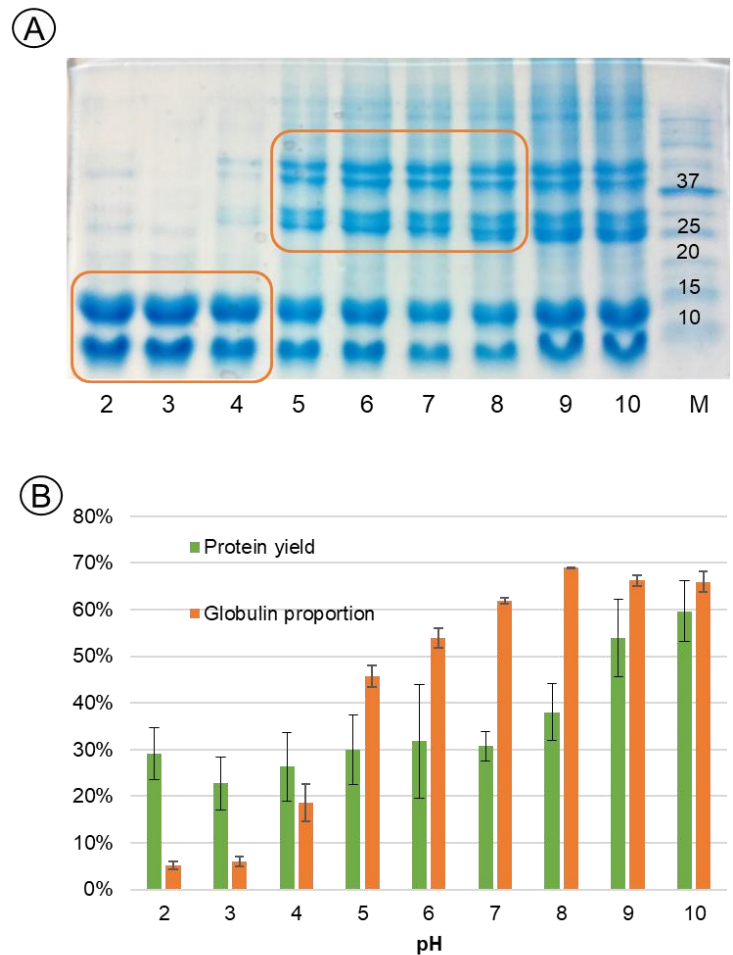


Figure 33: Preliminary experiments to assess the working domain of pH for the sequential extraction of globulin and albumin from rapeseed meal. A) Screening of pH on the extraction of rapeseed proteins by SDS-PAGE and, B) histogram of the protein yield (Kjeldahl N x5.35) and proportion of globulin (SE-HPLC) in rapeseed extracts.

To date, no study reports the specific albumin and globulin extraction yields. This information would be interesting to better understand rapeseed protein extractabilities. Furthermore, extraction yield is not the sole criteria to consider. The phenolic complexation and the solubility of the extracted protein after purification were also important criteria. Indeed, phenolic irreversible complexation has an impact on isolate color (Sosulski, 1979) and protein solubility is well known to condition their functional properties. The influence of the extraction conditions on these criteria was also not considered to our knowledge.

II.3.2 Impact of the pH and NaCl concentration on albumin, globulin and total protein yield

Two factors were identified as being particularly influential on the extractability of proteins: pH and NaCl addition. To better understand the extraction of rapeseed proteins, the overall yield was studied, along with the specific yields of globulin and albumin.

II.3.2.1 DoE matrix description and model validity

Tableau 6 shows the matrix used for the DoE. It corresponded to a Composite Faced-Centered Design, which means that the operating conditions remained inside the defined domains of each factor. Regarding the pH, it was chosen to carry out experiments between pH 5.5 and 8.5 to have high yields and moderate phenolic oxidation. The NaCl added to the medium was chosen between 0 and 0.2 M because of an industrial constraint: high concentrations of chloride would alter the steel tanks. A total of 12 experiments was carried out with 4 center points. The models were refined so that only significant factors were kept.

Tableau 6 shows the details of the models with the values of the coefficients and the p-value from the ANOVA analysis. The insignificant coefficients of the main factors were kept if quadratic terms were found significant.

All models were found with a correlation coefficient R^2 higher than 0.7 which corresponded to good models. The lack-of-fit test showed that, for all models, the mathematical equation fitted well the experimental data. The only exception was the solubility with a value of 4.3%.

Tableau 6 : Matrix for the design of experiments for the study of the sequential extraction of globulin and then albumin and related results.

N°	pH	NaCl content	Raw protein yield (%)	Globulin yield (%)	Albumin yield (%)	Phenolic complexation	Solubility of purified proteins at pH 5 (%)
1	5.5	0	25.99	19.84	35.22	10.13	92.72
2	8.5	0	58.91	61.38	55.19	17.82	79.99
3	5.5	0.2	50.35	17.26	99.98	12.65	98.29
4	8.5	0.2	48.82	33.55	71.74	21.76	91.45
5	5.5	0.1	36.31	16.29	66.33	7.87	103.57
6	8.5	0.1	63.02	52.94	78.15	14.57	93.86
7	7	0	31.00	27.49	36.27	8.92	88.21
8	7	0.2	48.48	20.94	89.79	14.65	98.13
<i>Center points</i>							
9	7	0.1	55.58	31.32	91.97	15.93	92.26
10	7	0.1	55.28	31.37	91.14	8.84	94.53
11	7	0.1	46.08	23.96	79.26	8.39	94.04
12	7	0.1	43.37	24.16	72.18	12.28	94.26
MODELLING RESULTS – ANOVA ANALYSIS							
R²			0.742	0.887	0.878	0.752	0.852
Q²			0.659	0.831	0.642	0.529	0.768
Reproducibility			0.686	0.862	0.793	0.339	0.970
ANOVA F-test p-value			0.010	<0.001	0.003	0.027	0.001
Lack of fit test			0.468	0.545	0.567	0.874	0.043*

Values followed by a * indicate a p-value superior to 10% for the ANOVA F-test or a value inferior to 5% for the lack of fit test.

Tableau 7 : Coefficients for the studied models

Coefficient	Raw protein yield (%)			Globulin yield (%)			Albumin yield (%)			Phenolic complexation			Solubility of purified proteins at pH 5		
	Coded	Uncoded	p	Coded	Uncoded	p	Coded	Uncoded	p	Coded	Uncoded	p	Coded	Uncoded	p
Constant	46.93	-43.74	10 ⁻⁹	1.44	0.51	10 ⁻¹⁴	79.83	-16.78	10 ⁻⁷	12.70	-0.86	10 ⁻⁸	95.43	109.68	10 ⁻¹³
pH	7.15	12.20	0.01	0.16	0.14	10 ⁻⁵	0.44	8.43	0.88*	2.16	1.95	10 ⁻³	-3.59	-3.24	10 ⁻³
NaCl content	3.91	454.83	0.09	-0.05	-0.74	0.023	16.60	1089.97	10 ⁻⁴	1.16	-17.67	0.06	3.32	124.33	10 ⁻³
pH*pH	/	/	/	/	/	/	/	/	/	/	/	/	/	/	/
NaCl*NaCl	/	/	/	/	/	/	-8.26	-1514.00	0.03	0.91	167.17	0.14*	-2.16	-396.67	0.03
pH*NaCl	-4.70	-57.42	0.03	/	/	/	-6.57	-80.35	0.03	/	/	/	/	/	/

The coded values correspond to the coefficients when the considered function is centered and scaled. The uncoded values correspond to the actual values to calculate the chosen response

Values followed by a * indicate a p-value superior to 10%.

11.3.2.2 Response surface analysis

Figure 34 A shows the contour plots of the raw protein yield (Kjeldahl N x5.35). The yields were observed between 26 and 63% with the highest yields at high pH or high content of salt in the medium. At low NaCl concentration, a strong positive effect of the pH on protein extraction was noticed. Indeed, the extraction yield went from around 25% to more than 60% at pH 5.5 and 8.5 respectively. This effect was consistent with what was observed in Figure 33 B. At low pH, a positive effect of the NaCl concentration was observed (the yield was increased by a factor 2 at pH 5.5 from 0 to 0.2 M NaCl). This effect of NaCl concentration tends to decrease while increasing the extraction pH. It even seemed to have a negative impact at pH 8.5. Eventually, at high NaCl concentration, the pH had a low impact on protein extraction (around 50% extraction yield whatever the pH). The positive effect of the NaCl on protein extraction around neutrality was already observed with sunflower meal (Pickardt, et al., 2009) and rapeseed meal (Gerzhova, Mondor, Benali, & Aider, 2016). These two authors attributed this to a salting in effect of NaCl on protein solubility.

In other words, NaCl would increase the protein affinity for the liquid phase. But it was hard to understand the reverse effect of NaCl at high pH. Furthermore, other phenomena were likely to occur like ionic interactions between protein and the solid phase.

To better understand this complex system of extraction, the effect of pH and NaCl on specific globulin and albumin extraction yield were also studied by DoE (Figure 34 B and C). Figure 34 B shows the contour plot of the globulin yield that was observed between 16 and 61%. The largest part of the area showed yields below 50%. Hence, globulin extractability was rather low in the domain. A positive effect of the pH, whatever the NaCl concentration was observed. A negative effect of NaCl concentration more important under basic conditions was also noticed. Hence, except for the positive effect of the pH at low NaCl concentration classically observed elsewhere, globulin extractability looked decorrelated with the total protein extractability pattern. The low extractability observed at pH 5.5 was explained by the closeness to their pI where they are in a precipitated state. The increase of pH also increased their net charge (it becomes negative) increasing their solubility in the aqueous phase. The negative effect of salts on globulin extractability was more surprising. Indeed, globulin are known to be highly soluble in salted solutions (Osborne, 1897). Furthermore, rapeseed globulin was reported as more soluble when salt was added (Cheung, Wanasundara, & Nickerson, 2014). This would suggest that compounds co-extracted in the aqueous phase act on globulin solubility. Phytate are known to be extracted from rapeseed meal at pH ranging from 4 to 10 (De Rham & Jost, 1979). They are also known to limit protein solubility. At pH above 6, globulin charge becomes negative, like phytate. But up to pH 9, lysine or arginine residues remain positively charged. So, it was possible that

salt decreased overall phytate – globulin repulsion favoring interactions with lysine and / or arginine that reduced their solubility.

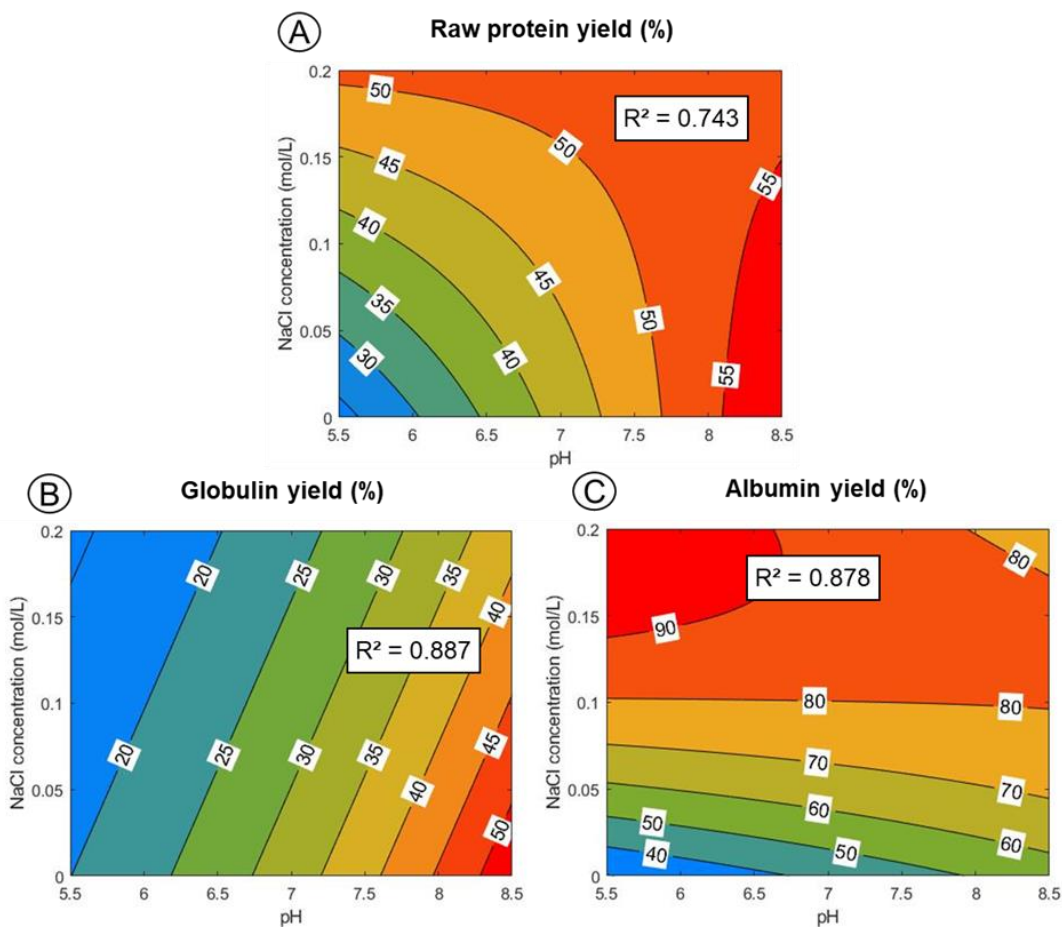


Figure 34 : Contour plots of the yields in raw protein, globulin, and albumin in the extracts.

Figure 34 C shows the contour plot for the albumin yield. Obviously, the pH had a positive effect on albumin extraction without NaCl (the albumin yield ranged from 35% to 55% at pH 5.5 and 8.5 respectively). This positive effect of the pH was suppressed with NaCl. The strong positive effect of NaCl was also remarked. It was most important at low pH (from less than 40% to over 90% at pH 5.5 from 0 to 0.2 M NaCl). Interestingly, a recent study demonstrated that phytate – lysozyme interactions around neutrality strongly reduced its solubility (Bye, Cowieson, Cowieson, Selle, & Falconer, 2013). This limitation of solubility was clearly eliminated by salt (mainly up to 0.2 M NaCl). Lysozyme and rapeseed albumin are very close in terms of molecular size and charge. So, a first hypothesis would be that ionic interactions between albumin (positively charged over the whole pH range) and phytate explains albumin extractability pattern. Besides, rapeseed meal is known to contain hemicellulose with glucuronic or galacturonic acid (Pustjens, Schols, Kabel, & Gruppen, 2013). The pKa of the carboxylic function is around 3.5. Hence this would expose negative charge at the solid surface in the

pH range of the study. Another explanation of the extractability pattern would be ionic bounds between albumin and the solid phase.

Hence, the impact of NaCl concentration on total protein extraction at low pH was mainly explained by the particular behavior of albumin. The similar extraction yield (around 55%) at 0.2 M NaCl whatever the pH and pH 8.5 whatever the NaCl concentration was due to compensating effect of the operating conditions on albumin and globulin extraction.

II.3.3 Impact of the pH and NaCl content on the phenolic complexation

Figure 35 A shows the contour plot of the irreversible phenolic complexation to proteins in the liquid phase. This criterion was taken into consideration because such complexation lead to undesirable brown color of protein isolates. The result was expressed in term of a protein peak UV signal ratio at 280 nm and 325 nm.

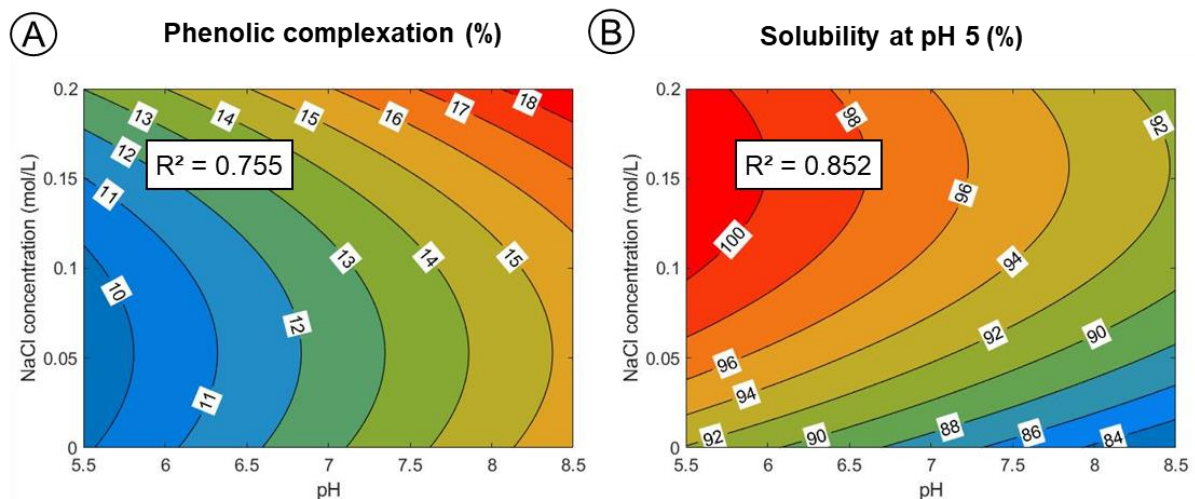


Figure 35 : Contour plots of the phenolic complexation in the extract and B) of the solubility of the purified proteins at pH 5.

Values were found between 10 and 18%. As expected, the pH was found influential: at high pH, there was a higher phenolic irreversible complexation. Indeed, at alkaline pH, phenolic compounds are more subjected to oxidation phenomenon and covalent bonding between phenolic compounds and proteins occur with the polyphenol oxidized form (Sosulski, 1979). The most visible result of this reaction was a change in the color of the solution. Surprisingly, the addition of salt in the extraction medium also leads to an increase in the phenolic complexation: at pH 5.5, it increased from 10 to 13% and at pH 8.5 it was between 16 and 18%. NaCl should not directly interfere with protein – oxidized polyphenol reaction. Hence, ionic interactions would probably be involved. The main rapeseed phenolic compound is sinapine, which results from the esterification of sinapic acid and choline. Hence sinapine has a positively-charged quaternary amine. The irreversible association between oxidized

polyphenols and protein occurs with positively charged lysine side chains (Naczki, Amarowicz, Sullivan, & Shahidi, 1998) (Fagerquist, Sultan, & Carter, 2012). So, electrostatic repulsions should partly protect the irreversible complexation. This protective effect could be possibly reduced by NaCl explaining the promoting effect on the complexation.

II.3.4 Impact of the pH and NaCl content on the solubility of purified proteins

Figure 35 B shows the contour plot of the solubility at pH 5 of the purified proteins. This particular pH was chosen because it corresponds to a large range of applications in the food industry. Overall, the solubility was high, between 84 and 100%. The increase of the pH allowed lower solubility of the proteins whereas the addition of salt improved the proportion of soluble proteins. The lowest solubility was found at pH 8.5 with no salt, whereas the highest solubility was at pH 5.5 with 0.2 M NaCl. The solubility seemed correlated to the proportion of albumin in the extracted proteins. Previous study showed that globulin have a limited solubility at pH 5 (Cheung, Wanasundara, & Nickerson, 2014). On the contrary, the same team showed that albumin was soluble over the pH between 3 and 7 (Cheung, Wanasundara, & Nickerson, 2015) So, albumin would promote globulin solubility around their isoelectric point.

II.3.5 Determination of optimal operating conditions

To be used in the food industry as an ingredient, the protein isolate should have a low phenolic complexation with a high solubility. Moreover, to be profitable, the extraction yields should be high. To find optimal operating conditions, constraints were established: more than 45% of protein yield, less than 13% phenolic complexation in the extract, and more than 95% of solubility of the purified proteins. The 13% limit for phenolic complexation was chosen because, beyond this value, the color of protein isolates was considered unsuitable.

Figure 36 shows the area of operating conditions that fulfilled all constraints. It was located between pH 5.5 and pH 7-7.2 and between 0.09 and 0.2 M of NaCl. The optimal condition was established at pH 6.8 and 0.1 M because this was the condition of the acceptable area having the less NaCl consumption. This choice could be discussed for being on a borderline of this area. It was subjectively as a prevision of a scaling-up of the process to industrial scale where the presence of chloride ions would alter the steel tanks.

In this condition, the protein extraction yield was predicted at 45.7%. The usual way to increase extraction yields is to carry out a second extraction step, in other operating conditions. Recently, Fetzer *et al.* described a multiple extraction step process to increase the extraction yield. In this case, the first extraction was carried out at pH 5.7 or 5.8 and then at pH 12 (Fetzer, Herfellner, Stabler,

Menner, & Eisner, 2018). From our results, such a basic condition would have two drawbacks. The first one is a strong phenolic complexation. The second one is that it would increase the protein composition in globulin that would be detrimental to the isolate solubility at pH 5. In the chosen operating conditions, around 30% of albumin should remain in the solid. Figure 34 showed that albumin extractability at pH 2 was around 100%. So, a second step of extraction at pH 2 would increase the overall extraction protein up to around 60% and it would even increase the isolate solubility at pH 5 and would not darken the color of the isolate.

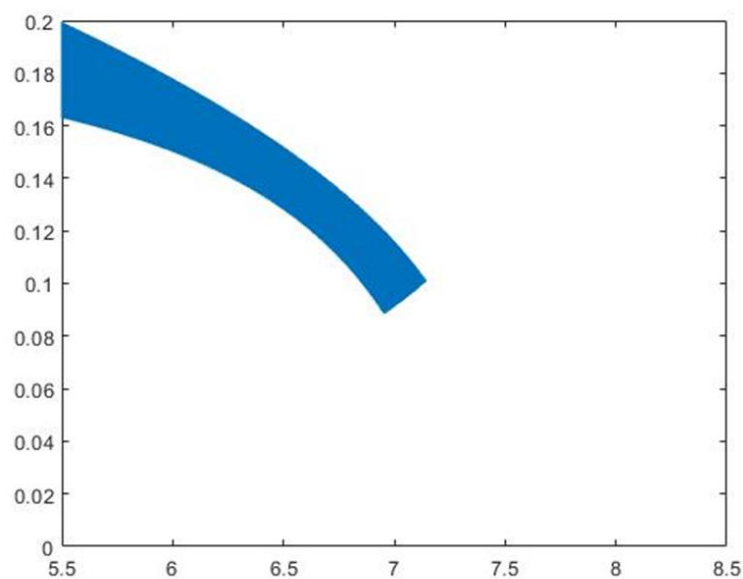


Figure 36 : Area of pH and quantity of NaCl where the constraints were all respected: overall yield higher than 45%, solubility at pH 5 higher than 95% and phenolic complexation lower than 13%.

II.3.6 Application of the two extraction steps process

This two-step sequence was implemented. It resulted in a yield of $47.3 \pm 2\%$ for the first extraction, and it was increased to $60.7 \pm 1.5\%$ thanks to the second step. This was close to what was obtained by combining a neutral extraction and a pH 12 extraction (Fetzer, Herfellner, Stähler, Menner, & Eisner, 2018). Then, the product was characterized to verify that it complies with the previous definition of the desirable product.

Figure 37 shows other results of the characterization of the final product. Figure 37 A shows a photograph of the powder that was light and creamish. Figure 37 B shows the solubility of the protein product over the pH range 3 to 9. It was found with more than 95% at pH 3, 5 and 9, and at 92% at pH 7. As a consequence, the pool of both extracts did not alter the solubility of the final product. Figure 37 C shows the chromatograms from the SE-HPLC method at 280 and 325 nm. It was used to determine a phenolic complexation of $10 \pm 1.2\%$ in the final protein product. The analysis at 280 nm

was also used to determine the composition of the final product: it was balanced in globulin and albumin with respective proportions of 46 and 54%.

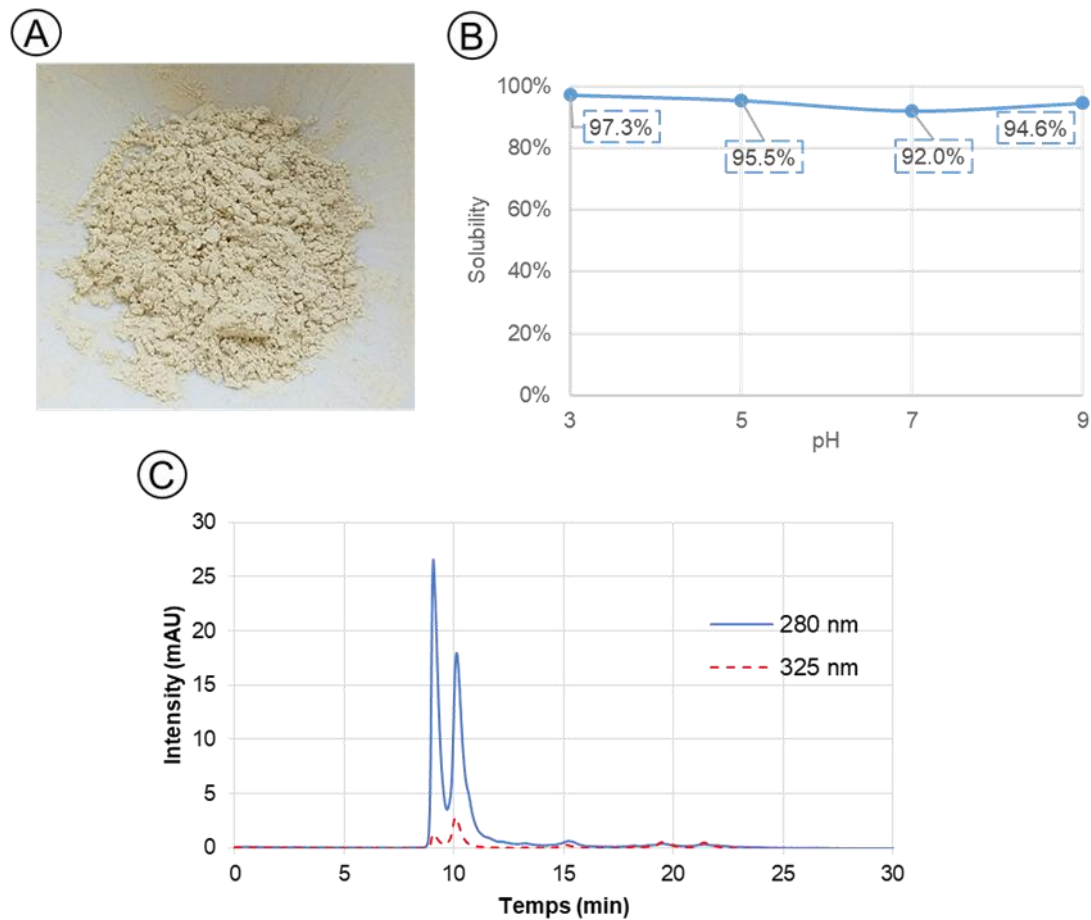


Figure 37 : Results of the two extraction steps process and characterization of the final product A) Solubility of the purified protein, B) photography of the light powder obtained after freeze-drying; C) chromatograms at 280 (blue) and 325 nm (red) by SE-HPLC.

II.3.7 Conclusion

The effect of the pH and NaCl on the protein extraction, irreversible phenol complexation and protein solubility at pH 5 was investigated with a cold-press rapeseed meal by DoE. The highest yields were observed at pH higher than 8 and with moderate use of salt: it corresponded to more than 40% of globulin extracted and up to 80% of the albumin in the meal. However, this area also corresponded to high phenolic complexation and to the lowest solubility. That is why multicriteria optimization by the use of constraints resulted in an optimal condition at pH 6.8 with 0.1 M of NaCl. Nevertheless, yields were found below 60%, which is a typical yield found in articles where extraction of rapeseed protein is carried out under alkaline conditions. To increase the overall yield, it was chosen to carry out a second extraction step at pH 2, where albumin yield was observed. This allowed to increase the yield to 60%. The final product was beige and soluble over the pH range of 3 to 9, which corresponded to a very promising product for the food industry.

Acknowledgement

Authors wish to thank the company Olead, Pessac, France, for providing the meals used in this study.

Authors disclose no potential conflict of interest.

II.4 Contribution de l'article

L'article détaille l'influence du pH et de l'ajout de sel sur l'extraction des protéines issues de tourteaux de colza. Un criblage de pH a mis en évidence les différences de rendements et de composition globuline / albumine des extraits. En particulier, la gamme de pH entre 5,5 et 8,5 permet d'extraire un mélange des deux protéines. L'ajout de sel a été étudié entre des concentrations de 0 à 0,2 M. La modélisation, par la méthodologie des plans d'expériences, a permis la génération de modèles sur les rendements en protéines brutes, en globulines et en albumines. De plus, des critères ayant une influence sur la qualité des protéines finales ont aussi été modélisés : la complexation phénolique et la solubilité des protéines purifiées à pH 5.

La définition de contraintes sur ces critères a permis de définir un ensemble de conditions opératoires générant un produit acceptable. Un ensemble de paramètres a été choisi pour minimiser la quantité de sel ajoutée : pH 6,8 avec 0,1 M de sel. Cette extraction permet d'atteindre un rendement de 47 % ce qui reste faible par rapport aux procédés classiques. L'ajout d'une étape d'extraction, à pH 2, permet d'extraire une fraction riche en albumine. Ainsi, en cumulant les deux étapes, le rendement d'extraction est de 60 %.

Les protéines ainsi obtenues sont solubles à plus de 90 % sur la gamme de pH entre 3 et 9. De plus, la couleur du produit final est beige clair, ce qui en fait un ingrédient potentiellement intéressant pour une utilisation en alimentation humaine.

III Un procédé séquentiel pour la production séparément des globulines et des albumines

Dans l'article qui précède, le criblage de pH d'extraction a mis en évidence deux zones permettant d'extraire sélectivement chacune des protéines majoritaires : entre les pH 2 et 4, seules les albumines sont extraites, alors qu'entre les pH 5,5 et 8,5, les globulines sont majoritaires. Ainsi, plutôt que d'utiliser la seconde extraction comme un moyen d'augmenter le rendement d'extraction global, le procédé de deux extractions séquentielles pourrait être utilisé afin de produire séparément des globulines et des albumines. L'intérêt d'un tel procédé réside dans l'exploitation des protéines en tant qu'ingrédient, car les albumines et globulines ont des propriétés fonctionnelles différentes. En effet,

les albumines sont très solubles dans l'eau et sont capables de former des mousses et des émulsions. Les globulines ont aussi des propriétés émulsifiantes mais aussi gélifiantes. Ces différences font que les protéines du colza pourraient être utilisées dans des applications alimentaires très diverses. Cette idée a déjà été émise par l'entreprise canadienne Burcon, avec ses deux produits : Puratein, riche en globulines et Supertein, en albumines. Ces deux fractions sont obtenues par une extraction à pH neutre avec du sel, puis les protéines sont purifiées et séparées via une micellisation à froid (Wanasundara J. , McIntosh, Perera, Withana-Gamage, & Mitra, 2016). Cependant, le début du présent chapitre met en avant une voie permettant d'éviter cette étape de séparation, en utilisant la deuxième étape pour extraire sélectivement les albumines.

Un procédé reposant sur la sélectivité d'extraction des protéines a déjà été mis en œuvre en 2008. Puisque les albumines sont extraites sélectivement entre les pH 2 et 4, la première extraction a été réalisée dans ces conditions. La seconde était à pH basique (World Patent No. 144939 A1, 2008). Cependant, il a été montré par la suite que le passage à pH acide des globulines entraîne leur altération probablement irréversible, ce qui limiterait leur extractibilité par la suite (Perera, McIntosh, & Wanasundara, 2016). C'est pourquoi la séquence inverse a été envisagée : commencer avec une extraction à pH neutre permettant d'extraire les globulines, suivie d'une extraction à pH acide pour extraire les albumines.

III.1 Détermination de conditions opératoires optimales

Les résultats de modélisation présentés précédemment (pH 5,5 – 8,5, NaCl 0 – 0,2 M) peuvent être utilisés avec un nouvel objectif : celui de maximiser la teneur en globulines extraites. La Figure 38 représente la surface de réponse pour la proportion en globulines après la première étape d'extraction. La proportion en globulines est comprise entre 30 et 65 %. Comme expliqué précédemment, cette répartition sur l'axe des pH provient du fait que le point isoélectrique des globulines soit d'environ 5. La proportion la plus élevée en globulines est observée à pH 8,5 sans ajout de sel. Cette zone correspond également à de hauts rendements en protéines totales, avec plus de 55 %, d'après la Figure 34.

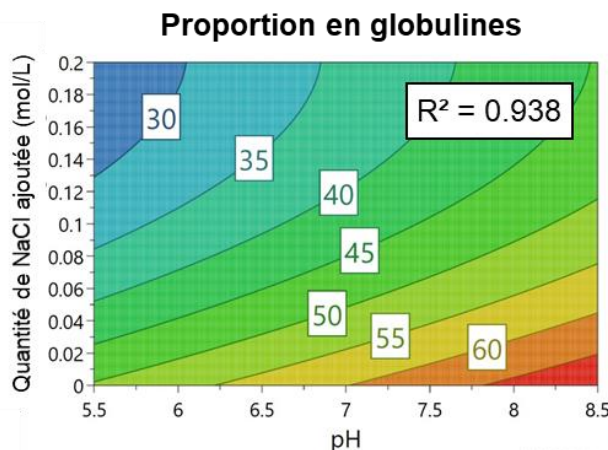


Figure 38 : Surface de réponse de la proportion en globulines en fonction du pH et de la quantité de NaCl ajoutée.

Comme précédemment, il est choisi de réaliser la seconde étape d'extraction à pH 2, sans ajout de sel.

III.2 Mise en œuvre du procédé

La Figure 39 est un schéma bilan de l'application du procédé. La première étape d'extraction à pH 8,5 a abouti à un rendement de 59 % de protéines dont 65 % de globulines. De par le rendement élevé, une importante quantité de protéines est produite. Ensuite, les protéines ont été purifiées par ultrafiltration avec 5 diavolumes en eau salée (0,5 M NaCl) et 5 diavolumes en eau. Le produit final a une pureté de 63 % (N x 5.35) et une couleur brune. Malgré la présence d'impuretés, la solubilité de ce concentrat a été évaluée comme le montre la Figure 40. La solubilité a été observée à plus de 85 % au pH 3 et 9 mais elle était de 59 et 65 % au pH 5 et 7.

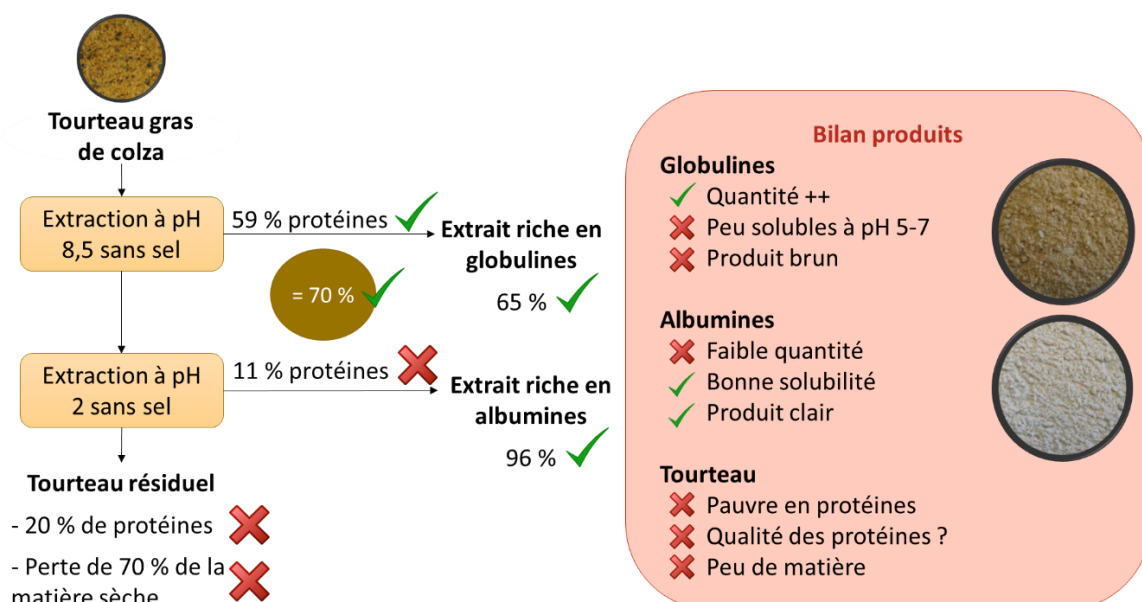


Figure 39 : Bilan des conditions optimales du procédé d'extraction séquentiel par rapport à la qualité des produits

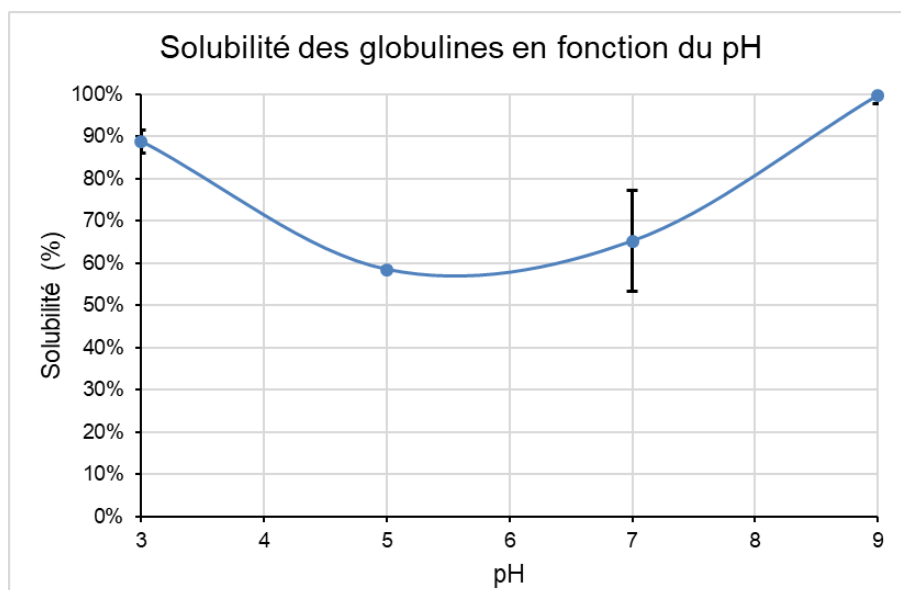


Figure 40 : Solubilité du concentrat de globulines en fonction du pH

Concernant les albumines, l'extraction à pH 2 a donné un rendement de 11 % en protéines, ce qui est relativement faible. Néanmoins l'extrait obtenu contenait 96 % d'albumines. Après purification, dans les mêmes conditions que les globulines, une pureté de 56 % a été obtenue (N x 5.35). La poudre obtenue était jaune pâle. Cependant, à cause du faible rendement, une faible quantité d'albumines a été obtenue, mais dont la solubilité était élevée sur l'ensemble de la gamme de pH. Ainsi, sur les deux étapes, un rendement global de 70 % est observé ce qui répond au cahier des charges d'un rendement supérieur à 60 %.

Concernant le tourteau résiduel, une faible quantité a été obtenue suite aux deux extractions avec une perte de 66 % de sa matière sèche. La teneur en protéines de ce tourteau est de 17,8 % ce qui correspond à une déplétion du titre en protéine de 80 %. Ce co-produit est donc généré en faible quantité avec une faible teneur en protéines dont la qualité a probablement été altérée par le passage à pH acide.

IV Conclusion du Chapitre III

La première voie de production de protéines de colza étudiée au cours de la thèse repose sur une séquence de deux extractions : la première à pH neutre et la seconde à pH acide. Deux scénarios ont été envisagés pour valoriser différemment ce concept innovant par rapport aux données de la littérature. Le premier a consisté à utiliser la deuxième extraction comme un moyen pour produire un unique produit protéique aux qualités désirables pour l'industrie alimentaire : une couleur claire, une solubilité élevée et ceci avec un rendement d'extraction de 60 %. Ce premier scénario est donc une solution intéressante. De plus, par rapport aux données de la littérature, de hauts rendements sont

observés de par l'utilisation d'un tourteau gras, c'est-à-dire qui n'a pas été exposé à de hautes températures.

Le second scénario a consisté à produire deux produits : l'un riche en globulines et l'autre, en albumines. Malgré un succès sur la possibilité de mettre en œuvre le procédé, cette voie de production a cependant montré des limites. Lors de la première extraction, un mélange constitué en majorité de globulines est obtenu avec un haut rendement d'extraction. Cependant, lors de cette étape des albumines sont aussi extraites. Ce qui fait que la seconde extraction, à pH acide et sélective des albumines, a un faible rendement. Après purification, les produits sont considérablement différents avec une grande quantité de globulines brunes et modérément solubles et une faible quantité d'albumines de couleur claire. De plus, le tourteau résiduel est largement appauvri en protéines et devient difficilement valorisable en alimentation animale.

Une extrapolation du procédé pourrait être envisageable concernant l'étape d'extraction. Cependant, la purification a été réalisée par précipitation des protéines avec du sulfate d'ammonium. Cette méthode est pratique à l'échelle laboratoire mais elle n'est pas pertinente pour une échelle industrielle puisqu'elle nécessite une grande quantité de sulfate d'ammonium. Et le surnageant, chargé de sels, représenterait un volume d'effluent important.

Il est à noter que lors de la seconde étape d'extraction, sur un tourteau déjà appauvri en protéines, un rendement de 10 % a été obtenu. Ceci est équivalent à des résultats de la littérature sur des tourteaux industriels (Nioi, Kapel, Rondags, & Marc, 2012) (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018). De plus, le tourteau récupéré en fin de ce procédé séquentiel est de moindre valeur. C'est pourquoi, un procédé alternatif a été étudié : l'extraction des protéines de colza à pH acide pour générer un isolat d'albumines et un tourteau résiduel, riche en globulines.

CHAPITRE IV.

MODELISATION DE L'EFFET

DES CONDITIONS

OPERATOIRES SUR

L'EXTRACTION DES

ALBUMINES

I Introduction

Le Chapitre III a mis en évidence des différences dans les rendements et la qualité des produits riches en chacune des deux fractions protéiques. Celui riche en globulines était abondant mais certaines caractéristiques étaient peu désirables pour des applications alimentaires, avec notamment une solubilité moindre par rapport à celle des albumines aux pH 5 et 7 et une couleur foncée du produit fini. Quant au produit riche en albumines, il était manufacturé en plus faibles quantités mais ses caractéristiques en font un excellent candidat pour des applications alimentaires variées avec une couleur claire et une solubilité élevée sur une large gamme de pH. Or, dans le chapitre précédent, il a également été démontré la possibilité d'extraire sélectivement les albumines du colza à pH acide.

Cette extraction sélective a été peu étudiée dans la littérature scientifique, si ce n'est au travers de trois études, dont un brevet (World Patent No. 144939 A1, 2008) (Nioi, Kapel, Rondags, & Marc, 2012) (Perera, McIntosh, & Wanasundara, 2016). Ces travaux, réalisés sur des tourteaux industriels déshuilés à l'hexane, ont montré que de faibles rendements étaient obtenus. De plus, en comparaison avec les données de la littérature, les résultats du Chapitre III ont confirmé le fort intérêt d'utiliser un tourteau de colza peu exposé aux traitements thermiques usuels liés à l'étape de déshuilage, afin d'obtenir de hauts rendements d'extraction. Ainsi, l'utilisation d'un tourteau gras de colza pour une extraction sélective des albumines permettrait de générer des protéines d'intérêt pour l'alimentation humaine et en plus grande quantité que précédemment.

Afin d'optimiser l'extraction des albumines, la démarche mise en œuvre se décompose en plusieurs étapes comme le montre la Figure 41. D'abord, trois facteurs ont été identifiés comme ayant un impact sur l'extraction : le pH, l'ajout de sel dans le milieu et la température. Ainsi, deux plans d'expériences identiques ont été réalisés mais à deux températures différentes, 20 et 55 °C, pour comparer les modèles obtenus. Des algorithmes génético-évolutionnaires ont été appliqués à ces modèles afin de tracer le front et le domaine de Pareto qui correspondent aux solutions non-dominées. Enfin, un outil d'aide à la décision a permis de classer l'ensemble des solutions non-dominées grâce à la modélisation des règles de classement du décideur.

Les résultats ont ainsi fait l'objet de deux manuscrits. Le premier concerne l'optimisation de l'extraction sélective des albumines à 20 °C et applique l'ensemble de la démarche d'optimisation. Le second compare l'extraction sélective à deux températures pour en étudier les effets sur le procédé en lui-même et aussi sur les protéines purifiées. L'objectif de ces travaux est la valorisation du tourteau de colza au travers de la génération d'un isolat d'albumines, à destination de l'industrie alimentaire, et d'un tourteau résiduel, pour l'alimentation animale. Ainsi, la première partie de ce Chapitre

présente un manuscrit soumis au Journal Food Chemistry et la seconde repose sur un article soumis au Journal *Process Biochemistry*.

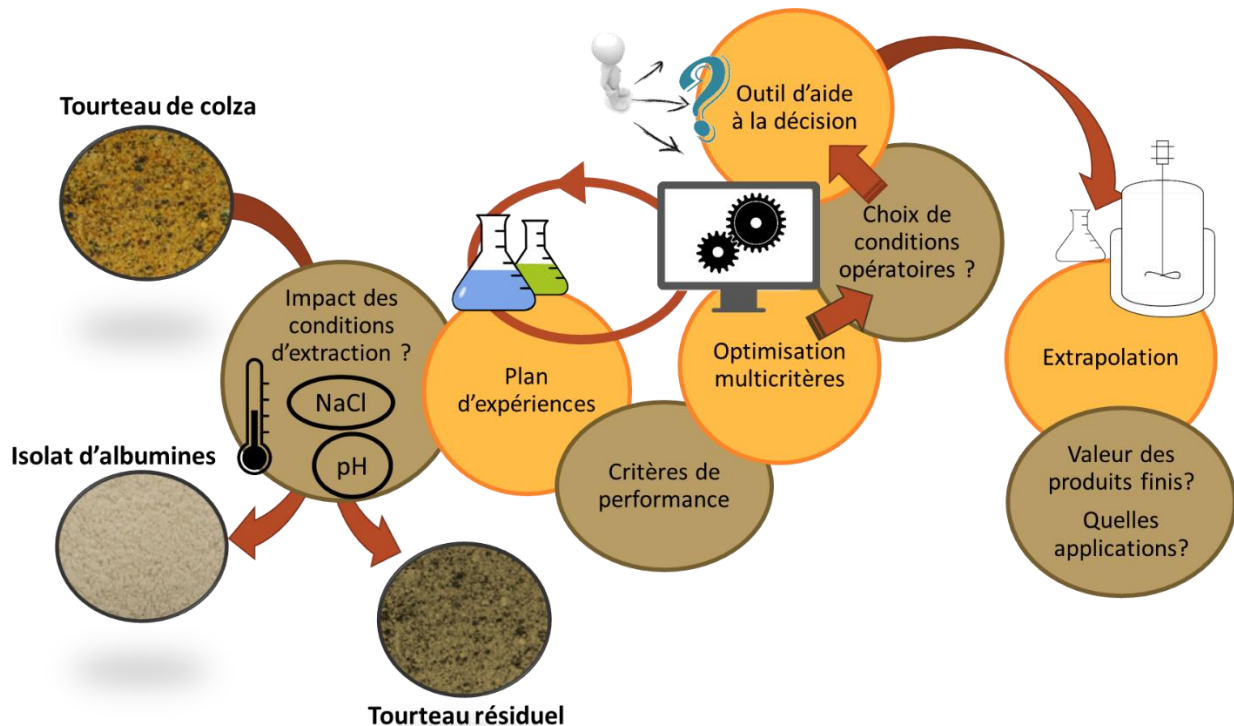


Figure 41 : Schéma de synthèse de l'étude de l'extraction sélective des albumines de colza

II Manuscript: Multicriteria optimization of albumin extraction from rapeseed meal for both the food and feed industry

II.1 Introduction

In 2016, 68 million tons of rapeseed were produced in the world (FAOStats, 2019). The solid residue generated after pressing the oil (rapeseed meal), is rich in fibers (28 to 32%) and proteins (30 to 40%). Rapeseed proteins have a balanced composition in amino acids with notable contents of lysine and sulfuric amino acids which were reported as deficient in other plant resources (Guéguen , Walrand, & Bourgeois, 2016). This makes the meal particularly suitable for feed application. To improve the rapeseed meal value in this field, the level of antinutritional molecules like glucosinolate were strongly reduced by varietal selection (Shahidi, Canola and Rapeseed, 1990). But the high content of phytates still causes limitation in the meal nutritional properties.

Beside interesting nutritional properties, rapeseed proteins have great functional properties (Aider & Barbana, 2011). Hence, many studies reported rapeseed isolate production processes for human

nutrition. Rapeseed proteins are mainly composed of two storage proteins. 50 to 60% are 11 S proteins, called globulin (cruciferin) with a high molecular weight of 300 to 350 kDa and an isoelectric point (pI) around 5 (Cheung, Wanasundara, & Nickerson, 2014). The second main protein fraction are 2 S proteins, called albumin (napin) in a proportion of 35% to 40%. Its molecular weight is 12 to 15 kDa and its pI is close to 10 - 11 (Lonnerdal & Janson, 1972). Isolate production processes are based on a protein extraction step in aqueous solutions and a purification step to reach a protein purity higher than 90% (on a dry basis). The extraction of rapeseed proteins is impacted by two main conditions: extraction pH and ionic strength. The solubility of nitrogen was reported as the highest when pH values were higher than 10 (Tzeng, Diosady, & Rubin, 1990). That is why many researchers worked under alkaline conditions. The addition of salt improves the protein solubility because of the ions forming a layer at the protein-water interface (Deak, Murphy, & Johnson, 2006). The purification processes mostly used are precipitation. Either by chilling liquid extracts with calcium chloride or by adjusting the pH around globulin isoelectric point (Chen & Rohani, 1992) (Akbari & Wu, 2015). These purification processes led to massive albumin losses (because they remain in the liquid phase) and the isolate produced often have an undesirable brownish color due to irreversible interaction with phenolic compounds, occurring under basic conditions (Sosulski, 1979). Furthermore, the solubility of proteins in water is poor at acidic or mild basic pH which limits its application field (Pedroche, et al., 2004). Eventually, the solid residue after protein extraction (around 50% of the initial meal dry matter) is very poor in nitrogen and consequently cannot be used for feed applications anymore.

Rapeseed albumin has a very good foaming capability and was reported soluble in water whatever the pH (Perera, McIntosh, & Wanasundara, 2016) (Akbari & Wu, 2015). Some studies reported the possibility to extract albumin selectively under acidic conditions (Nioi, Kapel, Rondags, & Marc, 2012). It was shown that the equilibrium was reached after 15 min. The pH yielding the better extraction yield was obtained at pH 2 with a ratio of solid/liquid of 1/9. The albumin from acid extraction (purified by size-exclusion chromatography) revealed to have a high denaturation temperature (102 °C). They also showed large amounts of α -helix (37%) in their secondary structure. This revealed particularly interesting since a high-quality protein isolate would be obtained with a remaining solid residue with high nitrogen content that might be still used for animal nutrition. Furthermore, the aqueous extraction step would probably lead to a reduction in phytate content, even increasing its potential value.

However, the observed extraction yield of albumin was rather poor (around 10%) (Nioi, Kapel, Rondags, & Marc, 2012). This low yield was probably linked to the production process of the meals. Like most of the studies regarding rapeseed meal protein extraction, the above-mentioned work was

carried out with a hexane-defatted meal. However, the meal defatting step by hexane after seed pressing implies a heat treatment for solvent removal. This most probably affected the protein extraction yield (Fetzer, Herfellner, Stähler, Menner, & Eisner, 2018). Furthermore, the effect of NaCl on albumin extraction yield and quality has never been reported to our knowledge.

The present work was based on the hypothesis that the use of a cold-pressed rapeseed meal that was not deoiled with hexane, would allow higher yield and high protein quality while allowing the recovery of a low phytate meal for feed applications. To do so, the effect of pH and NaCl concentration on albumin extraction yields and content in the liquid phase were investigated by design of experiments (DoE). The effect of these conditions on the content in proteins and phytic acid in the residual solid was also considered. The statistical analysis of the DoE was used to better understand the complex extraction system in the acidic pH range. DoE models were also integrated in an original multicriteria optimization methodology including a decision-making tool to select an optimal set of operating conditions. Finally, an albumin isolate obtained in the optimal extraction condition was produced for further characterization (color, structure, solubility and techno-functional properties).

II.2 Material and methods

II.2.1 Raw material

The rapeseed meal used was provided by Olead (Pessac, France): after pressing off the seeds, the meal was flaked and a step of dehulling was carried out. Partial characterization was determined as follow: 32% (w/w) (N x 5.35) proteins, 3.5% (w/w) total dietary fibers, 8.7% (w/w) of carbohydrates, 9.6% (w/w) water, < 1% (w/w) polyphenols and 15.4% (w/w) oil on a dry matter basis.

Chlorhydric acid was prepared from a 37% stock solution from Carlo Erba (Val-de-Reuil, France). Acetonitrile used for the Size-Exclusion High Performance Liquid Chromatographic (SE-HPLC) method was provided by Biosolve BV (Valkenswaard, the Netherlands)/ UltraPure Water was generated by a system Purelab Ultra by Elga (High Wycombe, United Kingdom). Trifluoroacetic acid (TFA) was from ThermoFisher (Waltham, Massachusetts, USA)

II.2.2 Protein extraction

Extractions were carried out in a jacketed bioreactor of 200 mL allowing the regulation of the temperature in the medium at 20 °C. 15 g of rapeseed meal were poured in 120 mL water. NaCl was added at a concentration between 0 and 0.2 M. Then, pH was adjusted between 2 and 4 with HCl 1 M. A final adjustment with water occurred to have a final ratio of solid:liquid of 1:9. After 60 min, the liquid extract was separated from the remaining solid by centrifugation (15 minutes at 15000 x g). Then, the liquid was filtered through a cellulose filter paper from Fisher Scientific (Pittsburgh,

Pennsylvania, USA). Then the liquid phase was clarified by microfiltration with a plane Pellicon XL membrane (Millipore, Burlington, Massachusetts, USA) with 0.22 μm pore with a 55 cm^2 surface. The microfiltration system used, Cogent μScale tangential filtration system, was also from Millipore. For each extract, a Volume Reduction Factor (VRF) of 10 was applied.

The microfiltered extract was then characterized for its composition in proteins, phytic acid, and proportion of albumin / globulin. The remaining meal was also characterized for proteins and phytic acid contents.

II.2.3 Analytics

II.2.3.1 Determination of protein content in solid and liquid phases

The analyses were achieved according to the official AOAC 979.09 Kjeldahl method. Samples were prepared with 4 mL of sulfuric acid (Sigma-Aldrich, Saint-Louis, USA) and a spatula of catalyst Cu-Se (AppliChem, Gatersleben, Germany). The mineralization step was carried out at 450 $^{\circ}\text{C}$ for 90 minutes in a Büchi SpeedDigester K-439 (Rungis, France). The distillation step (Kjelflex K-360, Büchi, Rungis, France) was done with sodium hydroxide 32% (VWR, Radnor, Pennsylvania, USA) and boric acid (Sigma-Aldrich, Saint-Louis, USA) at 3% w/V. The titration was done using a 0.01 M HCl (Honeywell, Morris Plains, New Jersey, USA) in a Titrino Plus 877 (Metrohm, Herisau, Switzerland).

II.2.3.2 Quantification of albumin proportion in liquid phases

Albumin proportion in liquid phases was assessed by Size-Exclusion High Performance Liquid Chromatography according to (Defaix, et al., 2019).

The chromatographic system was a UFLC system from Shimadzu (Kyoto, Japan) and the column was a Biosep-SEC-s2000 300 x 7.8 mm, 5 μm , from Phenomenex (Torrance, California, USA). Eluent was 45% acetonitrile, 55% Ultra-Pure water and 0.1% of TFA. Elution was done at 0.6 mL / min flow during 30 minutes. Samples were filtered on 0.22 μm membranes from Pall (Port Washington, New York, USA) beforehand. The sample volume was 5 μL and the detection was set to 214 nm. Data were extracted and chromatograms were integrated manually using the spreadsheet software Excel from Microsoft (Redmond, Washington, USA).

II.2.3.3 Determination of phytic acid content in solid phases

Phytic acid (PA) content was determined on solid samples with an adaptation of the method developed by Garcia Estepa *et al.*, (Garcia-Estepa, Guerra-Hernandez, & Garcia-Villanova, 1999). The first step of this method consisted in an extraction step at a pH close to 1 during at least 2 hours. After a centrifugation (15 minutes, 15 000 x g), the supernatant was recovered, filtered (0.22 μm membranes from Pall, (Port Washington, New York, USA)) and mixed with equal volume of the

following solutions: 20 mM iron chloride (Sigma-Aldrich, Saint-Louis, USA), 20% sulfosalicylic acid (Sigma-Aldrich, Saint-Louis, USA), and 0.4 M HCl + 10% Na₂SO₄ (Sigma-Aldrich, Saint-Louis, USA) solutions. Fe³⁺ and sulfosalicylic acid form a purple complex together that is also able to complex with PA. A thermal treatment at 100 °C during 20 minutes is carried out to precipitate the complexed PA. After cooling and centrifuging, the supernatant is recovered and the precipitate is washed three times with water to retrieve the free Fe³⁺ ions. Then, a titration was carried out with ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich, Saint-Louis, USA). The equivalent volume gave the quantity of free Fe³⁺ ions. Knowing the total Fe³⁺ content, it was possible to determine the quantity of Fe³⁺ involved in the precipitate with PA. In the literature, it is usually admitted that 6 phosphorus bind to 4 ions Fe³⁺. Moreover, it was supposed that one molecule of PA contains 6 phosphorus. Finally, the mass of PA was determined and the PA content was expressed relatively to dry matter.

II.2.3.4 Determination of the dry matter

For this last analysis, 1 mL of liquid sample was placed on an aluminum cup. The mass of the cup was recorded before placing the sample, after placing it and after 24 hours in an oven at 110 °C. The dry matter was expressed as:

$$\text{Dry matter (\%)} = \frac{\text{Mass of dry sample}}{\text{Mass of wet sample}} \quad (\text{Eq. 22})$$

II.2.4 Modelling and optimization

To study albumin extraction, a methodology using several mathematical tools was used.

II.2.4.1 Use of the design of experiments methodology

Two factors were studied: (i) pH with values between 2 and 4 to have a specific extraction of albumin; (ii) NaCl content from 0 to 0.2 M of NaCl. The matrix was a Central Composite Face-centered (CCF) design generated by the software Modde (Umetrics, Umeå, Sweden) and it required a minimum of 11 experiments including 3 repetitions of the center points.

The performance criteria were: (i) the albumin proportion; (ii) the protein yield before microfiltration; (iii) the protein yield after microfiltration; (iv) the albumin yield; (v) the phytic acid content in the residual meal; (vi) the protein content in the residual meal.

The albumin proportion was determined using the SE-HPLC method:

$$\text{Albumin proportion} = \frac{\text{Albumin content}}{\text{Total protein content}} \quad (\text{Eq. 29})$$

The protein yields before and after microfiltration were determined by converting the nitrogen content of the extract determined with the Kjeldahl method:

$$\text{Protein yield} = \frac{\text{Mass of nitrogen converted into protein in the extract (N x 5.35)}}{\text{Theoretical mass of protein in the meal (N x 5.35)}} \quad (\text{Eq. 30})$$

The albumin yield was calculated as follow:

$$\text{Albumin yield} = \frac{\text{Mass of albumins in the microfiltered extract (SE-HPLC method)}}{\text{Theoretical mass of albumins in the meal}} \quad (\text{Eq. 31})$$

The theoretical albumin content in the meal was chosen as 35% of the total proteins in the meal.

The remaining phytic acid (PA) in the meal was expressed as the mass of phytic in the dry mass of the meal minus the remaining salt:

$$\text{PA in the meal} = \frac{\text{Mass of phytic acid}}{\text{Dry mass of meal} - \text{mass of remaining salt}} \quad (\text{Eq. 32})$$

The protein content in the remaining meal was determined with the Kjeldahl method (N x 5.35).

II.2.4.2 Validity of the models

To interpret the obtained models, an analysis was carried out based on indicators determined by the software Modde (Umetrics, Umeå, Sweden). Five indicators were particularly important to assess the robustness of a model: (i) R^2 represented the correlation between modeled and experimental values; (ii) Q^2 was the ability of the model to predict more or less accurately new data, the difference with R^2 should be less than 0.3 to result in a good model; (iii) the Reproducibility represented the variation that may occur between two identical experiments; (iv) an ANOVA test was carried out to determine if the models were reliable and the result was represented by a p-value; (v) another ANOVA was used to statistically verify if the model lacks of fit.

II.2.4.3 Multicriteria optimization

Multicriteria optimization relied upon a Matlab (MathWorks, Natick, Massachusetts, USA) program developed internally in the laboratory. In this study, it was chosen to optimize three antagonist criteria. Polynomial models of each performance criteria were built using DoE software. A genetical-evolutionary algorithm, exploiting the Pareto's Domination Concept, was used to identify the Pareto's Front and Pareto's Zone of the problem, *i.e.* the set of all acceptable trade-offs of the performances. This algorithm was developed using the metaphor of Darwin's natural selection. In a large population, each set of operating conditions represented an *individual* which was characterized by its phenotype, depicted by two chromosomes (genotype) and a recessive/dominance index. Performances of each individual were assessed by a polynomial model and the multicriteria optimization index was defined as the number of Pareto's dominations undergone by each individual in the population. Cross-over and mutation operators on homozygote and heterozygote individuals

coupled with elitist selection were applied to lead the population, along generation, towards a non-dominated population.

The algorithm thus resulted in a set of non-dominated solutions which can be represented either in the operating conditions space (Pareto's Zone) or in the performance criteria space (Pareto's Front). The ultimate solution was then selected in the Pareto's Zone using Decision Engineering.

To select the most appropriate operating conditions according to the preference of the decision maker between criteria, a decision support tool was applied. First, it consisted in modeling the preferences of the decision maker and then, applying them on the whole Pareto's Front to rank the criteria set. In this approach, the first step consisted in randomly selecting a small (5 to 10) sample of significant performance set in the Pareto's Front and have them ranked by an expert of the process, called the decision-maker. The ranking was analyzed using the *Rough Set* approach: it is a mathematical tool that consists in finding relationships, that will be called rules, between the ranking of the decision-maker and the data. It is used to determine if some preference rules or non-preference rules can explain the proposed ranking (Pawlak, 1982). If they existed, these rules were then applied on the criteria set of the whole Pareto's Front. Each performance trade-off was thus ranked according to the number of times it would have been *preferred* or *non-preferred* to other trade-off of the Pareto's Front. The *best* criteria combination had to be selected in the first ranked Pareto's solution.

II.2.5 Purification of albumin

The whole extraction and purification steps were carried out on one set of conditions: at 20 °C, pH 2 and 0.11 M NaCl. 250 g of rapeseed meal was used to generate more than 10 g of protein powder.

After microfiltration, the obtained extract was treated on an ultrafiltration system Äkta Flux 6 (GE Healthcare, Chicago, Illinois, USA) with a 3 kDa and 4800 cm² hollow fibers cartridge (GE Healthcare, Chicago, Illinois, USA) at a 2 L/min feed flow and a transmembrane pressure of 1.5 bars at room temperature. Five diavolumes were carried out using salted water at a concentration of 0.5 M of NaCl and maintaining the pH with HCl 1M. Then, four diavolumes were achieved with ultrapure water only.

The final retentate was freeze-dried (Shanghai Pharmaceutical Machinery Co., LTD type GZL-0.5).

The final product, rich in albumin, had the following composition (relatively to the dry matter): 91.5% (w/w) of proteins, 5.1% (w/w) of water and 6.0% (w/w) of phytic acid.

II.2.6 Albumin isolate characterization

II.2.6.1 Purity and solubility

A 5 g.L⁻¹ mL solution was prepared from lyophilized proteins. The Kjeldahl analysis (N x 6.25) of this solution gave the purity of the lyophilized proteins:

$$\text{Protein purity} = \frac{\text{Mass of protein by Kjeldahl's method}}{\text{Mass of lyophilised proteins}} \times 100 \quad (\text{Eq. 8})$$

A 20 mL aliquot was used for each point of pH. The pH of the aliquot was adjusted to 3, 5, 7 or 9 with HCl or NaOH solutions at a concentration of 0.1 M. The added volumes were recorded. A centrifugation step took place for 15 minutes at 15000 x g at room temperature. The supernatant was recovered and filtered through a 0.22 µm filter and its protein content was measured with the Kjeldahl method. Each analysis was duplicated. Solubility was calculated as the ratio of the mass of protein in the aliquot (the change in volume was taken into account) on the mass of protein in the initial solution.

$$\text{Solubility}_{\text{pH } x} = \frac{\text{mass of protein in the aliquot}}{\text{mass of protein in the initial solution}} \quad (\text{Eq. 24})$$

II.2.6.2 Color analysis

The color of the protein isolate was determined in aqueous solution. The analysis in liquid was carried out on a Lovibond PFX195 (Tintometer, Amesbury, United Kingdom). Samples were prepared as 1% (w/w) solution of protein powder in phosphate buffer 0.5 M at pH 7. The system measured the color of the sample on the L.a.b scale. The term L related to the quantity of black / white, a to green / red and b to blue / yellow.

II.2.6.3 Circular Dichroism (CD)

The secondary structure of albumin was analyzed using far-UV between 180 and 260 nm. Albumin solutions at 1 mg.mL⁻¹ were prepared in a 0.5 M phosphate buffer at pH 7. The spectral analyses were carried out at room temperature with a Chirascan plus from Applied Photophysics (Leatherhead, United Kingdom). The samples were prepared in cells of 0.1 mm path length. For each sample, the analysis was carried out three times, and an average was traced. The software Pro-Data Chirascan from Applied Photophysics (Leatherhead, United Kingdom) was used to process the data.

The same system was used to assess tertiary structures using near-UV spectra between 250 and 320 nm. A quartz cell of 10mm path length was used.

II.2.6.4 Dynamic Light Scattering (DLS)

DLS was used in order to determine the aggregation state of albumin under several conditions of pH. Protein solutions at 1 mg.mL⁻¹ were prepared in a 0.5 M phosphate buffer at pH 7. Quartz cells of 10

mm path length were used. The analyses were carried out with a Zetasizer Nano – S from Malvern Instruments (Malvern, United Kingdom). Each analysis was carried out in a 3 mm cell, repeated three times and an average trace was determined.

II.3 Results and discussion

II.3.1 Effect of pH and NaCl on albumin extractions and remaining meal composition

The effect of extraction pH and NaCl concentration on albumin yield and composition after extraction and clarification by microfiltration was studied by Design of Experiments (DoE). The influence of protein extraction conditions on the residual meal composition in terms of protein and phytic acid was also considered. Contrary to most of rapeseed meal extraction studies, the starting meal used was a pressed meal. Hence, no hexane deoiling and desolventing steps were performed after seeds pressing. Tableau 8 shows the matrix of experiments and the results obtained for each experiment. The pH range used (pH 2 to 4) was selected because it allows a selective extraction of albumin (Nioi, Kapel, Rondags, & Marc, 2012). To our knowledge, the effect of NaCl on albumin extraction in this pH range has never been studied. We chose to limit our study to 0.2 M of NaCl due to corrosive effect of salts on steel under acidic conditions. The other operating conditions (solid:liquid and extraction duration) were fixed at 1:9 (w/w) and 60 min respectively also on the basis of Nioi *et al* (Nioi, Kapel, Rondags, & Marc, 2012).

Tableau 8 : Matrix for the design of experiments at 20 °C for the study of albumin extraction and ANOVA analysis of the main indicators used to validate the models.

Experiment number	pH	NaCl content	Raw protein yield (%)	<u>Albumin proportion (%)</u>	Protein losses during MF (%)	<u>Albumin yield (%)</u>	<u>Phytic acid in the residual meal (%) DM)</u>	Protein in the residual meal (%)
1	2	0	36.43	98.03	34.25	57.53	3.88	30.83
2	4	0	26.30	91.58	25.62	40.08	3.84	32.11
3	2	0.2	37.08	98.02	33.19	54.83	3.25	30.83
4	4	0.2	41.41	86.30	31.68	61.37	2.54	27.90
5	2	0.1	32.94	97.73	34.68	53.27	3.73	30.95
6	4	0.1	33.42	90.41	28.44	51.31	2.78	29.98
7	3	0	30.79	97.18	25.43	51.23	4.51	32.85
8	3	0.2	32.18	96.67	29.47	53.89	4.63	31.17
Center points								
9	3	0.1	31.11	96.26	23.01	55.61	4.61	30.28
10	3	0.1	34.58	96.09	28.79	57.40	4.00	30.53
11	3	0.1	31.37	95.56	27.97	53.63	4.87	31.66
MODELLING RESULTS – ANOVA ANALYSIS								
R2			0.794	0.940	0.851	0.847	0.824	0.825
Q2			0.218	0.896	0.675	0.563	0.674	0.588
Reproducibility			0.761	0.982	0.335	0.852	0.667	0.667
ANOVA F-test p-value (if < 5%, the model is statistically good)			0.030	<0.001	0.012	0.003	0.005	0.019
Lack of fit test (if > 5% the model has no lack of fit)			0.410	0.200	0.988	0.418	0.695	0.616

The underlined responses correspond to the selected criteria for the multicriteria optimization.

II.3.1.1 Validity of the models

Tableau 9 shows significant model coefficients from the polynomial regression achieved by Modde software. The significance of the terms was assessed with a 90% confidence interval and non-significant terms were removed to improve the reliability of the models. Furthermore, two responses were transformed in negative logarithm: the proportion of albumin and the albumin yield. Doing so, non-linear regression R^2 were higher than 0.8. Hence, most of the observed variations can be explained by the model. Moreover, the ANOVA test to assess the lack of fit always resulted in p-values higher than 0.1. As a consequence, all the models can be reliably used to interpret the obtained data.

Tableau 9 : Coefficients for the performance models and their p-values

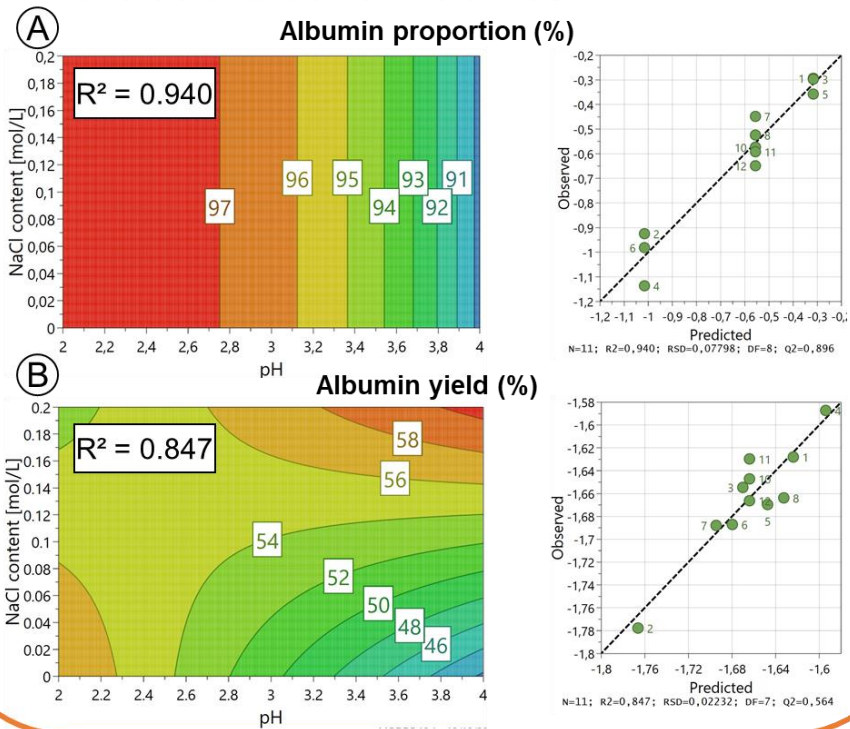
Transformation	<u>Albumin proportion (%)</u>		<u>Albumin yield (%)</u>		<u>Phytic acid in the residual meal (% DM)</u>	
	-10log(100-Y)	-10log(100-Y)	-10log(100-Y)	-10log(100-Y)	Value	p
Coefficient	Value	p	Value	p	Value	p
Constant	-0.48	10^{-7}	-1.48	10^{-15}	-5.04	10^{-8}
pH	0.30	10^{-6}	-0.07	0.11	6.87	0.11
NaCl content	/	/	-1.32	0.01	-2.99	0.10
pH*pH	-0.11	0.05	/	/	-1.19	10^{-3}
NaCl * NaCl	/	/	/	/	/	/
pH* NaCl	/	/	0.54	10^{-3}	/	/

The underlined responses correspond to the selected criteria for the multicriteria optimization. The confidence level for the significance of the coefficients was 90%.

II.3.1.2 Effect of pH and NaCl on albumin composition and yield in the extract

Figure 42 A shows the modeled albumin composition in the microfiltered extract as a function of the pH and NaCl concentration of the extraction step. Whatever the extraction conditions applied albumin represented more than 88% of the soluble proteins. The maximum was 98% at pH 2. SDS-PAGE analysis prior clarification of the extract clearly showed a large proportion of albumin (data not shown). This indicated that albumin selective extraction observed by others with hexane defatted meal also occurred with a pressed meal (Nioi, Kapel, Rondags, & Marc, 2012). Globulin is known to precipitate irreversibly in aqueous extracts at acidic pH while albumin remain soluble. This difference of stability was hypothesized to be the cause of the selective albumin extraction (Perera, McIntosh, & Wanasundara, 2016). Obviously, this hypothesis still applied in this study. As a consequence, the poor stability of globulin under acidic conditions was not a result of the steps of deoiling and desolvanting of the meal but it seemed like an inherent property of rapeseed globulin.

Characterization of the extract



Characterization of the remaining meal

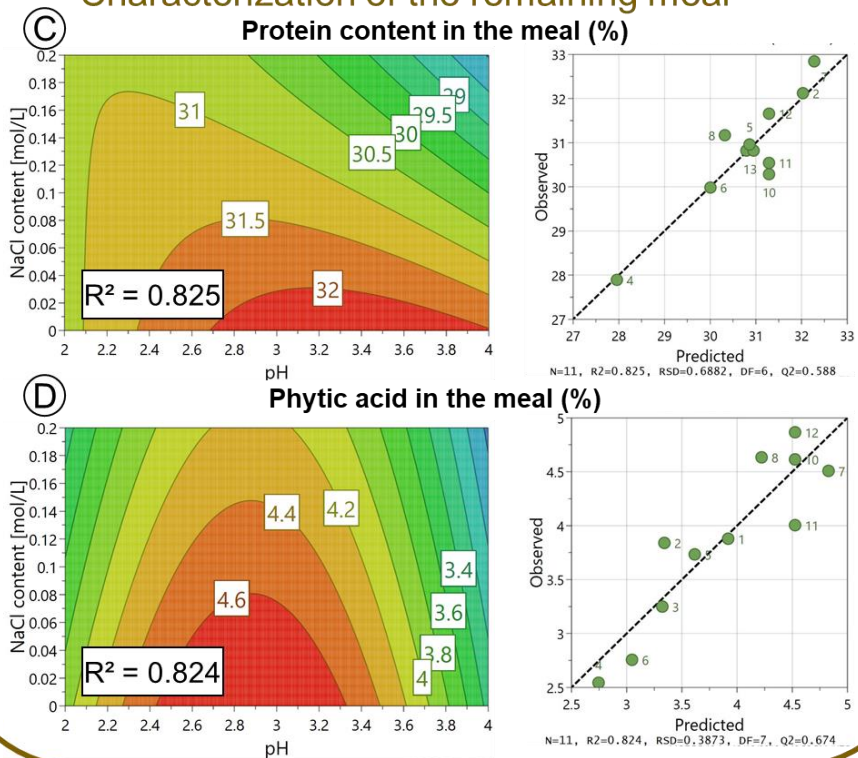


Figure 42 : Response Surface of the predicted performances and plots of the observed values as functions of predicted values for the performance criteria regarding a) and b) the extraction of albumin and c) and d) the remaining meal.

Figure 42 B shows the response surface of albumin yields as a function of the extraction operating conditions (still after microfiltration). This criterion ranged from 44% to 59%. It is interesting to note that the observed yields with the pressed meal were much higher than the 10% albumin yield reported at pH 2 with an hexane defatted-meal (Nioi, Kapel, Rondags, & Marc, 2012). Three zones were clearly distinct. From pH 2 to 2.5, the albumin yield was high (> 55%) and almost independent of NaCl. For pH above 2.5 at NaCl concentration below 0.1M, the yield was very sensitive to both pH and NaCl. In the same range of pH at higher NaCl concentration, albumin yield was poorly dependent of the pH but sensitive to NaCl. The maximum yield was observed in this area, at pH 4, 0.2 M NaCl with a value of 61% (experimental) or 59% (predicted).

Rapeseed albumin is known to be particularly soluble over a large pH range (pH 2 to 11). Hence, what was observed at low ionic strength was a limitation of albumin extraction between pH 2.5 to 4. This limitation should imply ionic interactions. Over this pH range, albumin have a largely positive net charge because of their basic isoelectric point (pH 10 – 11). Hence, an element from the solid phase having a negative function between pH 2.5 and 4 might limit albumin extraction. Rapeseed meal hemicellulose is poorly soluble in water and contains uronic acids with pKa around 3.2-3.5 (Pustjens, Schols, Kabel, & Gruppen, 2013) (Bhattacharyya & Rohrer, 2012). Considering the charge of uronic acids as a function of the pH, it can be hypothesized that ionic interactions between albumin and hemicellulose explain the limitation of albumin extraction at pH from 4 to around 3. The increase in albumin extraction observed in this pH range with NaCl concentration probably suppressed this ionic bonding. At pH around 2, uronic acids should be predominantly in an uncharged form. Hence no ionic interaction would limit the extraction, explaining the high extraction yield and its independence toward NaCl.

Interestingly, the maximum loss of proteins (around 33%) during the microfiltration (MF) was also observed with extraction pH below 2.5, whatever the NaCl content,

Figure 43. At pH 3-4 and NaCl concentration of 0.2 M the protein loss was less than 28%. Hence, the slight decrease in albumin yield observed in the pH 2 area would be due to a protein loss during the MF. At pH around 3, it was reported that rapeseed albumin has a considerably higher hydrophobic surface than at higher pH (Perera, McIntosh, & Wanasundara, 2016). Such observations were also made with other low molar weight globular proteins like α -lactalbumin increasing protein oil binding capacity (Matsumura, Mitsui, Dickinson, & Mori, 1994). The meal used in our study still contained oil (15.2% of the dry matter). So, the increase of protein loss in MF at pH 2 might be due to the formation of albumin – oil aggregates.

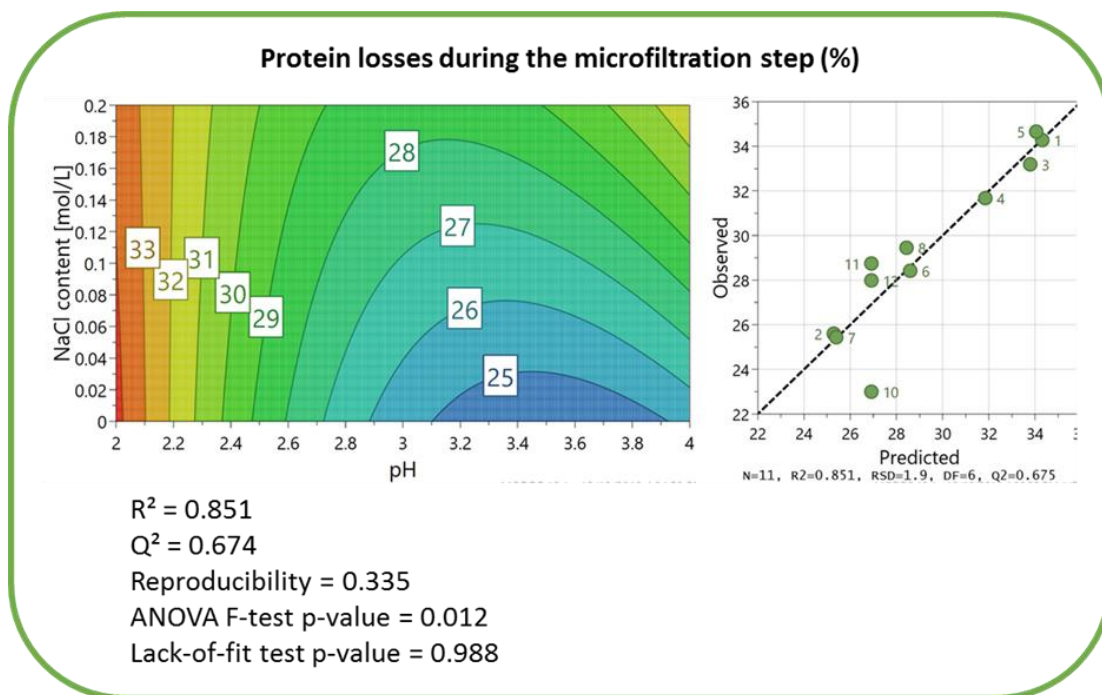


Figure 43 : Response Surface of the predicted performance and plot of the observed values as a function of predicted values for the protein losses during the microfiltration step (determined by the Kjeldahl method).

II.3.1.3 Effect of pH and NaCl on the residual meal composition

Protein and phytic acid content are key parameters for assessing the value of a plant meal in animal nutrition. Tableau 8 shows that protein content remained high whatever the extraction conditions (28 - 32% w/w on a dry matter basis). As expected from albumin yields, Figure 42 C shows that the protein content was lowest at pH 2 and pH 3-4 with a high NaCl concentration. The slight discrepancies in trends were probably due to the fact that the extraction of the total dry matter as a function of the operating conditions may not perfectly follow the protein extraction (Gerzhova, Mondor, Benali, & Aider, 2016). In any case, the protein content in the residual meal remained similar to the protein content in the initial meal. Hence, this meal was still of interest for the feed industry as for the protein content.

Figure 42 D shows that phytic acid (PA) content in the meal increases from pH 4 to 3.4 – 2.5 and decreases toward more acidic conditions. The lowest value observed was 2.8% when pH was 4 with 0.1 M NaCl. Globally, the NaCl content had a poor impact on the PA content in the meal except in the pH area around pH 3 where a high PA content was observed. The lowest phytate solubility in water is known to be around pH 4 (De Rham & Jost, 1979). This probably explains the increase in PA content from pH 4 to 3. Surprisingly, the PA content in the residual meal diminished at pH below 2.5 and at around 3 with high NaCl concentrations. This clearly indicated a promotion of PA extraction under these conditions. This operating condition area also corresponded to a high albumin extraction. Interestingly, PA interacts strongly with proteins by electrostatic interactions (Ali, Ippersiel, Lamarche,

& Mondor, 2010) (Kaspchak, Igarashi Mafra, & Rogério Mafra, 2018). In this pH range, PA phosphoric groups are largely negative (Martin & Evans, 1986) so the decrease of PA content in the meal might be explained by a PA - albumin co-extraction.

II.3.2 Multicriteria optimization

Three criteria were chosen for multicriteria optimization: the albumin proportion, the albumin yield and the PA content in the residual meal. Such an optimization is complicated since these criteria are antagonists. To do so, an original methodology based on our multicriteria genetical-evolutionary algorithm was used (Muniglia, Kiss, Fonteix, & Marc, 2004).

II.3.2.1 Determination of the Pareto fronts and domains

The first step of the methodology was to define Pareto's Front and Pareto's domain which are depicted in Figure 44. Each point on the graphs corresponds to a non-dominated solution, *i.e.* an acceptable trade-off of performances: where increasing one criterion in the set would automatically alter another performance criterion. Two zones can be distinguished: the first was located at pH 2 with salt from 0 to 0.2 M. It corresponded to moderate protein and albumin yields with a high proportion of albumin in the microfiltered extract. The second zone was located between pH 3.3 and pH 4 with 0.2 M NaCl. In that area, the yields were close to maximum values but with a lesser proportion of albumin. Hence, the next step consisted in choosing among the non-dominated solutions, each one representing the best acceptable trade-off. The dots in the Pareto front and Pareto domain on Figure 44 are coloured according to the ranking of the non-dominated solutions after using the decision-making tool. Hence, the population was divided into groups to highlight the ranks of the solutions.

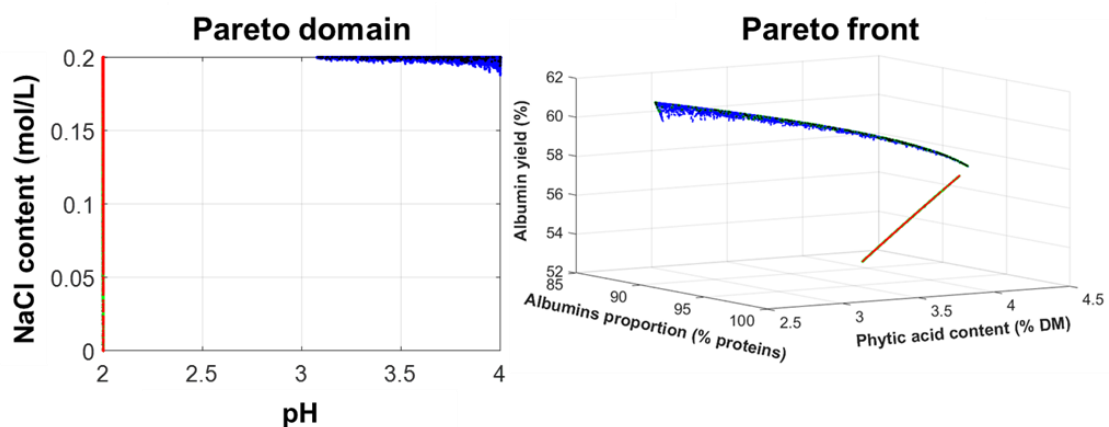


Figure 44 : Pareto front and Pareto domain resulting from the multicriteria optimization of albumin extraction according to three criteria: the albumin yield, the albumin proportion and the phytic acid content in the meal after extraction. The colors represent the ranking of the solution after applying the decision tool: in red are the first 10%, in green the solutions ranked between 10 and 30%, in black between 30 and 70% and in blue between 70 and 100%.

II.3.2.2 Use of a decision-making tool

Four samples with five individuals each were given to the expert for ranking. The decision-making algorithm generated 6 rules of preference. Figure 44 shows the ranking of the solutions of the whole population with different colors. The red dots represent the solutions ranked in the first 10% of the population, among which the multicriteria solution has to be selected. In terms of operating conditions, the area of pH 2 with more or less salt was preferable according to the rules governing the preference of the expert. Among these solutions, one was chosen at pH 2 with 0.11 M of NaCl. This value of NaCl content was chosen because of its robustness as it was surrounded by other points also ranked in the same 10% of the preferable solutions which ensured similar outcomes in case of small variations during the process.

The optimal operating conditions were applied at a larger scale to produce a sufficient quantity of protein for further characterization (functional and structural). The experimental performances of the process were: an albumin yield of 48.64% (55.62% predicted), an albumin proportion of 97.31% (97.91% predicted) and a PA content in the meal of 4.86% (3.59% predicted). Variations were observed between the model and the scaled-up process. The main reason would probably be due to the scaling-up where some parameters such as stirring were not exactly the same. Also, there was a large difference regarding the phytic acid content in the meal that was determined through a colorimetric method, which probably induced additional bias.

II.3.2.3 Evaluation of the protein solubility and color in aqueous solution

To be used in the food industry, the solubility of the proteins is essential to define the field of application of the protein ingredient. The color in solution is also determinant. Figure 45 A shows the solubility of the albumin isolate. For all four points of pH studied, the solubility was higher than 80%. At pH 3 and 9 the solubility was the highest with respectively 98.3% and 98.5%. At pH 7 it was 88.2% and the minimum of solubility was observed at pH 5 with 83%. It was previously reported that albumin has a high solubility (> 90%) whatever the pH (Cheung, Wanasundara, & Nickerson, 2015). A possible explanation for this slight lack of solubility at pH 5 might be related to the phytic acid content in the isolate (around 6% of the isolate on a dry matter basis). Indeed, phytic acid has been reported as interacting with proteins and thus decreasing their solubility (Bye, Cowieson, Cowieson, Selle, & Falconer, 2013). A detailed study on the solubility of rapeseed proteins with addition of various PA content was reported and showed a minimum of solubility at pH 5 when 5% PA was added (Kroll, 1991).

The color of the albumin in solution is also an important parameter in the food industry as it may change the color of the final product. Because of associations with phenolic compounds, proteins from

plants are usually coloured (in powder or in solution) (Xu & Diosady, 2002). In the present study, the analysis of the pH 7 solution resulted in a set of values in the L*a*b scale: 92*-4.2*24.3. This corresponded to a creamish solution. Again, these results were encouraging toward a possible use of rapeseed albumin in food thanks to a light colouring in solution.

Figure 45 B shows the chromatograms of the proteins at pH 7 analyzed at several wavelengths. At 214 nm, there was only one peak meaning that the isolate contained pure albumin. The chromatograms at 280 nm and 325 nm were used to show the complexation with phenols. Indeed, proteins do not absorb at 325 nm but polyphenols absorb at both wavelengths. As a consequence, this result showed that the color previously observed was probably due to irreversible phenolic complexation with proteins (Sosulski, 1979).

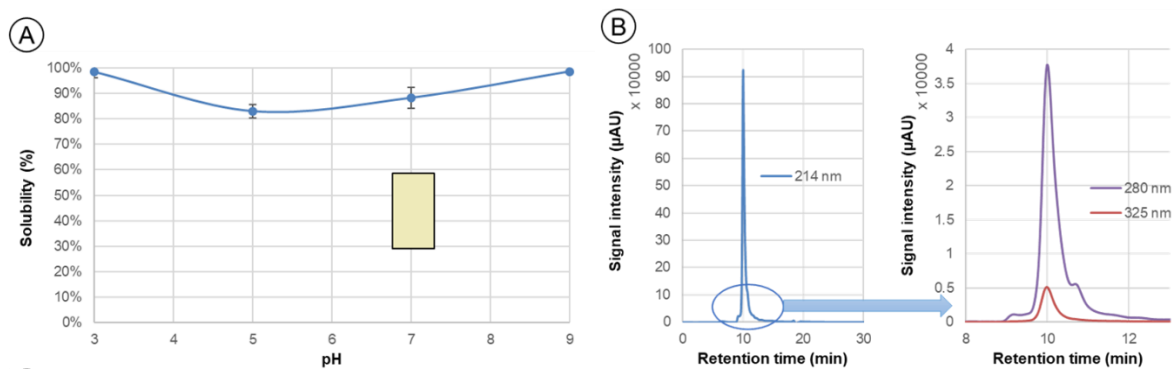


Figure 45 : Solubility of the albumin isolate as a function of the pH. The rectangle below the point at pH 7 shows the color of the isolate in a 1% solution diluted in a phosphate buffer pH 7.

11.3.2.4 Evaluation of the structural properties

Structural analysis of proteins was also carried out to characterize albumin. The first technique used was circular dichroism (CD) which gave information regarding the secondary structure of albumin. Figure 46 shows the results obtained for far-UV and near-UV CD for the albumin at pH 7. The far-UV spectra, Figure 46 A, has an overall shape consistent with a majority of α -helix (Kelly & Price, 2000). The signal was converted to determine the proportion of each type of secondary structures. It was found that the albumin product was composed of 49.7% of α -helix, 7.8% β -sheets, 14.3% β -turn and 21.3% of random coil. These values were different from what was previously reported by (Perera, McIntosh, & Wanasundara, 2016) who found only 27.5% of α -helix at pH 7. A possible explanation would be that the seeds used originated from different cultivar and thus, the amino acid sequences of the albumin may differ.

The near-UV spectra, Figure 46 B, showed a clear signal was obtained where three peaks were visible: tryptophan was around 290 nm, the tyrosine peak spreads between 275 and 282nm and the phenylalanine is located around 255 to 270 nm (Kelly & Price, 2000). These results were different from

the profile observed by Perera *et al* (Perera, McIntosh, & Wanasundara, 2016) with a peak for the phenylalanine more intense in the present study. However, these profiles were similar and did not show major conformational change.

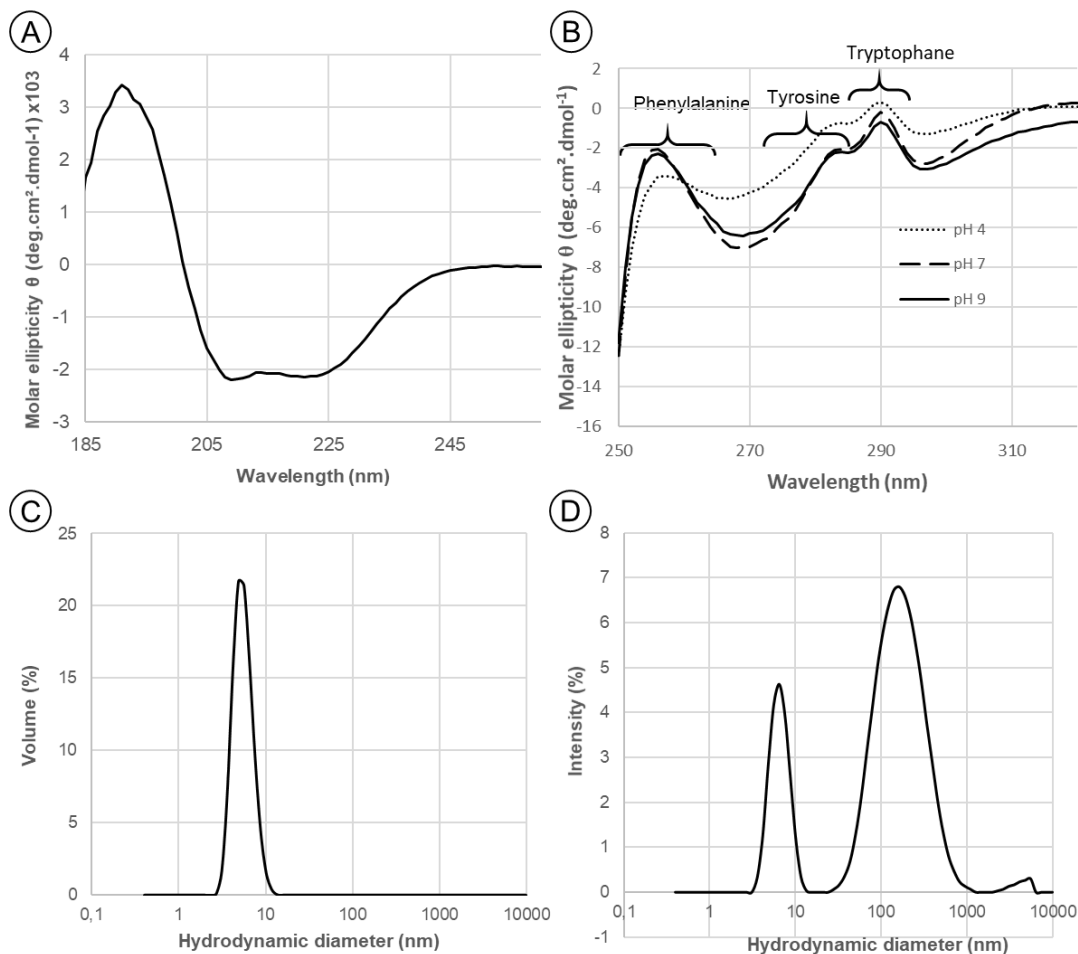


Figure 46 : Structural analyses by Circular Dichroism (CD) of rapeseed albumin A) shows the spectra observed by CD far-UV at room temperature and at pH 7. B) is the spectra in CD near-UV at room temperature at pH 4,7 and 9. C) and D) show the DLS repartition respectively in volume and in intensity as functions of the diameter of the particles.

Figure 46 C and Figure 46 D show the repartition in size of the particles according to the volume of the particle or to the intensity. As visible on Figure 46 C, there was only one peak visible when considering the volume of the particles which meant that there was only one species of particle in the medium. It was then possible to determine the average hydrodynamic diameter of the particle of 5.60 ± 0.20 nm.

Figure 46 D shows the size repartition of the observed particles regarding the intensity of the signal. Because only one population was visible according to the volume, this representation showed the size of potential aggregates formed with only albumin. Two main peaks were observed: the first one was located at the same place than the peak observed with volumes, hence it corresponded to the albumin not aggregated. The second peak was above 150 nm of diameter meaning that the proteins formed soluble aggregates. This aggregation state could be the result of hydrophobic interactions between

albumins. To check that out, it would have been of interest to carry out zeta-potential analyses to assess the surface hydrophobicity of the proteins.

II.4 Conclusion

The main objective of the present study was to rationally optimize the production process of rapeseed albumin. Two operating parameters were considered: the pH and the NaCl content. Both the produced albumin and residual meal were characterized. The methodology of DoE was successfully applied to obtain the polynomial relationship between operating conditions and performances of the extraction process. The use of multicriteria optimization allowed the determination of all acceptable trade-offs between albumin yield and composition in the extract and phytic acid content in the residual meal. A decision-making tool was applied to choose the best set of operating conditions according to the decision-maker preferences. The resulting operating conditions were experimentally carried out. Finally, the albumin isolate was characterized relatively to both the functional and the structural properties and showed promises for a future use in food industry: solubility above 80%, a light color in solution and the formation of small aggregates remaining soluble. The residual meal remained valuable for the feed industry. The proposed process is an enhancement from the existing procedures because it allowed the production of two valuable products: a protein isolate with high quality and a meal with a lower phytic acid content and an equivalent protein content compared to the initial meal. To continue this work, the second step of the production of proteins, the purification by membrane processes, should be optimized as well.

Acknowledgements

Authors wish to thank the company Olead, Pessac, France, for providing the meals used in this study. We are also grateful to Alexandre Kriznik for providing his time and expertise regarding structural analysis of proteins. Authors disclose no potential conflict of interest.

II.5 Contribution de l'article

Dans cet article, les influences du pH (entre 2 et 4) et de l'ajout de sel (de 0 à 0,2 M) ont été investiguées. Trois critères de performance étaient majeurs dans cette étude : la proportion en albumines, le rendement en albumines et le titre en acide phytique dans le tourteau résiduel. Les optimums de ces critères ont été trouvés dans des conditions de pH et de sel ajoutés bien distinctes : la proportion en albumines est de plus de 97 % à pH 2 sans sel, le rendement en napines est le plus élevé à pH 4 et en présence de 0,2 M de NaCl alors que la concentration minimale en acide phytique dans le tourteau est obtenue, soit à pH 2 avec 0,2 M de sel, soit à pH 4 avec 0,2 M de sel. Ainsi, la démarche rationnelle d'optimisation a été mise en œuvre. La modélisation des règles de préférences du décideur a orienté le choix des conditions opératoires du procédé à pH 2 avec 0,11 M de NaCl. Ainsi, un rendement d'extraction des albumines de 49 % a été obtenu avec une pureté de 97 % en cette protéine.

Dans ces conditions, les protéines extraites ont été purifiées par microfiltration et ultrafiltration pour générer deux produits. Un isolat d'albumines avec une solubilité de plus de 80 % sur la gamme de pH entre 3 et 9. La poudre est de couleur crème, ce qui est explicable par une complexation phénolique visible par la méthode CLHP-ES. Un tourteau avec un taux de protéines de 29,5 % et une teneur en acide phytique de 4,9 %, contre 6,8 % dans le tourteau initial, ce qui est favorable pour une utilisation en nutrition animale. Ainsi, un procédé extrapolable a été développé permettant de valoriser 70 à 80 % de la matière mise en œuvre dans le procédé.

Un autre facteur est susceptible d'affecter significativement l'extraction des albumines : la température, qui est reconnue pour aider l'extractibilité de molécules. Cependant, les protéines sont susceptibles de se dénaturer si la température est trop élevée ce qui aboutirait à une perte de la valeur nutritionnelle et fonctionnelle du tourteau. Le manuscrit suivant, soumis au journal *Process Biochemistry*, présente les résultats obtenus par comparaison de l'extraction des albumines à 20 et 55 °C.

III Manuscript: Influence of the extraction temperature on the quality of albumins from rapeseed meal

III.1 Introduction

Rapeseed is one of the largest cultures worldwide with an annual production of 70 million tons between 2013 and 2016 (FAOStats, 2019). The meal remaining from oil extraction is an industrial commodity used for animal nutrition due to its high content in proteins (30 to 40%) and balanced amino acid profile. Rapeseed proteins are mainly composed of the two major storage proteins: globulin and albumin. Globulin (11S protein) and albumin (2S protein) constitute respectively 50 – 60% and 35% of the rapeseed proteins. Globulin (cruciferin) is a multimeric protein of 300 to 350 kDa and its isoelectric point (pI) is close to 5 (Cheung, Wanasundara, & Nickerson, 2014). The albumin (napin) has a lower molar weight (12 to 15 kDa) and is a strongly basic protein (pI = 10 – 11) (Lonnerdal & Janson, 1972).

The solubility of albumin was found very high over a wide range of pH values with more than 90% (Cheung, Wanasundara, & Nickerson, 2015). On the contrary, globulin is usually poorly soluble under mild acid conditions. Furthermore, albumin was found to have an Emulsifying Activity Index higher than globulin and formed smaller droplets of oil (Krause & Schwenke, 2001). They were also reported to have better foaming capacity and stability (Akbari & Wu, 2015). Furthermore, the amino acid profile of these proteins is well balanced, especially regarding the sulfur amino acids (methionine, cysteine) that are lacking in other plant resources, such as pulses (Singh N., 2017). Hence, rapeseed albumin is of a great interest for food applications as isolate products.

Several studies reported a selective extraction of rapeseed albumin under acidic conditions (pH 2 – 4) (World Patent No. 144939 A1, 2008) (Nioi, Kapel, Rondags, & Marc, 2012). This selective extraction was explained by an alteration of globulin in an acidic medium while albumin remain largely soluble (Perera, McIntosh, & Wanasundara, 2016). Interestingly, beside the extraction of a highly valuable protein fraction, such an extraction would yield a residual meal still rich in proteins (mainly globulin) and poor in hydrosoluble antinutritional (phytic acid). Hence, this meal would still be of great interest for animal nutrition contrarily to residual meal obtained after total protein extraction. The main drawback of this process is the low extraction yield of albumin (close to 10%) (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018).

Several studies demonstrated that NaCl can increase the protein extraction from sunflower or fenugreek seeds (Pickardt, et al., 2009), (Feyzi, Varidi, Zare, & Varidi, 2015). In 2015, a study on rapeseed extraction showed that the addition of up to 0.15 M NaCl also increased the protein

extraction yields (Das Purkayastha & Lata Mahanta, 2015). Besides, other studies reported the positive effect of relatively high temperature on plant proteins extraction yield. Thus, an extraction process of soy proteins showed that a temperature of 50°C instead of 4 or 25°C increased the protein extraction yield by up to 24% (Lopes Barbosa, Lajolo, & Genovese, 2006). However, heat and NaCl concentration could impact protein structure, solubility or functional properties.

The first step of this study was to investigate the effect of the pH and the NaCl on the albumin yield and albumin composition in the aqueous extract. The nitrogen and phytic acid content in the residual solid phase were also considered. This was achieved by Design of Experiments (DoE). To assess the impact of the temperature, it was chosen to perform two DoE, one at 20°C and one at 55°C. In a second step, polynomial models from DoE were used in an original multicriteria procedure based on genetical-evolutionary algorithms in order to establish the optimal set of pH and NaCl concentration at both 20°C and 55°C. Then, in a third part, albumin was extracted in optimal conditions at 20°C and 55°C and purified by ultrafiltration in order to characterize them in terms of aggregation state, secondary and tertiary structures, and functional properties (solubility, foaming, emulsifying, color).

III.2 Material and methods

III.2.1 Raw material

The meal used was a cold-pressed dehulled rapeseed meal provided by Olead (Pessac, France). No de-oiling was carried out. Its proximal composition was: 32% (w/w) (N x 5.35) proteins, 3.5% (w/w) total dietary fibers, 8.7% (w/w) of carbohydrates, 9.6% (w/w) water, < 1% (w/w) polyphenols and 15.4% (w/w) oil on a dry matter basis. Chlorhydric acid was prepared from a 37% stock solution from Carlo Erba (Val-de-Reuil, France). Acetonitrile used for the Size-Exclusion High Performance Liquid Chromatographic (SE-HPLC) method was provided by Biosolve BV (Valkenswaard, the Netherlands) / ultrapure water was generated by a system Purelab Ultra by Elga (High Wycombe, United Kingdom). Trifluoroacetic acid (TFA) was from ThermoFisher (Waltham, Massachusetts, USA).

III.2.2 Protein extraction

Extractions were carried out in a temperature-controlled bioreactor of 200 mL at 20 or 55°C. 15 g of rapeseed meal was poured in 0 to 0.2 M NaCl solutions. The pH was adjusted between 2 and 4 with HCl 1 M and the final ratio solid:liquid was 1:9 (w/w). After 60 min, the liquid extract was separated from the remaining solid by centrifugation (15 minutes at 15000 x *g*). Then, the liquid was filtered through a cellulose filter paper from Fisher Scientific (Pittsburgh, Pennsylvania, USA). It was further clarified by microfiltration with a 0.22 µm membrane (Pellicon XL, Millipore, Burlington, Massachusetts, USA) plugged to a Cogent µScale tangential filtration system (also from Millipore). The

operation was carried out at room temperature and lasted until a volume reduction factor (VRF) of 10. The microfiltered extract was characterized for its albumin content (to determine a yield), and proportion of albumin / globulin. The remaining meal was freeze-dried and further characterized in terms of protein and phytic acid contents.

III.2.3 Analytics

III.2.3.1 Determination of protein content in solid and liquid phases

The analyses were achieved according to the official AOAC 979.09 Kjeldahl method. Samples were prepared with 4 mL of sulfuric acid (Sigma-Aldrich, Saint-Louis, USA) and catalyst Cu-Se (AppliChem, Gatersleben, Germany). The mineralization step was carried out at 450°C for 90 minutes in a Büchi SpeedDigester K-439 (Rungis, France). The distillation step (Kjelflex K-360, Büchi, Rungis, France) was done with sodium hydroxide 32% (w/v) (VWR, Radnor, Pennsylvania, USA) and boric acid (Sigma-Aldrich, Saint-Louis, USA) at 3% w/v. The titration was done using a 0.01 M HCl (Honeywell, Morris Plains, New Jersey, USA) in a Titrino Plus 877 (Metrohm, Herisau, Switzerland).

III.2.3.2 Quantification of albumin proportion in liquid phases

Albumin proportion in liquid phases was assessed by Size-Exclusion High Performance Liquid Chromatography according to (Defaix, et al., 2019). The chromatographic system was a UFLC system from Shimadzu (Kyoto, Japan) and the column was a Biosep-SEC-s2000 300 x 7.8 mm, 5 µm, from Phenomenex (Torrance, California, USA). The mobile phase was 45% (v/v) acetonitrile, 55% (v/v) ultrapure water and 0.1% (v/v) of trifluoroacetic acid (TFA). Elution was achieved at 0.6 mL / min during 30 minutes. Samples were filtered on 0.22 µm membranes from Pall (Port Washington, New York, USA) prior injection. The injection volume was 5 µL and the detection was set to 214 nm. Data were extracted and chromatograms were integrated manually using the spreadsheet software Excel from Microsoft (Redmond, Washington, USA).

III.2.3.3 Assessment of the phenolic complexation in the liquid phase

The SE-HPLC method previously described was also used in order to assess the irreversible complexation of phenolic compounds to proteins. It was expressed as the area under the albumin peak at 325 nm divided by the same at 280 nm.

III.2.3.4 Determination of phytic acid content in the solid phase

Phytic acid (PA) content was determined on freeze dried solid samples after protein extraction according to Garcia Estepa with slight modifications (Garcia-Estepa, Guerra-Hernandez, & Garcia-Villanova, 1999). The first step consisted in an extraction by a 0.4 M HCl solution with 10% (w/v) Na₂SO₄ (Sigma-Aldrich, Saint-Louis, USA) during at least 2 hours. After extraction, a centrifugation was

achieved (15 minutes, 15000 *g*). The supernatant was recovered, filtered (0.22 μm membranes from Pall, (Port Washington, New York, USA)) and mixed with equal volume of the following solutions: 20 mM iron chloride (Sigma-Aldrich, Saint-Louis, USA), 20% (w/v) sulfosalicylic acid (Sigma-Aldrich, Saint-Louis, USA), and 0.4 M HCl + 10% (w/v) Na_2SO_4 solutions to form a complex with PA. A thermal treatment at 100°C for 20 minutes was carried out to precipitate the complexed PA. After cooling and centrifuging (30 minutes, 15000 *g*), the supernatant was recovered and the precipitate was washed three times with water. Then, a titration was carried out with ethylenediaminetetraacetic acid (EDTA) (Sigma-Aldrich, Saint-Louis, USA). Finally, the mass of PA was determined according to (Garcia-Esteva, Guerra-Hernandez, & Garcia-Villanova, 1999) and the PA content was expressed relatively to protein mass or dry matter.

III.2.4 Modelling and optimization

III.2.4.1 Design of Experiments (DoE)

The effect of two operating conditions (factors) were studied: (i) pH (2 – 4) and (ii) NaCl concentration (0 M – 0.2 M) at 20 and 55°C by DoE. Central Composite Face-centered (CCF) matrixes composed of 11 experiments including 3 repetitions of the center points were used. The performance criteria were: (i) the albumin proportion; (ii) the albumin yield; (iii) the protein content in the residual meal; (iv) phytic acid content in the meal.

The albumin proportion was determined using the SE-HPLC method:

$$\text{Albumin proportion} = \frac{\text{Albumin content}}{\text{Total protein content}} \quad (\text{Eq 29})$$

The albumin yield was calculated as follow:

$$\text{Albumin yield} = \frac{\text{Mass of albumins in the microfiltered extract (SE-HPLC method)}}{\text{Theoretical mass of albumins in the meal}} \quad (\text{Eq.31})$$

The theoretical albumin content in the meal was chosen as 35% of the total proteins in the meal.

The remaining PA in the meal was expressed as the mass of phytic in the dry mass of the meal minus the remaining salt:

$$\text{PA in the meal} = \frac{\text{Mass of phytic acid}}{\text{Dry mass of meal} - \text{mass of remaining salt}} \quad (\text{Eq. 32})$$

The protein content in the remaining meal was determined with the Kjeldahl method ($\text{N} \times 5.35$).

III.2.4.2 Validity of the models

The software Modde (Umetrics, Umeå, Sweden) was used to define polynomial models for each response. Moreover, each model was refined by eliminating non-significant coefficients. Non-significant coefficients of the main factors were kept when at least one interaction coefficient was significant. The analysis of the coefficients was carried out using a 90% confidence interval. Some responses were transformed using a negative logarithm function.

Then validity of the models was assessed through several indicators: R^2 is the correlation coefficients; Q^2 represents the ability of the model to predict new data. A statistical test was also carried out to assess whether there was a lack-of-fit for each model. The ANOVA tests for the correlation and the lack-of-fit test were carried out with a confidence interval of 95%.

III.2.5 Albumin isolate production

Albumin isolates were produced from optimal extraction conditions at 20 and 55°C to study their structure and functionalities. 250 g of meal were introduced at a ratio solid:liquid of 1:9 (w/w). At 20°C, the conditions were pH 2 and 0.11 M NaCl. At 55°C, it was pH 3.75 and 0.2 M NaCl. After extraction a centrifugation step was carried out at 15000 *g* for 30 minutes (Lynx 6000 centrifuge, Thermo Fisher Scientific, Waltham, Massachusetts, USA). Extracts were microfiltered on a ProScale system (Millipore, Burlington, Massachusetts, USA) with a 0.1 m² Pellicon XL membrane with a 0.22 µm pore size membrane (Millipore, Burlington, Massachusetts, USA) at a reduction volume factor of 10.

Albumin extracts were purified by tangential diafiltration (Äkta Flux 6, GE Healthcare, Chicago, Illinois, USA) using a 3 kDa hollow fibers cartridges of 4800 cm² (GE Healthcare, Chicago, Illinois, USA). The feed flow rate applied was 1 L.min⁻¹ and the transmembrane pressure was 1.5 bars (at room temperature). The first step of the procedure consisted in five diavolumes with a feed solution of 0.5 M of NaCl maintaining the pH in the feed tank at 2 (for the extraction at 20°C) or at 3.75 (for the process at 55°C) with HCl 1M. Then, four diavolumes of ultrapure water were applied to remove salts. Final retentates were freeze-dried (LTD type GZL-0.5, Shanghai Pharmaceutical Machinery Co., Shanghai, China). The 20°C albumin had the following composition (relatively to the dry matter): 91.6% (w/w) of proteins, 5.1% (w/w) of water. The 55°C albumin were composed of 92.6% (w/w) of proteins, 7.1% (w/w) of water.

III.2.6 Functional properties assessment

III.2.6.1 Foaming capacity and stability

The protocol was adapted from (Aluko & McIntosh, 2001). A 20 mL solution at 1% (w/v) was prepared from lyophilized isolates in a 0.5 M phosphate buffer at pH 7 and placed in a 100 mL plastic cylindrical container. The solution was then homogenized with an Ultraturax at 10000 rpm for 5 minutes. Foaming capacity, FC, was expressed as:

$$FC = \frac{\text{Volume of foam}}{\text{Initial volume of liquid}} \times 100 \quad (\text{Eq. 25})$$

The foaming stability was the proportion of foam remaining after 5, 15, 30, 60 and 120 minutes.

$$\text{Foaming stability}_{x \text{ minutes}} = \frac{\text{Volume of foam after } x \text{ minutes}}{\text{Initial volume of foam}} \quad (\text{Eq. 26})$$

III.2.6.2 Emulsifying capacity and stability

The protocol was adapted from (Naczki, Diosady, & Rubin, 1985). A 5 mL solution at 5% (w/v) was prepared from lyophilized proteins in a 0.5 M phosphate buffer at pH 7. The solution was then homogenized with an Ultraturax at 10000 rpm for 30 seconds. Then 2.5 mL of rapeseed oil was added and the mixture was homogenized for another 30 seconds with an Ultraturax at 10000 rpm. Again, 2.5 mL of rapeseed oil was added and homogenized 90 seconds with an Ultraturax at 10000 rpm. The emulsion thus formed was transferred in a 15 mL conical tube and centrifuged 5 minutes at 1100 g at room temperature. Emulsifying capacity, EC, was expressed as:

$$EC = \frac{\text{Volume of emulsion after centrifugation}}{\text{Volume of emulsion before centrifugation}} \times 100 \quad (\text{Eq. 27})$$

Emulsion stability, ES, was assessed by heating the 15 mL conical tube for 15 minutes at 85°C. After cooling, a centrifugation step was carried out at room temperature for 5 minutes at 1100 g.

$$ES = \frac{\text{Volume of emulsion after heating}}{\text{Volume of emulsion before heating}} \times 100 \quad (\text{Eq. 28})$$

III.2.6.3 Purity and solubility

A 5 g.L⁻¹ solution was prepared from lyophilized proteins. The Kjeldahl analysis of this solution gave the purity of the lyophilized proteins:

$$\text{Protein purity} = \frac{\text{Mass of protein by Kjeldahl's method}}{\text{Mass of lyophilised proteins} \times \text{Dry Matter (\%)}} \times 100 \quad (\text{Eq.8})$$

A 20 mL sample was used for each point of pH. The pH was adjusted to 3, 5, 7 or 9 with 0.1 M NaOH or HCl. Precipitated proteins were separated by centrifugation (15 minutes, 10000 g at room

temperature). The supernatant was recovered and filtered through a 0.22 μm filter. Each analysis was duplicated. Solubility at a given pH was calculated as the ratio of protein amount in the supernatant on the protein amount in the starting sample.

$$\text{Solubility}_{\text{pH } X} = \frac{\text{mass of protein in the aliquot}}{\text{mass of protein in the initial solution}} \quad (\text{Eq. 24})$$

III.2.6.4 Color analysis

The color of the protein isolate was determined in aqueous solution. The analysis in liquid was carried out on a Lovibond PFX195 (Tintometer, Amesbury, United Kingdom). Samples were prepared as 1% (w/v) solution of protein powder in phosphate buffer 0.5 M at pH 7. The system was able to measure the color of the sample on the L.a.b scale. The term L related to the quantity of black / white, a to green / red and b to blue / yellow.

III.2.7 Structural analyses

III.2.7.1 Circular Dichroism (CD)

The secondary structures of proteins were analyzed using far-UV between 180 and 260 nm. Protein solutions at 1 $\text{mg}\cdot\text{mL}^{-1}$ were prepared in a 0.5 M phosphate buffer at pH 7. The spectral analyses were carried out at room temperature with a Chirascan plus from Applied Photophysics (Leatherhead, United Kingdom). For the acquisition of spectra at a given temperature, samples were prepared in cells of 0.1 mm path length. For each sample, the analysis was carried out three times, and an average was traced. The software Pro-Data Chirascan from Applied Photophysics (Leatherhead, United Kingdom) was used to process the data. The same system was used for assessing tertiary structures using near-UV spectra between 250 and 320 nm. A quartz cell of 10 mm path length was used.

III.2.7.2 Dynamic Light Scattering (DLS)

DLS was used to determine the aggregation state of proteins. Protein solutions at 1 $\text{mg}\cdot\text{mL}^{-1}$ were prepared in a 0.5 M phosphate buffer at pH 7. Quartz cells of 10 mm path length were used. The analyses were carried out with a Zetasizer Nano – S from Malvern Instruments (Malvern, United Kingdom). Each analysis was carried out in a 3 mm cell, repeated three times and an average trace was determined.

III.3 Results and discussion

III.3.1 Impact of temperature, pH and NaCl concentration on rapeseed albumin extraction

III.3.1.1 DOE matrixes and choice of operating conditions

Two DoE were carried out, at 20 and 55°C, in order to assess the influence of the pH and the NaCl content on the extraction of albumin. 55°C temperature was chosen because it classically limits microbial contaminations. The values of pH were chosen between 2 and 4 according to Nioi *et al.* (Nioi, Kapel, Rondags, & Marc, 2012). The NaCl content was studied between 0 and 0.2 M because higher content would damage steel tanks commonly used in the food industry. Other parameters such as, the duration of extraction (60 minutes) and the solid-liquid ratio (1:9) were chosen based on the work of Nioi *et al.* (Nioi, Kapel, Rondags, & Marc, 2012). The performance criteria were the albumin yield and the albumin proportion. Both matrix of experiments and the obtained results are shown in Tableau 10.

Tableau 10 : Matrix for the design of experiments at 20°C for the study of albumin extraction.

Experim ent number	pH	NaCl cont ent	Albumin proportion (%)		Albumin yield (%)		Protein in the residual meal (%)		Phytic acid in the meal (%)	
			20°C	55°C	20°C	55°C	20°C	55°C	20°C	55°C
1	2	0	98.03	92.67	57.53	64.13	30.83	29.06	3,88	5,62
2	4	0	91.58	95.76	40.08	55.40	32.11	29.82	3,84	5,20
3	2	0.2	98.02	93.28	54.83	63.13	30.83	31.48	3,25	5,40
4	4	0.2	86.30	95.83	61.37	76.76	27.90	27.92	2,54	2,75
5	2	0.1	97.73	92.69	53.27	63.20	30.95	29.93	3,73	5,78
6	4	0.1	90.41	96.43	51.31	70.69	29.98	31.64	2,76	3,58
7	3	0	97.18	96.65	51.23	65.01	32.85	30.92	4,51	4,14
8	3	0.2	96.67	96.65	53.89	68.82	31.17	34.00	4,63	4,62
<i>Center points</i>										
9	3	0.1	96.26	96.38	55.61	67.29	30.28	28.03	4,61	3,71
10	3	0.1	96.09	96.9	57.40	67.05	30.53	33.51	4,00	4,10
11	3	0.1	95.56	96.02	53.63	62.22	31.66	29.89	4,87	3,68
R²			0.940	0.965	0.847	0.840	0.825	0.297	0.824	0.821
Q²			0.896	0.944	0.563	0.566	0.588	-0.200	0.674	0.402
Reproducibility			0.982	0.929	0.852	0.723	0.667	-0.200	0.667	0.944
ANOVA F-test p-value			<0.001	<0.001	0.003	0.004	0.019	0.816*	0.005	0.020
Lack of fit test p-value			0.173	0.793	0.418	0.654	0.616	0.705	0.695	0.122

Values followed by a * indicate a p-value superior to 5% for the ANOVA F-test.

III.3.1.2 Validity of the models

After linear regression, models were refined by removing some of the non-significant terms.

Tableau 11 shows the coefficients used for each response and their values. Doing so, the two selected criteria (albumin extraction yield and composition) showed high R^2 with values higher than 0.84 which correspond to models with a good correlation. Moreover, analysis of the lack-of-fit was carried out and the p-values obtained were higher than 0.05 meaning no lack-of-fit for these responses.

Tableau 11 : Coefficients for the studied models of both DOE at 20 and 55°C.

	Albumin proportion (%)		Albumin yield (%)		Phytic acid in the residual meal	
Transformation	-10log(100-Y)		-10log(100-Y)		/	
Coefficient	Value	p	Value	p	Value	p
Constant	-0.48	10^{-7}	-1.48	10^{-15}	-5.04	10^{-8}
pH	0.30	10^{-6}	-0.07	0.11	6.87	0.11
NaCl	/	/	-1.32	0.01	-2.99	0.10
pH*pH	-0.11	0.05	/	/	-1.19	10^{-3}
NaCl*NaCl	/	/	/	/	/	/
pH*NaCl	/	/	0.54	10^{-3}	/	/
DoE at 55°C						
Transformation	/		/		/	
Coefficient	Uncoded	p	Uncoded	p	Value	p
Constant	73.14	10^{-20}	72.34	10^{-12}	11.43	10^{-6}
pH	14.02	10^{-6}	-3.53	0.09	-4.36	0.01
NaCl	/	/	-127.4	10^{-3}	13.11	0.15*
pH*pH	-2.08	10^{-6}	/	/	0.67	0.08
NaCl*NaCl	/	/	/	/	/	/
pH*NaCl	/	/	55.90	10^{-3}	-5.58	0.08

* The p-value is higher than 10 % but it was kept because the coefficient of an interaction involving the NaCl content was significant.

III.3.1.3 Influence of the pH, NaCl content and temperature on the proportion of albumins extracted

Figure 47 a) shows the response surface for the proportion of albumin in the extracts at 20°C. A significant influence of the pH is observed with values decreasing from 98% at pH 2 to less than 90% at pH 4. There was no effect of the addition of NaCl on this response. The effect of the pH can be explained by an alteration of globulin under more acidic pH even limiting their extraction (Perera, McIntosh, & Wanasundara, 2016).

Figure 47 b) shows the response surface of the proportion of albumin at 55°C. The values ranged from 93 to 96.5%. As for the 20°C extraction, no effect of the addition of NaCl was noticed. Contrary to the response surface at 20°C, the lowest values (93%) was found at pH 2 whereas the highest value was 96.5% between pH 3 and 3.7. These slight discrepancies could be explained by a slightly better

solubilization of globulin at high temperature. Indeed, Radwan *et al.* reported that the nitrogen contained in rapeseed meal was more soluble at 55°C than at 25°C. The same study also showed a shift of the minimum of solubility from pH 4 – 5 at 25°C (which corresponds to the pI of globulin) to pH 7 at 55°C (Radwan & Lu, 1976). In any cases, the extraction revealed to be selective toward albumin.

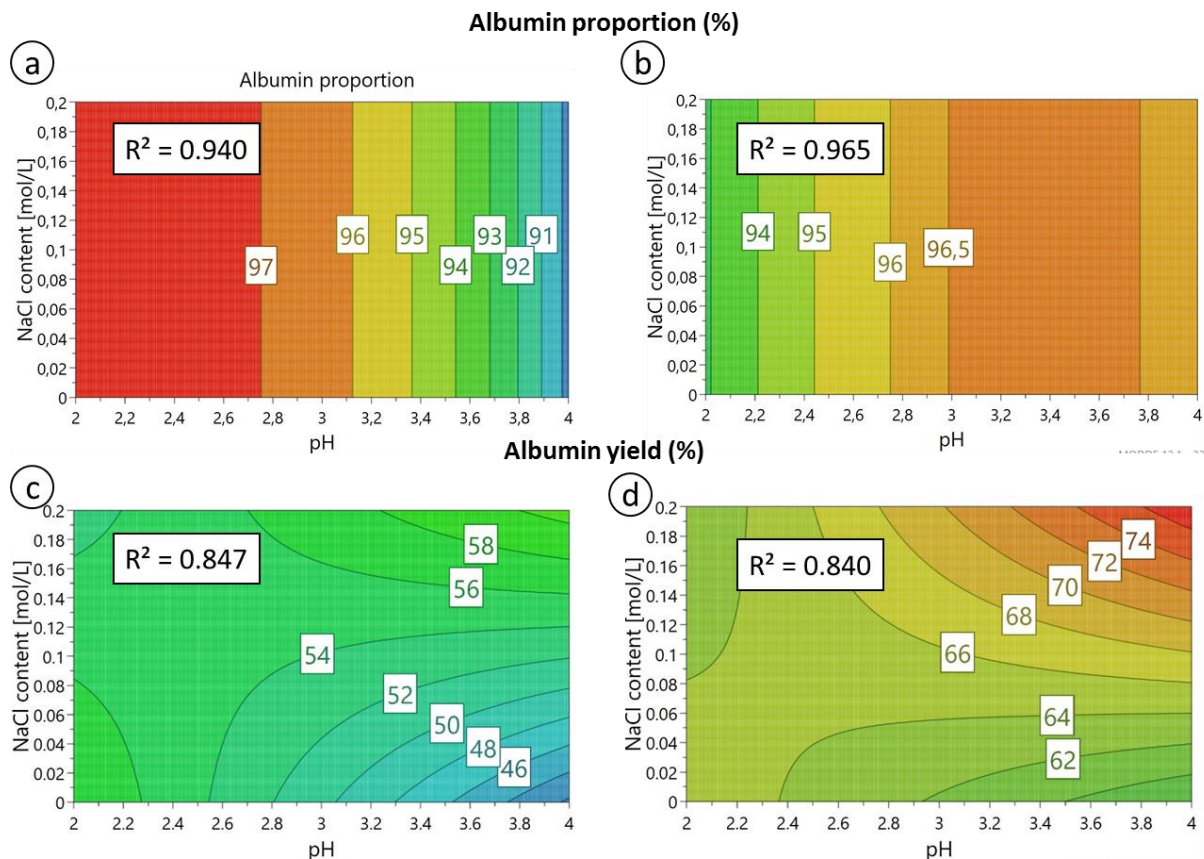


Figure 47 : Contour plots of the performance criteria regarding the extraction of albumin. a) and b) show the albumin proportions, respectively at 20 and 55°C; c) and d) show the albumin yields respectively at 20 and 55°C.

III.3.1.4 Influence of pH, NaCl content and the temperature on the albumin yield

Figure 47 c) represents the response surface of the albumin yield determined by the SE-HPLC method at 20°C. This surface can be divided into three parts. Firstly, at pH between 2 and 2.5, there was no impact of the pH and NaCl concentration on the yields that were high, between 54 and 56%. Secondly, at pH higher than 2.5 and with low NaCl concentrations (below 0.1 M), the increasing pH resulted in lower yields, from 54 to 40%. The yields were higher when more NaCl was added to the medium, from 40 to 54%, at pH 4. Thirdly, at pH higher than 2.5 and high NaCl content, the pH had little impact and the addition of salt contributed to an increase of the yields to 60%. As a consequence, electrostatic interactions seem to be involved in the solubilization of proteins. Albumin are positively charged over this range of pH, which means that negatively charged functions on the solid phase would limit their extraction. Rapeseed meal are known to be composed of insoluble hemicellulosic polymers that

contain glucuronic or galacturonic acids. Carboxylic acid groups of these carbohydrate could be involved in these limiting electrostatic interactions.

A striking difference was observed from previously reported work by Nioi *et al.* who found around 10% in albumin yield (Nioi, Kapel, Rondags, & Marc, 2012). As reported by Fetzer *et al.* such a discrepancy could be explained by the process used to obtain the rapeseed meal. Indeed, Nioi *et al.* worked with an industrial toasted meal (hot pressing, hexane deoiling and hot desolventation). In Fetzer *et al.* it was shown that the solubility of proteins was reduced from approximately 35% to 12% at pH 4 between a de-oiled cold-pressed meal and a toasted meal (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018). Moreover, in the present study, the cold-pressed meal was used as is whereas it was deoiled in the work of Fetzer *et al.* This deoiling step with solvent involved an exposition of the meal to 60°C which probably contributed to a decrease of the solubility of proteins as well. As a consequence, because of the mild condition of production of the meal in the current study, the values of the yields were higher.

Nevertheless, at 20°C the albumin yields were observed at 60% at most, even when electrostatic interaction cannot limit the extraction. The microfiltration losses were quantified and explained only 15-20% of the albumin losses (data not shown). Hence another limitation occurred. A possible explanation would be that precipitated globulins in the pores of the solid would act as a barrier for the extraction of albumin. This effect seemed enhanced at pH 2 where we suppose that globulins are in a further denatured state.

Figure 47 d) shows the surface for the albumin yields at 55°C. A shape similar to that at 20°C was observed but with higher yields, between 60 to 76%. This effect can most probably be attributed to a favored extraction of proteins through increased the diffusion coefficients from the solid phase to the liquid phase. Also, Radwan *et al.* suggested a conformational change of proteins when the temperature was high (Radwan & Lu, 1976). Following this hypothesis, the arrangement of the globulin would differ. This would result in a less limiting effect thus towards the extraction of more albumin.

III.3.2 Establishment of the optimal pH and NaCl concentration at 20°C and 55°C

The objective of the extraction step was to extract albumin at the highest yield with a satisfying albumin composition and quality. The phytic acid and protein content in the residual solid was also important because it will define the value of this side-product for feed applications. The overall protein content in the meal ranged from 28 to 34% (w/w) (Tableau 10). This corresponded to the usual range

of protein content in meal. So, this criterion was considered satisfying whatever the extracting conditions.

Phytic acid (PA) has been recognized over the years as an antinutrient that impacts the meal nutritional property negatively (Cheryan & Rackis, 1980). The content of PA in the starting meal was 6.8% (w/w). In the remaining meal it was found lower values: at 20°C, between 2.8 and 4.9% (w/w) of the dry matter of the meal and, at 55°C, between 2.8 and 5.8% (w/w), Figure 48. The effect of the operating conditions on PA content was very close at both temperatures, except toward pH 2 where the PA content was more important at 55°C. The results regarding the PA content in the meal were particularly hard to interpret. Indeed, different phenomenon like PA solubility, overall dry matter solubility or ion pair formation depending on the temperature, pH and NaCl content can modulate the extraction. In any case, an acceptable limit of 4% PA in the residual meal was considered acceptable.

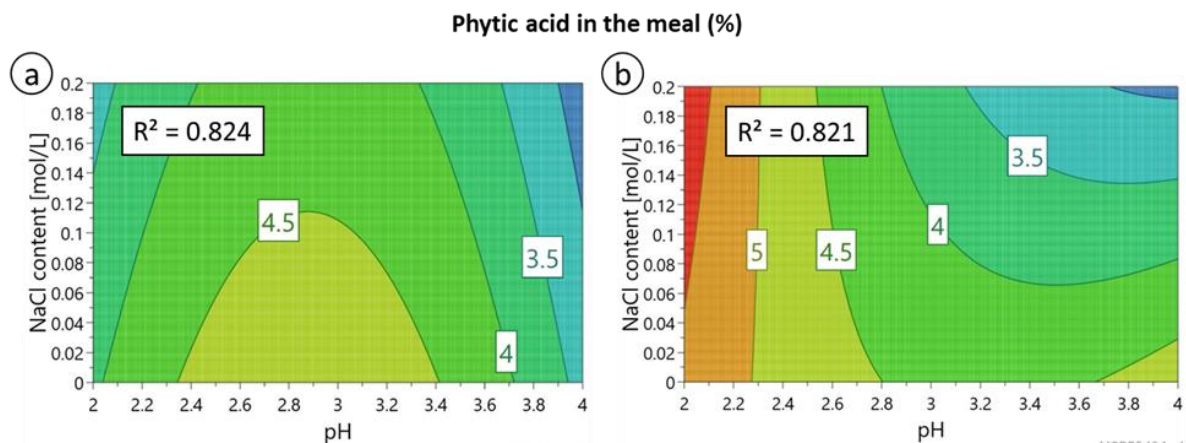


Figure 48 : Contour plots of the phytic acid content in the meal after extraction of albumin: a) from the Design of Experiments (DOE) at 20°C. and b) from the DoE at 55°C.

Regarding the performances of the extraction, albumin composition thresholds were established at 95% of albumin whatever the temperature. Because extraction yields were found very different, the albumin yield limits were established at 55% and 70% at 20°C and 55°C respectively. Finally, the phytic acid content in the meal was chosen lower than 4%. Figure 49 shows the operating conditions areas fitting these constraints (in blue at 20°C and in orange at 55°C). It was remarkable that the two areas corresponded to very different operating condition sets. The acceptable conditions at 20°C were located below pH 2.3 with up to 0.13 M of NaCl. On the contrary, the acceptable conditions at 55°C were found when pH was above 3 with 0.13 to 0.2 M of NaCl.

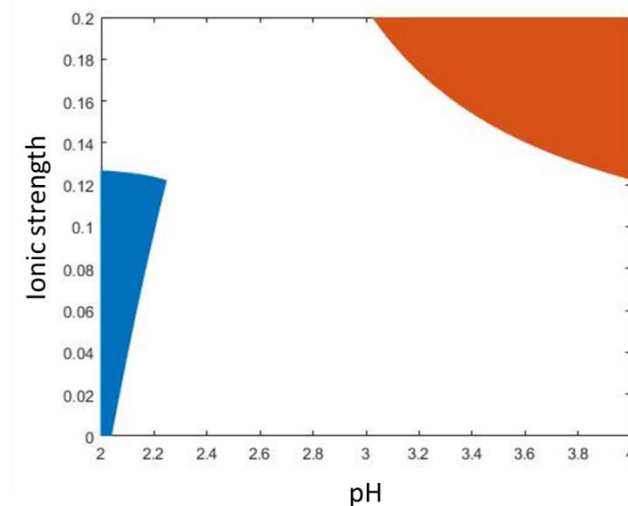


Figure 49 : Graph of the desirable operating conditions to produce rapeseed albumin. The area in blue corresponds to an albumin yield higher than 55%, an albumin proportion higher than 95% and a phytic acid content in the meal lesser than 4%. The orange area corresponds to an albumin yield higher than 70%, an albumin proportion higher than 95% and a phytic acid content in the meal lesser than 4%.

As a consequence, two sets of operating conditions were chosen inside these areas: at 20°C, pH 2 with 0.11 M of NaCl and, at 55°C, pH 3.75 with 0.2 M NaCl. Even though both conditions were chosen for their high yields, the expected value at 55°C (75.6%) is much higher than expected at 20°C (55.4%) by nearly 20%. However, the temperature may have an impact on the quality of the final product: indeed, heat may alter the structure, color or properties of napins. To check that out, albumin extracted in these two conditions were purified by tangential filtration and characterized. The experimental yields for both albumins were close to the model: at 20°C, the value predicted by the model was 55.4% and the experimental value was 48.6% which correspond to a relative difference of -12.6%. At 55°C, it was predicted 75.6% and found 74.7%, which corresponds to a relative difference of -1.2%. Also, the albumin proportion was accurately predicted with, at 20°C, 97.9% predicted versus 97.3% experimentally (-0.6% of relative difference), and, at 55°C, 96.6% predicted for 95.4% observed (-1.3% of relative difference). Regarding the phytic acid content, it was found respectively at 20 and 55°C, 4.86% (w/w) and 4.74% (w/w) which was higher than expected by 35% and 59%. The main explanation for this deviation is the variability of the method used to determine the phytic acid content. Nevertheless, the phytic acid content was significantly lower than the starting material at 6.8% (w/w).

III.3.3 Characterization of albumin obtained at 20°C and 55°C

III.3.3.1 Evaluation of the structural properties

Figure 50 a) shows albumin circular dichroism spectra obtained in far-UV. The spectra of albumin extracted at both 20°C and 55°C were found very close. The overall shape of the spectra was consistent with a majority of α -helix (Kelly & Price, 2000). The proportion of each type of secondary structure

was determined for each spectrum at pH 7. It was found that the albumin extracted at 20°C were composed of 49.7% of α -helix, 7.8 β -sheets, 14.3% β -turn and 21.3% of random coil. Albumin extracted at 55°C were slightly different with 44.2% of α -helix, 9.9 β -sheets, 15.3% β -turn and 22.9% of random coil. These values were different from what was previously reported by (Perera, McIntosh, & Wanasundara, 2016) who found only 27.5% of α -helix at pH 7. A possible explanation would be that the seeds used originated from different cultivar and thus, the amino acid sequence of the albumin would differ. Another possibility could originate from the process to produce rapeseed meal: in the study of Perera *et al.* the meal was de-oiled with hexane followed by a desolvating step, whereas the present study used a cold-pressed meal. According to Fetzer *et al.* the meal production process has a great influence upon the quality, and particularly the solubility, of the proteins (Fetzer, Herfellner, Stabler, Menner, & Eisner, 2018). These differences are likely to explain the variation observed from the literature.

The near-UV spectra, Figure 50 b, gives information regarding the aromatic amino acids: tryptophan is usually visible at 290 nm, the tyrosine peak spreads between 275 and 282nm and the phenylalanine is located around 255 to 270 nm (Kelly & Price, 2000). The overall shape of the near-UV spectra was very similar between both albumin products. There was however a slight difference regarding the tyrosine area: at 20°C, the signal was stronger than at 55°C. This can be interpreted as a difference of conformation in the surrounding of the tyrosine residues. This may be related to the variation observed in the composition of secondary structures observed with far-UV CD.

Figure 50 c) show the repartition in size of the particles according to volume of the particle. For both albumin, there was only one peak visible when considering the volume of the particles which means that there was only one species of particle in the medium. It was then possible to determine the average diameter of the particle. The observed diameter for the albumin produced at 55°C was higher than for the albumin produced at 20°C (5.60 ± 0.20 nm at 20°C and 6.34 ± 0.39 nm at 55°C). A statistical analysis showed a significant difference in the buffer at pH 7 with a p-value of 4.3%.

Figure 50 d) shows the size repartition of the observed particles regarding the intensity of the signal. Because only one population was visible regarding the volume, this second representation showed the size of potential aggregates formed with only albumin. For the albumin at 20°C, two main peaks were observed: the first one was located at the same hydrodynamic diameter than the peak observed with volumes, hence it corresponds to the albumin not aggregated; the second peak was above 150 nm of diameter meaning that proteins formed an aggregate. A similar phenomenon was observed for the albumin produced at 55°C with two peaks. In this case, the aggregates were observed at a value close to 300 nm. Hence it would seem that the albumin produced at 55°C formed bigger aggregates

than the albumin 20°C. This new data, in relationship with the CD results indicates that there were indeed differences in structures between both albumin products.

Perera *et al.* also studied their albumin by near UV-CD (Perera, McIntosh, & Wanasundara, 2016). Even though the overall shape was similar, the spectra obtained in the present study were slightly different with more visible peaks in both the tyrosine and the phenylalanine areas. However, these results were difficult to interpret as is. That is why, further investigation was carried out through the study of the amino acid profile and through the functional properties to determine whether these differences were significant.

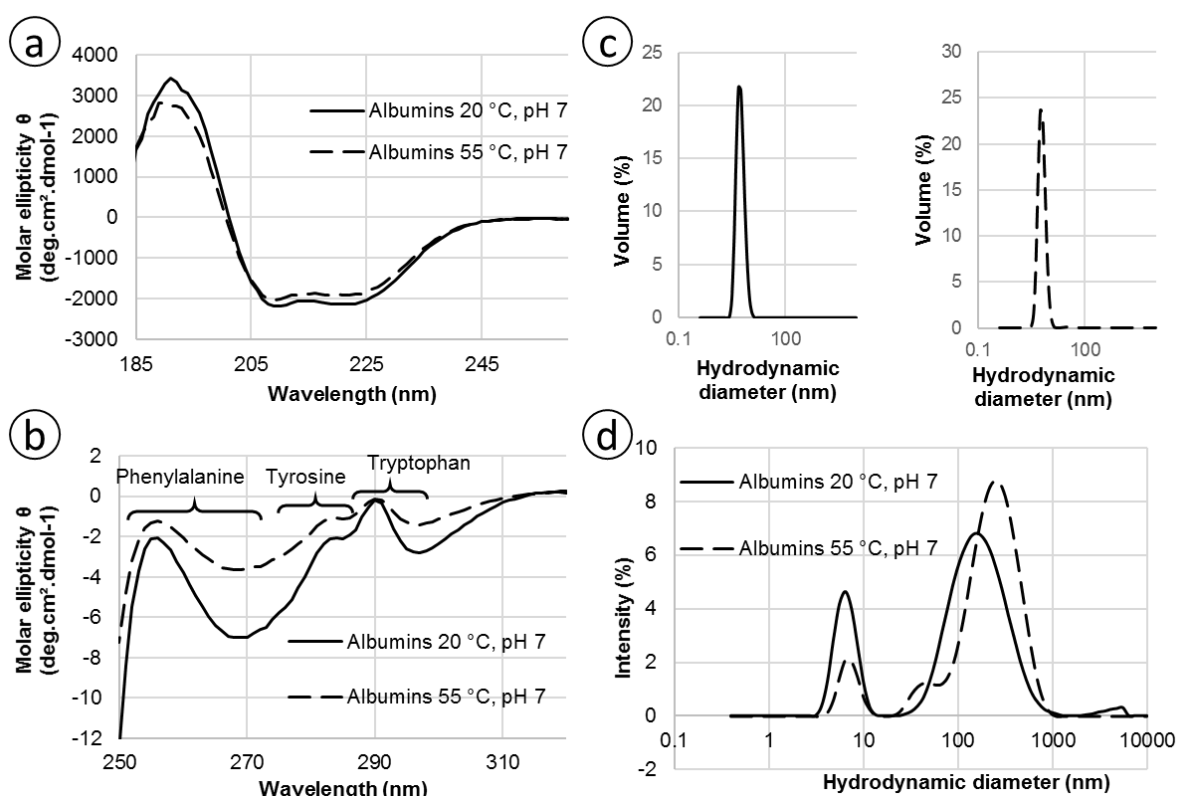


Figure 50 : Structural analyses by Circular Dichroism for both albumin products from 20 and 55°C processes. a) shows the spectra observed by CD far-UV at room temperature at pH 7. b) shows spectra in CD near-UV at room temperature at pH 7. c) and d) show the DLS repartition for the albumin at 20°C and at 55°C respectively in volume and in intensity as functions of the diameter of the particles at pH 7.

III.3.3.2 Amino acid composition

Tableau 12 shows the values for the aminograms of both albumin products. When compared to each other, the rapeseed albumin showed differences regarding the content of several amino acids. Six amino acids were identified with more than 10% of relative difference between both products: hydroxyproline, threonine, aspartic acid, proline, tyrosine and cysteine. The difference in the tyrosine residue may explain the result observed in the structural analysis. Hence, a hypothesis can be

advanced that there may be a difference in the isoforms of rapeseed albumin extracted according to the temperature of the extraction.

Another interesting comparison would be with the recommendations of the FAO for the essential amino acid in human nutrition. On the 10 amino acids present in the recommendations of the World Health Organization (WHO) (World Health Organization, 2007), only three were found in insufficient quantities: phenylalanine (with a relative difference of 28 and 31% respectively for the albumin at 20°C and at 55°C from the values from WHO), isoleucine (with 6 and 8% relative difference) and leucine (with 1 and 6% relative difference). It is to note that, as previously described in the literature, lysine and sulfur amino acids (methionine and cysteine) were present in excess from the WHO references.

Tableau 12 : Aminograms of albumin produced at 20°C and at 55°C in comparison with the World Health Organization (WHO) recommendations

Amino acid	Albumin 20°C Content (g / 100 g protein)	Albumin 55°C Content (g / 100 g protein)	WHO Content (g / 100 g protein)
Tryptophane	1.07 ± 0.11	1.04 ± 0.10	0.6
Hydroxyproline	0.118 ± 0.024	0.198 ± 0.04	/
Threonine	3.13 ± 0.44	3.42 ± 0.48	2.3
Aspartic acid	2.78 ± 0.39	3.50 ± 0.49	/
Serine	3.53 ± 0.49	3.63 ± 0.51	/
Lysine	7.48 ± 1.05	7.36 ± 1.03	4.5
Valine	4.03 ± 0.56	3.99 ± 0.56	3.9
Proline	7.98 ± 1.12	7.44 ± 1.04	/
Alanine	3.8 ± 0.53	3.84 ± 0.54	/
Phenylalanine	2.75 ± 0.39	2.63 ± 0.37	3.8
Isoleucine	2.81 ± 0.39	2.77 ± 0.39	3
Glycine	4.41 ± 0.62	4.56 ± 0.64	/
Tyrosine	1.43 ± 0.20	1.54 ± 0.22	/
Arginine	5.55 ± 0.78	5.62 ± 0.79	/
Leucine	5.82 ± 0.81	5.56 ± 0.78	3.8
Histidine	3.39 ± 0.47	3.29 ± 0.46	3
Glutamic acid	24.2 ± 3.40	23.0 ± 3.20	/
Méthionine	1.92 ± 0.27	1.83 ± 0.26	1.6
Cysteine + Cystine	3.90 ± 0.55	3.5 ± 0.49	0.6

III.3.3.3 Evaluation of the functional properties

Figure 51 a) shows albumin solubility at pH between 2 and 9. As expected, both albumin products were largely soluble in the whole pH range. To our knowledge, no plant protein isolate revealed as soluble on a pH range that wide. Indeed, most proteins have an isoelectric point between pH 5 and 7 and a very low solubility in this area. For example, a study compared the solubility of isolates from hemp and soybean: for both isolates it was shown a very low solubility for pH from 2 to 6-7 in

relationship with a pI close to pH 5 (Tang, Ten, Wang, & Yang, 2006,). This considerably reduces the field of applications of such isolates and confirms the great interest of rapeseed albumin for food applications.

At pH 5 the solubility was slightly lowest than expected (83 and 63% for the processes at 20°C and 55°C respectively). Interestingly, phytic acid, known to limit protein solubility, remained associated to albumin (Bye, Cowieson, Cowieson, Selle, & Falconer, 2013). Furthermore, the 55°C albumin isolate contained more PA than the 20°C isolate (6.8% at 55°C versus 6.0% at 20°C). This difference would be explained by a higher association of phytic acid under acidic conditions.

Figure 51 b) shows the Emulsifying Activity (EA) of both albumin isolates. For the albumin produced at 20°C it was found $32.7\% \pm 2.5$ whereas it was $30.1\% \pm 1.3$ for the process at 55°C. The two values were found statistically identical with respect to, for albumin at 20°C and 55°C, 33 and 30% ($p = 11.5\%$). For both isolates, the emulsions were perfectly stable after heating. The EA was found lower than previously reported by (Naczki, Diosady, & Rubin, 1985). In this study, the protein concentration of the initial solution was higher which is a possible explanation for the better EA observed. Similarly, *Pedroche et al.* observed higher EA but the emulsion stability was lower than observed in the present study (Pedroche, et al., 2004).

Figure 51 c) shows that the foaming capacity of albumin extracted at 20°C was found $277\% \pm 31$ whereas it was $319\% \pm 26$ at 55°C. But this difference was not significant ($p = 8.5\%$). Fig. 3 d) shows that both foams were as stable (65% of the initial volume of foam remaining after 120 minutes). The results regarding the foaming capacity and stability were comparable to the experiments reported by (Pedroche, et al., 2004) even though their initial solution was more concentrated. It was also reported foaming capacities of products from pulses in a range of 6 to 80%, which was less than observed in the present study (Sathe & Salunkhe, 1981). Compared to other products reported in the literature, rapeseed albumin has a good potential as food ingredient.

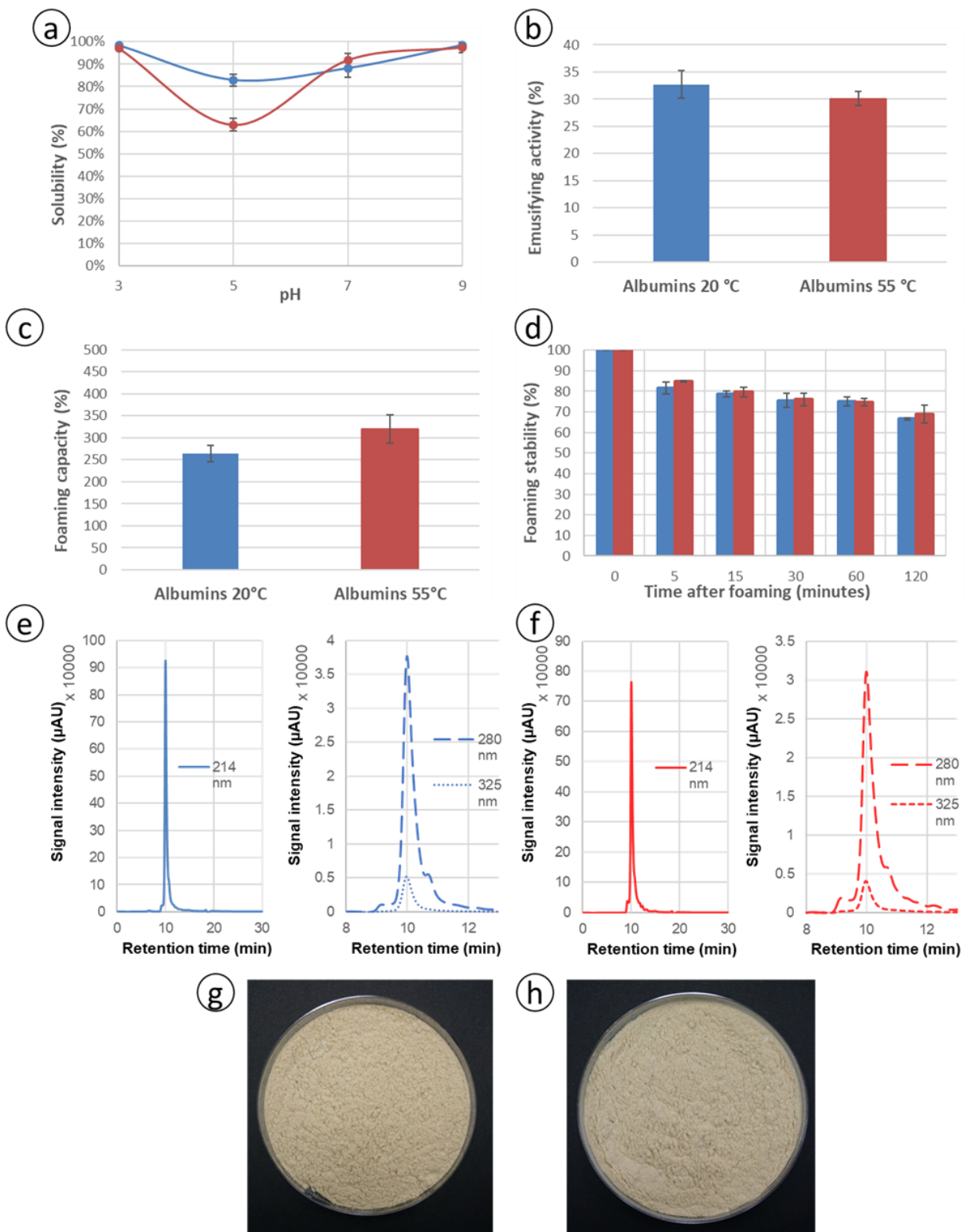


Figure 51 : Functional properties of rapeseed albumin isolates. Albumin produced according to the process at 20°C are represented in blue, the albumin produced at 55°C in red. a) shows the solubility of albumin at various pH; b) represents the emulsifying activity; c) is the foaming capacity and d) is the foaming stability; e) and f) show the chromatograms of both proteins at 20 and 55°C at 214 nm and 280 and 325 nm; g) and h) show photographs, respectively of the albumin at 20 and 55°C.

The color of the albumin in solution is also an important parameter in the food industry as it may change the color of the final product. Both albumin solutions had similar colors: pale brown-yellow. To quantify this color, 1% of both albumin solutions were analyzed and defined according to the L*a*b scale. It was found respective average values of $92^* - 4.2^* 24.3$ and $96.3^* - 3.3^* 23.1$ for the albumin at 20°C and the albumin at 55°C. These values were very similar and only the term a was found statistically different. This means that there was a slight difference in the color with a slightly browner shade for the albumin at 20°C. It is also to note the high values for the term L meaning a light solution. This color is supposedly due to the presence of phenolic compounds, which are components known to interact with proteins and to take part in the coloration of a product even at very low concentration (Sosulski, 1979). To check that out, the phenolic complexation was investigated thanks to the SE-HPLC method at two wavelengths: 280 and 325 nm. Figure 51 e) and f) show the chromatograms at the three wavelengths, 214, 280 and 325 nm. For both albumin products, it was found that the signal of polyphenols (at 325 nm) represented 13.3% of the signal at 280 nm. Figure 51 g) and h) show photographs of albumin produced at 20 and 55°C as powders. Both proteins were found with a light-yellow color. As a result, it seems that both processes allow the production of albumin with similar functional properties of interest for the formulation of food.

III.4 Conclusion

The present study objective was to rationally study the production process of rapeseed albumin. Three operating conditions were chosen: pH, NaCl content and the temperature. Both the albumin produced and the residual meal were characterized. The methodology of DoE was successfully applied to obtain equations linking the operating conditions and the performances of the process. It was found that extraction at 55°C resulted in higher extraction yields than at 20°C. Constraints were defined in order to rationally choose the best operating conditions for each temperature. Then, the structural and functional properties were compared. Very similar results were found between both albumin products. Some structural variations were observed between both albumin with no significant impact on the functional properties that were comparable for foaming, emulsifying and color. As a consequence, it seems that an extraction process at 55°C is favorable because it significantly increased the albumin yield while maintaining the selectivity and preserving the functional properties of the isolate. Moreover, another interesting feature of this process is that a meal with around 30% remaining protein was generated, which is similar to the starting material, but with less phytic acid (4% versus 7%).

Acknowledgements

Authors wish to thank the company Olead, Pessac, France, for providing the meals used in this study. We are also grateful to Alexandre Kriznik for providing his time and expertise regarding structural analysis of proteins. Authors disclose no potential conflict of interest.

III.5 Contribution de l'article

Les résultats présentés dans l'article permettent de comparer les performances d'extraction des protéines dans un milieu à 20 °C et à 55 °C. Deux plans d'expériences, un par température, ont été réalisés et ont permis de générer des équations mathématiques entre le pH et la concentration en NaCl avec les critères de performance suivants : le rendement en protéines, la proportion en albumines et la teneur en acide phytique dans le tourteau. Le choix de contraintes sur ces derniers a permis de délimiter, pour chaque température, une zone de conditions opératoires acceptables.

Une condition opératoire par température a été choisie dans ces deux zones pour extraire les protéines et les purifier par ultrafiltration : à 20 °C à pH 2 avec 0,11 M de NaCl et à 55 °C à pH 3,75 avec 0,2 M de NaCl. Les rendements obtenus étaient respectivement de 49 % et 75 %. Les analyses des propriétés fonctionnelles et structurales des deux protéines ainsi purifiées ont montré des résultats similaires. Le procédé d'extraction sélective des albumines est donc favorable à 55 °C, sans altération de la qualité des protéines produites.

III.6 Extrapolation du procédé

Le procédé a été extrapolé par le Groupe Avril. A partir de 150 kg de tourteau, 4,9 kg d'albumines ont été produites. Afin d'abaisser la teneur en acide phytique dans les napines finales, une phytase a été utilisée au cours du procédé. Cela a permis d'atteindre une teneur en acide phytique dans les albumines de 2,7 % (m/m) sur la matière sèche. Un intérêt supplémentaire d'utiliser une phytase est que les albumines et l'acide phytique sont moins susceptibles d'interagir et de former des complexes rendant les albumines moins solubles. Une pureté de 96 % a été obtenue et la solubilité des protéines était de plus de 99 % aux pH 3, 5, 7 et 9. Concernant le tourteau, une teneur en protéines de 27,2 % (N x 5,35) a été obtenue et une proportion d'acide phytique de 4,3 % (m/m) sur la matière sèche.

IV Conclusion du Chapitre IV

Le procédé développé, à partir de l'extraction sélective des albumines, est prometteur, du point de vue des rendements d'extraction et de la qualité des protéines. Il génère deux produits valorisables : un tourteau riche en globulines et utilisable en alimentation animale et un ingrédient protéique aux

caractéristiques (couleur, solubilité, fonctionnalités) désirables pour l'alimentation humaine. De plus, les étapes d'extraction et de purification ont été effectivement extrapolées de l'échelle laboratoire à l'échelle du mètre cube.

L'utilisation de procédés membranaires implique la consommation de grands volumes d'eau ainsi que de sel ce qui génère des effluents liquides. C'est pourquoi, l'optimisation de l'étape de purification est particulièrement importante afin de limiter les quantités d'effluent à traiter.

Les albumines, étant proches des protéines allergènes de la moutarde, font que des études sont requises avant de les mettre sur le marché. Dans cette optique, le Groupe Avril fait partie d'un projet collaboratif avec le Laboratoire BIA (Biopolymères, Interactions, Assemblages). Ces études permettront d'évaluer si les protéines végétales issues du colza présentent un risque allergène pour le consommateur.

Par ailleurs, il est essentiel d'évaluer la digestibilité des albumines du colza afin de les utiliser en alimentation humaine. En effet, la qualité d'un aliment tient également dans sa capacité à être correctement assimilé par l'organisme afin de bénéficier du profil en acides aminés équilibré des protéines du colza. Ceci fait l'objet d'une thèse en cours au LRGP avec le Groupe Avril.

Concernant le tourteau, des études de digestibilité sont également envisagées afin d'estimer le gain obtenu dans l'utilisation d'un tourteau à faible teneur en acide phytique par rapport à un tourteau classique. L'objectif principal de ces études sera d'évaluer le potentiel d'utilisation du tourteau en alimentation animale *a posteriori* de l'extraction des albumines.

**CONCLUSION GENERALE ET
PERSPECTIVES**

I Conclusion

Les protéines de colza sont étudiées par de nombreuses équipes à travers le monde pour exploiter leurs propriétés nutritionnelles et fonctionnelles qui en font d'excellents candidats pour une utilisation en alimentation humaine. Les travaux menés au cours de la thèse ont contribué à faire progresser l'état des connaissances sur le sujet. L'état de l'art sur l'extraction des protéines de colza a montré des limites, notamment méthodologiques. En effet, la plupart des études se sont concentrées sur l'extraction des protéines pour atteindre de hauts rendements sans considération de la qualité du tourteau ni des proportions d'albumine et de globuline extraites.

C'est pourquoi une nouvelle méthode d'analyse a tout d'abord été développée, applicable aux protéines de colza, mais plus globalement aux protéines de plantes oléoprotéagineuses. Cette méthode, par CLHP-ES, permet de quantifier les teneurs des fractions protéiques (albumine ou globuline), de manière fiable et rapide. Elle pourrait servir en complément ou bien en substitution de la méthode de Kjeldahl.

Cet outil analytique a permis d'enrichir les connaissances acquises au cours du développement de stratégies originales d'extractions. Ainsi, deux voies principales ont été mises en œuvre pour produire des protéines issues du tourteau de colza. La première stratégie a été orientée vers une maximisation de l'extraction des protéines par une séquence de deux extractions. Les plans d'expériences et les outils de modélisation ont été utilisés afin de lever les verrous méthodologiques de la littérature. Ce procédé en deux étapes permet d'atteindre de hauts rendements d'extraction, qui dépassent la cible du cahier des charges (60 %). Deux scénarios ont été envisagés. Le premier aboutit à un unique ingrédient protéique, un mélange d'albumines et de globulines. La couleur et la solubilité de ce produit sont compatibles avec pour une utilisation en industrie alimentaire. Le second scénario repose sur l'observation de gamme de pH permettant d'extraire majoritairement l'une ou l'autre des fractions protéiques. Ainsi, la combinaison d'une extraction à pH légèrement alcalin et d'une seconde à pH acide permet de générer deux produits, le premier, en grande quantité, est riche en globuline et l'autre, en faible quantité, est composé d'albumine. Cependant, la caractérisation des produits a montré que les globulines présentent un potentiel de valorisation moins élevé que les albumines, notamment en raison de sa faible solubilité. Dans les deux cas, une faible quantité de tourteau résiduel appauvri en protéines est récupérée à l'issue du procédé.

C'est pourquoi, une seconde stratégie a été élaborée. Etant donné que les albumines ont le plus haut potentiel de valorisation et qu'il est possible de les extraire sélectivement, leur extraction a été étudiée. De nouveau, les plans d'expériences associés aux outils d'optimisation multicritères ont permis de dégager des conditions opératoires optimales de pH, de teneur en sel et de température pour maximiser l'extraction des albumines. Ainsi, il a été montré que la température est un paramètre particulièrement

influent puisque le procédé à 55 °C a permis d'obtenir des rendements d'extraction plus élevés qu'à 20 °C tout en conservant les caractéristiques fonctionnelles des protéines. De légères variations ont été observées au niveau structural ainsi qu'au niveau de la séquence en acides aminés mais sans conséquence sur la fonctionnalité des protéines. Le procédé à 55 °C a ainsi permis d'extraire 75 % des albumines, alors que des travaux précédents rapportaient des rendements de l'ordre de 10 % (Nioi, Kapel, Rondags, & Marc, 2012). De plus, le produit final était caractérisé par une couleur claire, une solubilité élevée à tous les pH et des propriétés moussantes et émulsifiantes d'intérêt pour une utilisation en tant qu'ingrédient.

Au cours des travaux de thèse, un tourteau gras non déshuilé a été utilisé. Ceci est une innovation par rapport aux procédés classiques de la littérature, où un tourteau industriel est souvent utilisé. Comme le montre une étude récente (Fetzer, Herfellner, Stäbler, Menner, & Eisner, 2018), l'utilisation de tourteaux peu exposés à des températures élevées, comme le tourteau gras, permet une augmentation significative des rendements d'extraction des protéines.

II Perspectives

Les travaux réalisés au cours de cette thèse ont permis de développer des procédés originaux aboutissant à des isolats aux propriétés et aux qualités d'intérêt pour l'alimentation humaine. En particulier, le procédé d'extraction sélective des albumines et de purification par ultrafiltration/diafiltration a été extrapolé avec succès à l'échelle du pilote industriel à partir de 75 kg de tourteau, soit environ un mètre cube d'extrait liquide : les rendements et les propriétés de l'isolat obtenu étaient comparables à ce qui a été observé à l'échelle du laboratoire.

Une première perspective de ce travail serait d'améliorer et de poursuivre la caractérisation des isolats, notamment des albumines qui présentent un plus grand potentiel de valorisation. Pour ce faire, les protéines pourraient être incorporées dans des matrices alimentaires afin d'évaluer leur impact sur la texture, la couleur et le goût de produits comme des yaourts, des biscuits ou des émulsions comme la mayonnaise. Dans le cas d'un isolat de protéines totales, composé d'albumines et de globulines, il serait intéressant de caractériser ses propriétés gélifiantes afin d'élargir le champ d'applications. Des études sensorielles devraient permettre de cibler plus précisément les secteurs de marché les plus adaptés pour ces isolats de protéines de colza. Toutes ces études seraient à réaliser en comparaison avec des protéines d'origine variées : qu'elles soient d'origine végétale, comme le soja, le pois ou le blé, ou d'origine animale comme le lait ou les œufs.

Par ailleurs, du fait de la résistance des albumines à la protéolyse gastro-intestinale, des études de digestibilité des isolats doivent être entreprises (Bos, et al., 2007). Compte-tenu de la similarité de structure et de séquence avec les albumines de moutarde (*Brassica juncea*), qui sont connues pour être allergènes, il est essentiel de réaliser une évaluation du potentiel allergénique des albumines de colza

avant la mise sur le marché des isolats. Ces études sont actuellement menées avec des partenaires académiques dans le cadre de projets financés par l'interprofession des oléoprotéagineux et l'Agence Nationale de la Recherche (ANR).

En plus de l'huile et du tourteau, l'intégration des isolats de protéines en tant que produit dans la bioraffinerie industrielle du colza permettrait une plus grande valorisation de la graine. Les travaux menés au cours de la thèse ont permis d'identifier une condition d'extraction des albumines permettant de générer un tourteau résiduel post-extraction avec une teneur en protéines satisfaisante pour une utilisation en alimentation animale et une teneur en acide phytique réduite. Ainsi, la valeur du tourteau résiduel pour l'alimentation animale est au moins égale à celle du tourteau avant extraction. Toutefois, une validation des performances en alimentation animale doit être réalisée (études de digestibilité *in-vitro* avec et *in-vivo*).

Le procédé de production des isolats de protéines, composé d'une extraction solide-liquide et d'une purification des protéines par ultrafiltration/diafiltration, nécessite des volumes d'eau et des quantités de sel très importants. L'utilisation de ce dernier est indispensable au cours de la purification pour dissocier les protéines des molécules anti-nutritionnelles, comme l'acide phytique ou les phénols, permettant finalement de les éliminer et d'obtenir des protéines de qualité. C'est pourquoi, une autre perspective de ce travail consisterait à optimiser le procédé d'ultrafiltration/diafiltration afin de limiter les quantités d'eau et de sel. Les paramètres à prendre en compte au cours de la purification sont très nombreux et une démarche uniquement expérimentale risquerait d'être longue et fastidieuse à mettre en œuvre. Ainsi, la démarche d'étude mise en œuvre au cours de la thèse pourrait être appliquée. Dans un premier temps, les outils de modélisation pourraient être utilisés afin d'établir des corrélations entre les taux de rétention des molécules d'intérêt (les protéines, l'acide phytique et les phénols) et le flux de perméat avec les conditions opératoires de l'ultrafiltration/diafiltration (force ionique, température, pH). Dans un second temps, ces paramètres pourraient être inclus dans des équations de bilans de matière et de flux dans le but de déterminer l'optimum de fonctionnement *via* les algorithmes génético-évolutionnaires. Cet optimum viserait à minimiser les ressources utilisées en fonction de la quantité d'isolat produit. Les performances de différentes configurations de procédés (système multi-étages d'ultrafiltration/diafiltration, contre-courant, recyclage partiel des perméats dilués pour l'extraction etc.) pourraient ainsi être évaluées numériquement avant d'être validées expérimentalement puis extrapolées.

Cependant, même après optimisation, des volumes d'effluents liquides salés seront nécessairement générés. Ces derniers contiennent des molécules hydrosolubles co-extraites avec les protéines et dont la masse molaire est suffisamment faible pour traverser la membrane d'ultrafiltration. Parmi ces molécules, les composés phénoliques, comme l'acide sinapique et l'acide phytique, présentent un potentiel de

valorisation important. Aujourd'hui, 100 g d'acide phytique sous forme de sel de sodium coûte environ 180 € et 100 g d'acide sinapique pur coûte 2800 €. Des procédés d'adsorption ou membranaires peuvent être envisagés pour les capturer et les valoriser. Une démarche de modélisation et de simulation de ces étapes associées aux étapes de purification et d'extraction des protéines doit permettre d'évaluer la pertinence d'un tel procédé. Cette problématique fait actuellement l'objet d'une thèse réalisée au LRGP sur les composés phénoliques.

L'ensemble de ces perspectives permettrait de valoriser pleinement les co-produits de la bioraffinerie du colza. De plus, il sera alors possible de réaliser une analyse technico-économique complète afin d'établir la viabilité d'une telle bioraffinerie originale et la place des produits sur leurs marchés respectifs. En complément, une Analyse du Cycle de Vie (ACV), selon les normes ISO 14040 et ISO 14044, pourra être conduite afin d'évaluer l'impact environnemental de l'ensemble des étapes de production. Cette ACV prend en compte tous les aspects du procédés en termes d'énergie et de matière consommées ainsi que les émissions de déchets et de gaz qui contribueraient à une pollution de l'environnement. La réalisation d'une ACV pour une industrie est essentielle car elle constitue un outil permettant d'identifier des améliorations du procédé tout en conservant une viabilité économique de ce dernier.

Références

Agence Française de Sécurité Sanitaire des Aliments (AFSSA). (2007). Apport en protéines : consommation, qualité, besoins et recommandations. Récupéré sur <https://www.anses.fr/fr/system/files/NUT-Ra-Proteines.pdf>

Aider, M., & Barbana, C. (2011). Canola proteins: composition, extraction, functional properties, bioactivity, applications as a food ingredient and allergenicity - A practical and critical review. *Trends in Food Science & Technology* 22, 21-39.

Akbari, A., & Wu, J. (2015). An integrated method of isolating napin and cruciferin from defatted canola meal. *Food Science and Technology* 64, 308-315.

Ali, F., Ippersiel, D., Lamarche, F., & Mondor, M. (2010). Characterization of low-phytate soy protein isolates produced by membrane technologies. *Innovative Food Science and Emerging Technologies* 11, 162–168.

Aluko, R., & McIntosh, T. (2001). Polypeptide profile and functional properties of defatted meals and protein isolates of canola seeds. *Journal of the Science of Food and Agriculture* 81, 391-396.

ANSES Agence nationale de sécurité sanitaire de l'alimentation, de l'environnement et du travail. (2019, Janvier 27). Ciqual, Table de composition nutritionnelle des aliments. Récupéré sur Ciqual, Table de composition nutritionnelle des aliments: <https://ciqual.anses.fr/#/cms/download/node/20>

Bangur, R., Batey, I., McKenzie, E., & MacRitchie, F. (1997). Dependence of Extensograph Parameters on Wheat Protein Composition Measured by SE-HPLC. *Journal of Cereal Science*, 25, 237–241.

Bhattacharyya, L., & Rohrer, J. (2012). Appendix 2: dissociation constants (pKa) of common sugars and alcohols. Dans *Applications of Ion Chromatography for Pharmaceutical and Biological Products*, First Edition (pp. 455-456). John Wiley & Sons.

Bodin, A., Framboisier, X., Alonso, D., Marc, I., & Kapel, R. (2015). Size-exclusion HPLC as a sensitive and calibrationless method for complex peptide mixtures quantification. *Journal of Chromatography B*, Vol. 1006, 71-79.

Borisjuk, L., Neuberger, T., Schwender, J., Heinzl, N., Sunderhaus, S., Fuchs, J., . . . Rolletschek, H. (2013). Seed Architecture Shapes Embryo Metabolism in Oilseed Rape. *The Plant Cell*, Vol. 25, 1625–1640.

Bos, C., Airinei, G., Mariotti, F., Benamouzig, R., Bérot, S., Evrard, J., . . . Gaudichon, C. (2007). The Poor Digestibility of Rapeseed Protein Is Balanced by Its Very High Metabolic Utilization in Humans. *The Journal of Nutrition*, 594-600.

Boukroufa, M., Sicaire, A.-G., Fine, F., Larré, C., Le Goff, A., Solé Jamault, V., . . . Chemat, F. (2017). Green Sonoextraction of Protein from Oleaginous. *Molecules*, 22, 80, 1-14.

Bye, J., Cowieson, N., Cowieson, A., Selle, P., & Falconer, R. (2013). Dual Effects of Sodium Phytate on the Structural Stability and Solubility of Proteins. *Journal of Agricultural and Food Chemistry*, 61, 290–295.

Canola Council. (2009). *Canola Meal Feed Industry Guide*, 4th edition. Winnipeg: Canola Council of Canada.

Canola council of Canada. (2015). *Canola meal feeding guide*. Winnipeg, Manitoba, Canada: Canola council of Canada.

Chabanon, G. (2005). *Hydrolyses enzymatiques d'isolats protéiques issus de tourteaux de colza : cinétique, modélisation, caractérisation et fonctionnalité des peptides*. Thèse de doctorat.

Chen, M., & Rohani, S. (1992). Recovery of Canola Meal Proteins by Precipitation. *Biotechnology and Bioengineering*. Vol. 40, , 63-68 .

Cheryan, M., & Rackis, J. (1980). Phytic acid in Food Systems. *Critical Reviews in Food Science and Nutrition* 13:4, 297-335.

Cheung, L., Wanasundara, J., & Nickerson, M. (2014). The Effect of pH and NaCl Levels on the Physicochemical and Emulsifying Properties of a Cruciferin Protein Isolate. *Food Biophysics*, 9, 105–113.

Cheung, L., Wanasundara, J., & Nickerson, M. (2015). Effect of pH and NaCl on the Emulsifying Properties of a Napin Protein Isolate. *Food Biophysics*, 10, 30–38.

Chiou, S.-H., & Wu, S.-H. (1999). Evaluation of commonly used electrophoretic methods for the analysis. *Analytica Chimica Acta* 383 , 47-60.

Chung, M., Lei, B., & Li-Chan, E. (2005). Isolation and structural characterization of the major protein fraction from NorMan flaxseed (*Linum usitatissimum* L.). *Food Chemistry* 90, 271–279.

Crouch, M., & Sussex, I. (1981). Development and storage-protein synthesis in *Brassica napus* L. embryos in vivo and in vitro. *Planta*, 153, 64-74.

Das Purkayastha, M., & Lata Mahanta, C. (2015). Statistically designed optimal process conditions for recuperation of protein from rapeseed meal. *Journal of Food Science and Technology*, 52, 6, 3203–3218.

De Rham, O., & Jost, T. (1979). Phytate-protein interactions in soybean extracts and low-phytate soy protein products. *Journal of Food Science*, Vol. 44, No. 2, 596-600.

Deak, N., Murphy, P., & Johnson, L. (2006). Effects of NaCl Concentration on Salting-in and Dilution During Salting-out on Soy Protein Fractionation. *Journal of Food Science C: Food Chemistry and Toxicology* Vol. 71, Nr. 4, 247-254.

Defaix, C., Aymes, A., Albe Slabi, S., Basselin, M., Mathé, C., Galet, O., & Kapel, R. (2019). A new size-exclusion chromatography method for fast rapeseed albumin and globulin quantification. *Food Chemistry*, Vol. 287, 151-159.

Direction de l'information légale et administrative (Premier ministre). (Consulté le 16 janvier 2019). Code de la consommation. Sous-section 2 : Information relative à la présence de substances ou produits provoquant des allergies ou intolérances. Articles R412-12 à R412-16 et annexe IV.

EFSA (European Food Safety Authority) Panel Contaminants Food. (2016). Erucic acid in feed and food. *EFSA Journal*, 1-173.

Faergemand, M., & Qvist, K. (1997). Transglutaminase: effect on rheological properties, microstructure and permeability of set style acid skim milk gel. *Food Hydrocolloids*, Vol. 11 no. 3, 287-292.

Fagerquist, C., Sultan, O., & Carter, M. (2012). Possible Evidence of Amide Bond Formation Between Sinapinic Acid and Lysine-Containing Bacterial Proteins by Matrix-Assisted Laser Desorption/Ionization (MALDI) at 355 nm. *Journal of the American Society for Mass Spectrometry*, 23, 2102-2114.

FAOStats. (2019). Food and Agriculture Organization, of the United Nations. Consulté le April 3rd, 2018, sur <http://www.fao.org/faostat>

Fetzer, A., Herfellner, T., & Eisner, P. (2019). Rapeseed protein concentrates for non-food applications prepared from pre-pressed and cold-pressed press cake via acidic precipitation and ultrafiltration. *Industrial Crops & Products*, Vol.132, 396–406.

Fetzer, A., Herfellner, T., Stäbler, A., Menner, M., & Eisner, P. (2018). Influence of process conditions during aqueous protein extraction upon yield from pre-pressed and cold-pressed rapeseed press cake. *Industrial Crops and Products* 112, 236-246.

Feyzi, S., Varidi, M., Zare, F., & Varidi, M. (2015). Fenugreek (*Trigonella foenum graecum*) seed protein isolate: extraction optimization, amino acid composition, thermo and functional properties. *Journal of the Science of Food and Agriculture*, 95, 3165–3176.

Fine, F., Abert Vian, M., Fabiano Tixier, A.-S., Carre, P., Pages, X., & Chemat, F. (2013). Les agro-solvants pour l'extraction des huiles végétales issues de graines oléagineuses. *Oilseeds & fat Crops and Lipids*, 20 (5), 1-6.

Food and Agriculture Organization of the United Nation (FAO). (2019, January 17). FAOSTAT. Récupéré sur <http://www.fao.org/faostat/>

Garcia-Esteva, R., Guerra-Hernandez, E., & Garcia-Villanova, B. (1999). Phytic acid content in milled cereal products and breads. *Food Research International* 32 , 217-221.

Georgiou, C., Grintzalis, K., Zervoudakis, G., & Papapostolou, I. (2008). Mechanism of Coomassie brilliant blue G-250 binding to proteins: A hydrophobic assay for nanogram quantities of proteins. *Analytical and bioanalytical chemistry*. 391, 1, 391-403.

Gerzhova, A., Mondor, M., Benali, M., & Aider, M. (2016). Study of total dry matter and protein extraction from canola meal as affected by the pH, salt addition and use of zeta-potential/turbidimetry analysis to optimize the extraction conditions. *Food Chemistry* 201, 243–252.

Gibney, M., Lanham-New, S., Cassidy, A., & Vorster, H. (2009). *Introduction to Human Nutrition*, Second Edition. Oxford: Wiley-Blackwell.

Gonzalez-Pérez, S., & Vereijken, J. (2007). Sunflower proteins: overview of their physicochemical, structural and functional properties. *Journal of the Science of Food and Agriculture* 87, 2173–2191.

GraceVydac. (2002). *The Handbook of Analysis and Purification of Peptides and Proteins by Reversed-Phase HPLC*.

Guéguen , J., Walrand, S., & Bourgeois, O. (2016). Les protéines végétales : contexte et potentiels en alimentation humaine. *Cahiers de nutrition et de diététique* 51, 177—185.

Hald, C., Dawid, C., Tressel, R., & Hofmann, T. (2019). Kaempferol 3-O-(2'''-O-sinapoyl- β -sophoroside) Causes the Undesired Bitter Taste of Canola/Rapeseed Protein Isolates. *Journal of Agricultural and Food Chemistry*, Vol. 67, 372-378.

Hoffmann, B., Münch, S., Schwägele, F., Neusüss, C., & Jira, W. (2017). A sensitive HPLC-MS/MS screening method for the simultaneous detection of lupine, pea, and soy proteins in meat products. *Food Control*, Vol. 71, 200-209.

Hu, Z.-Y., Hua, W., Zhang, L., Deng, L.-B., Wang, X.-F., Liu, G.-H., . . . Wang, H.-Z. (2013). Seed Structure Characteristics to Form Ultrahigh Oil Content in Rapeseed. *PLoS ONE*, Vol. 8, no. 4, 1-10.

INRA, CIRAD, & Association Française de Zootechnie. (2019). Tourteau de colza, huile 5-20 %. Récupéré sur [Feedtables.com](https://feedtables.com/fr/content/tourteau-de-colza-huile-5-20): <https://feedtables.com/fr/content/tourteau-de-colza-huile-5-20>

Jung, S., Rickert, D., Deak, N., Aldin, E., Recknor, J., Johnson, L., & Murphy, P. (2003). Comparison of Kjeldahl and Dumas Methods for Determining Protein Contents of Soybean Products. *Journal of the American Oil Chemists' Society*, Vol. 80, no. 12, 1169-1173.

Kaspchak, E., Igarashi Mafra, L., & Rogério Mafra, M. (2018). Effect of heating and ionic strength on the interaction of bovine serum albumin and the antinutrients tannic and phytic acids, and its influence on in vitro protein digestibility. *Food Chemistry* 252, 1–8.

Kelly, S., & Price, N. (2000). The Use of Circular Dichroism in the Investigation of Protein Structure and Function. *Current Protein and Peptide Science*, 1, 349-384.

Klockeman, D., Toledo, R., & Sims, K. (1997). Isolation and Characterization of Defatted Canola Meal Protein. *Journal of Agricultural and Food Chemistry*, Vol. 45, 3867-3870.

Kolster, P., Krechting, C., & Van Gelder, W. (1992). Quantification of Individual High Molecular Weight Subunits of Wheat Glutenin Using SD8-PAGE and Scanning Densitometry. *Journal of Cereal Science*, 15, 49-61.

Krause, J.-P., & Schwenke, K. (2001). Behaviour of a protein isolate from rapeseed (*Brassica napus*) and its main protein components — globulin and albumin — at air/solution and solid interfaces, and in emulsions. *Colloids and Surfaces B: Biointerfaces* 21, 29-36.

Kroll, J. (1991). Selected functional properties of detoxified rapeseed protein preparations effected by phytic acid. *Die Nahrung* 35, 6, 619-624 .

Kuipers, B., & Gruppen, H. (2007). Prediction of Molar Extinction Coefficients of Proteins and Peptides Using UV Absorption of the Constituent Amino Acids at 214 nm To Enable Quantitative Reverse Phase High-Performance Liquid Chromatography-Mass Spectrometry Analysis. *Journal of Agricultural and Food Chemistry*, 5445-5451.

Laemmli, U. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, Vol. 227, 680-685.

Li-Chan, E. (2004). Properties of proteins in food systems. *Dans Proteins in food processing* (pp. 19-43). Abington: CRC Press.

Lindeboom, N., & Wanasundara, P. (2007). Interference of phenolic compounds in *Brassica napus*, *Brassica rapa* and *Sinapis alba* seed extracts with the Lowry protein assay. *Food Chemistry*, Vol. 104, 30-38.

- Logie, J., & Milanova, R. (2010). United States Brevet n° US 2010/0063255 A1.
- Lonnerdal, B., & Janson, J.-C. (1972). Studies on Brassica seed proteins. I. The low molecular weight proteins in rapeseed. Isolation and characterization. *Biochimica et Biophysica Acta* 36198, 175-183.
- Lopes Barbosa, A., Lajolo, F., & Genovese, M. (2006). Influence of temperature, pH and ionic strength on the production of isoflavone-rich soy protein isolates. *Food Chemistry*, Vol. 98, 757-766.
- Lowry, O., Rosebrough, N., Farr, A., & Randall, R. (1951). Protein measurement with the Folin reagent. *Journal of Biological Chemistry*, 193, 265-275.
- Luykx, D., Cordewener, J., Ferranti, P., Frankhuizen, R., Bremer, M., Hooijerink, H., & America, A. (2007). Identification of plant proteins in adulterated skimmed milk powder by high-performance liquid chromatography—mass spectrometry. *Journal of Chromatography A*, Vol. 1164, 189-197.
- M**alvern Instruments. (Consulté en février 2019). DLS an introduction in 30 minutes. Malvern Instruments Technical Note. Récupéré sur <https://warwick.ac.uk/>
- Mariotti, F., Tomé, D., & Patureau Mirand, P. (2008). Converting Nitrogen into Protein—Beyond 6.25. *Critical Reviews in Food Science and Nutrition*, 177-184.
- Martin, C., & Evans, W. (1986). Phytic Acid-Metal Ion Interactions . II . The Effect of pH on Ca(II) Binding. *Journal of Inorganic Biochemistry* 27, 17-30.
- Matsumura, Y., Mitsui, S., Dickinson, E., & Mori, T. (1994). Competitive adsorption of α -lactalbumin in the molten globule state. *Food Hydrocolloids* Vol.8 no.6, 555-566 .
- Mavrotas, G. (2009). Effective implementation of the e-constraint method in Multi-Objective Mathematical Programming problems. *Applied Mathematics and Computation*, Vol. 213, 455–465.
- Mezzenga, R., Schurtenberger, P., Burbidge, A., & Michel, M. (2005). Understanding foods as soft materials. *Nature Materials*, Vol. 4, 729-740.
- Mills, E., Madsen, C., Shewry, P., & Wichers, H. (2003). Food allergens of plant origin - their molecular and evolutionary relationships. *Trends in Food Science & Technology*, 14, 145–156.
- Moore, J., DeVries, J., Lipp, M., Griffiths, J., & Abernethy, D. (2010). Total Protein Methods and Their Potential Utility to Reduce the Risk of Food Protein Adulteration. *Comprehensive Reviews in Food Science and Food Safety*, Vol. 9, 330-357.
- Mossé, J., Huet, J., & Baudet, J. (1985). The Amino Acid Composition of Wheat Grain as a Function of Nitrogen Content. *Journal of Cereal Science* 3, 115-130.

Mothes, R., Schwenke, K., Zirwer, D., & Gast, K. (1990). Rapeseed protein - polyanion interactions. Soluble complexes between the 2 S protein fraction (napin) and phytic acid. *Die Nahrung* 34, 4, 375-385.

Muniglia, L., Kiss, L., Fonteix, C., & Marc, I. (2004). Multicriteria optimization of a single-cell oil production. *European Journal of Operational Research* 153, 360–369.

Murray, E., Myers, C., & Barker, L. (1979). *Etats-Unis Brevet n° 4,169,090*.

Naczk, M., Amarowicz, R., Sullivan, A., & Shahidi, F. (1998). Current research developments on polyphenolics of rapeseed/canola: a review. *Food Chemistry*, Vol. 62, No. 4, 489-502.

Naczk, M., Diosady, L., & Rubin, L. (1985). Functional Properties of Canola Meals Produced by a Two-phase Solvent Extraction System. *Journal of Food Science*, Vol 50, 1685-1688.

Nicolai, T., Britten, M., & Schmitt, C. (2011). Beta-Lactoglobulin and WPI aggregates: Formation, structure and applications. *Food Hydrocolloids*, Vol. 25, 1945-1962.

Nioi, C., Kapel, R., Rondags, E., & Marc, I. (2012). Selective extraction, structural characterisation and antifungal activity assessment of napins from an industrial rapeseed meal. *Food Chemistry*, Vol. 134, Issue 4, 2149-2155.

Nwachukwu, I., & Aluko, R. (2018). Physicochemical and emulsification properties of flaxseed (*Linum usitatissimum*) albumin and globulin fractions. *Food Chemistry* 255, 216-225.

Osborne, T. (1897). The amount and properties of the proteins of the maize kernel. *Journal of the American Chemical Society*, 525-532.

Pawlak, Z. (1982). Rough Sets. *International Journal of Computer and Information Sciences* Vol. 11, No. 5, 341-356.

Pecora, R. (2000). Dynamic light scattering measurement of nanometer particles in liquids. *Journal of Nanoparticle Research*, Vol. 2, 123-131.

Pedroche, J., Yust, M., Lqari, H., Giron-Calle, J., Alaiz, M., Vioque, J., & Millán, F. (2004). Brassica carinata protein isolates: chemical composition, protein characterization and improvement of functional properties by protein hydrolysis. *Food Chemistry* 88, 337–346.

Pelegrine, D., & Gasparetto, C. (2005). Whey proteins solubility as function of temperature and pH. *Lebensmittel-Wissenschaft & Technologie*, Vol. 38, 77–80.

Perera, S., McIntosh, T., & Wanasundara, J. (2016). Structural Properties of Cruciferin and Napin of *Brassica napus* (Canola) Show Distinct Responses to Changes in pH and Temperature. *Plants*, 5, 36, 1-24.

Pickardt, C., Neidhart, S., Griesbach, C., Dube, M., Knauf, U., Kammerer, D., & Carle, R. (2009). Optimisation of mild-acidic protein extraction from defatted sunflower (*Helianthus annuus* L.) meal. *Food Hydrocolloids* 23 , 1966–1973.

Pinterits, A., & Arntfield, S. (2008). Improvement of canola protein gelation properties through enzymatic modification with transglutaminase. *Lebensmittel-Wissenschaft & Technologie*, Vol. 41, 128-138.

Purkayastha, M., Gogoi, J., Kalita, D., Chattopadhyay, P., Nakhuru, K., Goyary, D., & Mahanta, C. (2014). Physicochemical and Functional Properties of Rapeseed Protein Isolate: Influence of Antinutrient Removal with Acidified Organic Solvents from Rapeseed Meal. *Journal of Agricultural and Food Chemistry* 62, 7903–7914.

Pustjens, A., Schols, H., Kabel, M., & Gruppen, H. (2013). Characterisation of cell wall polysaccharides from rapeseed (*Brassica napus*) meal. *Carbohydrate Polymers* 98, 1650–1656.

Puumalainen, T., Poikonen, S., Kotovuori, A., Vaali, K., Kalkkinen, N., Reunala, T., . . . Palosuo, T. (2005). Napins, 2S albumins, are major allergens in oilseed rape and turnip rape. *The Journal of Allergy and Clinical Immunology*, Vol. 17, no. 2, 426-432.

Radwan , M., & Lu, B.-Y. (1976). Solubility of Rapeseed Protein in Aqueous Solutions. *Journal of the American Oil Chemists' Society*, Vol. 53, 142-144.

Rehder, A., Sulewska, A., Markedal, K., Sørensen , S., & Sørensen, J. (2017). Solubility of a cruciferin-rich protein product purified from rapeseed pressed cake (*Brassica napus* L.) by an aqueous processing method. *International Journal of Food Science and Technology* , 1-7.

Ritzel, B., Eheart, J., & Ranjithan, S. (1994). Using genetic algorithms to solve a multiple objective groundwater pollution containment problem. *Water resources Research*, Vol. 30, No. 5, 589-603.

Rodríguez-Nogales, J., Cifuentes, A., García, M., & Marina, M. (2007). Characterization of Protein Fractions from Bt-Transgenic and Non-transgenic Maize Varieties Using Perfusion and Monolithic RP-HPLC. Maize Differentiation by Multivariate Analysis. *Journal of Agricultural and Food Chemistry*, 3835-3842.

Rommi, K., Hakala, T., Holopainen, U., Nordlund, E., Poutanen, K., & Lantto, R. (2014). Effect of Enzyme-Aided Cell Wall Disintegration on Protein Extractability from Intact and Dehulled Rapeseed (*Brassica rapa* L. and *Brassica napus* L.) Press Cakes. *Journal of Agricultural and Food Chemistry*, 62, 7989–7997.

Rous, J.-F., & Groupe Avril. (2016). De l'oléochimie à la bioraffinerie : continuité de développement pour le secteur des corps gras ? *Oilseeds & fat Crops and Lipids*, 23 (5), 1-7.

Sáez-Plaza, P., Michałowski, T., Navas, M. J., Asuero, A. G., & Wybraniec, S. (2013). An Overview of the Kjeldahl Method of Nitrogen Determination. Part I. Early History,. *Critical Reviews in Analytical Chemistry* 43, 178-223.

Sarwar, G., Christensen, D., Finlayson, A., Friedman, M., Hackler, L., Mackenzie, S., . . . Tkachuk, R. (1983). Inter- and intra-laboratory variation in amino acid analysis of food proteins. *Journal of Food Science*, Vol. 48, 526-531.

Sathe, S., & Salunkhe, D. (1981). Functional Properties of the Great Northern Bean (*Phaseolus vulgaris* L.) Proteins: Emulsion, Foaming, Viscosity, and Gelation Properties. *Journal of Food Science*, Vol 46, 71-81.

Schweizer, M., Green, B., Segall, K., & Willardsen, R. (2013). US Brevet n° 8,541,046 B2.

Schwenke, K. (1994). Rapeseed proteins. Dans B. J. Hudson, *New and Developing Sources of Food Proteins* (pp. 281-306). Springer-Science+Business Media B.V.

Schwenke, K., Dahme, A., & Wolter, T. (1998). Heat-Induced Gelation of Rapeseed Proteins: Effect of Protein Interaction and Acetylation. *Journal of the American Oil Chemists' Society*, Vol. 75, no. 1, 83-87.

Schwenke, K., Raab , B., Plietz , P., & Damaschun, G. (1983). The structure of the 12 S globulin from rapeseed. *Nahrung*, 27, 165-175.

Shahidi, F. (1990). *Canola and Rapeseed*. New York: Springer Science+Business Media.

Shahidi, F., & Naczk, M. (1992). An Overview of the Phenolics of Canola and Rapeseed: Chemical, Sensory and Nutritional Significance. *Journal of the American Oil Chemists' Society*, Vol. 69, no. 9, 917-924.

Sicaire, A.-G., Vian, M., Fine, F., Joffre, F., Carré, P., Tostain, S., & Chemat, F. (2015). Alternative Bio-Based Solvents for Extraction of Fat and Oils: Solubility Prediction, Global Yield, Extraction Kinetics, Chemical Composition and Cost of Manufacturing. *International Journal of Molecular Sciences*, Vol.16 (4) , 8430-8453.

Singh, H. (2005). A study of changes in wheat protein during bread baking using SE-HPLC. *Food Chemistry* 90, 247–250.

Singh, N. (2017). Pulses: an overview. *Journal of Food Science and Technology*, 54, 4, 853-857.

Snatzke, G. (1968). Circular Dichroism and Optical Rotatory Dispersion - Principles and Application to the Investigation of the Stereochemistry of Natural Products. *Angew. Chem. internat. Edit. I* Vol. 7 (1968) / No. 1, 14-25.

Sosulski, F. (1979). Organoleptic and nutritional effects of phenolic compounds on oilseed protein products: A review. *Journal of the American Oil Chemists' Society* 56(8), 711-715.

Stepanchenko, N., Novikova, G., & Mosh, I. (2011). Protein Quantification. *Russian Journal of Plant Physiology* Vol. 58 No. 4, 737-742.

Tan, S., Mailer, R., Blanchard, C., & Agboola, S. (2011). Canola Proteins for Human Consumption: Extraction, Profile, and Functional Properties. *Journal of Food Science*, Vol. 76, R16-R28.

Tang, C.-H., Ten, Z., Wang, X.-S., & Yang, X.-Q. (2006). Physicochemical and Functional Properties of Hemp (*Cannabis sativa* L.) Protein Isolate. *J. Agric. Food Chem.* 54, 8945–8950.

Terres OléoPro. (2019). Alimentation animale. Récupéré sur Terres OléoPro: <https://www.terresoleopro.com/nos-produits/alimentation-animale>

Terres Univia. (2017). Sauvons le colza français. Terres Univia.

Terres Univia. (2019). Colza. Récupéré sur terresunivia.fr: <http://www.terresunivia.fr/cultures-utilisation/les-especes-cultivees/colza>

Tkachuk, R. (1969). Nitrogen-to-Protein Conversion Factors for Cereals and Oilseed Meals. *Cereal Chemistry*, 46, 419-423.

Tzeng, Y.-M., Diosady, L., & Rubin, L. (1990). Production of Canola Protein Materials by Alkaline Extraction, Precipitation, and Membrane Processing. *Journal of Food Science*, Volume 55, No. 4, 1147.

Tzeng, Y.-M., Diosady, L., & Rubin, L. (1988). Preparation of Rapeseed Protein Isolates Using Ultrafiltration, Precipitation and Diafiltration. *Canadian Institute of Food Science and Technology Journal*, Vol. 21, No. 4, 419-424.

UNdata. (2019, Août). UNdata | record view | Biodiesel. Récupéré sur UNdata: <http://data.un.org/Data.aspx?d=EDATA&f=cmID%3ABD>

Uruakpa, F., & Arntfield, S. (2005). Emulsifying characteristics of commercial canola protein–hydrocolloid systems. *Food Research International*, Vol.38, 659–672.

Validation of Analytical Procedures: Text and Methodology Q2(R1) . (2005). International Conference on Harmonisation of technical requirements for registration of pharmaceuticals for human use , (pp. 1-17).

Vioque, J., Sánchez-Vioque, R., Clemente, A., Pedroche, J., & Millán, F. (2000). Partially Hydrolyzed Rapeseed Protein Isolates with Improved Functional Properties. *Journal of the American Oil Chemistry Society*, Vol. 77, no. 4, 447-450.

Wanasundara, J. P., & McIntosh, T. C. (2008). World Patent No. 144939 A1.

Wanasundara, J., McIntosh, T., Perera, S., Withana-Gamage, T., & Mitra, P. (2016). How might oilseeds help meet the protein challenge? - Canola/rapeseed protein-functionality and nutrition. *Oilseeds & fats Crops and Lipids*, 23 (4) D407, 1-15.

Wingfield, P. (2001). Protein Precipitation Using Ammonium Sulfate. *Current Protocols in Protein Science*, Appendix, A.3F.1-A.3F.8.

World Health Organization. (2007). WHO Technical Report Series 935. Protein and amino acid requirements in human nutrition: report of a joint FAO/WHO/UNU expert consultation. Geneva, Switzerland: World Health Organization.

Wu, J., & Muir, A. (2008). Comparative Structural, Emulsifying, and Biological Properties of 2 Major Canola Proteins, Cruciferin and Napin. *Journal of Food Science*, Vol. 73, Nr. 3, 210-216.

Xu, L., & Diosady, L. (2002). Removal of phenolic compounds in the production of high-quality canola protein isolates. *Food Research International* 35, 23–30.

Yang, S.-Y., Huang, T.-K., Kuo, H.-F., & Chiou, T.-J. (2017). Role of vacuoles in phosphorus storage and remobilization. *Journal of Experimental Botany*, Vol. 68, no. 12, 3045–3055.

Young, V., & Pellett, P. (1994). Plant proteins in relation to human protein and amino acid nutrition. *The American Journal of Clinical Nutrition*, Vol. 59, no. 5, 1203S-1212S.

Développement de stratégies originales d'extraction des protéines du tourteau de colza pour la production d'isolats par une démarche de modélisation / optimisation multicritères

Les protéines sont des nutriments majeurs de l'alimentation : elles sont consommées via l'ingestion de produits qui en sont naturellement riches, ou bien à travers de formulations alimentaires, où les protéines jouent un rôle d'ingrédient. Depuis les années 1980, des recherches sont menées afin de valoriser les protéines du tourteau de colza dans l'alimentation humaine, en tant qu'ingrédient, alors qu'actuellement cette ressource est utilisée pour l'alimentation animale. Dans le cadre d'une démographie mondiale toujours croissante, la valorisation des protéines du colza dans l'alimentation humaine est investiguée comme un moyen de satisfaire les besoins mondiaux en protéines. Le tourteau de colza contient deux types de protéines de stockage, les albumines et les globulines, dont le profil est équilibré en acides aminés soufrés et en lysine. De plus, afin de les utiliser en tant qu'ingrédient, il est important que ces protéines soient solubles et fonctionnelles (formation de mousses, émulsions, gels). Les propriétés physico-chimiques des albumines et des globulines sont différentes, d'où l'intérêt d'en faire deux ingrédients distincts. L'objectif de la thèse a donc consisté au développement de stratégies d'extraction de ces protéines afin de les valoriser en tant qu'ingrédient, avec un grade d'isolat. Pour ce faire, une démarche reposant sur la modélisation via les plans d'expériences, et sur l'optimisation multicritères, via des algorithmes génético-évolutionnaires, a été mise en place afin de mieux appréhender les phénomènes en jeu au cours de l'extraction.

La première partie de la thèse a consisté au développement d'une méthode analytique permettant de quantifier individuellement les deux protéines. Cette méthode, par Chromatographie Liquide Haute Performance d'exclusion stérique, est originale et ne nécessite pas de calibration préalable. Il a été validé statistiquement qu'il s'agit d'un substitut viable aux références (Kjeldahl, densitométrie SDS-PAGE).

Les procédés classiques sont réalisés en conditions alcalines pour atteindre de hauts rendements d'extraction. Les protéines sont ensuite séparées et purifiées, afin d'obtenir deux isolats aux propriétés fonctionnelles distinctes. Cependant, ces étapes entraînent des pertes non négligeables et des coûts. Ainsi, la deuxième partie de la thèse présente une stratégie reposant sur l'extraction sélective des protéines selon le pH. Les outils de modélisation et d'optimisation ont permis d'identifier des conditions opératoires (pH, ajout de sel) aboutissant à un procédé en deux étapes d'extraction. Deux scénarios ont été envisagés : la production d'un unique isolat, mélange des deux protéines ; ou bien la production de deux isolats, d'albumines et de globulines. Une originalité de ce travail a été de considérer la qualité du tourteau résiduel. A l'issue du procédé, il est très appauvri en protéines et peu valorisable en alimentation animale.

C'est pourquoi, dans une troisième partie de la thèse, une nouvelle stratégie a été mise en œuvre pour optimiser l'extraction sélective des albumines à pH acide, afin de générer deux produits : un isolat d'albumines, pour l'alimentation humaine, et un tourteau riche en globulines, pour l'alimentation animale. L'utilisation d'outils de modélisation et d'optimisation multicritères ont permis d'étudier trois facteurs : le pH, la concentration en sel et la température du milieu d'extraction. Des rendements élevés ont été obtenus avec un procédé à 55 °C et l'extraction en conditions optimales a généré des albumines de colza aux propriétés structurales et fonctionnelles, comparables à celles issues d'un procédé à 20 °C. De plus, le tourteau résiduel conserve sa valeur avec un titre en protéines similaire à un tourteau initial, tout en contenant moins d'acide phytique, un antinutritionnel, co-extrait avec les albumines.

Mots clés : tourteau, protéines, colza, albumines, globulines, chromatographie, extraction, modélisation, optimisation multicritères.

Development of original strategies for the extraction of protein from rapeseed meal for the production of isolates by modelling and multicriteria optimization

Proteins are major nutrients for nutrition: they are consumed by the ingestion of products naturally rich in proteins, or by food formulation, where proteins are ingredients. Since the 1980s, research projects are carried out to increase the value of rapeseed meal and to use its proteins in human nutrition, as an ingredient, instead of being used for animal nutrition only. In the context of the ever-increasing demography, the use of rapeseed protein has been investigated as a mean to fulfill the needs in proteins worldwide. Rapeseed meal is composed of two main storage proteins, albumin and globulin, with an amino acid profile balanced regarding sulfuric and lysine residues. Moreover, to be used as ingredient, it is crucial that these proteins are soluble and functional (to form foams, emulsions, gels). The physico-chemical properties of albumin and globulin are different. That is why, it is of interest to make two distinct ingredients. The main objective of the thesis was to develop strategies to extract these proteins to use it as ingredients (isolate grade). To do so, an approach based on modelling by design of experiments, and on multicriteria optimization through genetical-evolutionary algorithms, was applied to better understand the phenomenon at play during extraction.

The first part of the thesis consisted in the development of an analytical method that would allow the separate quantification of both protein contents. This method, developed by High Performance Liquid Chromatography, Size-Exclusion, is original and is calibrationless. Statistics were used to prove the possible substitution of the reference techniques (Kjeldahl, SDS-PAGE densitometry) by this method.

Classical processes of extraction are carried out under alkaline conditions to reach high yields. Then, proteins are separated and purified to give two isolates with distinct functional properties. However, these steps imply non-negligible losses and costs. Hence, the second part of the thesis presents a strategy based on selective extraction of proteins with the pH. Modelling and optimization tools were used to identify operating conditions (pH, salt added) resulting in a two-step process. Two scenarios were considered: the production of a unique isolate, a mix of both proteins; or the production of two isolates, albumin and globulin-rich. An original feature of the work was the consideration of the quality of the residual meal. At the end of this process, the meal is very poor in proteins and thus, has low value for animal nutrition.

That is why, the third part of the thesis consisted in developing a new strategy to optimize the selective extraction of albumin at acidic pH, to yield two products: a protein isolate, for human nutrition, and a meal rich in globulins for animal nutrition. Modelling and optimization tools were used to assess the impact of three factors: pH, salt content and temperature of the extraction medium. High yields were obtained at 55°C, and the optimal conditions resulted in rapeseed albumin with comparable properties (functional and structural) than at 20°C. Moreover, the residual meal was still valuable with a protein content similar to the initial meal and with less phytic acid, an antinutritional molecule, co-extracted with albumin.

Keywords: proteins, rapeseed, meal, albumin, globulin, chromatography, extraction, modelling, multicriteria optimization