

**THESE
POUR LE DIPLOME D'ETAT
DE DOCTEUR EN PHARMACIE**

**Soutenue publiquement le 4 Juillet 2024
Par Mme Moum Leïla**

**LES DOMAINES EGF-LIKE DES PROTEINES MEMBRANAIRES
CLASSIFICATION STRUCTURALE, IMPLICATIONS
PHYSIOPATHOLOGIQUES ET INTERET THERAPEUTIQUE**

Membres du jury :

Président : Professeure Rebecca Deprez-Poulain - Docteur en pharmacie et Professeure des Universités à la faculté de pharmacie de Lille

Assesseurs : Professeure Réjane Lestrelin – Professeure des Universités à la faculté de pharmacie de Lille ; **Professeur Nicolas Lebègue** – Professeur des Universités à la faculté de pharmacie de Lille et Directeur de Thèse

Membre extérieur : Docteur Benoit Gosselin – Docteur en pharmacie, titulaire d'officine

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 1/9

REDACTION	VERIFICATION	APPROBATION
Audrey Hennebelle Assistante de direction	Cyrille Porta Responsable des Services	Delphine Allorge Doyen

Université de Lille

Président
Premier Vice-président
Vice-présidente Formation
ROBACZEWSKI
Vice-président Recherche
Vice-président Ressources humaines
Directrice Générale des Services

Régis BORDET
Etienne PEYRAT
Corinne

Olivier COLOT
Bertrand DÉCAUDIN
Anne-Valérie CHIRIS-FABRE

UFR3S

Doyen
Premier Vice-Doyen, Vice-Doyen RH, SI et Qualité
Vice-Doyenne Recherche
Vice-Doyen Finances et Patrimoine
Vice-Doyen International
Vice-Doyen Coordination pluriprofessionnelle et Formations sanitaires
Vice-Doyenne Formation tout au long de la vie
Vice-Doyen Territoire-Partenariats
Vice-Doyen Santé numérique et Communication
Vice-Doyenne Vie de Campus
Vice-Doyen étudiant

Dominique LACROIX
Hervé HUBERT
Karine FAURE
Damien CUNY
Vincent DERAMECOURT
Sébastien D'HARANCY
Caroline LANIER
Thomas MORGENROTH
Vincent SOBANSKI
Anne-Laure BARBOTIN
Valentin ROUSSEL

Faculté de Pharmacie

Doyen
Premier Assesseur et
Assesseur à la Santé et à l'Accompagnement
Assesseur à la Vie de la Faculté et
Assesseur aux Ressources et Personnels
Responsable des Services
Représentant étudiant
Chargé de mission 1er cycle
Chargée de mission 2eme cycle
Chargé de mission Accompagnement et Formation à la Recherche
Chargé de mission Relations Internationales
Chargée de Mission Qualité
Chargé de mission dossier HCERES

Delphine ALLORGE

Anne GARAT

Emmanuelle LIPKA
Cyrille PORTA
Honoré GUISE
Philippe GERVOIS
Héloïse HENRY
Nicolas WILLAND
Christophe FURMAN
Marie-Françoise ODOU
Réjane LESTRELIN

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 2/9

Professeurs des Universités - Praticiens Hospitaliers (PU-PH)

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
Mme	ALLORGE	Delphine	Toxicologie et Santé publique	81
M.	BROUSSEAU	Thierry	Biochimie	82
M.	DÉCAUDIN	Bertrand	Biopharmacie, Pharmacie galénique et hospitalière	81
M.	DINE	Thierry	Pharmacologie, Pharmacocinétique et Pharmacie clinique	81
Mme	DUPONT-PRADO	Annabelle	Hématologie	82
Mme	GOFFARD	Anne	Bactériologie - Virologie	82
M.	GRESSIER	Bernard	Pharmacologie, Pharmacocinétique et Pharmacie clinique	81
M.	ODOU	Pascal	Biopharmacie, Pharmacie galénique et hospitalière	80
Mme	POULAIN	Stéphanie	Hématologie	82
M.	SIMON	Nicolas	Pharmacologie, Pharmacocinétique et Pharmacie clinique	81
M.	STAELS	Bart	Biologie cellulaire	82

Professeurs des Universités (PU)

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
M.	ALIOUAT	El Moukhtar	Parasitologie - Biologie animale	87
Mme	AZAROUAL	Nathalie	Biophysique - RMN	85
M.	BERLARBI	Karim	Physiologie	86
M.	BERTIN	Benjamin	Immunologie	87
M.	BLANCHEMAIN	Nicolas	Pharmacotechnie industrielle	85
M.	CARNOY	Christophe	Immunologie	87
M.	CAZIN	Jean-Louis	Pharmacologie, Pharmacocinétique et Pharmacie clinique	86

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 3/9

M.	CUNY	Damien	Sciences végétales et fongiques	87
Mme	DELBAERE	Stéphanie	Biophysique - RMN	85
Mme	DEPREZ	Rebecca	Chimie thérapeutique	86
M.	DEPREZ	Benoît	Chimie bioinorganique	85
M.	DURIEZ	Patrick	Physiologie	86
M.	ELATI	Mohamed	Biomathématiques	27
M.	FOLIGNÉ	Benoît	Bactériologie - Virologie	87
Mme	FOULON	Catherine	Chimie analytique	85
M.	GARÇON	Guillaume	Toxicologie et Santé publique	86
M.	GOOSSENS	Jean-François	Chimie analytique	85
M.	HENNEBELLE	Thierry	Pharmacognosie	86
M.	LEBEGUE	Nicolas	Chimie thérapeutique	86
M.	LEMDANI	Mohamed	Biomathématiques	26
Mme	LESTAVEL	Sophie	Biologie cellulaire	87
Mme	LESTRELIN	Réjane	Biologie cellulaire	87
Mme	LIPKA	Emmanuelle	Chimie analytique	85
Mme	MELNYK	Patricia	Chimie physique	85
M.	MILLET	Régis	Institut de Chimie Pharmaceutique Albert Lespagnol	86
Mme	MUHR-TAILLEUX	Anne	Biochimie	87
Mme	PERROY	Anne-Catherine	Droit et Economie pharmaceutique	86
Mme	RIVIÈRE	Céline	Pharmacognosie	86
Mme	ROMOND	Marie-Bénédicte	Bactériologie - Virologie	87
Mme	SAHPAZ	Sevser	Pharmacognosie	86
M.	SERGHERAERT	Éric	Droit et Economie pharmaceutique	86
M.	SIEPMANN	Juergen	Pharmacotechnie industrielle	85

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 4/9

Mme	SIEPMANN	Florence	Pharmacotechnie industrielle	85
M.	WILLAND	Nicolas	Chimie organique	86

Maîtres de Conférences - Praticiens Hospitaliers (MCU-PH)

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
Mme	CUVELIER	Élodie	Pharmacologie, Pharmacocinétique et Pharmacie clinique	81
Mme	DANEL	Cécile	Chimie analytique	85
Mme	DEMARET	Julie	Immunologie	82
Mme	GARAT	Anne	Toxicologie et Santé publique	81
Mme	GENAY	Stéphanie	Biopharmacie, Pharmacie galénique et hospitalière	81
M.	GRZYCH	Guillaume	Biochimie	82
Mme	HENRY	Héloïse	Biopharmacie, Pharmacie galénique et hospitalière	80
M.	LANNOY	Damien	Biopharmacie, Pharmacie galénique et hospitalière	80
Mme	MASSE	Morgane	Biopharmacie, Pharmacie galénique et hospitalière	81
Mme	ODOU	Marie-Françoise	Bactériologie - Virologie	82

Maîtres de Conférences des Universités (MCU)

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
Mme	ALIOUAT	Cécile-Marie	Parasitologie - Biologie animale	87
M.	ANTHÉRIEU	Sébastien	Toxicologie et Santé publique	86
Mme	AUMERCIER	Pierrette	Biochimie	87
M.	BANTUBUNGI-BLUM	Kadiombo	Biologie cellulaire	87
M.	BERTHET	Jérôme	Biophysique - RMN	85
M.	BOCHU	Christophe	Biophysique - RMN	85
M.	BORDAGE	Simon	Pharmacognosie	86

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 5/9

M.	BOSC	Damien	Chimie thérapeutique	86
Mme	BOU KARROUM	Nour	Chimie bioinorganique	
M.	BRIAND	Olivier	Biochimie	87
Mme	CARON-HOUDE	Sandrine	Biologie cellulaire	87
Mme	CARRIÉ	Hélène	Pharmacologie, Pharmacocinétique et Pharmacie clinique	86
Mme	CHABÉ	Magali	Parasitologie - Biologie animale	87
Mme	CHARTON	Julie	Chimie organique	86
M.	CHEVALIER	Dany	Toxicologie et Santé publique	86
Mme	DEMANCHE	Christine	Parasitologie - Biologie animale	87
Mme	DEMARQUILLY	Catherine	Biomathématiques	85
M.	DHIFLI	Wajdi	Biomathématiques	27
Mme	DUMONT	Julie	Biologie cellulaire	87
M.	EL BAKALI	Jamal	Chimie thérapeutique	86
M.	FARCE	Amaury	Institut de Chimie Pharmaceutique Albert Lespagnol	86
M.	FLIPO	Marion	Chimie organique	86
M.	FRULEUX	Alexandre	Sciences végétales et fongiques	
M.	FURMAN	Christophe	Institut de Chimie Pharmaceutique Albert Lespagnol	86
M.	GERVOIS	Philippe	Biochimie	87
Mme	GOOSSENS	Laurence	Institut de Chimie Pharmaceutique Albert Lespagnol	86
Mme	GRAVE	Béatrice	Toxicologie et Santé publique	86
M.	HAMONIER	Julien	Biomathématiques	26
Mme	HAMOUDI-BEN YELLES	Chérifa-Mounira	Pharmacotechnie industrielle	85
Mme	HANNOTHIAUX	Marie-Hélène	Toxicologie et Santé publique	86
Mme	HELLEBOID	Audrey	Physiologie	86
M.	HERMANN	Emmanuel	Immunologie	87

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 6/9

M.	KAMBIA KPAKPAGA	Nicolas	Pharmacologie, Pharmacocinétique et Pharmacie clinique	86
M.	KARROUT	Younes	Pharmacotechnie industrielle	85
Mme	LALLOYER	Fanny	Biochimie	87
Mme	LECOEUR	Marie	Chimie analytique	85
Mme	LEHMANN	Hélène	Droit et Economie pharmaceutique	86
Mme	LELEU	Natascha	Institut de Chimie Pharmaceutique Albert Lespagnol	86
M.	LIBERELLE	Maxime	Biophysique - RMN	
Mme	LOINGEVILLE	Florence	Biomathématiques	26
Mme	MARTIN	Françoise	Physiologie	86
M.	MENETREY	Quentin	Bactériologie - Virologie	
M.	MOREAU	Pierre-Arthur	Sciences végétales et fongiques	87
M.	MORGENROTH	Thomas	Droit et Economie pharmaceutique	86
Mme	MUSCHERT	Susanne	Pharmacotechnie industrielle	85
Mme	NIKASINOVIC	Lydia	Toxicologie et Santé publique	86
Mme	PINÇON	Claire	Biomathématiques	85
M.	PIVA	Frank	Biochimie	85
Mme	PLATEL	Anne	Toxicologie et Santé publique	86
M.	POURCET	Benoît	Biochimie	87
M.	RAVAUX	Pierre	Biomathématiques / Innovations pédagogiques	85
Mme	RAVEZ	Séverine	Chimie thérapeutique	86
Mme	ROGEL	Anne	Immunologie	
M.	ROSA	Mickaël	Hématologie	
M.	ROUMY	Vincent	Pharmacognosie	86
Mme	SEBTI	Yasmine	Biochimie	87
Mme	SINGER	Elisabeth	Bactériologie - Virologie	87

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 7/9

Mme	STANDAERT	Annie	Parasitologie - Biologie animale	87
M.	TAGZIRT	Madjid	Hématologie	87
M.	VILLEMAGNE	Baptiste	Chimie organique	86
M.	WELTI	Stéphane	Sciences végétales et fongiques	87
M.	YOUS	Saïd	Chimie thérapeutique	86
M.	ZITOUNI	Djamel	Biomathématiques	85

Professeurs certifiés

Civ.	Nom	Prénom	Service d'enseignement
Mme	FAUQUANT	Soline	Anglais
M.	HUGES	Dominique	Anglais
Mme	KUBIK	Laurence	Anglais
M.	OSTYN	Gaël	Anglais

Professeurs Associés

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
M.	DAO PHAN	Haï Pascal	Chimie thérapeutique	86
M.	DHANANI	Alban	Droit et Economie pharmaceutique	86

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 8/9

Maîtres de Conférences Associés

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
M.	COUSEIN	Etienne	Biopharmacie, Pharmacie galénique et hospitalière	
Mme	CUCCHI	Malgorzata	Biomathématiques	85
M.	DUFOSSEZ	François	Biomathématiques	85
M.	FRIMAT	Bruno	Pharmacologie, Pharmacocinétique et Pharmacie clinique	85
M.	GILLOT	François	Droit et Economie pharmaceutique	86
M.	MITOUMBA	Fabrice	Biopharmacie, Pharmacie galénique et hospitalière	86
M.	PELLETIER	Franck	Droit et Economie pharmaceutique	86

Assistants Hospitalo-Universitaire (AHU)

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
M.	BOUDRY	Augustin	Biomathématiques	
Mme	DERAMOUDT	Laure	Pharmacologie, Pharmacocinétique et Pharmacie clinique	
Mme	GILLIOT	Sixtine	Biopharmacie, Pharmacie galénique et hospitalière	
M.	GISH	Alexandr	Toxicologie et Santé publique	
Mme	NEGRIER	Laura	Chimie analytique	

Hospitalo-Universitaire (PHU)

	Nom	Prénom	Service d'enseignement	Section CNU
M.	DESVAGES	Maximilien	Hématologie	
Mme	LENSKI	Marie	Toxicologie et Santé publique	

	LISTE GEREE	LG/FAC/001
FACULTE DE PHARMACIE	Enseignants et Enseignants-chercheurs 2023-2024	Version 2.2 Applicable au 02/01/2022
Document transversal		Page 9/9

Attachés Temporaires d'Enseignement et de Recherche (ATER)

Civ.	Nom	Prénom	Service d'enseignement	Section CNU
Mme	BERNARD	Lucie	Physiologie	
Mme	BARBIER	Emeline	Toxicologie	
Mme	COMAPGNE	Nina	Chimie Organique	
Mme	COULON	Audrey	Pharmacologie, Pharmacocinétique et Pharmacie clinique	
M.	DUFOSSEZ	Robin	Chimie physique	
Mme	KOUAGOU	Yolène	Sciences végétales et fongiques	
M.	MACKIN MOHAMOUR	Synthia	Biopharmacie, Pharmacie galénique et hospitalière	

Enseignant contractuel

Civ.	Nom	Prénom	Service d'enseignement
M.	MARTIN MENA	Anthony	Biopharmacie, Pharmacie galénique et hospitalière
M.	MASCAUT	Daniel	Pharmacologie, Pharmacocinétique et Pharmacie clinique
Mme	NDIAYE-BOIDIN	Maguette	Anglais
M.	ZANETTI	Sébastien	Biomathématiques

CYCLE DE VIE DU DOCUMENT

Version	Modifié par	Date	Principales modifications
1.0		20/02/2020	Création
2.0		02/01/2022	Mise à jour
2.1		21/06/2022	Mise à jour
2.2		01/02/2024	Mise à jour

UFR3S-Pharmacie

L'Université n'entend donner aucune approbation aux opinions émises dans les thèses ; celles-ci sont propres à leurs auteurs.



Remerciements

Aux membres et au président du jury :

Madame la Professeure Rebecca Deprez-Poulain

Ce sont en partie vos enseignements qui ont suscité mon intérêt pour le domaine de la recherche. Vous m'avez permis de réaliser mon premier stage de recherche au sein de votre laboratoire lors de ma troisième année de pharmacie. A l'époque, c'est ce stage qui m'a convaincu de m'orienter vers la filière industrie/recherche et aujourd'hui je continue ce parcours en réalisant une thèse universitaire de sciences. Alors, merci pour l'intérêt que vous m'avez porté et vos conseils qui m'ont permis de suivre cette voie. Merci de m'avoir soutenu dans mon cursus de pharmacien jusqu'aujourd'hui en me faisant l'honneur de présider le jury de ma thèse d'exercice.

Monsieur le Docteur Benoît Gosselin

Merci d'avoir accepté de faire partie de mon jury de thèse. Merci aussi pour ces années où j'ai eu le plaisir de travailler au sein votre officine à Lambersart, dans un environnement agréable et bienveillant. Merci pour votre bonne humeur et votre pédagogie.

Madame la Professeure Lestrelin

Merci de me faire l'honneur de faire partie de mon jury de thèse et de vous être intéressé à mon travail.

A mon directeur de thèse

Monsieur le Professeur Nicolas Lebègue

Merci d'avoir accepté d'être mon directeur de thèse une deuxième fois. Merci pour votre accompagnement depuis maintenant plus de 2 ans, ainsi que pour vos conseils qui me permettent de me former à la recherche.

Je souhaite remercier :

Ma famille qui m'a soutenu pendant toutes ces années,

Mon père. Merci pour tes sacrifices qui m'ont permis d'en arriver là aujourd'hui. **Ma mère.**

J'espère rendre hommage à la femme forte que tu étais et j'espère que tu es fière de moi.

Mes frères et sœurs : Myriam, Aicha, Lahoucen, Ali et Mohamed

Merci d'avoir été là pour moi depuis toujours. Merci d'avoir intégré **Simon, Affef, Mary et Daisy** dans notre famille et de m'avoir donné des neveux et nièces formidables ; **Keltoum, Söanne, Marwan, Kamil, Alma, Isaac** ; qui me comblent d'amour.

Merci à tous de constituer cette famille incroyable, je suis chanceuse de vous avoir tous à mes côtés.

Mes amis,

Ceux rencontrés pendant mes études de pharmacie

Nora Maupetit-Khali, ma voisine préférée, merci pour ton soutien inestimable qui m'a aidé à surmonter des moments très difficiles. Je te souhaite le meilleur dans ta nouvelle vie, Tabarnak !

Charlotte Quaeybeur, ma binôme de PACES et puis plus tard, ma binôme de TP. Celle avec qui tout a commencé ! Et bien sûr **Céline Merchez** et **Clotilde Terrier**, je suis reconnaissante de vous avoir eu à mes côtés pendant ces années d'études. Merci pour ces moments de vie partagés toutes les 5 et que je n'oublierai jamais. Je vous souhaite à toutes du bonheur et de la réussite.

Mes collègues de travail,

Marie Stoup ma binôme lors de mon stage de fin d'étude, un soutien sans faille avec qui j'ai le plaisir de continuer à travailler. Merci à **Émeline, Juliette, Maureen, Arthur, Robin, Alexandre, Justine, Amélie, Sergio** pour votre bonne humeur au travail.

Sans oublier mon amie du lycée, **Amélie Guilmet**, j'ai hâte de venir te voir et de découvrir une autre partie du monde avec toi.

Je pense aussi à **Darine, Grégory, Abdallah, Khalti, Mounia, Cheima, Daniela, Gary.**

Merci à vous tous pour avoir contribué à l'écriture de cette thèse par un soutien moral ou matériel.

Table des matières

PARTIE 1 : INTRODUCTION	19
I) STRUCTURE TRIDIMENSIONNELLE DES PROTEINES	19
1) Structure primaire	19
2) Structure secondaire	21
a. Interactions	21
b. Le graphique de Ramachandran	23
c. Hélice α et feuillet β	24
d. Fréquence d'apparition des acides aminés dans les hélices α et les feuillets β	25
e. Autres structures	26
3) Structure tertiaire	26
a. Repliement en motif structurel	27
b. Structure – fonction	27
c. Domaines protéiques	27
4) Structure quaternaire	28
II) ÉTUDE STRUCTURALE DES PROTEINES	30
1) Phylogénie	30
2) L'homologie structurale	31
a. Principe	31
b. Lien Structure-Fonction	31
3) Historique	32
a. Avènement des techniques de résolution structurale	32
b. Base de données	32
c. Utilisation de la bioinformatique	33
PARTIE 2 : LA STRUCTURE DES DOMAINES EGF-LIKE	34
I) STRUCTURE PRIMAIRE	34
1) Homologies de séquences	34
a. Identifications	34
b. Conservation des résidus	34
2) Classification	35
a. Domaines EGF classiques	36
b. Famille élargie : Domaines EGF-like à 8 cystéines	37
II) STRUCTURE SECONDAIRE ET TERTIAIRE	38
1) Homologie structurale	38
a. Comparaisons structurales	38
b. Domaines EGF-like classiques	39
c. Famille élargie	41

d.	Théorie évolutive	42
2)	Modifications co et post-traductionnelles	42
a.	O-glycosylation	43
b.	<i>N-glycosylation</i>	44
c.	β -hydroxylation	46
III)	STRUCTURE QUATERNAIRE	47
1)	Calcium – binding	47
2)	Les domaines EGF-like dans les protéines	48
a.	Fonction des domaines EGF-like dans les protéines	48
b.	Structure en mosaïque des domaines EGF-like	49
c.	Protéines contenant des domaines EGF-like	50
d.	Mécanismes protéolytiques des domaines EGF-like	51
	PARTIES 3 : PROTEINES MEMBRANAIRES CONTENANT DES DOMAINES EGF-LIKE	53
I)	PHYSIOPATHOLOGIE	53
1)	Développement et croissance	53
a.	HB-EGF, un facteur de croissance membranaire	53
b.	L'intégrine $\alpha\beta 8$ et son architecture unique	55
2)	Homéostasie et survie	56
a.	Les protéines NOTCH dans l'homéostasie vasculaire	56
b.	LDL-R dans le métabolisme hépatique des lipides	58
c.	Cripto-1, un modulateur crucial des voies de signalisation oncogéniques (89)	60
3)	Communication et immunité	61
a.	La superfamille des protéines EGF-TM7 dans la transduction du signal	61
b.	Les mucines membranaires	63
II)	INTERET EN THERAPEUTIQUE	65
1)	Intérêt dans le diagnostic	65
a.	Diagnostic mutationnel	65
b.	Biomarqueur thérapeutique	66
2)	Ciblage thérapeutique	67
a.	Ciblage thérapeutique indirect	67
b.	Ciblage thérapeutique direct	69
	CONCLUSION	72
	REFERENCES	77

Liste des figures

FIGURE 1 : LA LIAISON PEPTIDIQUE	20
FIGURE 2 : ANGLES DE ROTATION AUTOUR DE LA LIAISON PEPTIDIQUE	20
FIGURE 3: LE GRAPHIQUE DE RAMACHANDRAN.....	23
FIGURE 4 : STRUCTURES SECONDAIRES PRINCIPALES	24
FIGURE 5 : STRUCTURES SECONDAIRES NON REGULIERES	26
FIGURE 6 : ORGANISATION MODULAIRE DES DOMAINES PROTEIQUES DES MEMBRES DE LA FAMILLE DES RECEPTEURS LDL	29
FIGURE 7 : PRINCIPE DE L'ETUDE PHYLOGENETIQUE D'UNE PROTEINE	31
FIGURE 8 : NOMBRE DE STRUCTURES TRIDIMENSIONNELLES ENREGISTREES PAR AN DANS LA PROTEIN DATA BANK	33
FIGURE 9. ALIGNEMENT DE SEQUENCES DES DOMAINES EGF-LIKE BASE SUR LEUR STRUCTURE ET CLASSIFICATION	38
FIGURE 10 : POSITIONS DES CYSTEINES DANS LES DOMAINES EGF-LIKE.....	40
FIGURE 11. POSITION DES CYSTEINES SUPPLEMENTAIRES DES DOMAINES EGF-LIKE DE TYPE LAMININE ET INTEGRINE.....	42
FIGURE 12. STRUCTURES SECONDAIRES DE DOMAINES EGF-LIKE CONTENUS DANS LES PROTEINES..	45
FIGURE 13. IMPACT DE LA LIAISON DU CALCIUM AUX DOMAINES EGF DE LA FIBRILLINE SUR SA STRUCTURE.....	48
FIGURE 14. ARRANGEMENTS DES DIFFERENTS DOMAINES PROTEIQUES DE PROTEINES CONTENANT DES DOMAINES EGF-LIKE	49
FIGURE 15. REPRESENTATION SCHEMATIQUE DE LA PROTEINE S DE LA COAGULATION AVEC LA STRUCTURE SECONDAIRE DE SES DOMAINES EGF3 ET EGF4.....	51
FIGURE 16. LA PROTEINE HB-EGF	54
FIGURE 17. STRUCTURE DES SOUS UNITES ALPHA ET BETA DES INTEGRINES	56
FIGURE 18 : LES PROTEINES NOTCH	57
FIGURE 19. LE RECEPTEUR AU LDL	59
FIGURE 20. REPRESENTATION SCHEMATIQUE DE LA PROTEINE CRIPTO-1	60
FIGURE 21: REPRESENTATION SCHEMATIQUE D'ISOFORMES DE PROTEINES EMR.....	61
FIGURE 22: REPRESENTATION SCHEMATIQUE DE LA PROTEINE CD97.....	62
FIGURE 23. REPRESENTATION SCHEMATIQUE DE LA STRUCTURE DE LA PROTEINE ELTD1	63
FIGURE 24. INTERACTION MOLECULAIRE ENTRE LES PROTEINES HER2 ET MUC4.....	64
FIGURE 25. CONCEPT DE GREFFE MOLECULAIRE UTILISANT LES DOMAINES EGF- LIKE.....	67
FIGURE 26. FIXATION DU DOMAINE EGF5 DE LA THROMBOMODULINE A LA THROMBINE	68
FIGURE 27. PREUVE DE CONCEPT DU CIBLAGE THERAPEUTIQUE DES DOMAINES EGF DE MUC4 DANS LE CANCER DU PANCREAS.	71

Abbreviations

ADAM : A Disintegrin And Metalloproteinase

ADN : Acide désoxyribonucléique

AVC : Accident vasculaire cérébral

BCR-Abl : Fusion des gènes BCR (Breakpoint Cluster Region) et Abl (Abelson proto-oncogene)

CADASIL : Artériopathie cérébrale autosomique dominante avec infarctus sous-corticaux et leucoencéphalopathie

cb-EGF-like : Calcium-binding EGF-like (domaine EGF-like liant le calcium)

CD9, 53, 97 : Cluster de Differentiation 9, 53, 97

cEGF-like : domaine EGF-like du complément

CryoEM : Cryo-electron microscopy (microscopie électronique cryogénique)

CXCL12 : C-X-C motif chemokine ligand 12

CXCR4 : C-X-C motif chemokine receptor 4

EGF : Epidermal Growth Factor (Facteur de croissance épidermique)

EGF-like : de type EGF

EGFR : Epidermal Growth Factor Receptor (Récepteur du facteur de croissance épidermique)

EGF-TM7 : superfamille de protéines à domaines EGF-like et 7 domaines transmembranaire

ELTD1 : EGF, latrophilin, and seven transmembrane domain-containing protein 1

EMR : EGF-like module containing mucin-like hormone receptor-related

HB-EGF : Heparin-binding EGF-like growth factor (Facteur de croissance de type EGF liant l'héparine)

HER : Human epidermal growth factor receptor

hEGF-like : Human EGF-like (domaine EGF humain)

IGF : Insulin-like Growth Factor (Facteur de croissance de type insuline)

LDL : Low-Density Lipoprotein (Lipoprotéine de basse densité)

LDL-R : Low-Density Lipoprotein Receptor (Récepteur des lipoprotéines de basse densité)

LTBP-1 : Latent Transforming Growth Factor B Binding Protein 1 (Protéine de liaison du TGF- β latent 1)

MAPK : Mitogen-Activated Protein Kinase (Kinase activée par les mitogènes)

MUC : Mucine

PCSK9 : Proprotein convertase subtilisin/kexin type 9

PDB : Protein Data Bank

PI3K : Phosphoinositide 3-kinase

POFUT1 : Protein O-fucosyltransferase 1

POGLUT1 : Protein O-glucosyltransferase 1

POGLUT2/3 : Protein O-glucosyltransferase 2/3

Ras : Rat sarcoma virus (Famille de protéines GTPase)

SCUBE : Signal peptide-CUB-EGF-like domain-containing protein

TGF- α : Transforming Growth Factor α Facteur de croissance transformant β)

TGF- β : Transforming Growth Factor β (Facteur de croissance transformant β)

apoER2 : apolipoprotein E receptor 2

LRP : lipoprotein receptor related

MEGF7 : multiple EGF-like domains protein

FVII, FIX, FX : Facteurs de la coagulation VII, IX et X

MSP : Membrane scaffold protein

MMP : Métalloprotéase matricielle

PSI : Plexin-Semhorin-Integrin

DLL4 : Delta like ligand 4

LNG : Lunatic Fringe

PES : Domaine riche en résidus proline, acide glutamique et sérine

DSL : Delta-Serrate ligand

CBF1 : Centromere binding protein 1

YWTD : Domaine de répétition des résidus tyrosine, tryptophane, thréonine et acide aspartique

GPI : Glycophosphatidylinositol

AKT : Protéine Kinase B

DAPI : 4',6-diamidino-2-phénylindole, molécule se liant à l'ADN utilisée pour marquer le noyau des cellules

PARTIE 1 : INTRODUCTION

Les protéines sont des molécules essentielles à la vie, jouant un rôle clé dans de nombreuses fonctions biologiques. Elles sont composées d'une séquence spécifique d'acides aminés liés par des liaisons peptidiques, qui détermine leur structure tridimensionnelles (1). Cette structure tridimensionnelle est essentielle à leur fonction, car elle détermine les sites actifs impliqués dans la liaison à d'autres molécules ou dans la catalyse de réactions biochimiques. Les protéines ne sont pas uniformes dans leurs structures et leurs fonctions. Elles sont composées de différents domaines fonctionnels, qui sont des parties distinctes de la protéine ayant des fonctions spécifiques. Les domaines fonctionnels d'une protéine peuvent être responsables de la reconnaissance et de la liaison à d'autres molécules, de la catalyse de réactions chimiques, du transport de molécules ou de la régulation de processus cellulaires (2). L'étude approfondie des domaines fonctionnels des protéines est cruciale pour comprendre leur mécanisme d'action et leur impact sur les processus biologiques. De plus, cette connaissance est également précieuse dans le développement de médicaments ciblant spécifiquement certaines protéines ou domaines protéiques associés à des maladies. En somme, la diversité des protéines et de leurs domaines fonctionnels offre un vaste champ d'exploration pour la recherche en biologie structurale, en biochimie et en pharmacologie, ouvrant la voie à de nouvelles découvertes et applications médicales.

I) Structure tridimensionnelle des protéines

Elle est composée de quatre niveaux d'organisation : la structure primaire, la structure secondaire, la structure tertiaire et la structure quaternaire.

1) Structure primaire

La structure primaire des protéines fait référence à la séquence linéaire des acides aminés qui composent la protéine. Elle est déterminée par la séquence d'ADN qui code pour la protéine (3).

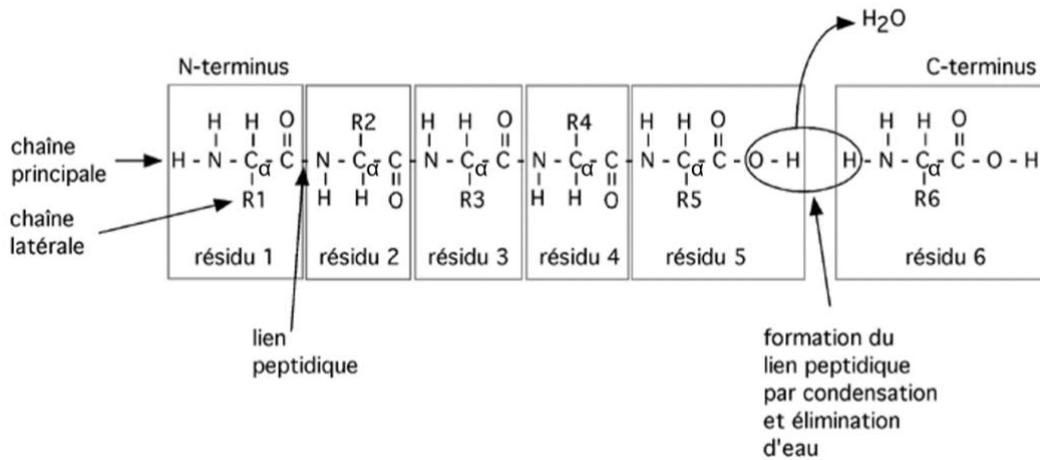


FIGURE 1 : LA LIAISON PEPTIDIQUE (adapté de (4))

Les acides aminés sont de petites molécules organiques composées d'un carbone chiral avec quatre substituants (**ANNEXE 1**). Chaque acide aminé est relié au suivant par une liaison peptidique covalente, formée lorsque le groupement carboxylique d'un premier acide aminé réagit avec le groupement amine d'un second (5). La chaîne polypeptidique n'est pas ramifiée et forme un unique filament étiré se repliant pour former la structure tridimensionnelle des protéines qui détermine leur fonction biologique. Par convention, on désigne le premier acide aminé de la chaîne comme étant celui dont le groupement amine reste libre ; cela constitue l'extrémité N-terminale. Le dernier résidu de la chaîne est identifié comme celui dont le groupement carboxylique reste libre; ceci correspond à l'extrémité C-terminale (3) (**FIGURE 1**).

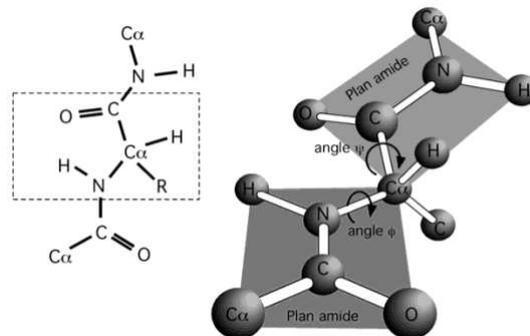


FIGURE 2 : ANGLES DE ROTATION AUTOUR DE LA LIAISON PEPTIDIQUE (adapté de (4))

Chaque liaison peptidique conserve sa rigidité, mais des rotations peuvent se produire de part et d'autre de cette liaison. L'angle de rotation d'un plan amide par rapport au carbone α N-terminal est désigné par l'angle ψ , tandis que l'angle entre le plan amide précédent et ce même carbone α est désigné par l'angle ϕ (**FIGURE 2**).

2) Structure secondaire

La structure secondaire des protéines se réfère à la manière dont une chaîne polypeptidique se replie localement dans l'espace pour former des motifs réguliers et répétitifs. Elle est due à des Interactions covalentes et non covalentes stabilisant l'état natif de la protéine.

a. Interactions

i. *Interactions covalentes*

En plus des liaisons peptidiques, des liaisons covalentes se forment entre les atomes de soufre de deux cystéines, appelés ponts disulfures.

ii. *Interactions non covalentes*

- Interactions électrostatiques : charge-charge entre résidus de charge inverse
- Interaction charge – dipôle : Lorsqu'une chaîne latérale ionisée entre en interaction avec le dipôle d'une molécule d'eau ou d'une liaison peptidique polaire
- Interaction hydrophobe : interactions entre des groupes hydrophobes qui excluent les molécules d'eau.
- Forces de Van der Waals : il s'agit de dipôles temporaires de faibles forces. Ces dipôles peuvent se former car les nuages électroniques des atomes individuels peuvent fluctuer, engendrant ainsi des dipôles instantanés.

Les interactions sont présentées dans le **TABLEAU 1**.

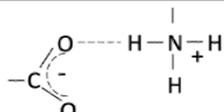
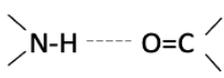
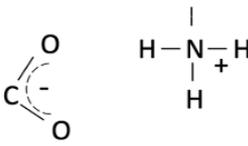
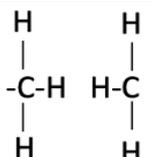
Interaction	Exemple	Dépendance à la distance	Distance typique	Energie libre (enthalpies de dissociation)
Liaison covalente	-C α -C	-	1,5 Å	356 kJ/mole
Pont disulfure	-Cys-S-S-Cys-	-	2,2 Å	167 kJ/mole
Liaison ionique		Atomes donneurs (ici N) et accepteurs (ici O) <3,5 Å	2,8 Å	12,5 – 17 kJ/mole
Liaison hydrogène		Atomes donneurs (ici N) et accepteurs (ici O) <3,5 Å	3,0 Å	2-6 kJ/mole
Interaction électrostatique de longue portée		Dépend de la constante diélectrique du milieu 1/r	Variable	Dépend de la distance et de l'environnement. Peut être très fort dans les régions non polaires mais très faible dans l'eau
Interaction de Van der Waals		Courte portée Chute rapidement au-delà de 4 Å 1 / r ⁶	3,5 Å	4 – 17 kJ/mole à l'intérieur des protéines Dépend de la taille du groupe chimique

TABLEAU 1 : LES INTERACTIONS CHIMIQUES STABILISANT LA CHAÎNE POLYPEPTIDIQUE (adapté de (6))

iii. Influence des paramètres environnementaux

Les modifications des paramètres physico-chimiques, tels que le pH, la température, la force ionique et la concentration en soluté, ont un impact profond sur la structure secondaire des protéines (7). Ces facteurs influencent les interactions hydrogène, les liaisons ioniques, et les forces de Van der Waals, essentielles pour la stabilité des hélices α et des feuillets β . Par exemple, une augmentation de la température peut provoquer la dénaturation des protéines en rompant les liaisons hydrogène, entraînant la perte de la structure secondaire (8). Les variations du pH peuvent modifier la charge des résidus acides et basiques (ANNEXE 1), altérant les interactions ioniques cruciales pour la stabilité de la structure secondaire (9). La force ionique affecte également la formation et la stabilité des ponts salins et des liaisons hydrogène, tandis que la présence de solutés peut perturber ou stabiliser la structure secondaire en modifiant l'environnement hydrophobe ou hydrophile des résidus protéiques (10–12). En outre, les facteurs environnementaux, tels que les radiations, les polluants chimiques, et les variations extrêmes de température, peuvent entraîner des mutations dans l'ADN qui codent pour

les protéines (13). Ces mutations peuvent conduire à des changements dans la séquence d'acides aminés, altérant ainsi la capacité de la protéine à former correctement sa structure secondaire.

b. Le graphique de Ramachandran

Le biologiste G.N. Ramachandran a remarqué en travaillant sur des modèles de polypeptides que certains angles ψ et ϕ ne pouvaient pas être combinés à cause de l'encombrement stérique. Le graphique de Ramachandran (FIGURE 3) est un outil graphique utilisé en biologie structurale pour visualiser les angles de torsion des liaisons peptidiques dans une protéine. Les régions permises sur le graphique de Ramachandran correspondent aux conformations stables et énergétiquement favorables, tandis que les régions interdites correspondent à des conformations stériquement impossibles ou hautement énergétiques (14). Son graphique permet aux chercheurs de comprendre et d'analyser la conformation spatiale des protéines.

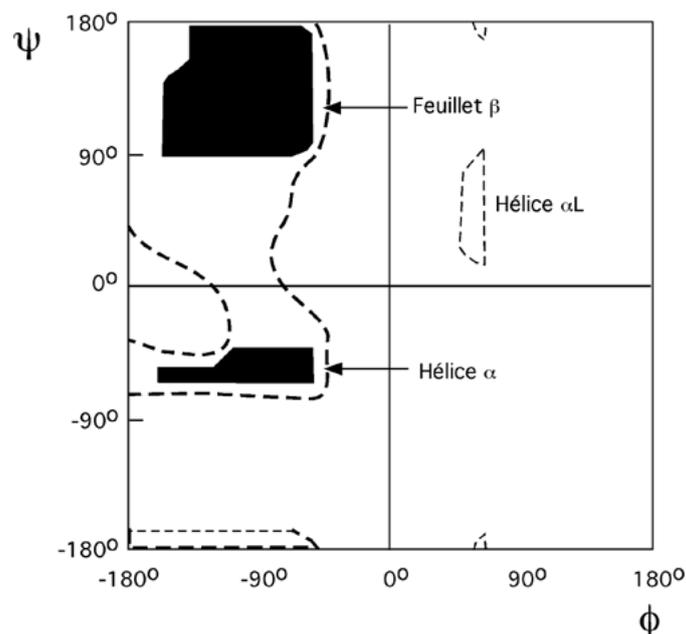


FIGURE 3: LE GRAPHIQUE DE RAMACHANDRAN (adapté de (15))

Il représente les angles ϕ sur l'axe des abscisses et les angles ψ sur l'axe des ordonnées. En traçant ces angles pour chaque résidu d'acide aminé dans une protéine donnée, le graphique de Ramachandran permet de visualiser les régions permises et interdites pour ces angles de torsion. Les régions toujours permises sont représentées en noir. Les régions non favorisées mais possiblement tolérées sont représentées en

pointillé. Ces angles influencent la manière dont la protéine se replie localement dans l'espace pour former des motifs de structure secondaire, les plus courants étant l'hélice α et le feuillet β .

c. Hélice α et feuillet β

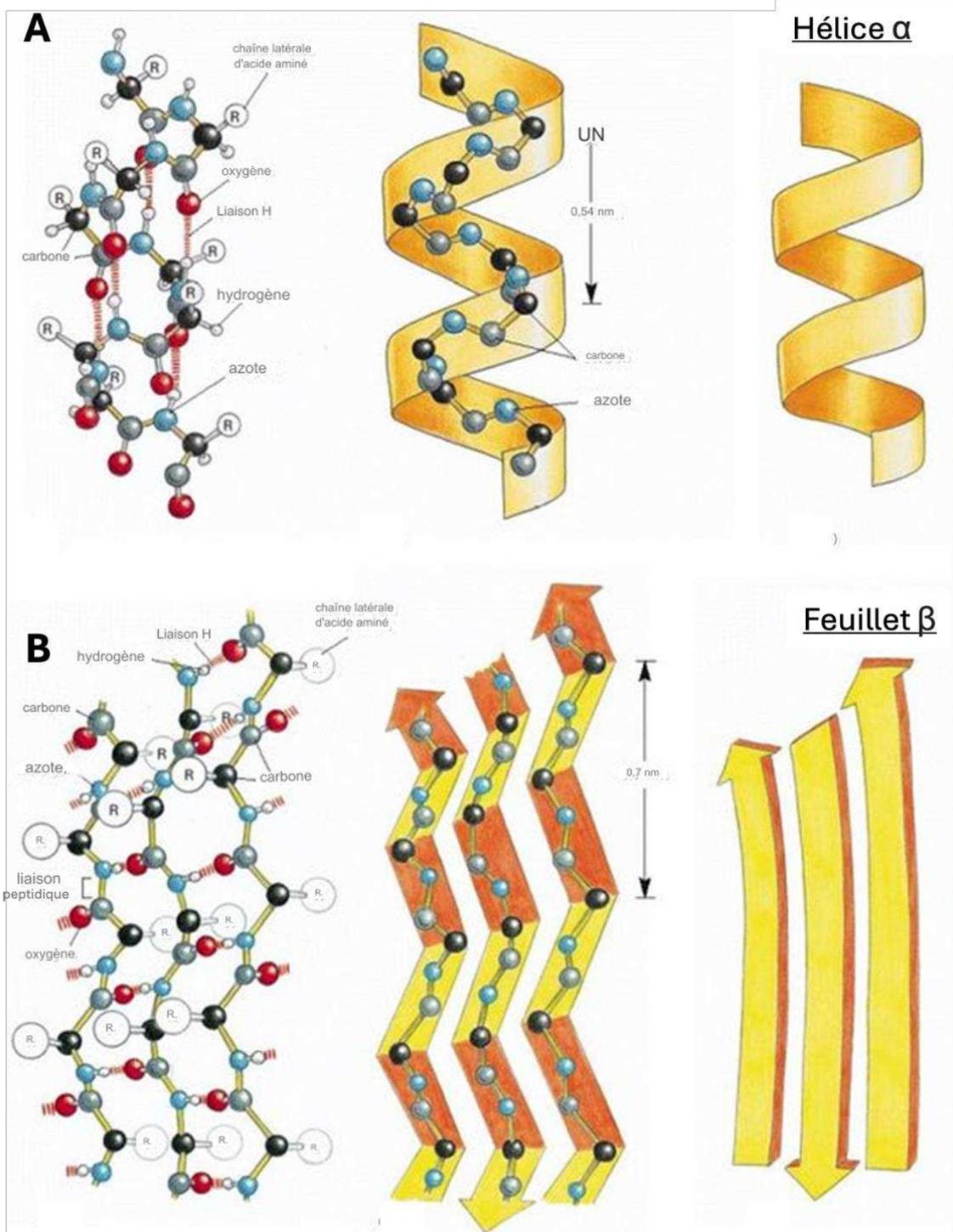


FIGURE 4 : STRUCTURES SECONDAIRES PRINCIPALES. A. HELICE α . B. FEUILLET β (adapté de (16))

i. Hélice α

Une hélice α résulte d'une succession d'angles ϕ et ψ de -57° et -47° (15) (FIGURE 4A). Le squelette polypeptidique s'enroule en spirale et les chaînes latérales pointent vers l'extérieur. Les liaisons hydrogène qui stabilisent la structure en hélice se forment entre les atomes d'hydrogène et les atomes d'oxygène des liaisons peptidiques, 4 positions plus loin. L'hélice α s'élève de 0,15 nm par résidu et de 0,54 nm à chaque tour. Elle compte 3,6 résidus par tour. Les liaisons hydrogène vont dans la même direction que l'hélice et par conséquent, la structure est plutôt élastique.

ii. Feuille β

Dans la structure de type feuille β , les liaisons hydrogène se forment entre les atomes d'oxygène et d'hydrogène des liaisons peptidiques de segments adjacents (FIGURE 4B). La chaîne polypeptidique forme des segments étirés qui se replient et s'alignent parallèlement ou antiparallèlement les uns par rapport aux autres, avec les chaînes latérales qui pointent vers l'intérieur. La distance entre deux résidus est de 0,335 nm et chaque chaîne est séparée de sa voisine de 0,465 nm. On parle de feuillets β parallèles quand les chaînes vont dans le même sens et d'antiparallèles quand elles vont dans des directions opposées. Comme les liaisons hydrogènes des feuillets β pointent perpendiculairement par rapport à la chaîne polypeptidique, ils ne donnent pas d'élasticité à la structure.

d. Fréquence d'apparition des acides aminés dans les hélices α et les feuillets β

Si certains résidus comme l'acide glutamique sont retrouvés plus fréquemment dans les hélices α que dans les feuillets β et que la cystéine est retrouvée plus fréquemment dans les feuillets β que dans les hélices α , il y a des résidus comme la leucine qui sont retrouvés dans les deux types de structures. Enfin, il y a d'autres acides aminés que l'on retrouve rarement dans les hélices α ou les feuillets β comme par exemple la proline (15) (ANNEXE 2).

e. Autres structures

La chaîne protéique peut contenir des régions moins structurées formant une discontinuité locale dans la structure qui ne se présentent ni sous forme d'hélices ni de feuillettes (FIGURE 5).

i. Tours et boucles

Ces motifs non répétitifs incluent entre autres les tours et les boucles. Les boucles sont des régions non structurées qui relient deux structures secondaires régulières. Elles sont généralement situées à la surface de la protéine dans des zones exposées aux solvants (1). En général, la glycine permet aux protéines d'effectuer des tours serrés et d'avoir des contacts étroits des chaînes polypeptidique.

ii. Renflement β

Certains feuillettes présentent une perturbation locale dans le réseau de liaison hydrogène par l'insertion de résidus supplémentaires appelé « renflement β ». En 1993 ces renflements β ont été décrits et classifiés par Chan et al.

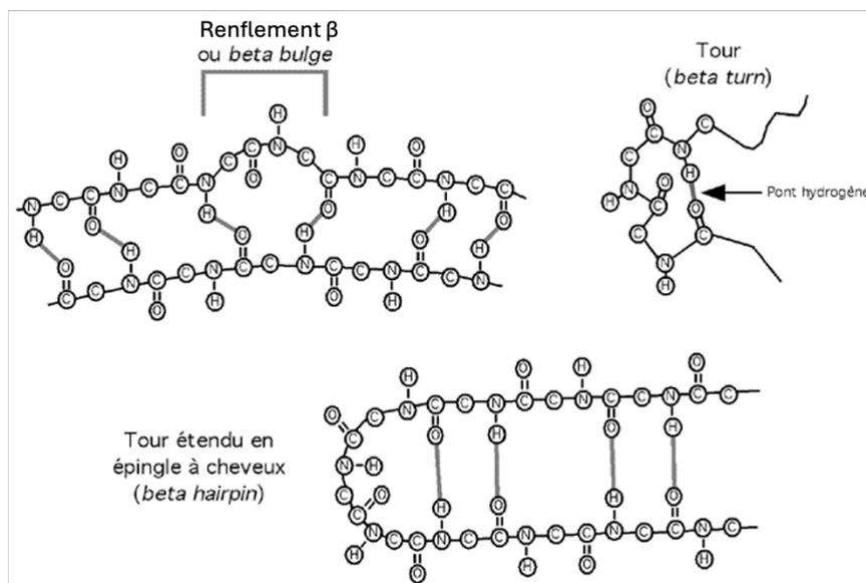


FIGURE 5 : STRUCTURES SECONDAIRES NON REGULIERES (adapté de (4))

3) Structure tertiaire

La structure tertiaire des protéines fait référence à l'organisation dans l'espace et à la combinaison des structures secondaires.

a. Repliement en motif structurel

Le repliement des protéines est stabilisé par des interactions entre les acides aminés éloignés dans la séquence (TABLEAU 1). Ces interactions impliquent des liaisons hydrophiles polaires externes telles que des liaisons hydrogène ou ioniques. Il existe des interactions hydrophobes internes entre les chaînes latérales non polaires des acides aminés. Ainsi, lors de la synthèse des protéines, la chaîne polypeptidique se replie spontanément en une forme tridimensionnelle spécifique, sous l'effet de ces interactions entre les différentes parties de la chaîne polypeptidique. La structure adoptée par la chaîne polypeptidique est généralement celle dans laquelle l'énergie libre est minimisée (16).

Chaque combinaison de structures secondaires différentes est connue sous le nom de motif structurel et jusqu'à présent des milliers de motifs structurels ont été identifiés parmi les protéines étudiées. Ces motifs, dont les conformations détaillées sont connues, sont recensés dans des bases de données (FIGURE 8) (17).

b. Structure – fonction

Un fait remarquable est que toutes les fonctions impliquant ces motifs structurels sont basées sur un principe commun : l'arrangement spatial des vingt acides aminés naturels qui peuvent composer la chaîne polypeptidique. Ainsi, toutes les fonctions des protéines dépendent de leur structure qui, à son tour, dépend des paramètres physiques et chimiques des séquences polypeptidiques (18).

c. Domaines protéiques

Certaines structures tertiaires forment des unités structurales et fonctionnelles stables conservées par la nature et qui peuvent être présentes dans différentes protéines. Ces unités structurellement indépendantes, appelées domaines, sont composées de 40 à 350 acides aminés environ et possèdent chacune une fonction propre (19) (FIGURE 6). Le noyau central d'un domaine peut être construit à partir d'hélices α , de feuillets β ou de diverses combinaisons de ces deux éléments structuraux fondamentaux. Ils sont retrouvés dans des protéines multidomaines qui sont des protéines contenant plusieurs domaines appartenant à des familles protéiques différentes (20).

Les protéines les plus petites se composent d'un seul domaine, tandis que des protéines plus grandes peuvent contenir plusieurs dizaines de domaines, souvent liés entre eux par de courtes chaînes polypeptidiques relativement peu structurées. Lorsqu'une protéine prend la forme d'un ensemble de plus d'une chaîne polypeptidique, la structure complète est désignée sous le nom de structure quaternaire (18).

4) Structure quaternaire

La structure quaternaire ne s'applique qu'aux protéines à plusieurs sous-unités, c'est-à-dire aux protéines constituées de plus d'une chaîne polypeptidique. Dans les protéines multimériques, les sous-unités peuvent s'assembler de différentes manières pour former des complexes protéiques.

Les changements dans la structure quaternaire peuvent réguler l'activité des protéines et jouer un rôle crucial dans divers processus biologiques, tels que la régulation de l'expression génique, la signalisation cellulaire et le métabolisme (22). Toutes les fonctions des protéines dépendent de leur structure et des paramètres physiques et chimiques. En outre, la diversité des fonctions des protéines est souvent attribuée à la présence de différents domaines fonctionnels. Ces domaines peuvent être des motifs conservés dans l'évolution, ou encore des structures spécifiques conférant à la protéine une fonction particulière. Par exemple, un domaine de liaison à l'ADN peut permettre à une protéine d'interagir spécifiquement avec des séquences d'ADN, régulant ainsi l'expression des gènes (23).

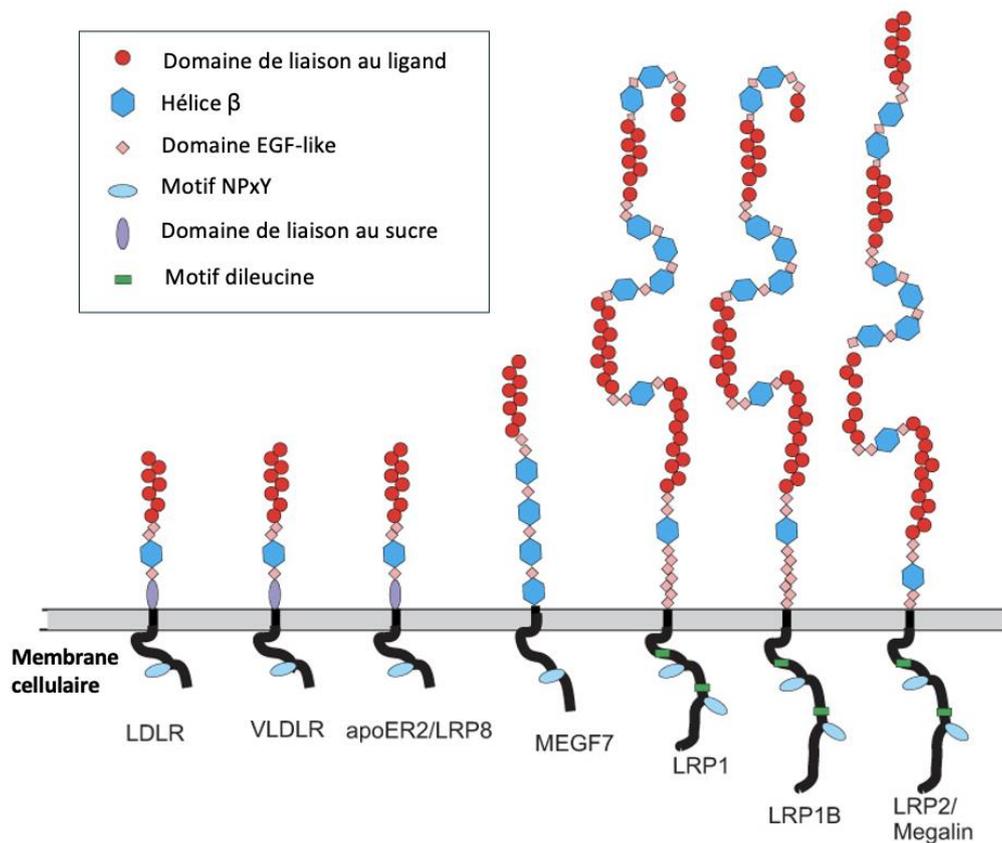


FIGURE 6 : ORGANISATION MODULAIRE DES DOMAINES PROTEIQUES DES MEMBRES DE LA FAMILLE DES RECEPTEURS LDL (adapté de (21))

L'étude approfondie de la structure tridimensionnelle d'une protéine permet généralement de comprendre son activité biologique ou d'interférer avec elle, par exemple dans le cas de la conception de médicaments. Dans les deux cas l'activité biologique dépend de détails atomiques et les modifications apparemment subtiles d'une structure peuvent avoir un impact considérable sur l'activité de la protéine.

Étant donné que chacun des 20 acides aminés peut théoriquement occuper n'importe-quelle position dans une chaîne protéique, il existe 20^n combinaisons possibles de chaînes polypeptidiques différentes dans une chaîne de n acides aminés. Cependant, seule une infime fraction de cet ensemble adopte une conformation tridimensionnelle unique et stable dans la cellule. Cette sélection rigoureuse est le résultat des pressions exercées par la sélection naturelle.

II) Étude structurale des protéines

1) Phylogénie

Les protéines dont la structure et l'activité biochimique varient de manière imprévisible ont peu de chances de contribuer à la survie de la cellule qui les héberge. Ainsi au fil des processus évolutifs, ces protéines sont éliminées par les pressions de sélection de la nature (24). Grâce à ce processus de sélection naturelle, la séquence d'acides aminés d'une protéine est ajustée de manière à favoriser une seule conformation extrêmement stable. De plus, cette conformation est finement adaptée sur le plan chimique pour permettre à la protéine de remplir une fonction catalytique ou structurale spécifique dans la cellule (6). Les protéines sont construites avec une telle précision que même de légères modifications dans la séquence d'acides aminés peuvent parfois perturber la structure globale de la molécule au point de compromettre la fonction (25–28).

De nombreuses protéines de grande taille montrent des signes évidents d'évolution par la jonction de domaines préexistants dans de nouvelles combinaisons, c'est-à-dire par un brassage des domaines (29). La phylogénie permet notamment de classer les protéines par famille car elle reconstruit les relations évolutives entre les différentes séquences protéiques. Les protéines (ou domaine protéique) ayant une origine évolutive commune sont appelées protéines homologues (**FIGURE 7A**). En comparant les différences et les similitudes des séquences de protéines homologues (**FIGURE 7B**), les chercheurs peuvent retracer l'histoire évolutive des protéines et des organismes qui les produisent.

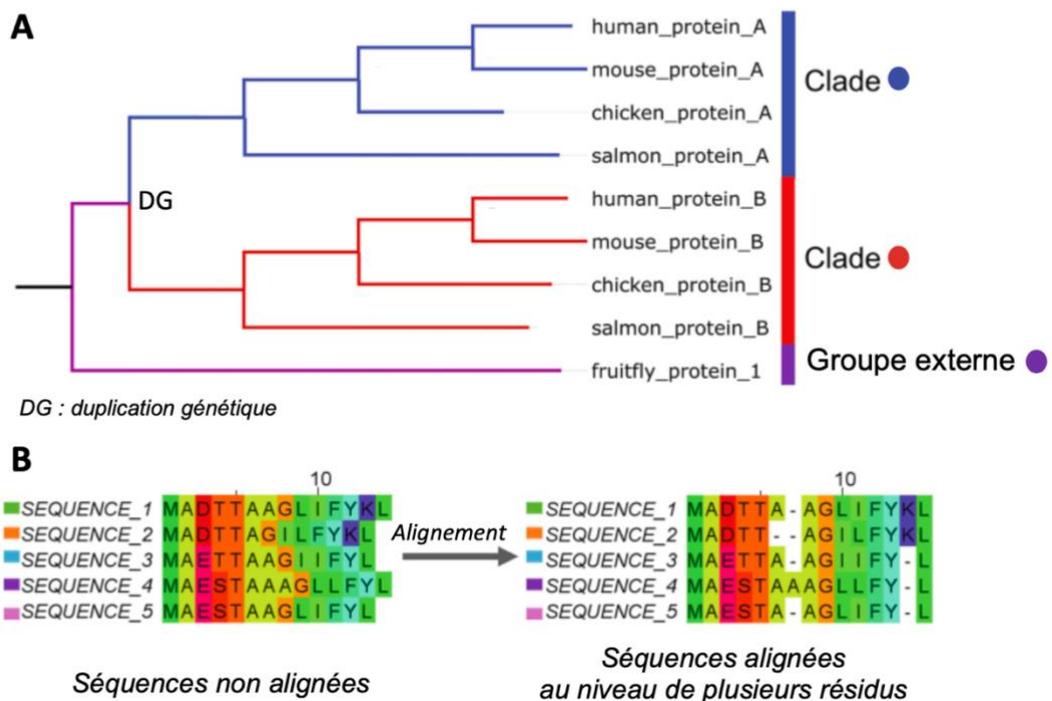


FIGURE 7 : PRINCIPE DE L'ETUDE PHYLOGENETIQUE D'UNE PROTEINE. A : ARBRE PHYLOGENETIQUE B : ALIGNEMENT DES SEQUENCES (adapté de (29))

2) L'homologie structurale

a. Principe

Les biochimistes définissent l'homologie entre les protéines à l'aide de structures 3D expérimentales et en comparant leurs conformations spatiales. Les protéines homologues ont généralement une structure tridimensionnelle similaire, reflétant leur origine évolutive commune, même si leurs séquences d'acides aminés peuvent avoir divergé de manière significative au fil du temps. La conformation structurale des protéines est généralement plus conservée que la séquence d'acides aminés ; par conséquent, lorsque deux protéines ont des structures tridimensionnelles significativement similaires, elles sont considérées comme homologues (30).

b. Lien Structure-Fonction

La fonction d'une protéine identifiée peut être anticipée en comparant la séquence avec celles de protéines déjà caractérisées. Étant donné que la séquence d'acides aminés détermine la structure de la protéine et que cette structure gouverne sa fonction biochimique, les protéines partageant une séquence d'acides aminés similaire ont

tendance à adopter une structure similaire et donc à remplir des fonctions biochimiques semblables, même si elles proviennent d'organismes très éloignés les uns des autres. Ainsi, actuellement, la détermination de la fonction d'une protéine nouvellement identifiée commence souvent par une recherche de protéines déjà répertoriées présentant des similitudes au niveau de leurs séquences d'acides aminés (31).

3) Historique

a. Avènement des techniques de résolution structurale

L'étude structurale des protéines a pris de l'importance à la fin des années 1960 et au début des années 1970, avec l'avènement des techniques expérimentales de résolution structurale comme la cristallographie à rayons X, la Résonance Magnétique Nucléaire (RMN) et la Cryo Microscopie Électronique (CryoEM). Alors que l'étude des séquences de protéines remonte au début du XXe siècle, la possibilité de déterminer les structures des protéines a donné une nouvelle dimension à la compréhension de la fonction, de l'évolution et des relations entre les protéines. La reconnaissance de l'homologie structurale en tant qu'aspect essentiel de l'analyse des protéines a commencé à émerger dans les années 1970 avec des articles de référence tels que "Evolutionary divergence and convergence in proteins" de Walter M. Fitch. Cet article mettait en évidence les limites de l'homologie de séquence pour comprendre l'évolution des protéines et proposait que les similitudes structurales puissent fournir des informations supplémentaires.

b. Base de données

En 1965, Margaret Oakley Dayhoff, professeure en chimie et bioinformatique a publié pour la première fois sa collection des 65 protéines connues dans l'Atlas of Protein Sequence and Structure, une version imprimée de sa base de données. Par la suite, les données ont été transférées sur bande magnétique, et aujourd'hui elles se trouvent en ligne, où les chercheurs continuent d'utiliser ces données pour identifier des milliers d'autres protéines (32). D'autres bases de données biomédicales ont vu le jour, notamment la Protein Data Bank (PDB), une collection collaborative de protéines et d'acides nucléiques lancée en 1971 et supervisée par la Banque Mondiale de Données sur les Protéines (WWPDB) (FIGURE 8).

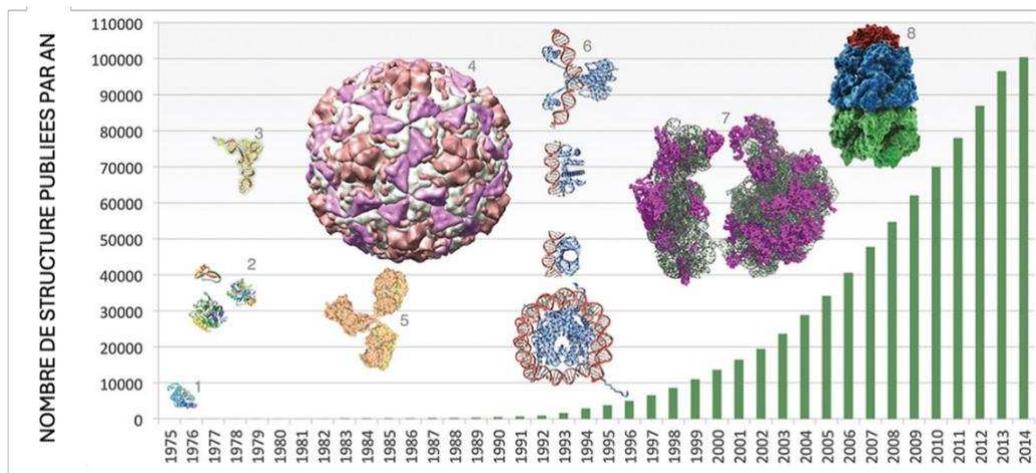


FIGURE 8 : NOMBRE DE STRUCTURES TRIDIMENSIONNELLES ENREGISTREES PAR AN DANS LA PROTEIN DATA BANK (adapté de (33))

c. Utilisation de la bioinformatique

L'un des premiers exemples de comparaison de bases de données structurales remonte aux travaux de Margaret O. Dayhoff, qui a été la première à utiliser des méthodes informatiques pour aligner les séquences de protéines. Toutefois, ce n'est que plus tard que les chercheurs ont commencé à comparer systématiquement les structures des protéines pour mieux comprendre leurs relations évolutives et leurs similitudes fonctionnelles.

En somme, la diversité des protéines et de leurs domaines fonctionnels offre un vaste champ d'exploration pour la recherche en biologie structurale, la biochimie et la pharmacologie, ouvrant la voie à de nouvelles découvertes et applications médicales.

PARTIE 2 : LA STRUCTURE DES DOMAINES EGF-LIKE

I) Structure primaire

1) Homologies de séquences

a. Identifications

Les domaines EGF-like sont des domaines protéiques extracellulaires de 30 à 40 acides aminés caractérisés par la présence de six résidus cystéine conservés à l'origine de la formation de trois ponts disulfures intradomains. La position des résidus cystéine et les ponts disulfures qui en résultent sont cruciaux pour la stabilité structurale du domaine EGF-like. Le schéma de connectivité disulfure est généralement C1-C3, C2-C4 et C5-C6 (34).

Leur topologie de séquence est décrite pour la première fois par Savage et Cohen en 1972 dans le facteur de croissance épidermique (EGF) (35). A l'époque ils identifient un petit polypeptide de 53 acides aminés, incluant six résidus cystéine formant trois ponts disulfures intramoléculaires. Cette structure de base a été par la suite identifiée dans d'autres protéines possédant des domaines du même type : les domaines de type EGF ou domaines « *EGF-like* », indiquant une conservation évolutive de cette topologie structurale.

En 1988, Appella examine en détail la structure et la fonction des domaines EGF-like dans diverses protéines, en particulier dans les récepteurs de l'EGF, les facteurs de coagulation sanguine, ainsi que certaines protéines de la matrice extracellulaire (36). Sa revue traite également de l'importance de ces domaines dans diverses fonctions biologiques, telles que la signalisation cellulaire et les interactions cellule-matrice.

b. Conservation des résidus

Les domaines EGF-like présentent certaines caractéristiques de structure primaire qui sont essentielles à leur fonction. La disposition et la liaison des résidus cystéines C3, C4 et C5 dans la séquence contribue à la géométrie spécifique nécessaire à la fonction du domaine (37).

Contrairement au reste du domaine, la séquence encadrée par la première cystéine (C1) et la deuxième cystéine (C2) ne présente qu'une très faible homologie. Dans les

domaines EGF-like cette partie varie en taille et en composition des acides aminés (38). L'alignement de séquence au niveau des cystéines a permis d'identifier certains résidus conservés dans la partie C-terminale. Les domaines EGF-like présentent généralement la séquence suivante :

(F/Y)-X-C4-(X/XX)-C5-(XX/XXX)-G-(F/Y)-X-(G/L)-(X)₂₋₇-C6-(Q/E) (37,38) (FIGURE 9).

Le schéma de liaison disulfure C1-C3, C2-C4, C5-C6 est typique de nombreux domaines EGF-like, mais il existe des variations qui sont apparues au cours de l'évolution (38–40). Des mutations ont pu entraîner un gain ou une perte de résidu(s) cystéine, ce qui a entraîné une connectivité disulfure différente pouvant conférer de nouvelles fonctions biologiques spécifiques (34) (ANNEXE 3).

Le regroupement des séquences de protéines en familles sur la base de similitudes (ou divergences) au niveau de la séquence joue un rôle crucial dans la compréhension des relations entre les protéines. Ce processus permet aux chercheurs d'identifier les protéines homologues, qui sont des protéines qui partagent un ancêtre commun (41). Les protéines homologues présentent souvent des fonctions et des caractéristiques structurales similaires, ce qui rend le regroupement des séquences de protéines en familles essentiel pour étudier l'évolution et la fonction des protéines (42).

2) Classification

En 1996, Downing distingue différents types de domaines EGF-like en fonction de leur capacité à lier le calcium et de leur séquence en acides aminés (43). Downing identifie trois catégories principales : les domaines EGF-like classiques, qui présentent une séquence conservée de cystéines formant des ponts disulfures stabilisant la structure ; les domaines EGF-like modifiés, caractérisés par des variations dans la séquence des cystéines ou dans le nombre de ponts disulfures ; et les domaines EGF-like liant le calcium (cbEGF), dotés d'une structure capable de lier le calcium et jouant un rôle clé dans des processus comme la coagulation sanguine.

En 1998, Bersch et ses co-auteurs identifient deux grands types de domaines EGF-like : les homologues du domaine EGF du facteur de croissance épidermique humain et les homologues du domaine EGF de la sérine protéase C1r du complément humaine (37).

En se basant en partie sur les travaux de Bersch, Wouters et ses co-auteurs ont développé de nouveaux motifs de séquence pour rechercher spécifiquement ces deux types de domaines EGF dans les bases de données de séquences SwissProt. Ainsi, il a contribué de manière significative à l'identification et à la classification des domaines EGF-like (38).

a. Domaines EGF classiques

i. Domaines classiques de l'EGF humain : hEGF-like (human EGF-like)

Ce type de domaine EGF possède généralement 8 à 9 acides aminés dans la partie C-terminale, entre la cinquième et la sixième cystéine. La glycine, située à 3 résidus en C-terminal de la dernière cystéine (C6-X-X-G) est très conservée et elle est généralement suivie d'un acide aminé chargé (R, Y, K, D, ou E) (FIGURE 9). Dans la partie N-terminale du domaine, la séquence entre la deuxième et la troisième cystéine comporte au moins 4 résidus : le deuxième résidu en N-terminal est une glycine ou une alanine et le troisième résidu en N-terminal est une glycine, un acide aspartique ou une glutamine (FIGURE 10). Dans certains cas, un résidu hydroxylé, sérine ou une thréonine, est retrouvé en position N-terminale de la troisième cystéine, dans la séquence C2-X-X-(G/D/Q)-(G/A)-(X/S/T)-C3 (FIGURES 9 et 10).

ii. Domaines apparentés au complément C1r : cEGF-like (C1r-EGF-like) de type 1 et 2

Les domaines de type cEGF ont la particularité de posséder 2 à 5 résidus supplémentaires (10 à 13 acides aminés) entre la cinquième et la sixième cystéine. Contrairement aux domaines de type hEGF, le troisième résidu en N-terminal de la sixième cystéine est un résidu volumineux (L, M, V, E ou I), généralement une leucine (L) (FIGURE 10).

Au sein des domaines de type cEGF, il existe des différences de séquence entre la deuxième et la troisième cystéine (C2-C3) qui ont permis d'identifier 2 sous types de domaine cEGF-like. Le premier sous type, appelé cEGF-like de type 1, comporte au moins 4 résidus entre la deuxième et la troisième cystéine, le deuxième résidu en N-terminal de la troisième cystéine est souvent une glycine (G), ce qui est commun aux domaines hEGF-like. Dans le sous type 2 la séquence possède généralement 3 résidus et la glycine (G) est remplacée par un résidu volumineux (L, M, V, E ou I) (FIGURE 9).

60 % des domaines EGF-like peuvent être classés dans les catégories hEGF-like et cEGF-like de type 1 et 2 (38). Cependant, la famille des domaines EGF-like est plus vaste et comprend diverses variantes structurales et fonctionnelles. Cette diversité est illustrée par les domaines EGF-like à 8 cystéines (40,44–46), qui diffèrent par leur configuration et leurs fonctions biologiques, offrant ainsi une complexité supplémentaire à l'architecture et aux rôles des protéines EGF-like.

b. Famille élargie : Domaines EGF-like à 8 cystéines

Les domaines EGF-like de type laminine et intégrine sont plus gros (40-60 résidus) et se différencient par la présence de 2 cystéines supplémentaires dans leur séquence, formant un pont disulfure interdomaine (FIGURE 11). Une cystéine supplémentaire se trouve à trois résidus C-terminal de la sixième cystéine (FIGURES 9, 10C ET 11). Cette cystéine forme un pont disulfure avec la cystéine d'un domaine EGF-like C-terminal en tandem. La position de la deuxième cystéine supplémentaire diffère dans les domaines EGF-like de type laminine et intégrine. Elle est la première cystéine dans les domaines EGF-like de type laminine et la quatrième cystéine dans les domaines EGF-like de type intégrine (FIGURE 10). La présence de ces deux cystéines ne modifie pas la connectivité canonique des cystéines homologues aux domaines hEGF-like et cEGF-like (38) (FIGURE 11). Les domaines EGF-like de type laminine et intégrine se distinguent aussi par le nombre de résidus entre la première et la deuxième cystéine. Les domaines EGF-like de type laminine ne comportent généralement qu'un seul résidu alors que les domaines EGF-like de type intégrine en comporte davantage, comme dans les domaines cEGF-like et hEGF-like (38,46,47)(FIGURE 9).

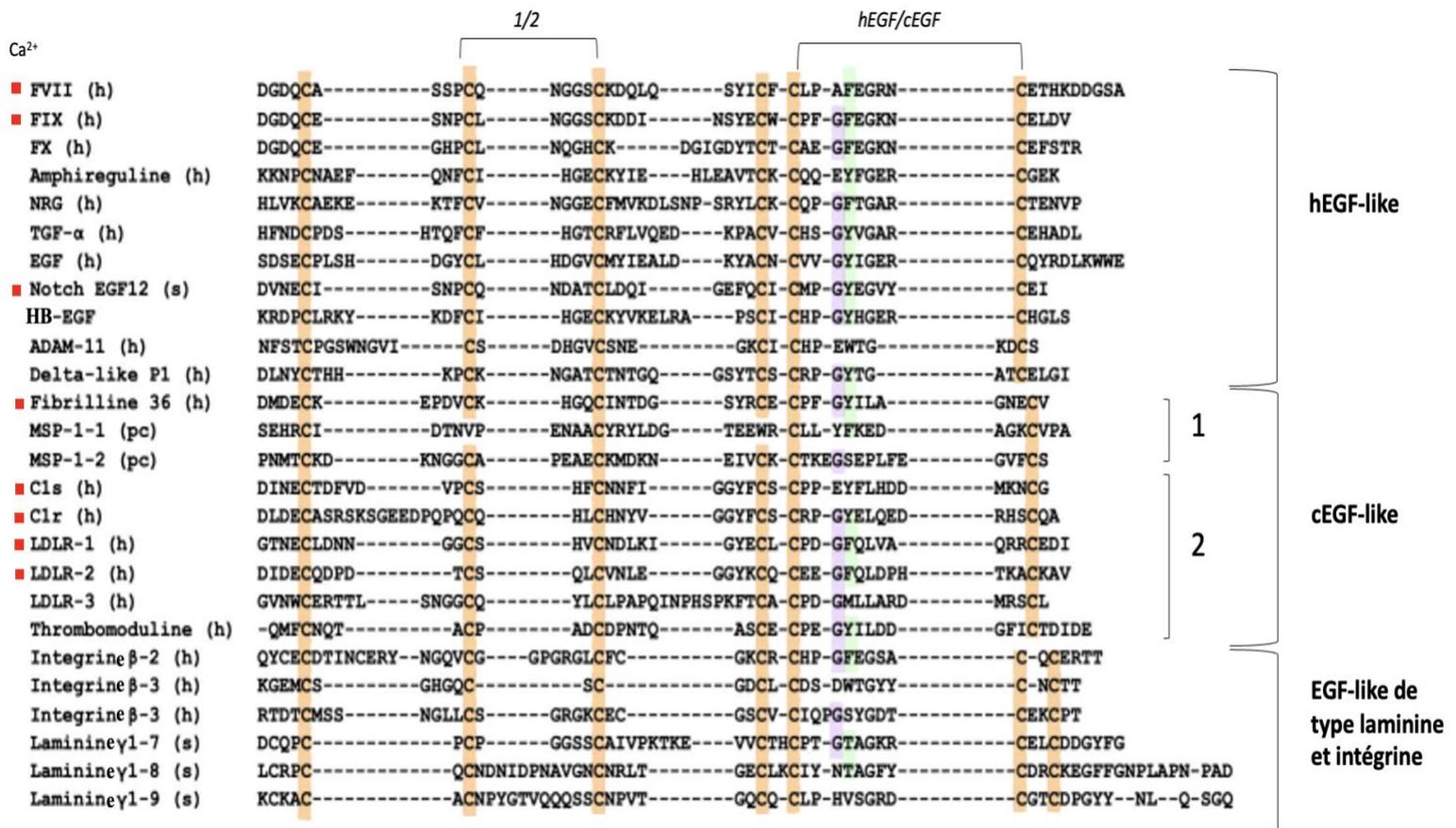


FIGURE 9. ALIGNEMENT DE SEQUENCES DES DOMAINES EGF-LIKE BASE SUR LEUR STRUCTURE ET CLASSIFICATION (adapté de (39))

Les domaines EGF-like sont retrouvés dans une grande variété de protéines avec une grande variété de motifs décrits. Dans certains cas, les domaines EGF-like présentent une grande similitude de séquence alors que dans d'autres cas ils peuvent avoir divergé considérablement. En outre, lorsque l'on compare les domaines EGF-like de protéines éloignées ou d'espèces différentes, le degré d'homologie peut être beaucoup plus faible et ce même parmi des protéines structurellement et fonctionnellement apparentées.

II) Structure secondaire et tertiaire

1) Homologie structurale

a. Comparaisons structurales

Dans les domaines EGF-like, malgré une faible homologie de séquence, on retrouve souvent un degré plus élevé d'homologie de structure en lien avec la conservation des

cystéines et la topologie de leurs ponts disulfures qui maintient la structure tridimensionnelle essentielle à leur fonction (48) (**ANNEXE 3**).

b. Domaines EGF-like classiques

i. Structure générale

Les domaines EGF-like présentent une structure conservée depuis la deuxième cystéine jusqu'à l'extrémité C-terminale de la chaîne polypeptidique. Cette partie du domaine est caractérisée par une alternance de brins β antiparallèles liés par des boucles de topologies variées et formant 2 feuillets β , le premier du côté N-terminal et le deuxième du côté C-terminal, nommés respectivement feuillet β majeur et mineur (48). Les deux brins du feuillet β majeur sont reliés par une boucle qui peut former un tour de type β II ou une boucle Ω lorsqu'il y a une insertion de résidus (37).

Le feuillet β majeur est stabilisé par un réseau régulier de liaisons hydrogènes inter-brins et se poursuit par une boucle formant un tour de type β II à partir de la glycine conservée (suivi de l'acide aminé aromatique en N-terminal) (38) (**FIGURE 12B**).

Le feuillet β mineur adjacent en C-terminal présente une irrégularité dans le motif des liaisons hydrogènes inter-brins par l'insertion d'un résidu sur le brin β C-terminal, formant une liaison hydrogène supplémentaire avec un résidu du brin β adjacent (37,38,48) (**FIGURE 10C**). Cette liaison hydrogène supplémentaire est à l'origine d'un renflement β causant une torsion du brin β C-terminal comprenant la dernière cystéine (38). Des liaisons hydrogènes formées entre un résidu polaire en N-terminal de la sixième cystéine et un résidu proche de la deuxième cystéine permettent de stabiliser le feuillet β mineur(37).

ii. Spécificités structurales

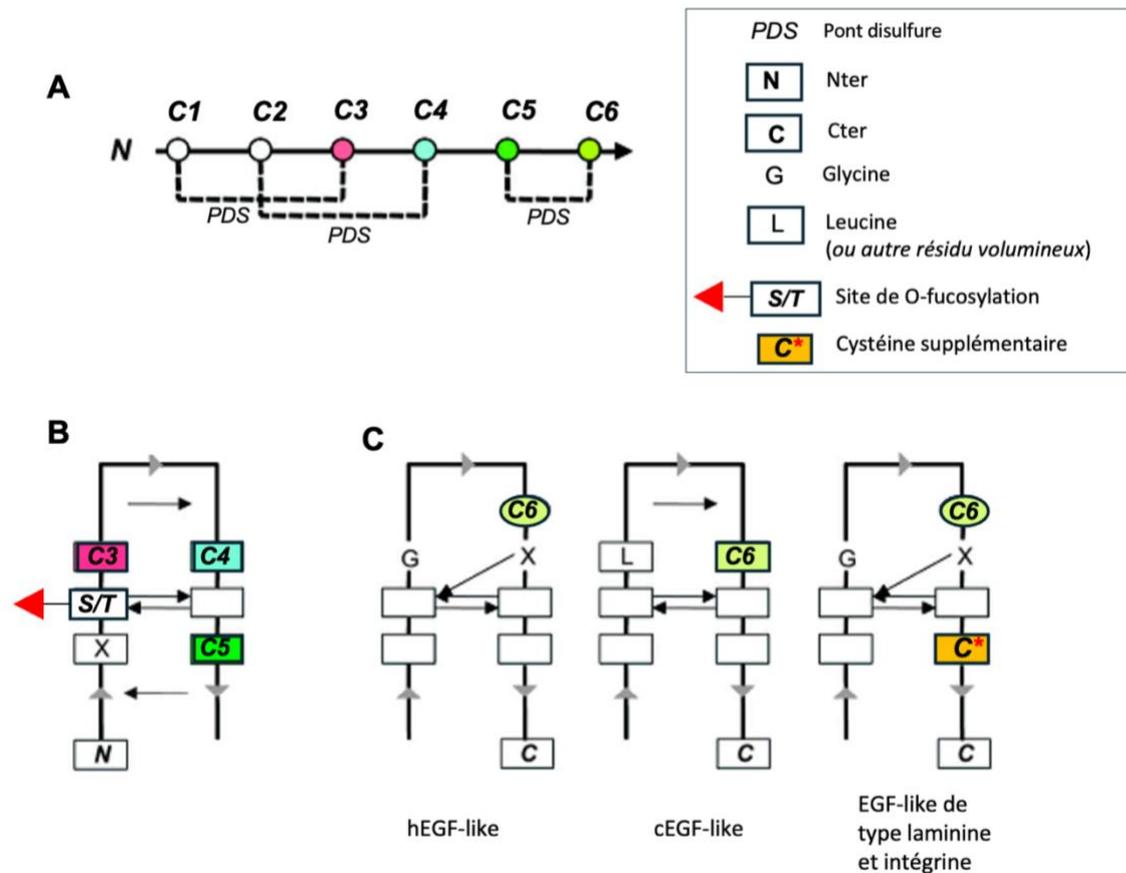


FIGURE 10 : POSITIONS DES CYSTEINES DANS LES DOMAINES EGF-LIKE. A) ARRANGEMENT DANS LA STRUCTURE PRIMAIRE. B) POSITION DES CYSTEINES DANS LE FEUILLET β MAJEUR. C) POSITIONS ALTERNATIVE DE LA SIXIEME CYSTEINE DANS LE FEUILLET β MINEUR (adapté de (38))

En comparant les structures des domaines EGF-like au niveau du feuillet β mineur, les chercheurs ont décrit une variation de forme et d'orientation ainsi qu'une variation du type de renflement β dans le feuillet β mineur (37,38). Ils expliquent cette différence par la nature du résidu en C-terminal du premier acide aminé formant des liaisons hydrogènes dans le brin N-terminal (acide aminé annoté X dans la classification de Chan et al.) (FIGURE 10C).

Dans les domaines hEGF-like ce résidu est généralement une glycine située à la tête d'épingle du feuillet (38). Cette glycine est la partenaire transversale d'un acide glutamique en N-terminal de la sixième cystéine. L'acide glutamique peut participer à la formation d'une liaison hydrogène avec le résidu X ce qui est à l'origine du renflement β (37).

Dans les domaines cEGF-like ce résidu est souvent une Leucine ou un acide aminé volumineux (M, V, E ou I) formant des interactions de type Vander Waals avec la cystéine

4 mais pas de liaison hydrogène avec la cystéine 6 (37,38). La position de ce résidu volumineux et ses propriétés sont responsables de la formation d'un renflement β de type S3 selon la classification de Chan et al.

Ainsi, lorsque l'on compare les structures des domaines EGF-like en alignant le réseau de liaisons hydrogènes du feuillet β mineur, on observe deux motifs structuraux différents. La conséquence de la variation du motif est la position alternative de la sixième cystéine dans les domaines hEGF-like et cEGF-like (**FIGURE 10C**). La cystéine 6 des domaines hEGF-like est déphasée d'un résidu et se trouve dans la tête d'épingle du feuillet alors que dans les homologues cEGF-like la cystéine est située dans le brin C-terminal (38) (**FIGURE 10C**). La cystéine est aussi en position alternative par rapport au renflement β du feuillet β mineur ce qui accentue la différence de forme et d'orientation de cette partie du domaine. Cette cystéine aurait donc une homologie de séquence mais pas d'homologie dans la structure des domaines EGF-like.

c. Famille élargie

Les domaines EGF-like de type laminine et intégrine adoptent une conformation de type hEGF-like (**FIGURE 10C**). Les deux cystéines supplémentaires décrites précédemment forment un pont disulfure supplémentaire entre deux domaines en tandem : entre la cystéine 1 et la cystéine 8 dans les domaines EGF-like de type laminine ; entre la cystéine 4 et la cystéine 8 dans les domaines EGF-like de type intégrine (34,38,46) (**FIGURE 11**).

La cystéine 7 (homologue de la sixième cystéine des domaines hEGF-like et cEGF-like) est contrainte par le pont disulfure au feuillet majeur de la même manière que dans les domaines hEGF-like. Le pont disulfure interdomaine ne semble pas impacter la forme et l'orientation du feuillet β mineur (38). Cette liaison disulfure participerait donc essentiellement au positionnement des domaines EGF-like les uns par rapport aux autres, formant dans certains cas une succession de domaines EGF-like en tandem (38,46,49).

En alignant les structures des domaines homologues au niveau des liaisons hydrogène du feuillet β mineur, Wouters décrit la position de la cystéine supplémentaire en C-terminal dans le même registre que la sixième cystéine des domaines cEGF-like.

d. Théorie évolutive

Dans sa revue, Wouters développe une théorie évolutive des domaines EGF-like pour expliquer les similitudes des domaines EGF-like à 8 cystéines aux domaines cEGF-like et hEGF-like. D'après lui, les domaines hEGF-like et cEGF-like proviendraient de la perte sélective d'une cystéine à l'extrémité C-terminale de la séquence ancestrale. En perdant la dernière cystéine, le domaine de type hEGF-like aurait conservé la connectivité de la cinquième et la sixième cystéine et ainsi la conformation du feuillet β mineur. En revanche dans les domaines cEGF-like, c'est l'avant dernière cystéine qui aurait été perdue ce qui aurait conduit à la capture de la première cystéine du domaine adjacent (FIGURE 10C).

Cela expliquerait entre autres la position alternative de la dernière cystéine, la formation d'une boucle plus longue, la présence d'un renflement β de type S3 et la conformation différente du feuillet β mineur dans l'espace. Outre, les données structurales, cette théorie s'appuie également sur la nature bimodale de la distribution des longueurs des boucles disulfures trouvées dans les domaines EGF-like (38).

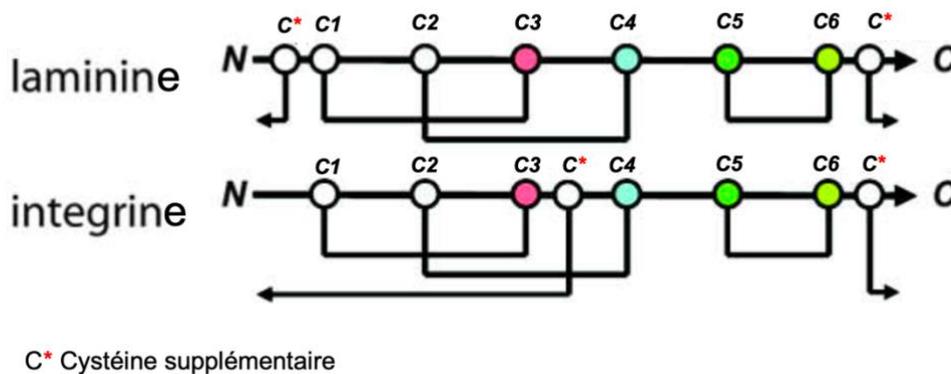


FIGURE 11. POSITION DES CYSTEINES SUPPLEMENTAIRES DES DOMAINES EGF-LIKE DE TYPE LAMININE ET INTEGRINE (adapté de (38))

Les variations dans la séquence s'accompagnent dans certains cas de modifications co- et post-traductionnelles influençant elles aussi la structure du domaine EGF-like.

2) Modifications co et post-traductionnelles

Plusieurs modifications de la séquence polypeptidique des domaines EGF-like par des sucres ont été décrites.

a. O-glycosylation

i. O-fucosylation

La O-glycosylation retrouvée majoritairement dans les domaines EGF-like est la O-fucosylation (38) qui consiste au couplage d'une molécule de fucose par la protéin-O-fucosyltransférase (POFUT1) sur la fonction hydroxylée d'un résidu sérine ou thréonine. Elle intervient au niveau de séquences consensus à 5 résidus entre 2 cystéines, retrouvés dans de nombreuses protéines dans une séquence de type C-X-X-X-X-(S/T)-C (**FIGURE 12A**). Dans les domaines EGF-like, ce fucose subit parfois une élongation par des glucosyltransférases, résultant à des motifs glycosidiques précis (50).

La O-fucosylation la plus décrite intervient dans la boucle disulfure C2-C3 des domaines hEGF-like, dans la séquence de type : C2-X-X-(D/Q)-(G/A)-(S/T)-C3 (51) (**FIGURE 10B**). Les résidus accompagnant l'acide aminé hydroxylé ont pour rôle de se lier à l'enzyme POFUT1 et initier la réaction. Les résidus voisins au site de fucosylation ont été décrits pour interagir avec l'enzyme. Les résidus X et la cystéine ont été décrit pour former des liaisons hydrogène avec des résidus de l'enzyme, la glycine quant à elle établit des liaisons de type Van Der Waals (52). De plus, la présence d'une sérine dans la boucle disulfure C1-C2 a été décrite pour faciliter l'interaction avec l'enzyme par la formation d'une liaison hydrogène avec une arginine. En revanche la boucle C3-C4 ne semble pas intervenir pas dans l'interaction avec POFUT1 (53).

Ensuite, le résidu O-fucosylé peut subir une élongation par l'ajout d'une molécule de N-acetylglucosamine par la β -1,3-N-acetylglucosaminyltransferase, Fringe (54). Ce motif peut être élongé d'avantage par l'ajout d'un galactose et un acide sialique pour former le tétrasaccharide sialylé $\text{sia}\alpha 2\text{-}3/6\text{Gal}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}3\text{fuc}\alpha$. Les enzymes à l'origine de ces réactions sont la β -1,4-galactosyltransferase et la $\text{sia}\alpha 2,3/\alpha 2,6$ sialyltransferase (50) (**FIGURE 12A**).

Jusqu'en 2017, les domaines EGF-like fucosylés décrits n'avaient été que des homologues hEGF-like comprenant 8 ou 9 résidus dans la boucle C5-C6 (55). Cependant en 2021 une étude a montré qu'un domaine de type cEGF-like, contenant 10 résidus dans sa boucle C5-C6 pouvait être O-fucosylé (50). Le résidu supplémentaire rendrait même possible la glycosylation simultanée d'autres résidus dans la boucle. Dans son étude Wouters évoque le fait que les domaines EGF-like O-fucosylés possèdent systématiquement 5 résidus formant une boucle de type $\beta 1$. Ainsi il semblerait que la

structure de l'épitope, déterminant important de la fucosylation (56), soit idéal dans certains domaines hEGF-like et cEGF-like de type 1.

Les domaines EGF-like des laminines subissent également la O-fucosylation qui est cruciale pour l'adhésion cellulaire et les interactions avec la matrice extracellulaire. De même, les intégrines présentent des sites de glycosylation influençant leur fonction et leur interaction avec les ligands extracellulaires (44,57).

ii. O-glycosylation

La O-glycosylation a été décrite sur la fonction hydroxyle de 2 sérines de la séquence des domaines hEGF-like. L'enzyme O-glycosyltransferase 1 (POGLUT 1) catalyse la réaction de couplage d'une molécule de glucose sur la fonction hydroxyle de la sérine dans la séquence C1-X-S-X-(P/A)-C2 (FIGURE 12A). Le monosaccharide peut être élongé par 2 xyloses par la β -D-glucoside- α 1,3-D-xylosetransferase et la α -D-glucoside- α 1,3-D-xylosetransferase. Moins souvent décrite, la réaction de couplage sur la sérine de la boucle disulfure C3-C4 dans la séquence C3-X-N-T-G-S-F-X-C4 est catalysée par la O-glycosyltransférase 2/3 (POGLUT 2/3) (50). L'élongation de ce motif n'est pas décrite dans la littérature (50).

iii. La O-GlcNAc-glycosylation

Dans les domaines EGF-like l'enzyme EOGT (EGF domain specific O-linked N-acetylglucosamine transferase) catalyse la réaction de couplage d'une molécule de N-Acétyleglucosamine sur le résidu sérine ou thréonine de la séquence C5-X-X-G-(F/Y)-(S/T)-G-X-X-C6. Le motif peut être élongé par un galactose puis un acide sialique (50,58) (FIGURE 12A)

b. N-glycosylation

La N-glycosylation est une opération biochimique co-translationnelle qui se produit dans le réticulum endoplasmique. Dans les domaines EGF-like, elle correspond à la fixation d'une arborescence sucrée de 14 sucres sur un acide aminé asparagine disponible dans la séquence N-X-(S/T)-C (où X représente n'importe quel acide aminé à l'exception de la proline pour des raisons d'encombrement stérique) (FIGURES 12A, 20 ET 21) (59). Cette réaction est catalysée par une enzyme membranaire de type oligosaccharyl-transférase, et débute en général par un N-acétyleglucosamine (51).

La glycosylation est décrite comme un mécanisme de protection vis-à-vis de la dégradation protéolytique (60). En effet, l'ajout de résidus sucrés impacte la structure : elle entraîne une stabilisation thermodynamique et cinétique des protéines (61).

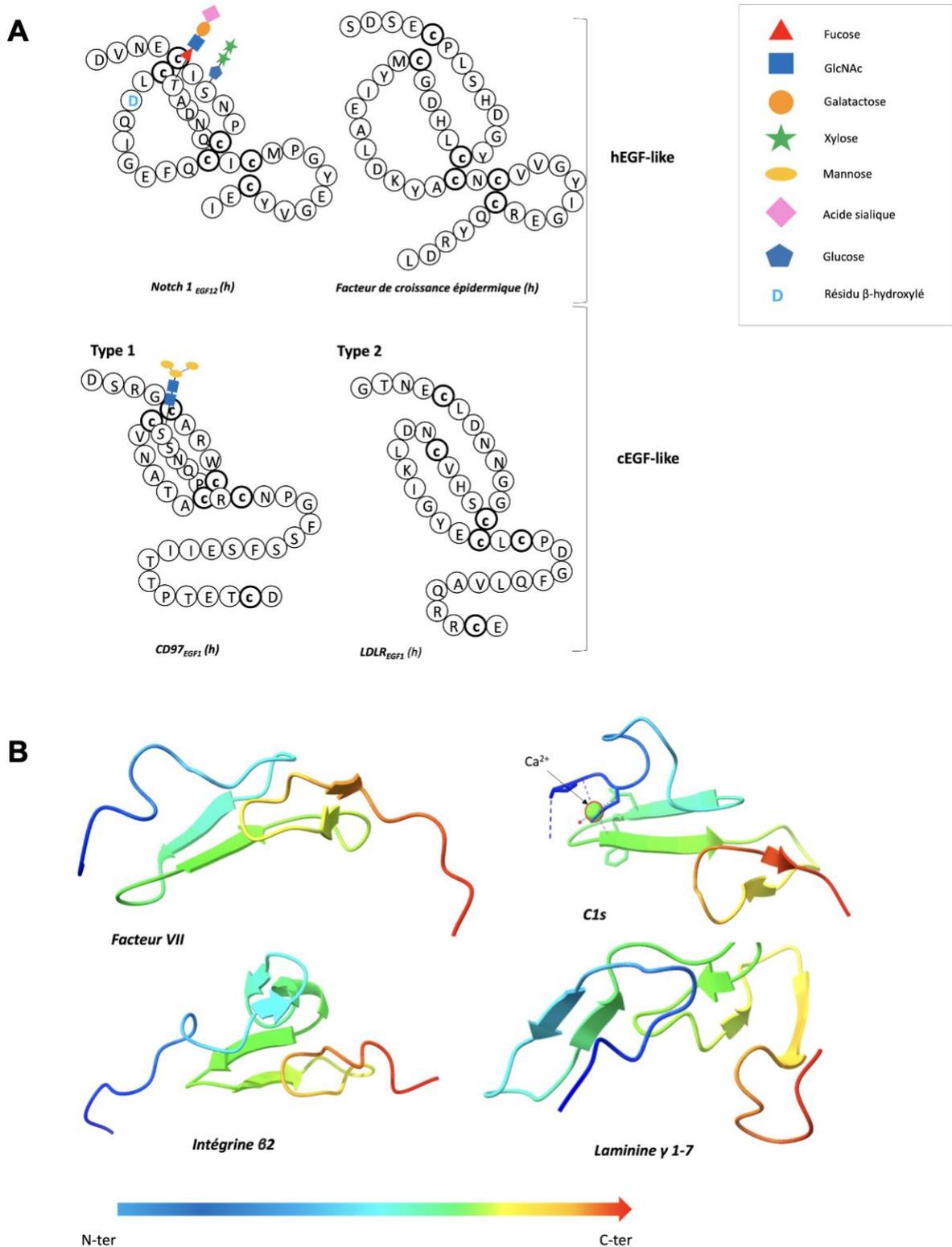


FIGURE 12. STRUCTURES SECONDAIRES DE DOMAINES EGF-LIKE CONTENUS DANS LES PROTEINES. A) STRUCTURES SECONDAIRES ET SITES DE MODIFICATIONS POST-TRADUCTIONNELLES. B) STRUCTURES 3D REPERTORIEES DANS LA PDB

c. β -hydroxylation

Cette modification post-traductionnelle n'a été décrite qu'à l'extrémité N-terminale des domaines EGF-like. Elle se produit sur un résidu acide aspartique dans les domaines hEGF-like et sur un résidu asparagine dans les domaines cEGF-like, dans la séquence C3-X-(D/N)-X-X-X-X-(F/Y)-C4, pour donner respectivement l'acide β -hydroxyaspartique et la β -hydroxyasparagine (38) (FIGURE 12A). Cette réaction est catalysée par des enzymes de type β -hydroxylase et nécessite la présence d'un résidu aromatique adjacent au résidu asparagine ou acide aspartique dans la structure en feuillet (62).

Il faut noter qu'il existe une proximité des sites de β -hydroxylation et de O-fucosylation dans les séquences qui suggère que les 2 modifications post-traductionnelles auraient une influence l'une sur l'autre (38) (FIGURE 9). Une hypothèse serait que ces sites de β -hydroxylation seraient des sites reconnus par les enzymes de type fucosyl-transférases pour réaliser la réaction de fucosylation sur le résidu sérine ou thréonine à proximité (52).

Généralement, la β -hydroxylation est impliquée dans la liaison au calcium, intervenant lui-même dans le mécanisme allostérique qui contrôle l'orientation du domaine EGF et plus généralement la structure quaternaire des protéines contenant des domaines EGF-like (63).

Ainsi, ces modifications covalentes de la structure ont un impact sur la structure et donc sur la fonction des domaines EGF-like.

III) Structure quaternaire

1) Calcium – binding

Beaucoup de domaines EGF-like sont décrits pour se lier au calcium à l'extrémité N-terminale de la chaîne polypeptidique (FIGURES 9, 12B, 18B, 19, 22). Les résidus impliqués dans l'interaction avec le calcium sont les résidus acide aspartique et glutamine de la séquence consensus du binding au calcium D-X-(D/N)-(Q/E)-C1. Le calcium interagit aussi fortement avec les chaînes latérales des résidus acide aspartique ou asparagine β -hydroxylés dans la boucle disulfure C3-C4 (64,65) (FIGURE 13). Les domaines EGF-like liant le calcium possèdent la même structure générale que ceux ne liant pas le calcium, en revanche, la boucle en N-terminal est tirée vers le feuillet β majeur (63). Ainsi, les interactions ioniques au calcium de part et d'autre de la boucle C1-C2 permettraient de la stabiliser et de mieux définir sa structure.

Outre la stabilisation du domaine, la liaison au calcium est cruciale pour l'interaction entre les modules d'une protéine contenant des domaines EGF-like en tandem. L'affinité des modules EGF-like isolés pour le calcium est moins bonne que pour les modules EGF-like en tandem. Par exemple, dans certaines protéines de la coagulation, cette affinité pour une paire de modules est 10 à 20 fois plus élevée que pour un module seul (63). Le calcium oriente successivement les modules EGF pour former une structure en forme de tige caractéristique, possédant une affinité 20 à 100 fois plus importante que pour un module seul. Finalement, le calcium agirait sur les domaines EGF-like en tandem comme la 4-hydroxyproline avec la protéine de collagène, ce qui souligne l'importance de cette interaction dans les protéines contenant des domaines EGF-like (FIGURE 13).

Les mutations des résidus cystéine et glycine en alanine peuvent être responsables d'un mauvais repliement de la protéine, entraînant une perte d'interaction similaire à ce qui est observé lors de la déplétion en calcium. Ces mutations augmentent la flexibilité du feuillet β majeur, le rendant distordu (66). En conséquence, la structure désordonnée due à la perte de la stabilisation par le calcium peut altérer les interactions des domaines EGF-like liant le calcium (63).

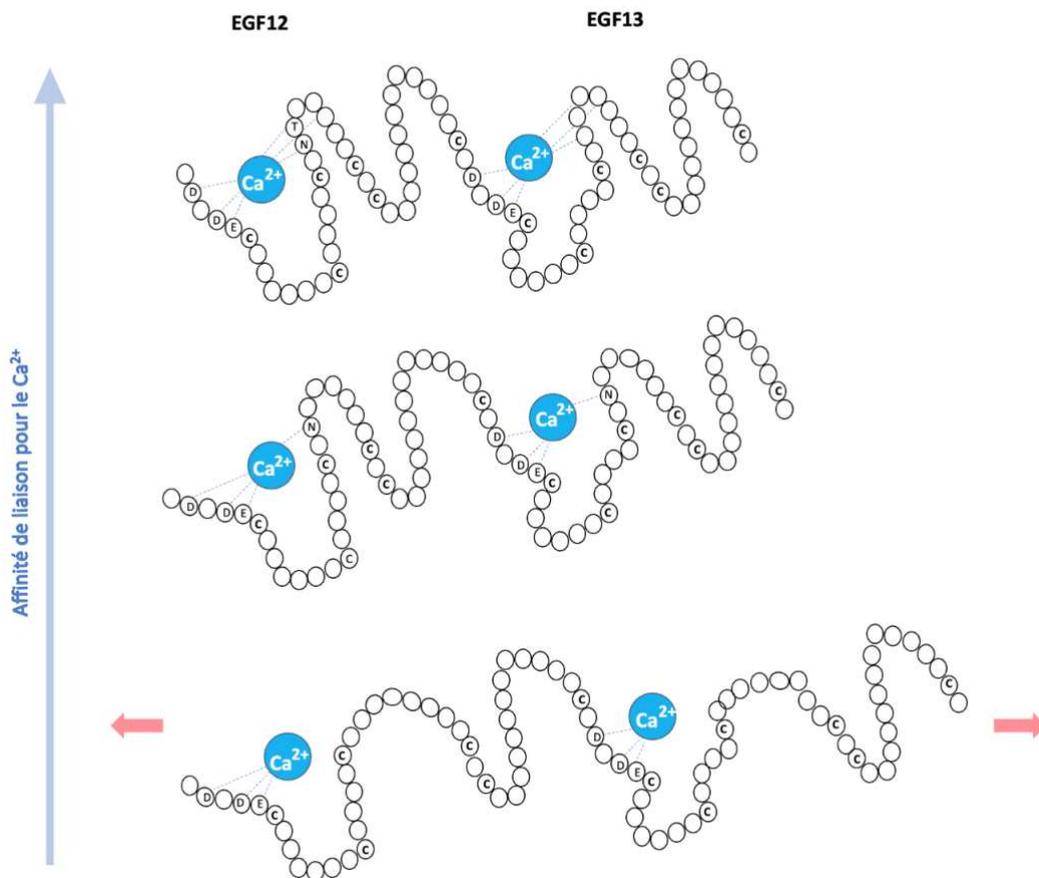


FIGURE 13. IMPACT DE LA LIAISON DU CALCIUM AUX DOMAINES EGF DE LA FIBRILLINE SUR SA STRUCTURE (adapté de (67))

2) Les domaines EGF-like dans les protéines

a. Fonction des domaines EGF-like dans les protéines

Les domaines EGF-like sont des unités modulaires que l'on trouve dans un large éventail de protéines extracellulaires (FIGURE 14) et qui peuvent exister sous forme de monomères, d'homodimères, d'hétérodimères ou même de multimères d'ordre supérieur de type structure en tandem. Les domaines EGF-like sont conservés au cours de l'évolution chez de nombreuses espèces, ce qui souligne leur importance dans la communication cellulaire et le maintien de la fonction des organismes multicellulaires (38). La divergence des domaines EGF-like a permis la diversification de la reconnaissance et de l'interaction moléculaire des protéines entre elles (68). En effet, les caractéristiques structurales essentielles et la variété des séquences retrouvées dans les domaines EGF-like leur permettent de se lier avec une grande affinité et spécificité à

d'autres protéines (FIGURE 12B). De nombreuses voies de signalisation, en particulier celles impliquées dans la croissance, la différenciation et le développement cellulaire, sont régulées par ces interactions (38).

En résumé, il semblerait que les domaines EGF-like soient intrinsèquement conçus pour les interactions. Leur stabilité structurale et la conservation évolutive de leurs sites d'interaction les rendent extrêmement efficaces dans la médiation des communications moléculaires qui sont cruciales pour la survie et l'homéostasie des organismes.

b. Structure en mosaïque des domaines EGF-like

De nombreuses protéines contiennent des domaines EGF-like en tandem et formant une mosaïque homogène cEGF-like ou hEGF-like dans leur structure. Par exemple, des protéines contenant exclusivement des domaines hEGF-like ont été décrites dans le développement embryonnaire (69), alors que d'autres ne contenant que des domaines cEGF-like interviennent dans la coagulation (70) (FIGURE 14).

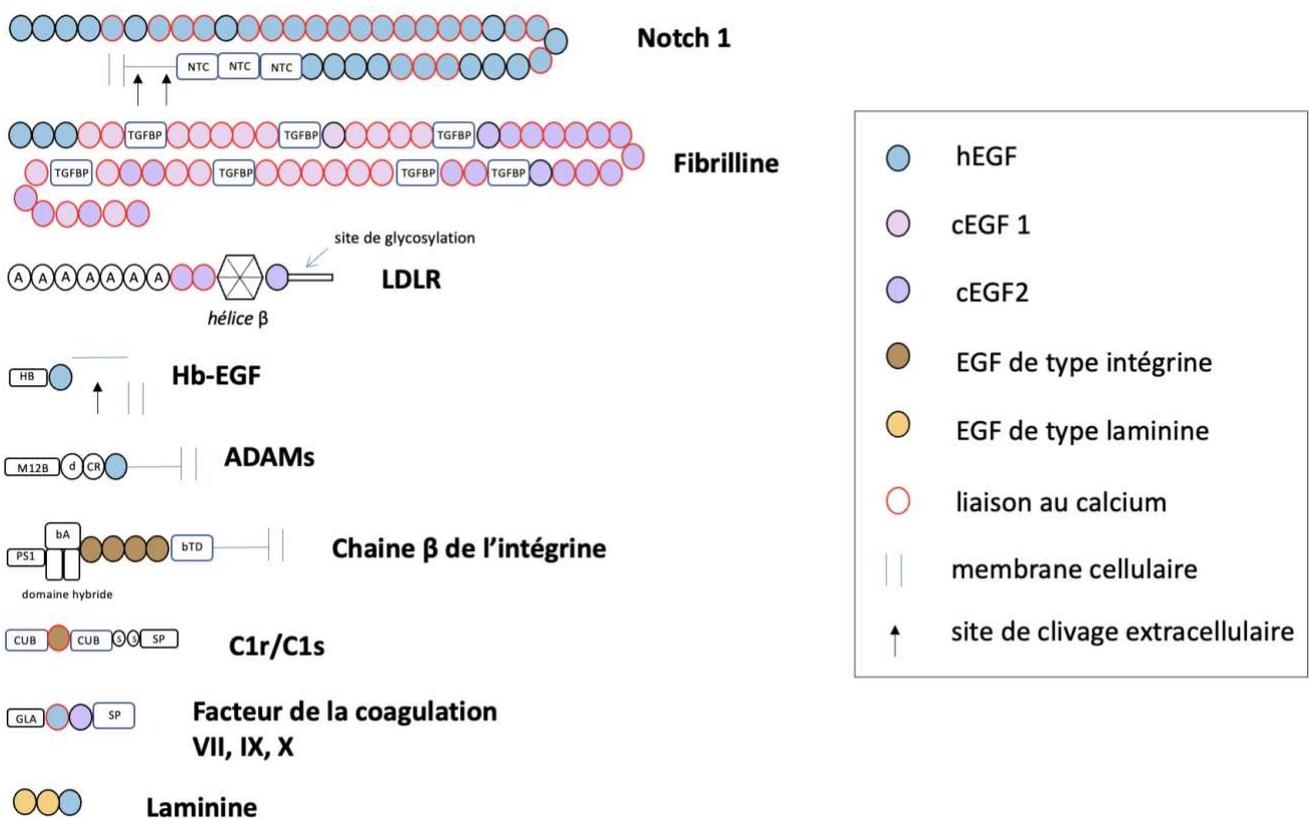


FIGURE 14. ARRANGEMENTS DES DIFFERENTS DOMAINES PROTEIQUES DE PROTEINES CONTENANT DES DOMAINES EGF-LIKE (adapté de (38))

Il existe aussi des mosaïques hétérogènes, contenant deux types de domaines EGF-like parmi les domaines de type hEGF-like, cEGF-like ou laminine EGF-like. Ces protéines sont généralement formées d'une structure bipartite avec les même types de domaines groupés entre eux à l'extrémité N-terminale ou C-terminale (38). Cet arrangement est par exemple souvent retrouvé dans les protéines de la matrice extracellulaire (40,44,46,71).

c. Protéines contenant des domaines EGF-like

La présence de ces domaines dans une protéine indique généralement un rôle dans la communication cellule-cellule ou cellule-matrice, souvent par le biais d'une interaction avec des récepteurs spécifiques à la surface de la cellule.

Dans la littérature, les domaines EGF-like sont décrits dans une grande variété de protéines aux fonctions diverses. En voici une liste non exhaustive :

- Les ligands des récepteurs aux facteurs de croissance à activité tyrosine kinases, en aval de la cascade de signalisation intracellulaire
Exemples : EGF, TGF- α , amphiréguline, neurégulines, HB-EGF (FIGURE 16), IGF
- Les protéines impliquées dans la cascade de la coagulation sanguine où les domaines de type EGF-like sont responsables de la médiation des interactions entre les divers composants du système de coagulation.
Exemples : facteurs de coagulation VII ; IX ; X, protéine C, protéine S (FIGURE 15), thrombomoduline (FIGURE 25)
- Les protéines d'adhésion cellulaire où les domaines EGF-like contribuent à l'interaction avec les ligands d'autres cellules, facilitant ainsi les processus d'adhésion cellule-cellule et cellule-matrice extracellulaire et signalisation.
Exemples : intégrines (FIGURE 17), laminines, protéines Notch (FIGURE 18), sélectines, cadhérines
- Les protéines impliquées dans la régulation du système immunitaire
Exemples : sous composants c1r du complément, fibrilline-1 (FIGURE 13), SCUBE, CD97 (FIGURE 21)

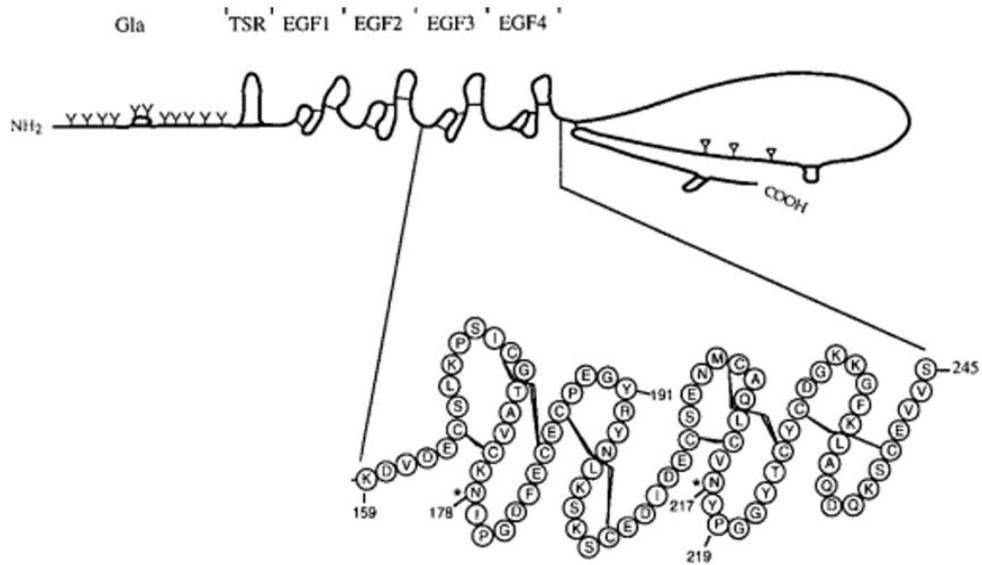


FIGURE 15. REPRESENTATION SCHEMATIQUE DE LA PROTEINE S DE LA COAGULATION AVEC LA STRUCTURE SECONDAIRE DE SES DOMAINES EGF3 ET EGF4 (adapté de (72))

La polyvalence des domaines EGF-like dans le milieu extracellulaire souligne leur importance dans le maintien de l'homéostasie cellulaire et des fonctions physiologiques appropriées. Les domaines sont retrouvés dans les protéines transmembranaires et dans les protéines solubles. D'ailleurs les protéines transmembranaires contenant des domaines EGF-like sont parfois les précurseurs d'effecteurs solubles qui une fois clivés par protéolyse sont libérés dans l'espace extracellulaire et agissent comme de véritables molécules de signalisation (72).

d. Mécanismes protéolytiques des domaines EGF-like

Les protéines membranaires contenant des domaines EGF-like peuvent subir un clivage membranaire ou intramembranaire (38) (FIGURES 14, 16 ET 18). Dans les protéines membranaires contenant des domaines cEGF-like et hEGF-like, les domaines hEGF-like sont généralement retrouvés à proximité de la membrane ce qui suggère leur implication dans la reconnaissance et la régulation du processus de clivage par les protéases. De plus, la transformation protéolytique des protéines contenant des domaines de type EGF-like est étroitement régulée par les interactions protéiques. En outre, les domaines hEGF-like sont aussi retrouvés dans la structure des métalloprotéases à zinc de la famille ADAM (69) (FIGURE 14).

L'hypothèse serait l'existence d'un mode de reconnaissance « semblable à semblable » entre les domaines hEGF-like des protéines transmembranaires et des protéases (38). Si les domaines apparentés au hEGF-like peuvent conférer des propriétés protéolytiques spécifiques aux protéines dont ils font partie, dans l'ensemble, la susceptibilité et le rôle dans la protéolyse seront très spécifiques de la fonction de la protéine et du contexte physiologique dans lequel la protéolyse se produit.

En conclusion, Les domaines EGF-like sont impliqués dans de nombreuses fonctions biologiques, notamment la signalisation intercellulaire et le développement. De plus, la diversité des protéines contenant des domaines EGF-like suggère un large éventail de fonctions biologiques qui restent encore à découvrir. Ainsi, la compréhension de leur structure et de leur mécanisme d'action continue d'être un domaine de recherche riche en potentiel pour de futures découvertes.

La fonction d'un domaine EGF-like peut varier considérablement en fonction de s'il est contenu dans une protéine soluble ou transmembranaire. Si les protéines solubles et les protéines transmembranaires peuvent toutes deux interagir avec les molécules du compartiment extracellulaire, les mécanismes d'interaction et les réponses cellulaires qui s'ensuivent diffèrent (73). Les protéines solubles diffusent librement pour interagir avec leurs partenaires protéiques, tandis que les protéines transmembranaires servent d'intermédiaires entre l'environnement extracellulaire et les voies de signalisation intracellulaires en transmettant des signaux à travers la membrane.

Les interactions des protéines contenant des domaines EGF-like avec d'autres molécules et leur capacité à influencer la communication cellulaire en font des cibles potentielles pour le développement de nouvelles thérapies. Le ciblage thérapeutique des protéines transmembranaires permet de mettre au point des traitements qui s'attaquent aux mécanismes moléculaires sous-jacents de la maladie avec une meilleure spécificité d'action améliorant les profils de sécurité du principe actif.

Dans la suite de cette thèse, nous nous intéresserons exclusivement aux protéines membranaires contenant des domaines EGF-like. Nous verrons d'abord leurs rôles physiologiques et physiopathologiques puis leurs intérêts en thérapeutique.

PARTIES 3 : PROTEINES MEMBRANAIRES CONTENANT DES DOMAINES EGF-LIKE

Comme pour les formes solubles, les domaines EGF-like ancrés à la membrane se sont avérés être impliqués dans la transmission des signaux et la régulation de la fonction des récepteurs, utilisés par les cellules dans la croissance et le développement (74). Les études ont montré que, comme pour les domaines EGF-like solubles, les domaines EGF-like ancrés à la membrane jouent un rôle critique dans le maintien de l'homéostasie tissulaire, la survie cellulaire, et la régulation de l'immunité (75,76). Les protéines membranaires contenant des domaines EGF-like induisent des mécanismes sophistiqués de régulation et des réponses cellulaires à divers stimuli externes. Il semblerait que les domaines EGF-like aient été exploités par les protéines membranaires au cours de l'évolution pour diversifier et spécialiser leurs fonctions. Leur implication dans ces processus cellulaires s'étend à leur impact délétère dans les conditions pathologiques, y compris le cancer et les maladies inflammatoires.

Une exploration plus poussée de leurs mécanismes et de leurs interactions promet d'améliorer notre compréhension des comportements cellulaires complexes et de mettre en lumière de nouvelles cibles thérapeutiques.

I) PHYSIOPATHOLOGIE

1) Développement et croissance

a. HB-EGF, un facteur de croissance membranaire

i. Physiopathologie

La protéine HB-EGF est composée d'un domaine de liaison à l'héparine et d'un domaine EGF-like indispensable pour la liaison aux récepteurs HER1 et HER4. Il existe sous une forme ancrée à la membrane, active et capable d'interagir avec ses récepteurs initiant ainsi une signalisation de type juxtacrine. Dans les épithéliums, cette interaction protège les cellules de l'anoïkose (75). HB-EGF ancrée à la membrane permet la survie des tissus épithéliaux après le détachement des cellules de la lame basale tout en maintenant un phénotype épithélial, ce qui prévient une évolution maligne. Cet effet cytoprotecteur contribue à la réparation et à la régénération des tissus. Contrairement

aux ligands solubles qui eux favorisent la dédifférenciation et la perte des marqueurs mésenchymateux (77), la forme ancrée de HB-EGF induit une réponse très différente sur les mêmes récepteurs.

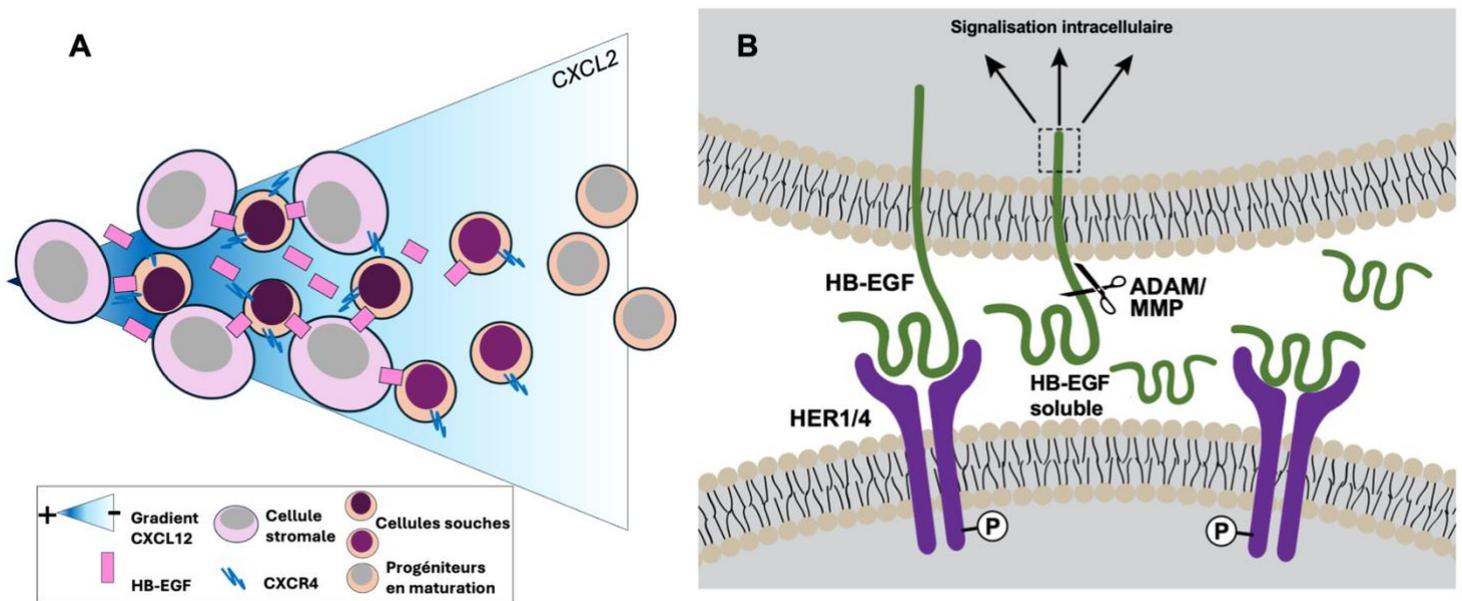


FIGURE 16. LA PROTEINE HB-EGF. A) SIGNALISATION JUXTACRINE, AUTOCRINE ET PARACRINE DE LA PROTEINE HB-EGF DANS LES NICHES HEMATOPOIETIQUES (adapté (76)) B) REPRESENTATION SCHEMATIQUE DE LA PROTEINE HB-EGF A LA MEMBRANE (adapté de (78))

La protéine joue un rôle crucial dans l'environnement hématopoïétique où elle est exprimée dans les niches des cellules souches de la moelle épinière, co-exprimée avec CXCL12 et son récepteur CXCR4. La liaison de CXCL12 à son récepteur induit le clivage de HB-EGF, activant HER1 par signal autocrine et paracrine, et déclenchant des cascades de signalisation conduisant à la différenciation et à la prolifération (FIGURE 16) (78).

Ce phénomène est largement observé dans les cellules immunitaires et favorise leur infiltration dans les tissus sains, contribuant ainsi aux maladies auto-immunes. Ce mécanisme est également détourné par les cellules souches cancéreuses pour former des niches permettant de cacher la tumeur, assurant ainsi la survie des cellules cancéreuses (79).

ii. HB-EGF dans les leucémies aiguës et chroniques

Dans les leucémies aiguës et chroniques, caractérisées par une leucocytose de cellules souches hématopoïétiques immatures et une déficience des fonctions

hématopoïétiques, une surexpression de HB-EGF, CXCR4 et CXCL12, induite par l'activation de signaux oncogènes tels que Ras et MAPK via BCR-Abl est observée (78).

La protéine CD9 joue un rôle crucial dans l'organisation des microdomaines membranaires et l'assemblage des complexes protéiques à la surface cellulaire. CD9, en se fixant à HB-EGF augmente l'activité mitogénique du signal juxtacrine, contribuant ainsi à la prolifération cellulaire (80). Cette interaction se produit principalement au niveau des jonctions adhérentes, où elle participe aux interactions cellule-cellule. D'ailleurs, une colocalisation de CD9 et HB-EGF a été observée dans ces jonctions adhérentes, en association avec l'intégrine $\alpha 3\beta 1$. En s'associant à ce complexe, l'intégrine prolonge l'interaction entre HB-EGF et CD9, renforçant ainsi d'avantage le signal juxtacrine (76). Ce mécanisme souligne l'importance de la coordination entre ces protéines dans la régulation des processus cellulaires cancéreux et par conséquent l'intérêt de les étudier.

b. L'intégrine $\alpha\beta 8$ et son architecture unique

i. Physiopathologie

L'intégrine $\alpha\beta 8$ possède une structure unique qui la distingue des autres intégrines, notamment par sa conformation étendue fermée (FIGURE 17). Cette particularité est due à la suppression d'un pont disulfure dans le domaine EGF-like 1, impactant la structure de la boucle C-terminale (81). Dans la sous unité β , le changement de conformation modifie l'interface d'interaction entre les domaines EGF1 et EGF2 contribuant à la flexibilité de la sous unité $\beta 8$ et par conséquent à l'extension de l'intégrine (FIGURE 17B) qui ne peut pas stabiliser sa forme repliée inactive. En conséquence, l'intégrine $\alpha\beta 8$ présente une affinité accrue pour ses ligands au niveau des domaines EGF-like (71). L'un de ces ligands est la protéine LTBP, dont la liaison avec $\alpha\beta 8$ entraîne le relargage de TGF- β , une cytokine de l'inflammation, sans nécessiter l'activation préalable de l'intégrine (82).

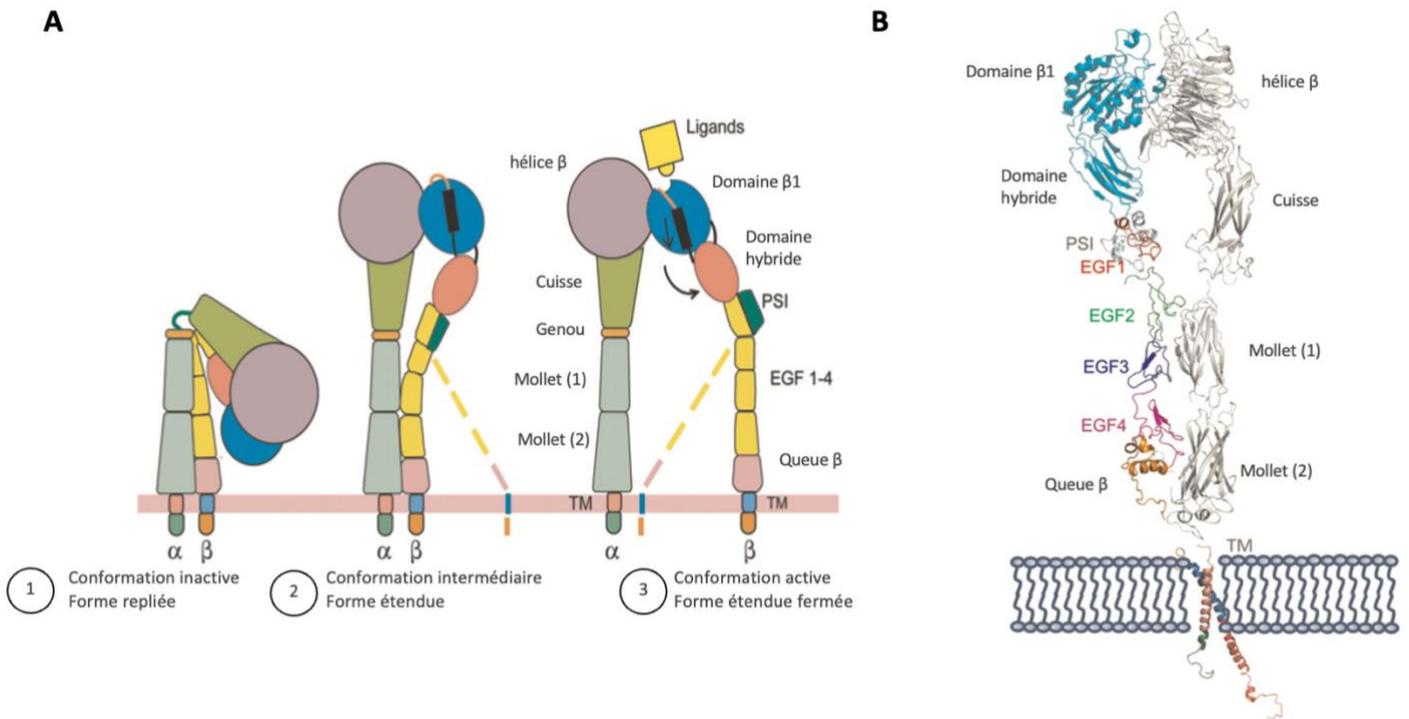


FIGURE 17. STRUCTURE DES SOUS UNITES ALPHA ET BETA DES INTEGRINES. A) REPRESENTATION SCHEMATIQUE DES DIFFERENTES CONFORMATIONS. B) STRUCTURE 3D DE LA FORME ETENDUE FERMEE DE L'INTEGRINE $\alpha\beta 8$. (adaptés de (81))

ii. L'intégrine $\alpha\beta 8$ dans la fibrose pulmonaire

L'intégrine $\alpha\beta 8$ est exprimée dans divers tissus, notamment le cerveau, le rein et le système nerveux central. Son implication pathologique est particulièrement notable dans le contexte de la fibrose.

Dans la bronchopneumopathie chronique obstructive (BPCO), le relargage de TGF- β par $\alpha\beta 8$ conduit à l'activation des fibroblastes du tractus respiratoire. Cela entraîne une surexpression de facteurs caractéristiques de la fibrose comme le collagène de type I et l'actine α du muscle lisse. Le relargage de la cytokine et son action paracrine entraîne en parallèle une réduction la prolifération des cellules épithéliales (82). La diminution de cette prolifération perturbe le remodelage du tissu épithélial, un facteur crucial dans la progression de la maladie.

2) Homéostasie et survie

a. Les protéines NOTCH dans l'homéostasie vasculaire

i. Physiopathologie

Les protéines NOTCH sont des récepteurs transmembranaires essentiels qui, une fois activés par leurs ligands déclenchent une cascade de signalisation intracellulaire

influençant de manière significative le destin cellulaire. En modulant l'équilibre entre la prolifération et la différenciation, les protéines NOTCH jouent un rôle crucial dans le maintien de l'homéostasie dans divers tissus, y compris la peau, l'intestin et le système immunitaire (69). Les domaines EGF-like présents dans les protéines NOTCH sont déterminants pour leur fonction.

Ces domaines permettent à NOTCH de se lier spécifiquement à ses ligands, tels que Delta et Jagged (26) (FIGURE 18). Cette interaction initie le clivage protéolytique de NOTCH et conduit à la libération de son domaine intracellulaire. Une fois libéré, ce domaine intracellulaire transloque vers le noyau où il régule l'expression de gènes cibles, orchestrant ainsi les processus de différenciation et de survie cellulaire.

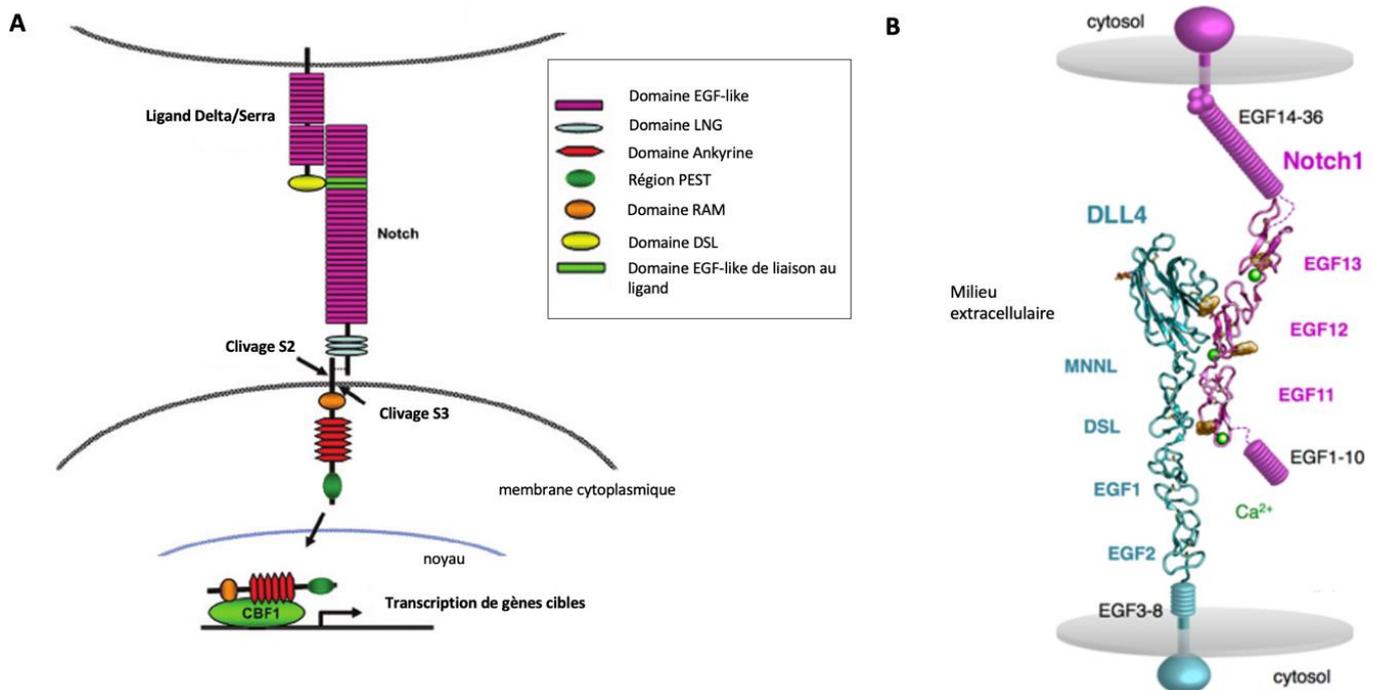


FIGURE 18 : LES PROTEINES NOTCH. A) REPERESNTATION SCHEMATIQUE DE LA VOIE DE SIGNALISATION NOTCH. B) STRUCTURE 3D DE LA PROTEINE NOTCH1 LIE A SON LIGAND DLL4 (adaptés de (83))

ii. Mutation de NOTCH3 dans la dégénérescence du tissu vasculaire

L'implication physiopathologique des protéines NOTCH, en particulier NOTCH3, dans le CADASIL (Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy) est cruciale pour comprendre l'homéostasie vasculaire et les dysfonctionnements associés.

NOTCH3 est exprimée à la surface des cellules musculaires lisses situées dans les parois des vaisseaux sanguins. Dans le CADASIL, il y a des mutations spécifiques dans les domaines EGF10 et EGF11 de NOTCH3, touchant une cystéine essentielle. Ces mutations entraînent un mauvais repliement du domaine EGF-like glycosylé, empêchant l'élongation normale de la molécule de fucose. Ce défaut structural conduit à une co-agrégation des protéines NOTCH3 et LTBP-1 (Latent TGF- β Binding Protein-1), entraînant le relargage aberrant de TGF- β .

Cette agrégation interfère avec la machinerie cellulaire, provoquant un dérèglement des fonctions cellulaires et une dégénérescence progressive du tissu vasculaire. La dégénérescence s'accompagne d'une inflammation et d'un épaissement des parois des vaisseaux (83). Ces altérations structurales et fonctionnelles augmentent le risque ischémique dans les petites artères cérébrales. En conséquence, les patients atteints de CADASIL subissent des accidents vasculaires cérébraux (AVC) récurrents. Ce mécanisme est aussi impliqué dans les déficits neurologiques et cognitifs caractéristiques de cette pathologie (84).

b. LDL-R dans le métabolisme hépatique des lipides

i. Physiopathologie

Les domaines EGF-like du récepteur LDL (LDLR) sont essentiels pour sa conformation et sa fonction. Ces domaines sont caractérisés par des ponts disulfures et des sites de liaison au calcium, qui jouent un rôle crucial dans la stabilisation de la structure du LDLR à pH neutre, favorisant ainsi une liaison optimale aux LDL (85).

Lors du transport du complexe LDL-LDLR vers les endosomes, l'environnement intracellulaire devient progressivement acide. À pH acide, les domaines EGF-like subissent un changement de conformation réversible (**FIGURE 19**). Cette altération de conformation est cruciale pour le relargage du LDL dans l'endosome, permettant la libération du cholestérol pour les besoins cellulaires. Après le relargage du LDL, le LDLR déchargé est recyclé vers la surface cellulaire. Ce mécanisme de recyclage est essentiel pour maintenir l'efficacité du LDLR dans la clairance continue du LDL du plasma sanguin, régulant ainsi les niveaux de cholestérol et contribuant à l'homéostasie lipidique (86). Un

dérèglement de ce processus de recyclage entraîne une accumulation de LDL dans le plasma sanguin. Ce mécanisme est largement décrit dans les hypercholestérolémies.

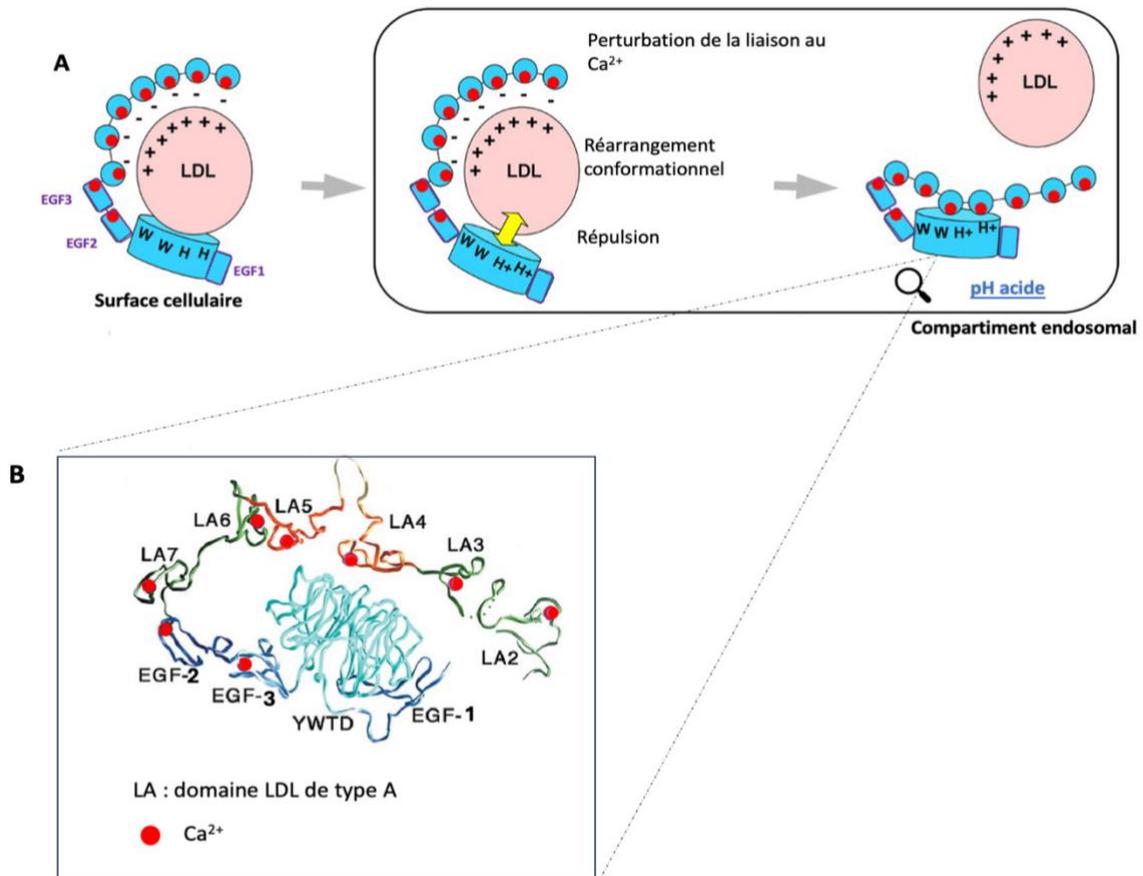


FIGURE 19. LE RECEPTEUR AU LDL. A) LIAISON ET LIBERATION DES LDL SUITE A UN CHANGEMENT CONFORMATIONNEL INDUIT PAR LE PH (adapté de (88)). B) STRUCTURE 3D DU RECEPTEUR LDL. (adapté de (87))

ii. Des mutations des domaines EGF-like du récepteur LDL dans des cas d'hypercholestérolémies familiales

Des mutations faux-sens au niveau d'une cystéine perturbent une boucle qui lie le calcium entraînant un défaut dans le changement conformationnel du récepteur lors des variations de pH dans la cellule. Ce défaut structural empêche le mécanisme de recyclage du LDL-R, qui ne parvient plus à être relargué à la membrane des hépatocytes et reste bloqué dans le réticulum endoplasmique. En conséquence, le nombre de récepteurs LDL fonctionnels à la surface des cellules est réduit, entraînant une accumulation de LDL-cholestérol libre dans le plasma sanguin (27).

Cette accumulation de LDL-cholestérol libre augmente significativement le risque d'athérosclérose et d'accidents vasculaires cérébraux (AVC) dans le contexte de l'hypercholestérolémie familiale.

c. Cripto-1, un modulateur crucial des voies de signalisation oncogéniques

i. Physiopathologie

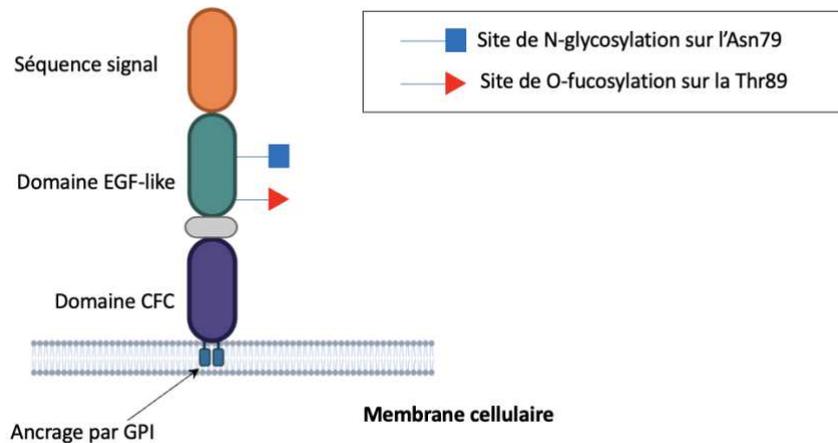


FIGURE 20. REPRESENTATION SCHEMATIQUE DE LA PROTEINE CRIPTO-1 (adapté de (89))

Cripto-1, une protéine cruciale dans la transformation oncogénique, nécessite la O-fucosylation de son domaine EGF-like pour exercer son activité (FIGURE 20). Elle joue un rôle significatif dans la phosphorylation du récepteur du facteur de croissance épidermique HER1.

Cripto-1 agit en tant que ligand de la protéine Glypican-1, une protéine membranaire impliquée dans la transduction du signal et l'activation de la kinase c-Src, qui à son tour phosphoryle les domaines de la protéine HER1. Cette phosphorylation d'HER1 par c-Src conduit à l'activation des voies de signalisation MAPK, PI3K, et AKT. Ces voies sont cruciales pour la survie cellulaire et possèdent des propriétés anti-apoptotiques, ce qui favorise la progression tumorale (90).

La surexpression de Cripto-1 a été observée dans de nombreux types de cancers, notamment le cancer colorectal, le cancer du sein, le cancer de l'estomac, le cancer pancréatique, le cancer des ovaires et le cancer des poumons. Normalement, Cripto-1 est exprimée à de faibles niveaux dans les tissus sains (91).

ii. Cripto-1 est associé à mauvais pronostic dans les cancers

L'association de Cripto-1 avec un mauvais pronostic est bien documentée. Par exemple, lors d'une infection par *Helicobacter pylori*, la surexpression de Cripto-1 suivant l'infection constitue un facteur de risque important pour la carcinogénèse

gastrique (90). En résumé, Cripto-1, par ses interactions et la modulation de voies de signalisation critiques, joue un rôle central dans le développement et la progression de divers cancers.

3) Communication et immunité

a. La superfamille des protéines EGF-TM7 dans la transduction du signal

i. Physiopathologie

Les protéines de la superfamille EGF-TM7 sont des récepteurs couplés à une protéine G, caractérisés par un domaine transmembranaire en hélice α qui traverse la membrane sept fois (FIGURE 21). Ces récepteurs possèdent une queue cytoplasmique qui propage le signal extracellulaire à l'intérieur de la cellule. Les protéines EGF-TM7 existent sous différentes isoformes, qui diffèrent principalement par le nombre et la composition des domaines EGF-like présents dans la région extracellulaire. Ces domaines sont essentiels pour les interactions avec divers ligands, et cette diversité modifie leur affinité et leur spécificité de liaison vis-à-vis des ligands. Ce mode de communication est particulièrement utilisé par les cellules immunitaires(92).(93)

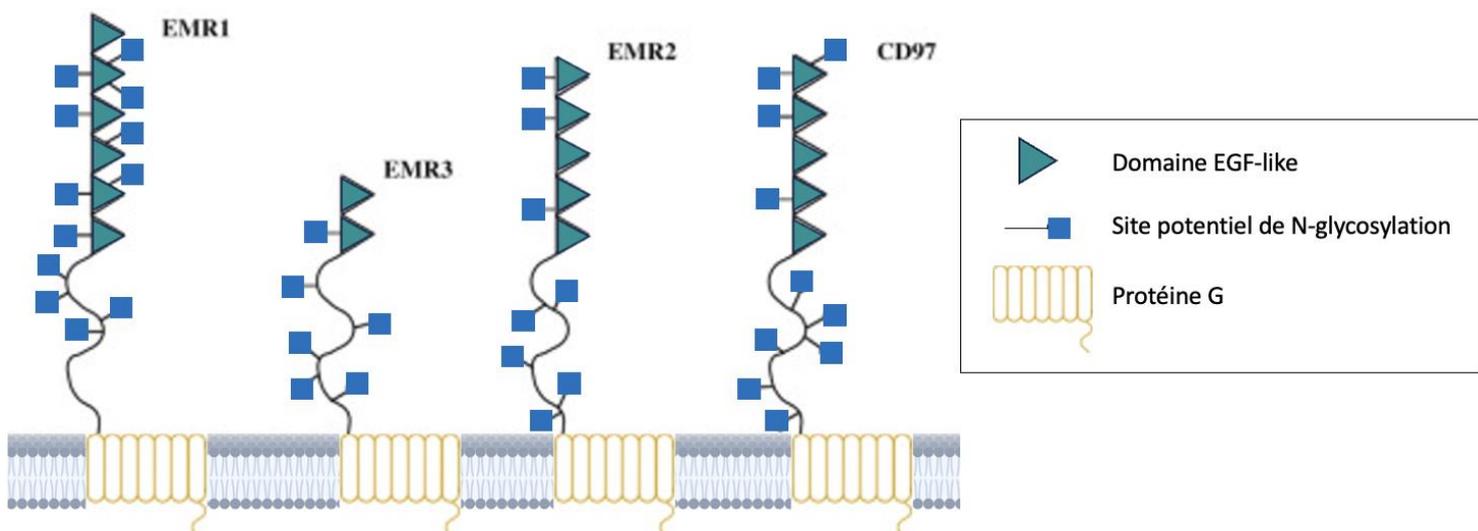


FIGURE 21: REPRESENTATION SCHEMATIQUE D'ISOFORMES DE PROTEINES EMR (adapté de (93))

ii. La protéine CD97

CD97 dans les maladies auto-immunes

Les protéines EMR constituent un sous-type de la superfamille EGF-TM7. Parmi elles, CD97 (aussi appelée ADGRE5 et EMR2) est exprimée à la surface des cellules

immunitaires activées (FIGURE 22). La fixation de son ligand CD55 au niveau de ses domaines EGF-like (94), favorise la migration des leucocytes en régulant leur adhésion aux composants de la matrice extracellulaire. La fixation de CD55 entraîne aussi une transduction du signal conduisant à l'expression de certains gènes de la progression cellulaire (94).

Dans les maladies auto-immunes, une surexpression de CD97 est observée, et son activation contribue à l'expression des composants du microenvironnement immunitaire. Par exemple, dans la polyarthrite rhumatoïde, CD97 est surexprimé dans le tissu synovial. La protéine forme un complexe avec la chondroïtine de la matrice extracellulaire ce qui participe à la sévérité de l'inflammation des articulations (95).

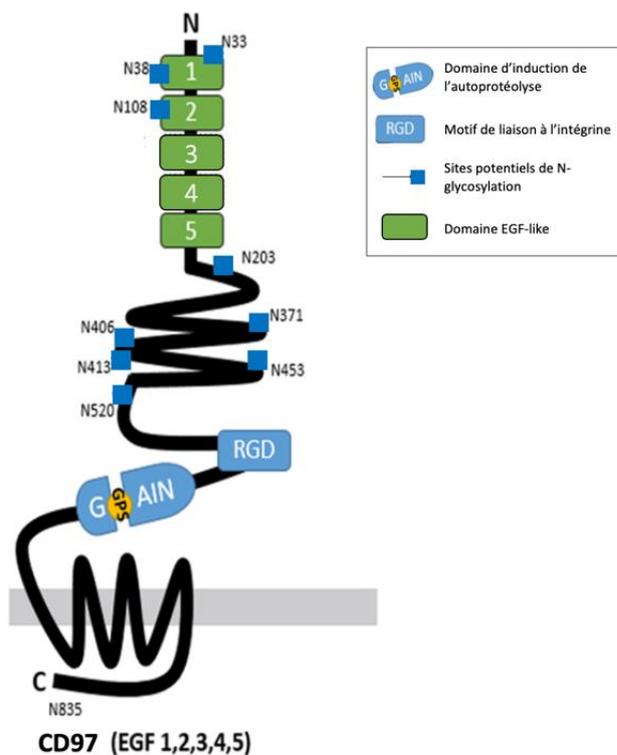


FIGURE 22: REPRESENTATION SCHEMATIQUE DE LA PROTEINE CD97 (adapté de (96))

iii. CD97 et ELTD1 dans les gliomes

Le gliome est une tumeur cérébrale provenant des cellules gliales du cerveau. Dans le glioblastome, une forme très agressive de cette tumeur cérébrale, l'activation de CD97 conduit à l'expression de métalloprotéinases qui dégradent la matrice extracellulaire, favorisant ainsi le caractère très invasif de cette tumeur (97).

Une des particularités dans le glioblastome est que la tumeur crée un environnement hypoxique qui induit l'expression de gènes pro-angiogéniques, dont la protéine ELTD1, un autre membre de la EGF-TM7 (FIGURE 23), qui favorise la croissance des vaisseaux sanguins pour améliorer l'apport en oxygène et en nutriments à la tumeur. Cela conduit à de l'angiogenèse, une caractéristique marquante du gliome et plus particulièrement du glioblastome (98).

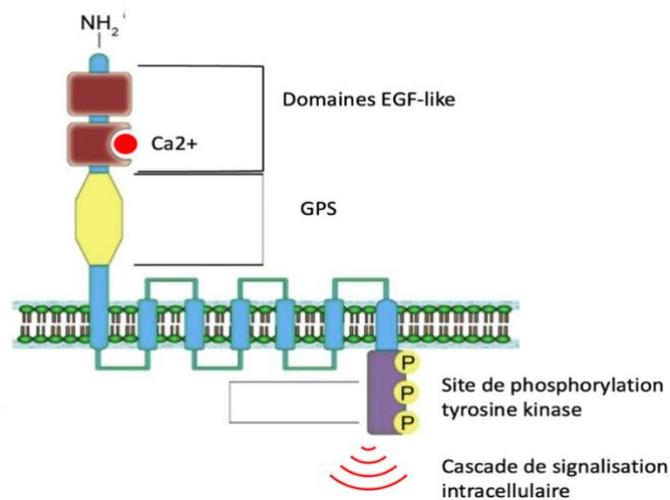


FIGURE 23. REPRESENTATION SCHEMATIQUE DE LA STRUCTURE DE LA PROTEINE ELTD1. (adapté de (97))

b. Les mucines membranaires

i. Physiopathologie

Les mucines membranaires forment une grande famille de O-glycoprotéines, synthétisées par les épithéliums, dont une partie sécrétée forme des gels pour leur protection. Elles jouent un rôle crucial dans la lubrification, la transduction de signaux et les interactions cellule-cellule et cellule-matrice. Parmi les mucines membranaires, certaines possèdent des domaines EGF-like, comme MUC3, MUC4 et MUC13 (99). Dans le cholangiocarcinome par exemple, MUC13 interagit avec HER1, entraînant l'activation des voies de prolifération et de survie des cellules cancéreuses. (100)

La mucine 4 (MUC4) participe à la protection des épithéliums intestinaux et gastriques. Cependant, dans certains cancers épithéliaux, la fonction de MUC4 est détournée. MUC4 interagit avec les récepteurs HER2 via ses domaines EGF-like, ce qui a pour conséquence la promotion de l'oncogenèse par l'activation de voies de signalisation intracellulaires (101). De plus, MUC4 empêche l'internalisation du

récepteur, prolongeant ainsi l'activation des voies de signalisation oncogéniques, conduisant à une agressivité tumorale accrue et à un caractère invasif de la tumeur (102) (FIGURE 19). Ce complexe protéique a été décrit dans les cancers du sein, des poumons et du pancréas (103–105).

ii. *MUC4 dans la résistance aux traitements anticancéreux*

MUC4 est responsable de la résistance à certaines thérapeutiques. Cela est dû à l'encombrement stérique causé par la partie supraglycosylée de MUC4, qui empêche les molécules thérapeutiques d'atteindre leur cible pharmacologique. Ce phénomène survient notamment avec la gemcitabine dans le cancer du pancréas (106) et le trastuzumab (herceptin) dans le cancer du sein (107) (FIGURE 24).

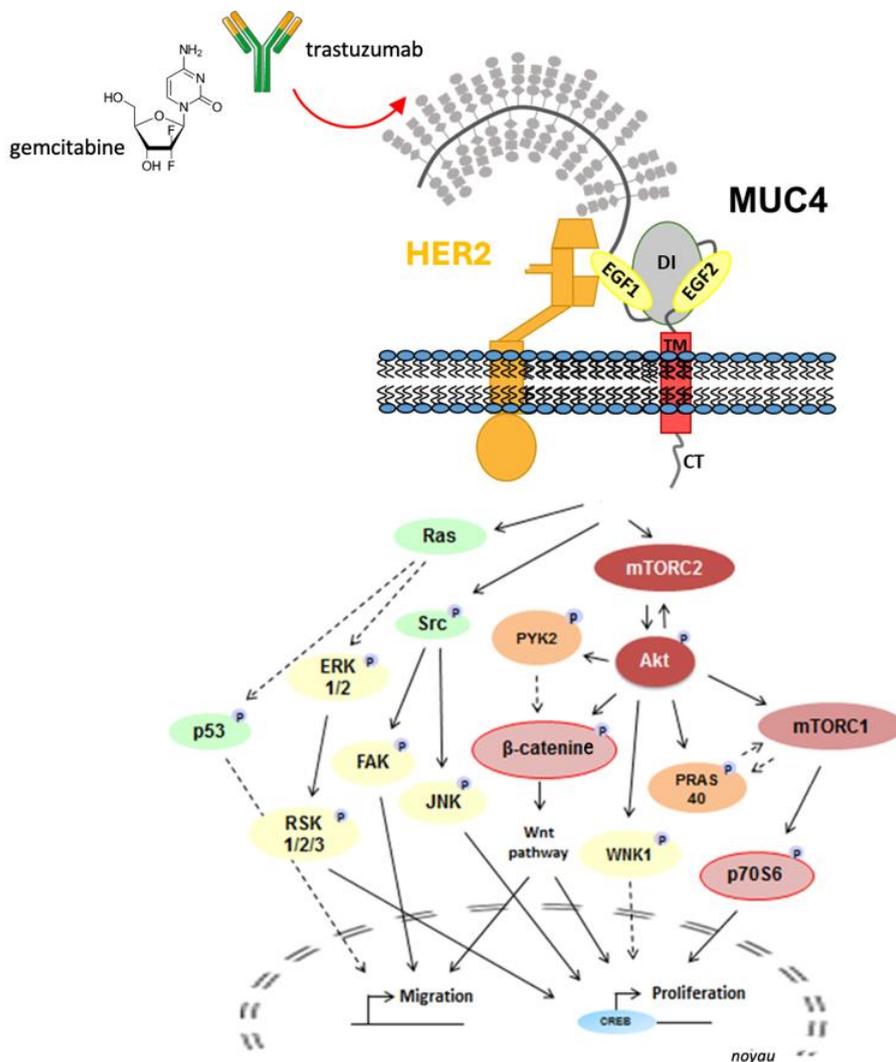


FIGURE 24. INTERACTION MOLECULAIRE ENTRE LES PROTEINES HER2 ET MUC4. A) REPRESENTATION SCHEMATIQUE DES RESISTANCES ET DES VOIES DE SIGNALISATION INTRACELLULAIRES ACTIVEES (non publié)

II) INTERET EN THERAPEUTIQUE

Les domaines EGF-like contenus dans les protéines membranaires présentent un intérêt thérapeutique dans le diagnostic des pathologies. De plus, le ciblage de ces protéines peut représenter une stratégie prometteuse dans le traitement de diverses maladies.

1) Intérêt dans le diagnostic

a. Diagnostic mutationnel

i. *Hypercholestérolémie familiale*

Les mutations des domaines EGF-like du récepteur LDL (LDLR) constituent un marqueur diagnostic crucial de l'hypercholestérolémie familiale, une condition génétique prédisposant les individus à des maladies cardiovasculaires précoces. Environ (51,7 %) des variants faux-sens identifiés dans le gène LDLR concernent spécifiquement les domaines EGF-like (27), ce qui souligne leur importance en clinique.

En tant que cause génétique primaire de l'hypercholestérolémie familiale, les mutations des domaines EGF-like du récepteur LDL permettent un diagnostic précoce de la maladie grâce à des tests génétiques. L'identification précise de ces mutations est essentielle non seulement pour confirmer le diagnostic, mais aussi pour adapter la stratégie thérapeutique à chaque patient (97). Cela favorise ainsi une approche de médecine personnalisée visant à réduire et à prévenir les complications cardiovasculaires.

ii. *Syndrome d'Alagille*

Le syndrome d'Alagille est une maladie orpheline causée par des mutations autosomiques dominantes des protéines JAG1 et NOTCH2 (108). Les mutations identifiées se situent souvent dans les domaines EGF-like des protéines. Ces mutations perturbent leur structure, affectant ainsi la liaison du ligand à son récepteur et l'activation de la cascade de signalisation intracellulaire à l'origine des processus de développement des tissus. La perturbation de ces signaux est responsable de manifestations cliniques variées incluant des anomalies hépatiques, cardiaques ou rénales (109).

Le syndrome d'Alagille est causé principalement par des mutations de la protéine JAG1 pour laquelle plus de 400 mutations ont été décrites, dont 70% sont des mutations

tronquant la protéine. Dans 1% des cas, la maladie est causée par des mutations de la protéine NOTCH2. Ces mutations sont associées à des malformations rénales plus fréquentes par rapport aux mutations de la protéine JAG1. Le diagnostic clinique est souvent suivi par un séquençage des gènes du patient pour confirmer la présence de mutations dans les protéines JAG1 ou NOTCH2. Dans certaines cohortes, le taux de détection des mutations de JAG1 est d'environ 94% (108).

b. Biomarqueur thérapeutique

i. ELTD1 dans le glioblastome

La protéine ELTD1 se distingue en tant que marqueur de l'angiogenèse dans les glioblastomes. Elle est utilisée comme biomarqueur diagnostique pour stratifier les patients atteints de glioblastomes en fonction son taux d'expression (110). Cette stratification permet d'identifier les patients les plus susceptibles de répondre favorablement aux thérapies anti-angiogéniques (111). De plus, ELTD1 peut également servir de biomarqueur pour évaluer la réponse thérapeutique, permettant ainsi des ajustements précoces et ciblés des thérapeutiques.

L'utilisation de ELTD1 dans cette double capacité diagnostique et de suivi thérapeutique contribuerait à une gestion plus personnalisée des glioblastomes, améliorant potentiellement les chances de survie des patients.

ii. Cripto-1 dans la progression tumorale

Cripto-1 est un biomarqueur thérapeutique important en raison de son implication dans les processus clefs de la progression tumorale. La protéine est impliquée dans la transition épithélio-mésenchymateuse et le maintien des cellules souches cancéreuses (112). Sa surexpression dans les cancers en fait un indicateur de pronostic intéressant. En effet, des niveaux élevés d'expression de Cripto-1 ont été corrélés à une progression plus sombre dans certains cancers.

Par exemple des niveaux élevés de Cripto-1 sont associés à une survie globale plus faible dans l'adénocarcinome pulmonaire, le carcinome épidermoïde œsophagien et le carcinome hépatocellulaire (113).

L'implication importante de Cripto-1 dans des processus cruciaux liés à la progression tumorale en fait un biomarqueur thérapeutique intéressant offrant également un potentiel pour le développement de thérapies efficaces (91).

2) Ciblage thérapeutique

a. Ciblage thérapeutique indirect

i. Utilisation des domaines EGF-like comme scaffold

Une stratégie de ciblage indirect des protéines membranaires contenant des domaines EGF-like consiste à utiliser ces domaines en tant que *scaffolds* (échafaudages) thérapeutiques. Les domaines EGF-like sont particulièrement intéressants en raison d'une structure à la fois flexible et stable (34). Cette stabilité peut être accrue par la capacité des domaines EGF-like à lier le calcium. De plus, la polyvalence des domaines EGF-like en font de bons candidats pour l'ingénierie moléculaire. Ils peuvent être modifiés pour créer de nouveaux peptides bioactifs utilisés en thérapeutique (FIGURE 25). Ces séquences peuvent être ajustées pour générer des analogues présentant des activités spécifiques.

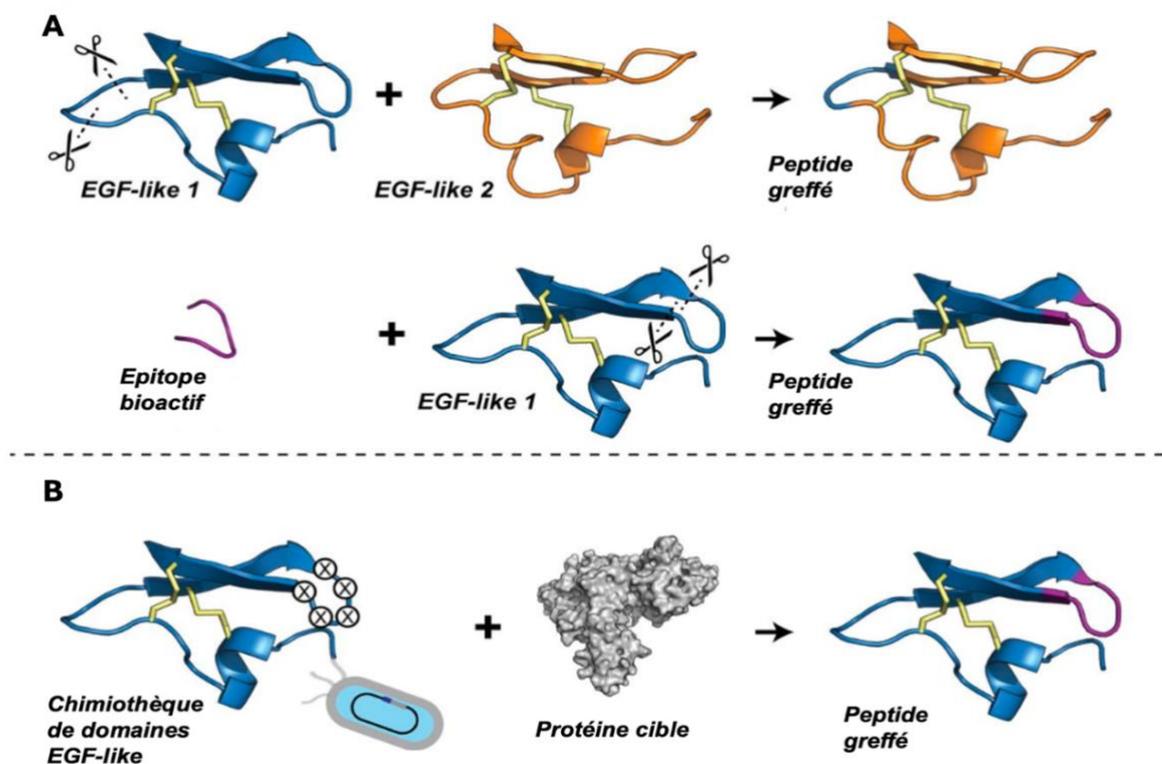


FIGURE 25. CONCEPT DE GREFFE MOLECULAIRE UTILISANT LES DOMAINES EGF-LIKE. A) FUSION DES DOMAINES EGF-LIKE ENTRE EUX OU A D'AUTRES EPITOPES BIOACTIFS. B) FUSION DES DOMAINES EGF-LIKE A LA SURFACE D'UN BACTERIOPHAGE POUR LE CRIBLAGE DE CIBLES PROTEIQUES (adapté de (34))

La thrombomoduline dans le traitement des maladies thromboemboliques

La thrombomoduline est une protéine endothéliale qui, en se liant à la thrombine, participe au maintien d'un microenvironnement anti-coagulant et anti-inflammatoire.

Des études de mutagenèse ont confirmé que les domaines EGF-like de la thrombomoduline sont essentiels pour moduler l'activité et la liaison à la thrombine (70). Plus précisément, ce sont le domaine EGF5 et une petite partie N-terminale du domaine EGF6 qui occupent l'interface d'interaction (114) (FIGURE 26). Cette interaction est cruciale pour l'activation de la protéine C et de l'inhibiteur de fibrinolyse.

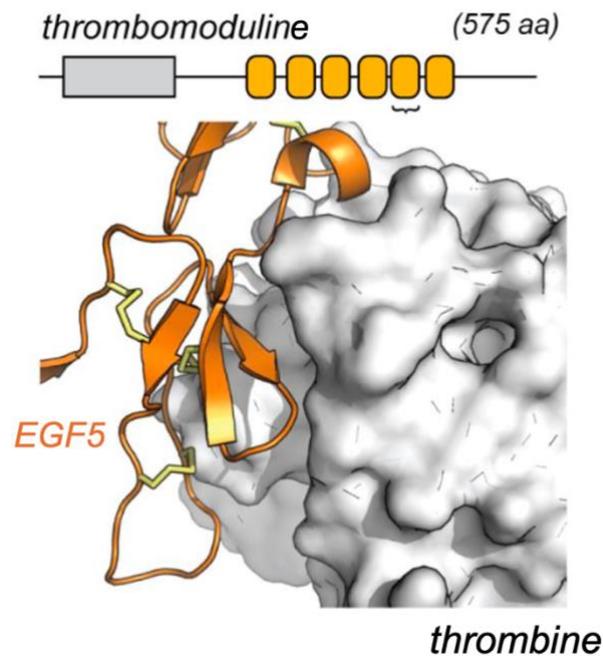


FIGURE 26. FIXATION DU DOMAINE EGF5 DE LA THROMBOMODULINE A LA THROMBINE (adapté de (34))

Un fragment recombinant de la thrombomoduline mimant l'interaction à la thrombine a été développé (102,103) pour traiter les troubles de la coagulation comme par exemple la coagulation intravasculaire disséminée (CIVD) (115), caractérisée par une coagulation excessive souvent causée par la septicémie. Les fragments basés sur les domaines EGF de la thrombomoduline, en particulier les domaines EGF4-6, ont été utilisés pour bloquer l'interaction thrombomoduline-thrombine (116,117).

Par conséquent cette stratégie peut s'appliquer aux maladies thromboemboliques, où la modulation de la coagulation et de l'inflammation est essentielle pour prévenir les complications.

ii. Blocage des voies de signalisation en amont

Le TNF α soluble est connu pour induire l'expression de la glycoprotéine MUC4 responsable des résistances au trastuzumab et la gemcitabine évoquées

précédemment. En bloquant le TNF α l'expression de MUC4 est diminuée ce qui permet la levée de l'encombrement stérique permettant le démasquage de l'épitope du trastuzumab et la pénétration intracellulaire de la gemcitabine (FIGURE 24) (113). De plus, l'association du trastuzumab et des médicaments bloquant le TNF α renforce la réponse immunitaire en modifiant l'environnement immunosuppresseur créé par les tumeurs et en favorisant une réponse immunitaire innée impliquant les lymphocytes NK et les macrophages (118).

b. Ciblage thérapeutique direct

Une stratégie thérapeutique prometteuse consiste à cibler spécifiquement les domaines EGF-like dans les protéines membranaires.

i. Cibler ELTD1 dans le glioblastome

Des anticorps ciblant spécifiquement la partie extracellulaire de la protéine ELTD1 contenant les domaines EGF-like ont été développés. En se liant, l'anticorps empêche au récepteur de fonctionner, bloquant ainsi les voies de signalisation intracellulaires. Ce processus, implique notamment la diminution de l'expression de la protéine NOTCH1 impliquée dans la différenciation cellulaire et la signalisation angiogénique. Cela a pour conséquence de réduire la densité des micro-vaisseaux au sein de la tumeur et de réduire la croissance et la prolifération des cellules cancéreuses (119,120).

ii. Cibler MUC4 dans le cancer du pancréas

Dans le contexte du cancer du pancréas, la néoexpression de la mucine MUC4 est corrélée à un pronostic sombre, avec une survie à cinq ans inférieure à 15 % en moyenne (121). MUC4 joue un rôle crucial dans la progression de la maladie par l'activation de voies de signalisation oncogéniques.

MUC4 active la formation d'un complexe quaternaire très agressif composé de MUC4, HER2 et de ses partenaires d'hétérodimérisation, HER3 et son ligand la neureguline (122). Cette interaction active et prolonge les voies de signalisation oncogéniques, entraînant une croissance tumorale et une incidence métastatique accrues.

Cibler MUC4 pour lever les résistances

Lorsqu'une tumeur devient résistante aux traitements, les options thérapeutiques, déjà limitées, se réduisent considérablement, rendant la gestion de la

maladie encore plus complexe. MUC4 est impliquée dans la résistance à la gemcitabine, l'un des traitements de première intention dans le cancer du pancréas. Dans d'autres contextes, MUC4 est également impliquée dans la résistance de certaines tumeurs surexprimant HER2 et traitées au trastuzumab (123) (FIGURE 24).

Des études préliminaires ont montré que le ciblage de MUC4 avec des anticorps monoclonaux (124) permet de réduire ces résistances. Cette approche pourrait restaurer la sensibilité des tumeurs à la gemcitabine et au trastuzumab, offrant ainsi de nouvelles options thérapeutiques pour les patients (125).

Ciblage des domaines EGF-like pour la mise au point d'inhibiteur d'interaction protéine-protéine

Des études préliminaires ont démontré la capacité d'un peptide ligand des domaines EGF1 et EGF2 de MUC4 à inhiber l'interaction avec HER2 (FIGURE 27A). Cette inhibition permet une resensibilisation de HER2 au trastuzumab (herceptin) en levant l'encombrement stérique de la partie supraglycosylée de MUC4. Des tests ont été menés dans une lignée cellulaire d'adénocarcinome pancréatique humain exprimant MUC4 (Pan-89). Les cellules sont traitées par le trastuzumab marqué par un fluorophore vert en combinaison avec le peptide ligand des domaines EGF1 et EGF2 (FIGURE 27B). Par cette approche, l'efficacité du trastuzumab couplé à un anticorps secondaire fluorescent est restaurée. Cette inhibition se traduit par une diminution significative de la prolifération des cellules cancéreuses pancréatiques et par une potentialisation de l'effet antiprolifératif du trastuzumab (FIGURE 27C).

Cette stratégie thérapeutique constitue une preuve de concept du ciblage thérapeutique des domaines EGF-like de MUC4. Elle pourrait être étendue à d'autres cancers exprimant le complexe MUC4-HER2. En ciblant spécifiquement les interactions entre MUC4 et HER2, il est possible d'intervenir à la fois sur les mécanismes de résistance et de prolifération tumorale.

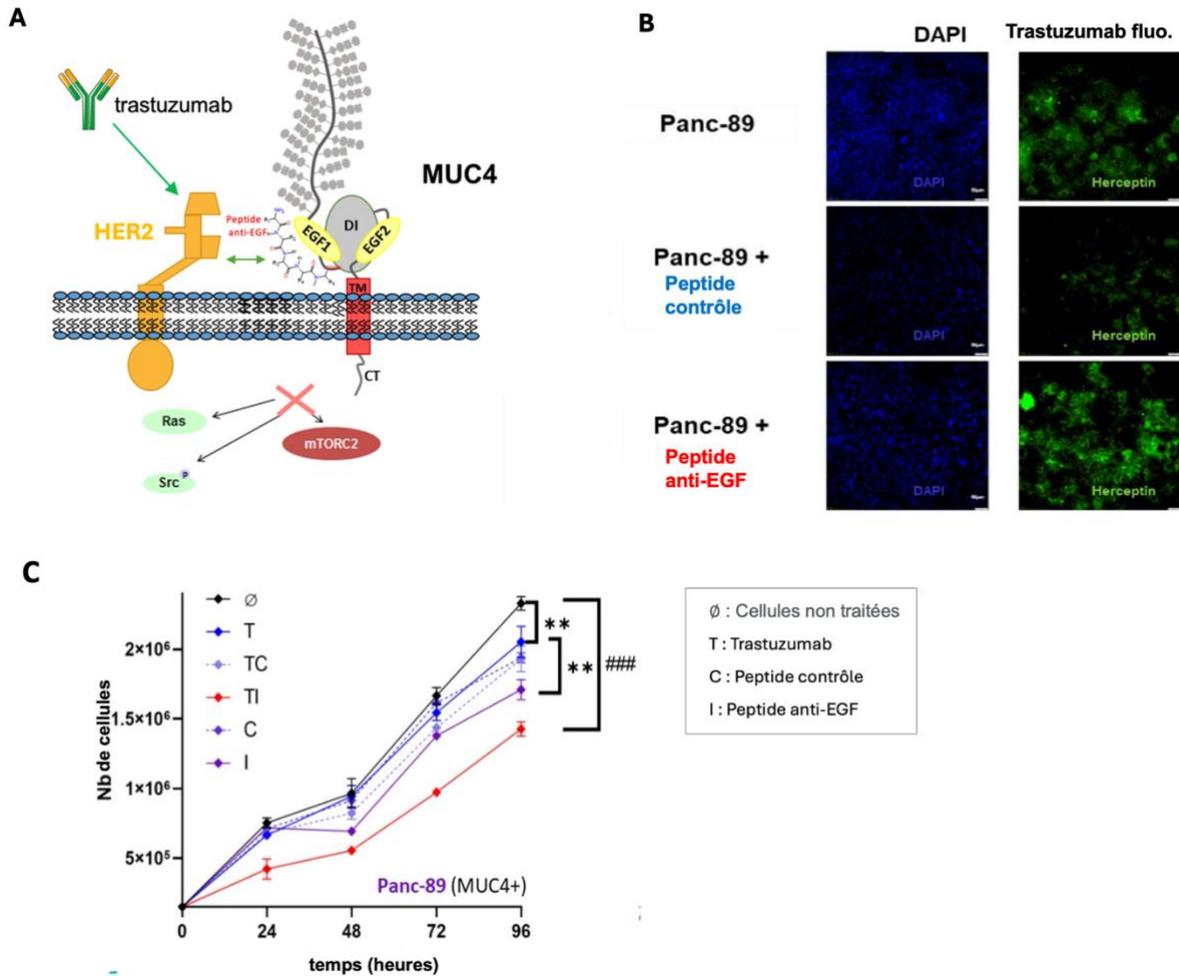


FIGURE 27. PREUVE DE CONCEPT DU CIBLAGE THERAPEUTIQUE DES DOMAINES EGF DE MUC4 DANS LE CANCER DU PANCREAS. A) MECANISME D'ACTION DU PEPTIDE INHIBITEUR DE L'INTERACTION ENTRE HER2 ET MUC4 (PEPTIDE ANTI-EGF). B) RESENSIBILISATION DE LA LIGNEE CELLULAIRE CANCEREUSE PANCREATIQUES PAN-89 EXPRIMANT MUC4 AU TRASTUZUMAB MARQUE PAR UN FLUOROPHORE. C) POTENTIALISATION DE L'EFFET ANTIPROLIFERATIF DU PEPTIDE ANTI-EGF (non publié)

CONCLUSION

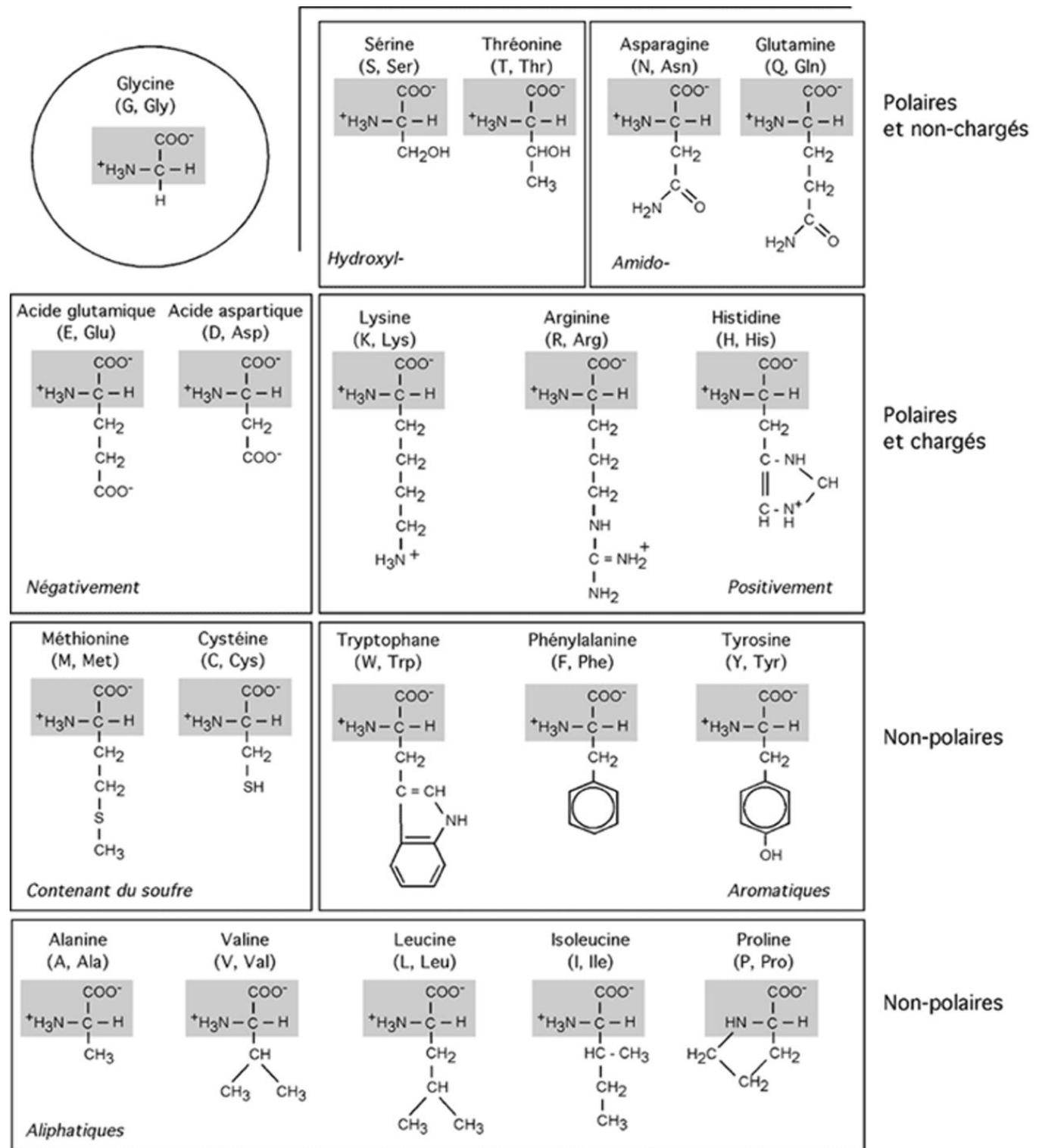
Les domaines EGF-like sont des motifs structuraux présents dans de nombreuses protéines membranaires. Ils jouent un rôle important dans les interactions protéine-protéine, la signalisation cellulaire et le contrôle de nombreux processus physiologiques. Au cours de l'évolution, ces domaines ont été retenus et optimisés pour diversifier les interactions protéines-protéines, jouer des rôles clés dans la signalisation cellulaire et participer à la régulation des fonctions physiologiques. En ciblant les domaines EGF-like, il serait possible de moduler des voies de signalisation aberrantes impliquées dans diverses pathologies, structurales y compris les cancers, les maladies cardiovasculaires et les troubles inflammatoires.

Les domaines EGF-like contenus dans les protéines membranaires offrent de nombreuses opportunités pour le développement de nouvelles thérapies ciblées, visant à interrompre les interactions critiques responsables de la pathologie, de la progression tumorale et la résistance aux traitements. Le ciblage des interactions protéines-protéines en thérapeutique permet une spécificité accrue. Les domaines EGF-like, en particulier, sont impliqués dans des processus biologiques critiques dérégulés dans différentes maladies cancéreuses et inflammatoires.

Leur forme globulaire, stabilisée par la présence des ponts disulfures rend difficile la conception de molécules capables de se lier spécifiquement à un seul domaine EGF-like. Toutefois il pourrait être envisagé de cibler une interface plus étendue, englobant plusieurs domaines EGF-like ou une interface hors de la zone d'interaction des domaines EGF-like, pour rechercher un mécanisme d'action allostérique.

ANNEXES

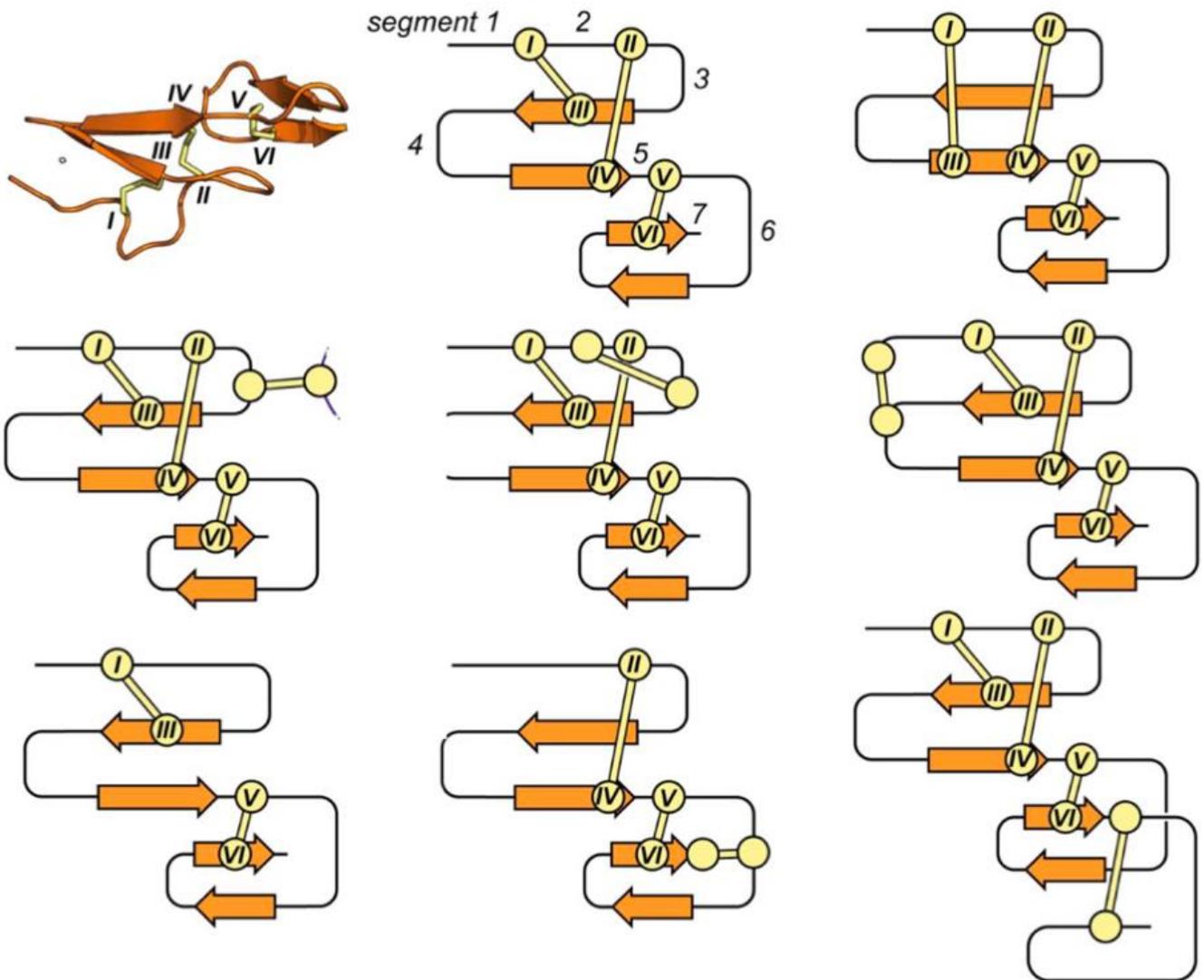
ANNEXE 1 : les acides aminés naturels (adapté de (4))



ANNEXE 2 : exposition des résidus dans les structures secondaires principales (adapté de (6))

Acide aminé	Hélice α			Feuillet β		
	Résidus exposés	Résidus enfouis	Résidus totaux	Résidus exposés	Résidus enfouis	Résidus totaux
V	0.83	0.89	0.91	2.31	1.57	2.00
I	0.96	1.01	1.04	2.02	1.39	1.79
L	1.16	1.27	1.28	1.18	0.93	1.15
M	1.03	1.29	1.26	1.01	0.84	1.01
P	0.48	0.41	0.44	0.49	0.42	0.40
A	1.43	1.37	1.41	0.48	0.72	0.75
C	0.63	0.85	0.85	1.24	1.07	1.36
F	0.88	0.99	1.00	1.50	1.10	1.4
Y	0.91	0.98	0.98	1.71	1.12	1.37
W	0.87	1.09	1.07	1.90	0.91	1.23
Q	1.34	1.21	1.26	0.96	0.82	0.72
S	0.74	0.80	0.76	0.86	0.85	0.81
T	0.72	0.84	0.78	1.58	1.08	1.21
N	0.74	0.77	0.73	0.71	0.76	0.63
H	0.90	0.85	0.87	1.15	0.98	0.99
D	0.91	0.73	0.82	0.61	0.76	0.55
K	1.25	1.13	1.17	1.14	0.98	0.76
E	1.51	1.25	1.39	0.89	0.86	0.65
R	1.31	1.13	1.21	1.27	0.82	0.85
G	0.28	0.59	0.44	0.41	0.81	0.67

ANNEXE 3 : topologies des ponts disulfures dans les domaines EGF-like (34)



REFERENCES

1. Richardson JS. The Anatomy and Taxonomy of Protein Structure. In: Anfinsen CB, Edsall JT, Richards FM, éditeurs. *Advances in Protein Chemistry* [Internet]. Academic Press; 1981 [cité 13 mai 2024]. p. 167-339. Disponible sur:
<https://www.sciencedirect.com/science/article/pii/S0065323308605203>
2. Cagiada M, Bottaro S, Lindemose S, Schenstrøm SM, Stein A, Hartmann-Petersen R, et al. Discovering functionally important sites in proteins. *Nat Commun*. 13 juill 2023;14(1):4175.
3. Sanvictores T, Farci F. Biochemistry, Primary Protein Structure. In: *StatPearls* [Internet]. Treasure Island (FL): StatPearls Publishing; 2024 [cité 13 mai 2024]. Disponible sur:
<http://www.ncbi.nlm.nih.gov/books/NBK564343/>
4. Biochimie des protéines BCM514 [Internet]. [cité 22 mai 2024]. Disponible sur:
<https://biochimiedesproteines.espaceweb.usherbrooke.ca/2a.html>
5. Gruzdev DA, Levit GL, Krasnov VP, Charushin VN. Carborane-containing amino acids and peptides: Synthesis, properties and applications. *Coord Chem Rev*. 15 avr 2021;433:213753.
6. Petsko GA, Ringe D. *Protein Structure and Function*. New Science Press; 2004. 220 p.
7. Privalov PL, Gill SJ. Stability of Protein Structure and Hydrophobic Interaction. In: Anfinsen CB, Edsall JT, Richards FM, Eisenberg DS, éditeurs. *Advances in Protein Chemistry* [Internet]. Academic Press; 1988 [cité 20 mai 2024]. p. 191-234. Disponible sur:
<https://www.sciencedirect.com/science/article/pii/S0065323308603770>
8. Schön A, Clarkson BR, Jaime M, Freire E. Temperature stability of proteins: Analysis of irreversible denaturation using isothermal calorimetry. *Proteins Struct Funct Bioinforma*. 2017;85(11):2009-16.
9. Di Russo NV, Estrin DA, Martí MA, Roitberg AE. pH-Dependent conformational changes in proteins and their effect on experimental pK(a)s: the case of Nitrophorin 4. *PLoS Comput Biol*. 2012;8(11):e1002761.
10. Tran HT, Pappu RV. Toward an Accurate Theoretical Framework for Describing Ensembles for Proteins under Strongly Denaturing Conditions. *Biophys J*. 1 sept 2006;91(5):1868-86.
11. Arakawa T, Tsumoto K. The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. *Biochem Biophys Res Commun*. 25 avr 2003;304(1):148-52.
12. Arakawa T, Timasheff SN. Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry*. 7 déc 1982;21(25):6545-52.
13. Surace AEA, Hedrich CM. The Role of Epigenetics in Autoimmune/Inflammatory Disease. *Front Immunol*. 2019;10:1525.
14. Hollingsworth SA, Karplus PA. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. *Biomol Concepts*. oct 2010;1(3-4):271-83.
15. Fujiwara K, Toda H, Ikeguchi M. Dependence of α -helical and β -sheet amino acid propensities on the overall protein fold type. *BMC Struct Biol*. 2 août 2012;12(1):18.
16. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *The Shape and Structure of Proteins*. In: *Molecular Biology of the Cell* 4th edition [Internet]. Garland Science; 2002 [cité 13 mai 2024]. Disponible

sur: <https://www.ncbi.nlm.nih.gov/books/NBK26830/>

17. Guruprasad K, Prasad MS, Kumar GR. Database of structural motifs in proteins. *Bioinforma Oxf Engl. avr 2000*;16(4):372-5.
18. Breda A, Valadares NF, Souza ON de, Garratt RC. Protein Structure, Modelling and Applications. In: *Bioinformatics in Tropical Disease Research: A Practical and Case-Study Approach* [Internet]. National Center for Biotechnology Information (US); 2007 [cité 13 mai 2024]. Disponible sur: <https://www.ncbi.nlm.nih.gov/books/NBK6824/>
19. Wang Y, Zhang H, Zhong H, Xue Z. Protein domain identification methods and online resources. *Comput Struct Biotechnol J.* 2 févr 2021;19:1145-53.
20. Aroul-Selvam R, Hubbard T, Sasidharan R. Domain Insertions in Protein Structures. *J Mol Biol.* 7 mai 2004;338(4):633-41.
21. Lillis AP, Van Duyen LB, Murphy-Ullrich JE, Strickland DK. LDL Receptor-Related Protein 1: Unique Tissue-Specific Functions Revealed by Selective Gene Knockout Studies. *Physiol Rev.* juill 2008;88(3):887-918.
22. Dehner C. Why Do Proteins Have Quaternary Structure: Non-allosteric Proteins. In: Wells RD, Bond JS, Klinman J, Masters BSS, Bell E, éditeurs. *Molecular Life Sciences: An Encyclopedic Reference* [Internet]. New York, NY: Springer; 2014 [cité 20 mai 2024]. p. 1-7. Disponible sur: https://doi.org/10.1007/978-1-4614-6436-5_21-1
23. Wang J, Zhang S, Lu H, Xu H. Differential regulation of alternative promoters emerges from unified kinetics of enhancer-promoter interaction. *Nat Commun.* 17 mai 2022;13(1):2714.
24. Bloom JD, Labthavikul ST, Otey CR, Arnold FH. Protein stability promotes evolvability. *Proc Natl Acad Sci.* 11 avr 2006;103(15):5869-74.
25. Martin C, Burgel PR. Carriers of a single CFTR mutation are asymptomatic: an evolving dogma? *Eur Respir J* [Internet]. 1 sept 2020 [cité 20 mai 2024];56(3). Disponible sur: <https://erj.ersjournals.com/content/56/3/2002645>
26. Lieber T, Wesley CS, Alcamo E, Hassel B, Krane JF, Campos-Ortega JA, et al. Single amino acid substitutions in EGF-like elements of notch and delta modify drosophila development and affect cell adhesion in vitro. *Neuron.* 1 nov 1992;9(5):847-59.
27. Galicia-Garcia U, Benito-Vicente A, Uribe KB, Jebari S, Larrea-Sebal A, Alonso-Estrada R, et al. Mutation type classification and pathogenicity assignment of sixteen missense variants located in the EGF-precursor homology domain of the LDLR. *Sci Rep.* 3 févr 2020;10(1):1727.
28. Khan S, Vihinen M. Spectrum of disease-causing mutations in protein secondary structures. *BMC Struct Biol.* 29 août 2007;7:56.
29. Kawashima T, Kawashima S, Tanaka C, Murai M, Yoneda M, Putnam NH, et al. Domain shuffling and the evolution of vertebrates. *Genome Res.* 8 janv 2009;19(8):1393-403.
30. Illergård K, Ardell DH, Elofsson A. Structure is three to ten times more conserved than sequence-- a study of structural response in protein cores. *Proteins.* 15 nov 2009;77(3):499-508.
31. Alberts B, Johnson A, Lewis J, Raff M, Roberts K, Walter P. *Analyzing Protein Structure and*

- Function. In: *Molecular Biology of the Cell* 4th edition [Internet]. Garland Science; 2002 [cité 20 mai 2024]. Disponible sur: <https://www.ncbi.nlm.nih.gov/books/NBK26820/>
32. Strasser BJ. Collecting, Comparing, and Computing Sequences: The Making of Margaret O. Dayhoff's Atlas of Protein Sequence and Structure, 1954–1965. *J Hist Biol.* 1 déc 2010;43(4):623-60.
 33. Protein Data Bank Archives its 100,000th Molecule Structure [Internet]. [cité 22 mai 2024]. Disponible sur: https://www.sdsc.edu/News%20Items/PR051414_pdb_100k.html
 34. Tombling BJ, Wang CK, Craik DJ. EGF-like and Other Disulfide-rich Microdomains as Therapeutic Scaffolds. *Angew Chem Int Ed Engl.* 6 juill 2020;59(28):11218-32.
 35. Savage CR, Inagami T, Cohen S. The Primary Structure of Epidermal Growth Factor. *J Biol Chem.* 1 déc 1972;247(23):7612-21.
 36. Appella E, Weber IT, Blasi F. Structure and function of epidermal growth factor-like regions in proteins. *FEBS Lett.* 11 avr 1988;231(1):1-4.
 37. Bersch B, Hernandez JF, Marion D, Arlaud GJ. Solution structure of the epidermal growth factor (EGF)-like module of human complement protease C1r, an atypical member of the EGF family. *Biochemistry.* 3 févr 1998;37(5):1204-14.
 38. Wouters MA, Rigoutsos I, Chu CK, Feng LL, Sparrow DB, Dunwoodie SL. Evolution of distinct EGF domains with specific functions. *Protein Sci Publ Protein Soc.* avr 2005;14(4):1091-103.
 39. Abe Y, Odaka M, Inagaki F, Lax I, Schlessinger J, Kohda D. Disulfide Bond Structure of Human Epidermal Growth Factor Receptor *. *J Biol Chem.* 1 mai 1998;273(18):11150-7.
 40. Stetefeld J, Mayer U, Timpl R, Huber R. Crystal Structure of Three Consecutive Laminin-type Epidermal Growth Factor-like (LE) Modules of Laminin γ 1 Chain Harboring the Nidogen Binding Site. *J Mol Biol.* 5 avr 1996;257(3):644-57.
 41. Penel S, Arigon AM, Dufayard JF, Sertier AS, Daubin V, Duret L, et al. Databases of homologous gene families for comparative genomics. *BMC Bioinformatics.* 16 juin 2009;10(6):S3.
 42. Turjanski P, Ferreira DU. On the Natural Structure of Amino Acid Patterns in Families of Protein Sequences. *J Phys Chem B.* 13 déc 2018;122(49):11295-301.
 43. Downing AK, Knott V, Werner JM, Cardy CM, Campbell ID, Handford PA. Solution Structure of a Pair of Calcium-Binding Epidermal Growth Factor-like Domains: Implications for the Marfan Syndrome and Other Genetic Disorders. *Cell.* 17 mai 1996;85(4):597-605.
 44. Arimori T, Miyazaki N, Mihara E, Takizawa M, Taniguchi Y, Cabañas C, et al. Structural mechanism of laminin recognition by integrin. *Nat Commun.* 29 juin 2021;12(1):4012.
 45. Hrabal R, Komives EA, Ni F. Structural resiliency of an EGF-like subdomain bound to its target protein, thrombin. *Protein Sci Publ Protein Soc.* févr 1996;5(2):195-203.
 46. Takagi J, Beglova N, Yalamanchili P, Blacklow SC, Springer TA. Definition of EGF-like, closely interacting modules that bear activation epitopes in integrin beta subunits. *Proc Natl Acad Sci U S A.* 25 sept 2001;98(20):11175-80.
 47. Engel J. EGF-like domains in extracellular matrix proteins: localized signals for growth and differentiation? *FEBS Lett.* 17 juill 1989;251(1-2):1-7.

48. Campbell ID, Bork P. Epidermal growth factor-like modules. *Curr Opin Struct Biol.* 1 juin 1993;3(3):385-92.
49. Panayotou G, End P, Aumailley M, Timpl R, Engel J. Domains of laminin with growth-factor activity. *Cell.* 13 janv 1989;56(1):93-101.
50. Pennarubia F, Germot A, Pinault E, Maftah A, Legardinier S. The single EGF-like domain of mouse PAMR1 is modified by O-Glucose, O-Fucose and O-GlcNAc. *Glycobiology.* 9 janv 2021;31(1):55-68.
51. Matsumoto K, Luther KB, Haltiwanger RS. Diseases related to Notch glycosylation. *Mol Aspects Med.* juin 2021;79:100938.
52. Pennarubia F. Analyses biochimiques et fonctionnelles de protéines cibles de POFUT1 [Internet] [phdthesis]. Université de Limoges; 2018 [cité 20 mai 2024]. Disponible sur: <https://theses.hal.science/tel-03065825>
53. Li Z, Han K, Pak JE, Satkunarajah M, Zhou D, Rini JM. Recognition of EGF-like domains by the Notch-modifying O-fucosyltransferase POFUT1. *Nat Chem Biol.* juill 2017;13(7):757-63.
54. Matsumoto K, Kumar V, Varshney S, Nairn AV, Ito A, Pennarubia F, et al. Fringe GlcNAc-transferases differentially extend O-fucose on endogenous NOTCH1 in mouse activated T cells. *J Biol Chem.* juill 2022;298(7):102064.
55. Pennarubia F, Pinault E, Maftah A, Legardinier S. In vitro acellular method to reveal O-fucosylation on EGF-like domains. *Glycobiology.* 1 mars 2019;29(3):192-8.
56. McMillan BJ, Zimmerman B, Egan ED, Lofgren M, Xu X, Hesser A, et al. Structure of human POFUT1, its requirement in ligand-independent oncogenic Notch signaling, and functional effects of Dowling-Degos mutations. *Glycobiology.* août 2017;27(8):777-86.
57. Ieguchi K, Fujita M, Ma Z, Davari P, Taniguchi Y, Sekiguchi K, et al. Direct binding of the EGF-like domain of neuregulin-1 to integrins ($\alpha v \beta 3$ and $\alpha 6 \beta 4$) is involved in neuregulin-1/ErbB signaling. *J Biol Chem.* 8 oct 2010;285(41):31388-98.
58. Varshney S, Stanley P. EOGT and O-GlcNAc on Secreted and Membrane Proteins. *Biochem Soc Trans.* 4 avr 2017;45(2):401.
59. Stacey M, Chang GW, Davies JQ, Kwakkenbos MJ, Sanderson RD, Hamann J, et al. The epidermal growth factor-like domains of the human EMR2 receptor mediate cell attachment through chondroitin sulfate glycosaminoglycans. *Blood.* 15 oct 2003;102(8):2916-24.
60. Moore SE. Oligosaccharide transport: pumping waste from the ER into lysosomes. *Trends Cell Biol.* nov 1999;9(11):441-6.
61. Shental-Bechor D, Levy Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc Natl Acad Sci U S A.* 17 juin 2008;105(24):8256-61.
62. Selander-Sunnerhagen M, Ullner M, Persson E, Teleman O, Stenflo J, Drakenberg T. How an epidermal growth factor (EGF)-like domain binds calcium. High resolution NMR structure of the calcium form of the NH₂-terminal EGF-like domain in coagulation factor X. *J Biol Chem.* sept 1992;267(27):19642-9.
63. Stenflo J, Stenberg Y, Muranyi A. Calcium-binding EGF-like modules in coagulation proteinases:

- function of the calcium ion in module interactions. *Biochim Biophys Acta*. 7 mars 2000;1477(1-2):51-63.
64. Valcarce C, Björk I, Stenflo J. The epidermal growth factor precursor. *Eur J Biochem*. 1999;260(1):200-7.
65. EGF-like domains with a C-x-D-x(4)-Y-x-C motif. *Open Access J Transl Med Res [Internet]*. 24 avr 2018 [cité 20 mai 2024];Volume 2(Issue 2). Disponible sur: <https://medcraveonline.com/OAJTMR/OAJTMR-02-00039.pdf>
66. Wang CK, Ghani HA, Bundock A, Weidmann J, Harvey PJ, Edwards IA, et al. Calcium-Mediated Allostery of the EGF Fold. *ACS Chem Biol*. 31 mai 2018;13(6):1659-67.
67. Haller S, Roitberg A, Dudley A. Steered molecular dynamic simulations reveal Marfan syndrome mutations disrupt fibrillin-1 cbEGF domain mechanosensitive calcium binding. *Sci Rep*. 8 oct 2020;10.
68. Lopez-Bigas N, De S, Teichmann SA. Functional protein divergence in the evolution of *Homo sapiens*. *Genome Biol*. 15 févr 2008;9(2):R33.
69. Zhou B, Lin W, Long Y, Yang Y, Zhang H, Wu K, et al. Notch signaling pathway: architecture, disease, and therapeutics. *Signal Transduct Target Ther*. 24 mars 2022;7(1):95.
70. Sadler JE. Thrombomodulin structure and function. *Thromb Haemost*. juill 1997;78(1):392-5.
71. Song G, Meng F, Luo BH. The $\beta 8$ integrin EGF domains support a constitutive extended conformation, and the cytoplasmic domain impairs outside-in signaling. *J Cell Physiol*. nov 2022;237(11):4251-61.
72. Herrlich A, Klinman E, Fu J, Sadegh C, Lodish H. Ectodomain cleavage of the EGF ligands HB-EGF, neuregulin1- β , and TGF- α is specifically triggered by different stimuli and involves different PKC isoenzymes. *FASEB J*. déc 2008;22(12):4281-95.
73. Momiji H, Hassall KL, Featherstone K, McNamara AV, Patist AL, Spiller DG, et al. Disentangling juxtacrine from paracrine signalling in dynamic tissue. *PLoS Comput Biol*. 13 juin 2019;15(6):e1007030.
74. Higashiyama S, Iwabuki H, Morimoto C, Hieda M, Inoue H, Matsushita N. Membrane-anchored growth factors, the epidermal growth factor family: beyond receptor ligands. *Cancer Sci*. févr 2008;99(2):214-20.
75. Singh AB, Sugimoto K, Harris RC. Juxtacrine activation of epidermal growth factor (EGF) receptor by membrane-anchored heparin-binding EGF-like growth factor protects epithelial cells from anoikis while maintaining an epithelial phenotype. *J Biol Chem*. 9 nov 2007;282(45):32890-901.
76. Nakamura K, Iwamoto R, Mekada E. Membrane-anchored heparin-binding EGF-like growth factor (HB-EGF) and diphtheria toxin receptor-associated protein (DRAP27)/CD9 form a complex with integrin alpha 3 beta 1 at cell-cell contact sites. *J Cell Biol*. juin 1995;129(6):1691-705.
77. Barberán S, Cebrià F. The role of the EGFR signaling pathway in stem cell differentiation during planarian regeneration and homeostasis. *Semin Cell Dev Biol*. mars 2019;87:45-57.
78. Vinante F, Rigo A. Heparin-binding epidermal growth factor-like growth factor/diphtheria toxin receptor in normal and neoplastic hematopoiesis. *Toxins*. juin 2013;5(6):1180-201.
79. Li H, Wu M, Zhao X. Role of chemokine systems in cancer and inflammatory diseases. *MedComm*. 2022;3(2):e147.

80. Sakuma T, Higashiyama S, Hosoe S, Hayashi S, Taniguchi N. CD9 Antigen Interacts with Heparin-Binding EGF-Like Growth Factor through Its Heparin-Binding Domain1. *J Biochem (Tokyo)*. 1 août 1997;122(2):474-80.
81. Song G, Luo B. Atypical structure and function of integrin $\alpha_v\beta_8$. *J Cell Physiol*. juill 2021;236(7):4874-87.
82. Nishimura SL. Integrin-mediated transforming growth factor-beta activation, a potential therapeutic target in fibrogenic disorders. *Am J Pathol*. oct 2009;175(4):1362-70.
83. Mizuta I, Nakao-Azuma Y, Yoshida H, Yamaguchi M, Mizuno T. Progress to Clarify How NOTCH3 Mutations Lead to CADASIL, a Hereditary Cerebral Small Vessel Disease. *Biomolecules*. 18 janv 2024;14(1):127.
84. Zhang X, Lee SJ, Wang MM. Hydrolysis of a second Asp-Pro site at the N-terminus of NOTCH3 in inherited vascular dementia. *Sci Rep*. 26 août 2021;11(1):17246.
85. Zhao Z, Michaely P. The epidermal growth factor homology domain of the LDL receptor drives lipoprotein release through an allosteric mechanism involving H190, H562, and H586. *J Biol Chem*. 26 sept 2008;283(39):26528-37.
86. Islam MM, Hlushchenko I, Pfisterer SG. Low-Density Lipoprotein Internalization, Degradation and Receptor Recycling Along Membrane Contact Sites. *Front Cell Dev Biol*. 24 janv 2022;10:826379.
87. Beglova N, Blacklow S. The LDL receptor: How acid pulls the trigger. *Trends Biochem Sci*. 1 juill 2005;30:309-17.
88. Huang S, Henry L, Ho YK, Pownall HJ, Rudenko G. Mechanism of LDL binding and release probed by structure-based mutagenesis of the LDL receptor[S]. *J Lipid Res*. 1 févr 2010;51(2):297-308.
89. Castro N, Rangel MC, Nagaoka T, Karasawa H, Salomon D, Bianco C. Cripto-1: At the Crossroads of Embryonic Stem Cells and Cancer. In 2011.
90. Bianco C, Salomon DS. Targeting the embryonic gene Cripto-1 in cancer and beyond. *Expert Opin Ther Pat*. déc 2010;20(12):1739-49.
91. Ishii H, Afify SM, Hassan G, Salomon DS, Seno M. Cripto-1 as a Potential Target of Cancer Stem Cells for Immunotherapy. *Cancers*. 20 mai 2021;13(10):2491.
92. Kwakkenbos MJ, Kop EN, Stacey M, Matmati M, Gordon S, Lin HH, et al. The EGF-TM7 family: a postgenomic view. *Immunogenetics*. janv 2004;55(10):655-66.
93. Stacey M, Lin HH, Hilyard KL, Gordon S, McKnight AJ. Human epidermal growth factor (EGF) module-containing mucin-like hormone receptor 3 is a new member of the EGF-TM7 family that recognizes a ligand on human macrophages and activated neutrophils. *J Biol Chem*. 1 juin 2001;276(22):18863-70.
94. Lin HH, Stacey M, Saxby C, Knott V, Chaudhry Y, Evans D, et al. Molecular analysis of the epidermal growth factor-like short consensus repeat domain-mediated protein-protein interactions: dissection of the CD97-CD55 complex. *J Biol Chem*. 29 juin 2001;276(26):24160-9.
95. Kop EN, Adriaansen J, Smeets TJM, Vervoordeldonk MJ, van Lier RAW, Hamann J, et al. CD97 neutralisation increases resistance to collagen-induced arthritis in mice. *Arthritis Res Ther*.

2006;8(5):R155.

96. Kleo K, Dimitrova L, Oker E, Tomaszewski N, Berg E, Taruttis F, et al. Identification of ADGRE5 as discriminating MYC target between Burkitt lymphoma and diffuse large B-cell lymphoma. *BMC Cancer*. 5 avr 2019;19(1):322.
97. Serban F, Artene SA, Georgescu A, Purcaru S, Tache D, Alexandru O, et al. Epidermal growth factor, latrophilin, and seven transmembrane domain-containing protein I marker, a novel angiogenesis marker. *OncoTargets Ther*. 16 déc 2015;8:3767.
98. Li J, Shen J, Wang Z, Xu H, Wang Q, Chai S, et al. ELTD1 facilitates glioma proliferation, migration and invasion by activating JAK/STAT3/HIF-1 α signaling axis. *Sci Rep*. 25 sept 2019;9(1):13904.
99. Corfield AP. Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim Biophys Acta*. janv 2015;1850(1):236-52.
100. Tiemin P, Fanzheng M, Peng X, Jihua H, Ruipeng S, Yaliang L, et al. MUC13 promotes intrahepatic cholangiocarcinoma progression via EGFR/PI3K/AKT pathways. *J Hepatol*. avr 2020;72(4):761-73.
101. Liberelle M, Jonckheere N, Melnyk P, Van Seuningen I, Lebègue N. EGF-Containing Membrane-Bound Mucins: A Hidden ErbB2 Targeting Pathway? *J Med Chem*. 28 mai 2020;63(10):5074-88.
102. Funes M, Miller JK, Lai C, Carraway KL, Sweeney C. The mucin Muc4 potentiates neuregulin signaling by increasing the cell-surface populations of ErbB2 and ErbB3. *J Biol Chem*. 14 juill 2006;281(28):19310-9.
103. Karg A, Dinç ZA, Başok O, Uçvet A. MUC4 expression and its relation to ErbB2 expression, apoptosis, proliferation, differentiation, and tumor stage in non-small cell lung cancer (NSCLC). *Pathol Res Pract*. 2006;202(8):577-83.
104. Stoup N, Liberelle M, Schulz C, Cavdarli S, Vasseur R, Magnez R, et al. The EGF Domains of MUC4 Oncomucin Mediate HER2 Binding Affinity and Promote Pancreatic Cancer Cell Tumorigenesis. *Cancers*. 16 nov 2021;13(22):5746.
105. Carraway KL, Price-Schiavi SA, Komatsu M, Jepson S, Perez A, Carraway CAC. Muc4/Sialomucin Complex in the Mammary Gland and Breast Cancer. *J Mammary Gland Biol Neoplasia*. 1 juill 2001;6(3):323-37.
106. Sagar S, Leiphakpam PD, Thomas D, McAndrews KL, Caffrey TC, Swanson BJ, et al. MUC4 enhances gemcitabine resistance and malignant behaviour in pancreatic cancer cells expressing cancer-associated short O-glycans. *Cancer Lett*. 10 avr 2021;503:91-102.
107. Hunter FW, Barker HR, Lipert B, Rothé F, Gebhart G, Piccart-Gebhart MJ, et al. Mechanisms of resistance to trastuzumab emtansine (T-DM1) in HER2-positive breast cancer. *Br J Cancer*. mars 2020;122(5):603-12.
108. Turnpenny PD, Ellard S. Alagille syndrome: pathogenesis, diagnosis and management | *European Journal of Human Genetics*. *Eur J Hum Genet*. mars 2012;20(3):251-7.
109. Fischetto R, Palmieri VV, Tripaldi ME, Gaeta A, Michelucci A, Delvecchio M, et al. Alagille Syndrome: A Novel Mutation in JAG1 Gene. *Front Pediatr*. 15 mai 2019;7:199.
110. Towner RA, Jensen RL, Colman H, Vaillant B, Smith N, Casteel R, et al. ELTD1, A Potential New

Biomarker for Gliomas. *Neurosurgery*. janv 2013;72(1):77-91.

111. Gerstner ER, Batchelor TT. Antiangiogenic Therapy for Glioblastoma. *Cancer J Sudbury Mass*. janv 2012;18(1):45-50.

112. Arnouk H, Yum G, Shah D. Cripto-1 as a Key Factor in Tumor Progression, Epithelial to Mesenchymal Transition and Cancer Stem Cells. *Int J Mol Sci*. 27 août 2021;22(17):9280.

113. Liu Q, Cui X, Yu X, Bian BSJ, Qian F, Hu X gang, et al. Cripto-1 acts as a functional marker of cancer stem-like cells and predicts prognosis of the patients in esophageal squamous cell carcinoma. *Mol Cancer*. 21 avr 2017;16(1):81.

114. Loghmani H, Conway EM. Exploring traditional and nontraditional roles for thrombomodulin. *Blood*. 12 juill 2018;132(2):148-58.

115. Saito H, Maruyama I, Shimazaki S, Yamamoto Y, Aikawa N, Ohno R, et al. Efficacy and safety of recombinant human soluble thrombomodulin (ART-123) in disseminated intravascular coagulation: results of a phase III, randomized, double-blind clinical trial. *J Thromb Haemost*. 1 janv 2007;5(1):31-41.

116. Hasegawa D, Nishida O. Individualized recombinant human thrombomodulin (ART-123) administration in sepsis patients based on predicted phenotypes. *Crit Care*. 24 juin 2019;23(1):231.

117. Mori S, Ai T, Sera T, Ochiai K, Otomo Y. Human Soluble Recombinant Thrombomodulin, ART-123, Resolved Early Phase Coagulopathies, but Did Not Significantly Alter the 28 Day Outcome in the Treatment of DIC Associated with Infectious Systemic Inflammatory Response Syndromes. *J Clin Med*. 27 sept 2019;8(10):1553.

118. Smirnova MG, Birchall JP, Pearson JP. TNF-ALPHA IN THE REGULATION OF MUC5AC SECRETION: SOME ASPECTS OF CYTOKINE-INDUCED MUCIN HYPERSECRETION ON THE IN VITRO MODEL. *Cytokine*. 1 nov 2000;12(11):1732-6.

119. Zalles M, Smith N, Saunders D, Saran T, Thomas L, Gulej R, et al. Assessment of an scFv Antibody Fragment Against ELTD1 in a G55 Glioblastoma Xenograft Model. *Transl Oncol*. mars 2020;13(3):100737.

120. Zalles M, Smith N, Ziegler J, Saunders D, Remerowski S, Thomas L, et al. Optimized monoclonal antibody treatment against ELTD1 for GBM in a G55 xenograft mouse model. *J Cell Mol Med*. janv 2020;24(2):1738-49.

121. Post E. Pancreatic Cancer Action Network. 2024 [cité 21 mai 2024]. Pancreatic Cancer Five-Year Survival Rate Increases to 13%. Disponible sur: <https://pancan.org/news/pancreatic-cancer-five-year-survival-rate-increases-to-13/>

122. Ramsauer VP, Carraway CC, Salas PJ, Carraway KL. Muc4 regulation of ErbB2/ErbB3 signaling. *Cancer Res*. 1 avr 2004;64(7_Supplement):1113.

123. Jonckheere N, Skrypek N, Van Seuning I. Mucins and tumor resistance to chemotherapeutic drugs. *Biochim Biophys Acta*. août 2014;1846(1):142-51.

124. Orzechowski C. Anti-MUC4 β antibodies as a novel therapeutic modality for the treatment of pancreatic cancer. Theses Diss [Internet]. 18 déc 2020; Disponible sur: <https://digitalcommons.unmc.edu/etd/487>

125. Urey C, Hilmersson KS, Andersson B, Ansari D, Andersson R. Development and In Vitro

Characterization of a Gemcitabine-loaded MUC4-targeted Immunoliposome Against Pancreatic Ductal Adenocarcinoma. *Anticancer Res.* 1 nov 2017;37(11):6031-9.

Université de Lille
UFR3S-Pharmacie
DIPLOME D'ETAT DE DOCTEUR EN PHARMACIE
Année Universitaire 2023/2024

Nom : Moum
Prénom : Leïla

Titre de la thèse : Les domaines EGF-like des protéines membranaires classification structurale, implications physiopathologiques et intérêt thérapeutique.

Mots-clés : biochimie, protéines, domaines EGF, EGF-like, protéines membranaires, interactions protéine-protéines

Résumé :

Les domaines EGF-like sont des domaines protéiques extracellulaires de 30 à 40 acides aminés caractérisés par la présence de six résidus Cystéine conservés à l'origine de la formation de trois ponts disulfures intradomains. Les domaines EGF-like présentent certaines caractéristiques de structure primaire qui sont essentielles à leur fonction. Ces particularités structurales sont conservées au cours de l'évolution chez de nombreuses espèces, ce qui souligne leur importance dans la communication cellulaire et le maintien de la fonction des organismes multicellulaires.

Les interactions des protéines contenant des domaines EGF-like avec d'autres molécules et leur capacité à influencer la communication cellulaire en font des cibles potentielles pour le développement de nouvelles thérapies. En effet, les domaines EGF-like contenus dans les protéines membranaires offrent de nombreuses opportunités pour le développement de nouvelles thérapies ciblées, visant à interrompre les interactions critiques responsables de la pathologie, de la progression tumorale et la résistance aux traitements.

Membres du jury :

Président : Professeure Rebecca Deprez-Poulain - Docteur en Pharmacie et Professeure des Universités à la faculté de pharmacie de Lille

Asseseurs : Professeure Réjane Lestrelin – Professeure des Universités à la faculté de pharmacie de Lille ; **Professeur Nicolas Lebègue** – Professeur des universités à la faculté de pharmacie de Lille et Directeur de Thèse

Membre extérieur : Docteur Benoit Gosselin – Docteur en pharmacie, titulaire d'officine