

# UNIVERSITÉ DE STRASBOURG

ÉCOLE DOCTORALE  
DES SCIENCES DE LA VIE ET DE LA SANTÉE  
IGBMC, CNRS UMR 7104

## Auguste Demenge

Soutenue le 5 mai 2023

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

Discipline : **Science de la vie et de la santé**

Spécialité : **Biophysique et biologie structurale**

---

## Analyse structurale et interactomique des ubiquitine-ligases E3, Ube3A et HERC2, impliquées dans le syndrome d'Angelman.

---

Thèse dirigée par :  
Dr. TRAVÉ Gilles

Directeur de recherches, UMR 7104 IGBMC  
CNRS (Illkirch-Graffenstaden, France)

Rapporteur et rapportice :  
Dr. WOLFF Nicolas

Directeur de recherches, Institut Pasteur  
(Paris, France)

Dr. BACHELERIE Françoise

Directrice de recherches, UMR-996 INSERM  
Paris-Saclay (Paris, France)

Autre membre du jury :  
Dr Murielle Masson

Chargée de recherches, UMR 7242 CNRS  
Université de Strasbourg (Illkirch-Graffenstaden,  
France)



# Remerciements

Tout d'abord, je tiens à remercier les membres du jury : Nicolas Wolff, Françoise Bachelier et Murielle Masson qui ont accepté de prendre de leur temps pour évaluer mon travail de thèse. Merci à Fabien Alpy et de nouveau à Murielle Masson qui ont été les membres de mon comité de suivi de thèse.

Je remercie l'Association Française du Syndrome d'Angelman et tout particulièrement Lara Harmann qui m'a invité à présenter mon travail lors des rencontres nationales de l'AFSA. Je remercie aussi la Fondation pour la Recherche Médicale qui a financé ma thèse.

Un grand merci à Gilles qui m'a accueilli au sein de son équipe de recherche pour deux stages puis pour ma thèse. Tu m'as permis de découvrir différentes techniques et m'a transmis la passion que tu as pour la science. Ça a toujours été un grand plaisir de travailler avec toi, de discuter de sciences, de musique et d'autres sujets.

Je remercie les membres de mon équipe. Anna, qui m'a encadré pendant mon premier stage et qui est surtout une super amie, Camille ma "buddy" avec qui chaque jour au labo était un bonheur, Alex et sa gentillesse infinie qui était toujours là pour répondre à mes questions, Irina qui m'a notamment enseigné la préparation du maté argentin, Eduardo avec qui j'ai travaillé avec un immense plaisir, André qui m'a transmis les secrets de la pêche aux cristaux, Goran mon camarade de thèse, Gergö et ses connaissances sans failles, Bogi et sa bonne humeur, Yves qui est toujours de bon conseil, Elodie et son énergie communicative et Alberto et ses conseils sur la vie. Je remercie aussi Laura et Albana qui ont illuminé le labo pendant leurs stages.

Merci à Susanne Schmidt et Anne Debant pour notre collaboration fructueuse. Je remercie aussi Pierre et Alastair de la plateforme de cristallographie ainsi que Pascal Eberling de la plateforme de production de peptide.

Un grand merci à mes amis avec qui j'ai passé beaucoup de temps et d'innombrables heures à jouer à Geoguessr : Alex, Charles, Léo, Max, Sasha, Sam, Sylwan et Tan Dat. Merci à Charlotte, Syrine et Géraldine pour les discussions autour d'un café et les parties

de tarot. Je fais aussi un coucou à Guillaume et Johanna : tellement de chemin parcouru depuis le DUT.

Enfin, je fais des gros bisous à mes parents et mes deux frères : Arsène et Victor.

# Résumé

## Introduction :

Mon sujet de thèse porte sur les syndromes d'Angelman, Angelman-like et dup15q. Ces trois syndromes sont des syndromes du neurodéveloppement. Ils sont caractérisés par un retard mental, une absence de langage, des troubles du sommeil et des crises d'épilepsie. Ces syndromes sont dus à des mutations, délétions ou problèmes d'empreinte parentale touchant les gènes Ube3a et/ou HERC2, portés par la même région chromosomique et produisant deux ubiquitine-ligases E3 qui interagissent entre elles.

Les protéines Ube3a et HERC2 sont deux ubiquitine-ligases à domaine HECT ("homologous to the E6AP Carboxyl Terminus"). Les ubiquitine-ligases interviennent dans le processus d'ubiquitination de protéines cibles. Une fois la protéine cible poly-ubiquitinée, celle-ci est dégradée par le protéasome. HERC2 est une ubiquitine-ligase de 4834 acides aminés, contenant 3 domaines RLD ("RCC1 like domain"). Les structures des domaines RLD1 et RLD3 sont connues (code pdb 4L1M et 3KCI).

La protéine Ube3a (aussi appelée E6AP) fait 852 acides aminés de long. Deux domaines sont structurellement connus. Le premier domaine connu, en position amino-terminale, est le domaine AZUL qui permet l'interaction de Ube3a avec le protéasome. Le second domaine connu, le domaine catalytique HECT, se trouve en position carboxy-terminale.

De plus, ces deux protéines sont capables d'interagir entre elles. Cette interaction a été décrite comme étant une interaction entre la région contenant le domaine RLD2 de HERC2 et une région comprise entre le 150ème et le 200ème acide aminé de Ube3a.

Ainsi, les objectifs de ma thèse peuvent se séparer en deux axes. Le premier axe est l'étude de l'interaction entre Ube3a et HERC2. Le but de cet axe de recherche est de comprendre cette interaction et de la décrire structurellement et de replacer cette interaction dans un réseau d'interactions. Le second axe de recherche se concentre sur la protéine Ube3a afin de mieux comprendre sa structure qui est, en grande partie, non connue.

## Résultats :

Tout d'abord nous avons identifié plus précisément la région de Ube3a interagissant avec le domaine RLD2 de HERC2. Pour cela, nous avons divisé cette région en sous-peptides chevauchants, dont les affinités pour le domaine RLD2 ont été mesurées par test holdup (une technique biochimique développée au sein de l'équipe) et par polarisation de fluorescence. Nous avons pu ainsi identifier une région minimale de 15 acides aminés de Ube3a permettant la liaison à HERC2.

Ensuite, nous avons cristallisé et résolu les structures du domaine RLD2 seul et du domaine RLD2 en complexe avec le peptide de Ube3a par diffraction des rayons X. Cette structure a montré que Ube3a possède le motif consensus DxDKDxD (x = acide aminé quelconque) qui se positionne sur la partie supérieure du domaine RLD2. Nous avons aussi identifié les résidus de RLD2 requis pour cette interaction. Nous avons effectué des analyses de séquences, montrant que HERC2 est présente dans tous les animaux jusqu'aux éponges, et que son domaine RLD2 contient les résidus de liaison au motif DxDKDxD jusqu'aux cnidaires (méduses). La protéine Ube3a est présente chez l'ensemble des animaux et des champignons. Cependant, le motif DxDKDxD de la protéine Ube3a n'est présent que chez les animaux "bilateria" (depuis les vertébrés jusqu'aux arthropodes). Ainsi, la plupart des animaux dotés d'un cerveau ou système nerveux central possèdent le motif DxDKDxD sur Ube3a et le domaine RLD2 de HERC2 comportant les résidus permettant l'interaction avec le motif DxDKDxD. Cependant, certaines espèces, telles que les nématodes ou les cnidaires, possèdent un domaine RLD2 pouvant interagir avec le motif DxDKDxD, mais ne possèdent pas ce motif dans la protéine Ube3a. Ceci nous a amené à penser qu'un tel motif pourrait exister dans des protéines autres que Ube3a.

Une recherche bioinformatique couvrant tout le protéome a alors permis d'identifier 12 protéines humaines contenant un motif DxDKDxD susceptible d'interagir avec RLD2. Nous avons ensuite mesuré l'affinité de peptides issus de ces différentes protéines pour le domaine RLD2 de HERC2. Parmi les protéines montrant une affinité significative pour RLD2 (de l'ordre du micromolaire) on retrouve Ube3a ainsi que PCM1, un composant du centrosôme impliqué dans le développement neuronal et déjà connu pour lier HERC2 ; MYT1, un facteur de transcription régulant le développement des neurones ; BAZ2B, un remodeleur de la chromatine impliqué dans le retard mental et l'autisme ; DOCK10, un co-facteur GEF de RhoGTPases requis pour la formation des épines dendritiques des neurones ; USP35, une dé-ubiquitineuse ; et ARID4a, un remodeleur de la chromatine directement impliqué dans la régulation épigénétique de l'empreinte parentale du gène codant pour Ube3a.

Par la suite, nous avons résolu les structures cristallographiques de RLD2 de HERC2

en complexe avec les peptides à motif DxDKxD issus des protéines DOCK10, PCM1, USP35, MYT1 et BAZ2B. Dans toutes ces structures, le motif DxDKxD est positionné exactement de la même façon que dans notre structure initiale de l'interface du complexe Ube3a-HERC2.

Ainsi, nous avons identifié un nouveau réseau d'interactions centré sur le domaine RLD2 de HERC2. La protéine ayant la plus grande affinité pour HERC2 est Ube3a. Nous pensons qu'en absence ou perturbation de l'activité d'Ube3a ou de HERC2 ce réseau est déséquilibré. Le dosage relatif des différents complexes HERC2-protéines à motifs DxDKxD est modifié, participant à des altérations du neurodéveloppement dans les syndromes d'Angelman, Angelman-like et Dup15q.

D'autre part, la structure de Ube3a entière reste inconnue alors qu'elle constituerait une information précieuse pour comprendre son rôle dans ces différents syndromes. Nous avons donc beaucoup travaillé à l'étude de la structure de la protéine Ube3a. La stratégie pour obtenir des informations structurales de Ube3a est basée sur la purification et la mise en place d'essais de cristallisation de Ube3a entière et de fragments de Ube3a.

Nous avons tout d'abord investi d'importants efforts pour optimiser sa purification. Cette purification nous permet notamment de séparer les monomères des multimères de Ube3a.

Nous avons obtenu des échantillons purs et monomériques d'Ube3a entière et de plusieurs de ses fragments, conçus sur la base de prédictions bioinformatiques. Les formes monomériques de ces différentes constructions n'ont pas permis d'obtenir des cristaux. Cependant, des cristaux ont été obtenus avec la forme multimérique de certaines de ces constructions, mais aucun n'a diffracté suffisamment pour permettre la résolution de la structure d'Ube3a.

Cependant, les purifications des différents fragments d'Ube3a nous ont permis de mieux comprendre la composition de Ube3a en domaines repliés et en régions non repliées ainsi que l'interaction de ces régions. Nous avons notamment produit des mutants de Ube3a comportant des sites protéolytiques (TEV) insérés dans les différentes régions non repliées afin de caractériser les différentes interactions interdomaines. Nos observations sont corroborées par la prédiction de structure de Ube3a récemment générée par le logiciel Alphafold. Toutefois, Alphafold ne prédit ni les complexes entre Ube3a et ses partenaires, ni les changements de conformation et d'activité ubiquitine-ligase qui en résultent.

Enfin, ces différents résultats ainsi que la prédiction d'Alphafold nous ont permis de créer une construction de Ube3a débarrassée de ses zones flexibles. Le but de cette construction appelée Ube3a compacte est de favoriser son étude par cristallographie ou par microscopie électronique. Cette construction se comporte très bien lors de la purification et ne forme pas de multimères. De plus, cette construction cristallise, ce qui n'est pas le cas

des formes monomériques des autres constructions.

## **Conclusion :**

Ces différentes recherches ont mené à une meilleure compréhension de l'interaction des protéines HERC2 et Ube3a. En effet, l'obtention de la structure de l'interaction entre le domaine RLD2 de la protéine HERC2 et un fragment de la protéine Ube3a a permis de mettre en évidence le motif DxDKxD indispensable à cette interaction. 8 autres protéines humaines peuvent interagir avec RLD2 via ce motif. Une collaboration avec une autre équipe travaillant sur la protéine Dock10 a permis de comprendre l'effet de ce motif notamment sur la formation des épines dendritiques des neurones.

Ainsi la protéine HERC2, via son domaine RLD2 est au centre d'un réseau d'interaction dont Ube3a est le meilleur interactant. La perte de la protéine HERC2 ferait disparaître ce réseau d'interactions et pourrait expliquer des effets liés au syndrome d'Angelman-like. La disparition de la protéine Ube3a entraînerait quant à elle un changement complet de la hiérarchie des interactions de ce réseau.

Les progrès faits sur la purification et la compréhension de Ube3a ont permis de produire, purifier et cristalliser une construction de la protéine sous forme monomérique. Ces progrès permettront de pouvoir étudier la structure de Ube3a.

# Summary

## Introduction :

My thesis topic is Angelman, Angelman-like and dup15q syndromes. These three syndromes are neurodevelopmental syndromes. They are characterized by mental retardation, lack of language, sleep disorders and seizures. These syndromes are due to mutations, deletions or parental imprinting problems affecting the Ube3a and/or HERC2 genes, carried on the same chromosomal region and producing two E3 ubiquitin ligases that interact with each other.

The Ube3a and HERC2 proteins are both HECT (homologous to the E6AP Carboxyl Terminus) domain ubiquitin ligases. Ubiquitin ligases are involved in the ubiquitination process of target proteins. Once the target protein is poly-ubiquitinated, it is degraded by the proteasome. HERC2 is a ubiquitin ligase of 4834 amino acids, containing 3 RLD domains ("RCC1 like domain"). The structures of the RLD1 and RLD3 domains are known (pdb code 4L1M and 3KCI).

The Ube3a protein (also called E6AP) is 852 amino acids long. Two domains are known structurally. The first known domain, at the amino-terminal position, is the Azul domain which allows the interaction of Ube3a with the proteasome. The second known domain, the catalytic domain HECT, is located in the carboxy-terminal position.

In addition, these two proteins are able to interact with each other. This interaction has been described as an interaction between the RLD2 domain containing region of HERC2 and a region between the 150th and 200th amino acid of Ube3a.

Thus, the objectives of my thesis can be separated into two axes. The first axis is the study of the interaction between Ube3a and HERC2. The goal of this research axis is to understand this interaction and to describe it structurally and to place this interaction in a network of interactions. The second research axis focuses on the Ube3a protein in order to better understand its structure which is, for the most part, unknown.

## Results :

First, we identified more precisely the region of Ube3a interacting with the RLD2 domain of HERC2. To do so, we divided this region into overlapping sub-peptides, whose affinities for the RLD2 domain were measured by holdup assay (a biochemical technique developed within the team) and by fluorescence polarization. We were able to identify a minimal region of 15 amino acids of Ube3a allowing the binding to HERC2.

Then, we crystallized and solved the structures of the RLD2 domain alone and the RLD2 domain in complex with the peptide of Ube3a by X-ray diffraction. This structure showed that Ube3a has the consensus motif DxDKDxD (x = any amino acid) which is positioned on the top of the RLD2 domain. We also identified the residues of RLD2 required for this interaction. We performed sequence analyses, showing that HERC2 is present in all animals up to sponges, and that its RLD2 domain contains the residues for binding to the DxDKDxD motif up to cnidarians (jellyfish). The Ube3a protein is present in all animals and fungi. However, the DxDKDxD motif of the Ube3a protein is only present in "bilaterian" animals (from vertebrates to arthropods). Thus, most animals with a brain or central nervous system have the DxDKDxD motif on Ube3a and the RLD2 domain of HERC2 containing residues that allow interaction with the DxDKDxD motif. However, some species, such as nematodes or cnidarians, have an RLD2 domain that can interact with the DxDKDxD motif, but do not have this motif in the Ube3a protein. This led us to believe that such a motif might exist in proteins other than Ube3a.

A proteome-wide bioinformatics search identified 12 human proteins containing a DxDKDxD motif that could interact with RLD2. We then measured the affinity of peptides from these different proteins for the RLD2 domain of HERC2. Among the proteins showing a significant affinity for RLD2 (in the micromolar range) are Ube3a and PCM1, a component of the centromere involved in neuronal development and already known to bind HERC2; MYT1, a transcription factor regulating neuronal development; BAZ2B, a chromatin remodeler implicated in mental retardation and autism; DOCK10, a GEF cofactor of RhoGTPases required for neuronal dendritic spine formation; USP35, a de-ubiquitinase; and ARID4a, a chromatin remodeler directly implicated in epigenetic regulation of the parental imprinting of the gene encoding Ube3a.

Subsequently, we solved the crystallographic structures of RLD2 of HERC2 in complex with the DxDKDxD motif peptides from DOCK10, PCM1, USP35, MYT1, and BAZ2B proteins. In all of these structures, the DxDKDxD motif is positioned exactly the same as in our initial structure of the Ube3a-HERC2 complex interface.

Thus, we have identified a new interaction network centered on the RLD2 domain of HERC2. The protein with the highest affinity for HERC2 is Ube3a. We believe that in the

absence or disruption of Ube3a or HERC2 activity this network is unbalanced. The relative dosage of the different HERC2-protein complexes with DxDKDxD motifs is altered, contributing to neurodevelopmental alterations in Angelman, Angelman-like and Dup15q syndromes.

On the other hand, the structure of the whole Ube3a remains unknown, although it would be a valuable information to understand its role in these different syndromes. We have therefore worked hard to study the structure of the Ube3a protein. The strategy to obtain structural information of Ube3a is based on the purification and the implementation of crystallization assays of whole Ube3a and fragments of Ube3a.

We first invested important efforts to optimize its purification. This purification allows us to separate monomers from multimers of Ube3a.

We obtained pure and monomeric samples of the whole Ube3a and of several of its fragments, designed on the basis of bioinformatics predictions. Monomeric forms of these different constructs did not yield crystals. However, crystals were obtained with the multimeric form of some of these constructs, but none diffracted sufficiently to allow resolution of the structure of Ube3a.

However, purifications of the different fragments of Ube3a have allowed us to better understand the composition of Ube3a into folded and unfolded regions as well as the interaction of these regions. In particular, we produced mutants of Ube3a with proteolytic sites (TEVs) inserted in the different unfolded regions in order to characterize the different interdomain interactions. Our observations are corroborated by the structure prediction of Ube3a recently generated by Alphafold software. However, Alphafold does not predict the complexes between Ube3a and its partners, nor the resulting changes in conformation and ubiquitin ligase activity.

Finally, these different results and the prediction of Alphafold allowed us to create a construct of Ube3a without its flexible zones. The aim of this construction called compact Ube3a is to facilitate its study by crystallography or electron microscopy. This construction behaves very well during purification and does not form multimers. Moreover, this construct crystallizes, which is not the case for the monomeric forms of the other constructs.

## **Conclusion :**

These different researches have led to a better understanding of the interaction of HERC2 and Ube3a proteins. Indeed, the structure of the interaction between the RLD2 domain of the HERC2 protein and a fragment of the Ube3a protein has allowed us to identify the DxDKDxD motif that is essential for this interaction. 8 other human proteins can interact

with RLD2 via this motif. A collaboration with another team working on the Dock10 protein allowed to understand the effect of this motif on the formation of dendritic spines of neurons.

Thus, the HERC2 protein, via its RLD2 domain, is at the center of an interaction network of which Ube3a is the best interactant. The loss of the HERC2 protein would make this interaction network disappear and could explain effects linked to the Angelman-like syndrome. The disappearance of Ube3a would lead to a complete change in the hierarchy of interactions in this network.

Progress in the purification and understanding of Ube3a has allowed the production, purification and crystallization of a monomeric construct of the protein. This progress will allow us to study the structure of Ube3a.

# Table des matières

<b>Remerciements</b>	<b>i</b>
<b>Résumé</b>	<b>iii</b>
<b>Summary</b>	<b>vii</b>
<b>Table des matières</b>	<b>xi</b>
<b>Liste des abréviations</b>	<b>xv</b>
<b>Liste des figures</b>	<b>xix</b>
<b>Liste des tableaux</b>	<b>xxii</b>
<b>I Introduction</b>	<b>1</b>
<b>1 L'ubiquitination</b>	<b>3</b>
1.1 Découverte de l'ubiquitination, historique . . . . .	3
1.2 L'ubiquitination : son mécanisme et ses différentes enzymes . . . . .	4
1.2.1 L'ubiquitine et les UBLs . . . . .	4
1.2.1.1 L'ubiquitine . . . . .	4
1.2.1.2 Les "ubiquitine-like" . . . . .	6
1.2.2 Activation de l'ubiquitine . . . . .	8
1.2.2.1 Structures et classes . . . . .	8
1.2.2.2 Mécanisme . . . . .	9
1.2.3 Transport de l'ubiquitine . . . . .	12
1.2.3.1 Structures et classes . . . . .	12
1.2.3.2 Mécanisme . . . . .	13
1.2.4 Transfert de l'ubiquitine . . . . .	14
1.2.4.1 Ubiquitine-ligases à domaine RING et à domaine U-BOX .	15

1.2.4.2	E3 à domaine U-box . . . . .	19
1.2.4.3	E3 à domaine HECT . . . . .	19
1.2.4.4	Domaine RBR . . . . .	22
1.3	Variété des chaînes d'ubiquitines . . . . .	23
1.3.1	Les différentes chaînes d'ubiquitines . . . . .	23
1.3.2	Rôles des différentes chaînes d'ubiquitines . . . . .	27
1.4	Domaines et autres enzymes relatives aux ubiquitines . . . . .	29
1.4.1	Les domaines de liaisons aux ubiquitines . . . . .	29
1.4.2	Les désubiquitinases . . . . .	29
1.5	Le protéasome . . . . .	31
1.5.1	Organisation structurale . . . . .	31
1.5.1.1	Le protéasome 20S . . . . .	32
1.5.1.2	Le protéasome 19S . . . . .	33
1.5.1.3	Les variations du protéasome . . . . .	33
1.5.2	Mécanisme de dégradation protéasomale . . . . .	33
<b>2</b>	<b>Les ubiquitine-ligases Ube3a et HERC2</b>	<b>35</b>
2.1	L'ubiquitine-ligase Ube3a . . . . .	35
2.1.1	Ube3a et le système ubiquitine-protéasome . . . . .	36
2.1.1.1	Le domaine HECT et l'ubiquitination . . . . .	36
2.1.1.2	Le domaine AZUL . . . . .	36
2.1.1.3	Interaction avec RAD23 . . . . .	38
2.1.2	Rôles de Ube3a . . . . .	38
2.1.3	Structure globale de Ube3a . . . . .	38
2.1.4	Ube3a et les cancers liés aux papillomavirus . . . . .	38
2.2	L'ubiquitine-ligase HERC2 . . . . .	41
2.2.1	Différents rôles identifiés de HERC2 . . . . .	42
2.2.2	Les domaines RLDs . . . . .	42
2.3	Interaction entre Ube3a et HERC2 . . . . .	44
2.4	La région chromosomique 15q11-q13 . . . . .	45
<b>3</b>	<b>Syndromes impliquant la région 15q11–q13</b>	<b>48</b>
3.1	Le syndrome d'Angelman . . . . .	48
3.1.1	Historique . . . . .	48
3.1.2	Les différents types du syndrome et les symptômes associés . . . . .	48
3.1.3	Pistes de traitements . . . . .	49
3.1.3.1	Thérapies géniques, cellulaires et enzymatiques . . . . .	50
3.1.3.2	Activation de l'allèle paternel . . . . .	51
3.1.3.3	Les traitements "Downstream" . . . . .	53

3.2	Les syndromes dits Angelman-like . . . . .	54
3.3	Le syndrome de Prader-Willi . . . . .	55
3.4	Le syndrome dup15q . . . . .	55
<b>4</b>	<b>Objectifs de la thèse</b>	<b>57</b>
<b>II</b>	<b>Résultats &amp; Discussions</b>	<b>59</b>
<b>1</b>	<b>Le domaine RLD2 et son interactome</b>	<b>61</b>
1.1	Avant-propos . . . . .	61
1.1.1	Introduction . . . . .	61
1.1.2	Contributions . . . . .	61
1.2	Abstract . . . . .	62
1.3	Introduction . . . . .	62
1.4	Results . . . . .	64
1.5	Discussion . . . . .	75
1.6	Material and methods . . . . .	78
1.7	Supplemental . . . . .	82
<b>2</b>	<b>Étude structurale de Ube3a</b>	<b>86</b>
2.1	Introduction . . . . .	86
2.1.1	Stratégie . . . . .	86
2.1.2	AlphaFold 2, un tournant? . . . . .	87
2.2	Contributions . . . . .	89
2.3	Résultats . . . . .	90
2.4	Discussion . . . . .	105
2.4.1	Structure de Ube3a . . . . .	105
2.4.2	Ube3a et son interaction avec E6 et RLD2 . . . . .	105
2.4.3	Multimérisation de Ube3a . . . . .	106
2.4.4	Activité ubiquitine-ligase de Ube3a et sa modulation . . . . .	108
2.4.5	Ube3a une proche parente des HERCs . . . . .	109
<b>III</b>	<b>Matériel &amp; méthodes</b>	<b>111</b>
<b>1</b>	<b>Matériel &amp; méthodes</b>	<b>113</b>
1.1	Expressions . . . . .	113
1.2	Purifications . . . . .	113
1.2.1	Purification de RLD2 . . . . .	113
1.2.2	Purification des constructions de Ube3a . . . . .	114

1.2.3	Purification de MBP-16E6 . . . . .	114
1.3	Cristallisation . . . . .	115
1.4	Mesures d'affinités par polarisation de fluorescence . . . . .	115
1.4.1	Génération des modèles par AlphaFold 2 . . . . .	116
1.4.2	Analyses phylogénétiques de séquences . . . . .	116
1.4.3	Mesure de surface accessible au solvant . . . . .	117
1.5	Séquences . . . . .	117
<b>IV</b>	<b>Conclusion &amp; Perspectives</b>	<b>121</b>
<b>1</b>	<b>Conclusion</b>	<b>123</b>
<b>2</b>	<b>Perspectives</b>	<b>126</b>
<b>V</b>	<b>Annexes</b>	<b>129</b>
<b>1</b>	<b>Annexe 1</b>	<b>131</b>
<b>2</b>	<b>Annexe 2</b>	<b>146</b>
<b>3</b>	<b>Annexe 3</b>	<b>154</b>
<b>VI</b>	<b>Bibliographie</b>	<b>157</b>

# Liste des abréviations

<b>AAV</b>	"Adeno-Associated Viruses"
<b>AD</b>	"Adenylation Domain"
<b>ADN</b>	acide désoxyribonucléique
<b>AMP</b>	Adenosine monophosphate
<b>AOS</b>	"Antisense oligonucleotides"
<b>APF1</b>	"Adenosine triphosphate dependent proteolysis factor 1"
<b>ARID4A</b>	"AT-Rich Interaction Domain 4A"
<b>ARIP4</b>	"Androgen Receptor-Interacting Protein 4"
<b>AS</b>	"Angelman Syndrome"
<b>ASD</b>	"Autism Spectrum Disorder"
<b>ATF</b>	"Activating transcription factor"
<b>ATP</b>	"Adenosine triphosphate"
<b>ATS</b>	"antisense DNA strand"
<b>AZUL</b>	"amino-terminal zinc-binding"
<b>Bard1</b>	"BRCA1 associated RING domain 1"
<b>BAZ2B</b>	"Bromodomain Adjacent To Zinc Finger Domain 2B"
<b>BIRC7</b>	"Baculoviral IAP repeat-containing protein 7"
<b>bmal1</b>	"basic helix-loop-helix ARNT like 1"
<b>BME</b>	Beta-Mercaptoethanol
<b>Brca1</b>	"breast cancer 1"
<b>BRcat</b>	"beginning RING catalytic"
<b>CASP</b>	"Critical Assessment of Structure Prediction"
<b>CCD</b>	"catalytic cysteine domain"

**CNS** "central nervous system"

**CP110** "Centriolar coiled-coil protein of 110 kDa"

**CRL** "cullin ring ligase"

**DOCK10** "Dedicator Of Cytokinesis 10"

**DUB** "Deubiquitinating enzyme"

**DUS** "domain of unknown structure"

**E4B** "ubiquitination factor E4B"

**E6AP** "E6-associated protein"

**E6C** E6 C-terminal

**E6N** E6 N-terminal

**H2A** histone 2A

**HECT** "Homologous to the E6-AP Carboxyl Terminus"

**HERC** "HECT and RLD domain containing E3 ubiquitin protein ligase 2"

**HPV** "Human Papillomavirus"

**HSC** "Hematopoietic Stem Cell"

**IBR** "In Between RING"

**IC** "imprinting centres"

**iDMRs** "imprinting-associated differentially methylated regions"

**IPR** "imprinting control region"

**IV** intraveineux

**Kb** kilobase

**kDa** kilodalton

**MBP** "Maltose Binding Protein"

**MJD** "Machado Joseph Diseases"

**MSA** Multiple Sequence Alignment

**MYT1** "myelin transcription factor 1 "

**NaCl** Chlorure de sodium

**NAE** NEDD8-activating enzyme

**NEDD** Neurodevelopmentally down-regulated

**OCA2** OCA2 Melanosomal Transmembrane Protein

**OTU** "Ovarian Tumor Protease"

**P53** "tumor protein 53"  
**PA200** "proteasome activator 200"  
**PAE** "Predicted Aligned Error"  
**PDB** "Protein Data Bank"  
**PDZ** "PSD-95, Dlg1, ZO-1"  
**pLDDT** "per-residue confidence score"  
**ppi** pyrophosphate  
**Psmc4** "26S proteasome non-ATPase regulatory subunit 4"  
**PV** Papillomavirus  
**RBR** "RING between-RING RING"  
**Rcat** "catalytic RING"  
**RCC1** "Regulator Of Chromosome Condensation 1"  
**RERE** "Arginine-glutamic acid dipeptide repeats protein"  
**RING** "Really Interesting New Gene"  
**RLD** "Regulator Of Chromosome Condensation 1 like domain"  
**RPL40** "Ubiquitin A-52 Residue Ribosomal Protein Fusion Product 1"  
**Rpn** particules régulatrices non-ATPases  
**RPS27A** "Ubiquitin And Ribosomal Protein S27a"  
**rpt** particules régulatrices AAA+ ATPases  
**SAE** "SUMO-activating enzyme"  
**SEC** "Size Exclusion Chromatography"  
**SMURF** "SMAD Specific E3 Ubiquitin Protein Ligase"  
**SNORD116** "Small Nucleolar RNA, C/D Box 116-1"  
**SNRPN** "Small nuclear ribonucleoprotein-associated protein N"  
**SNURF** "SNRPN Upstream Open Reading Frame"  
**TCEP** tris(2-carboxyéthyl)phosphine  
**TEV** "Tobacco Etch Virus"  
**Tris hcl** Tris Hydrochloride  
**Ub** Ubiquitine  
**UBA** "ubiquitin-activating enzyme"  
**UBA52** "Ubiquitin A-52 Residue Ribosomal Protein Fusion Product 1"

**UBA80** "Ubiquitin And Ribosomal Protein S27a"  
**UBC** "Ubiquitin Conjugation"  
**UBD** "Ubiquitin Binding Domain"  
**Ube3a** "Ubiquitin-protein ligase E3A"  
**UBIP** "ubiquitous immunopoietic polypeptide"  
**UBL** "Ubiquitin-like"  
**UBQLN** ubiquiline  
**UCH** "Ubiquitine C terminale Hydrolase"  
**UFD** "Ubiquitin Fold Domain"  
**USP** "Ubiquitin-specific proteases"  
**ZnF** "zinc finger"

# Liste des figures

Figure 1	Le mécanisme général de l'ubiquitination . . . . .	4
Figure 2	L'ubiquitine et ses différents patchs . . . . .	5
Figure 3	La structure de l'ubiquitine et des UBLs . . . . .	7
Figure 4	Les différentes classes d'enzymes E1 . . . . .	9
Figure 5	La cascade d'activation de l'ubiquitine par l'enzyme E1 . . . . .	10
Figure 6	La structure de l'enzyme E1 UBA1 . . . . .	11
Figure 7	Les différentes enzymes E2 . . . . .	13
Figure 8	La structure de l'enzyme E2 E2H . . . . .	14
Figure 9	La structure du domaine RING . . . . .	15
Figure 10	Les différents types d'ubiquitine-ligases à domaine RING . . . . .	17
Figure 11	L'interaction entre Ubch7 et le domaine RING de c-CBL . . . . .	18
Figure 12	La comparaison du domaine RING et du domaine U-Box . . . . .	19
Figure 13	Les différentes E3 à domaine HECT . . . . .	20
Figure 14	Les structures du domaine HECT aux différentes étapes de l'ubiquitination d'un substrat . . . . .	22
Figure 15	Les lysines de l'ubiquitine . . . . .	24
Figure 16	Les différentes chaînes d'ubiquitines . . . . .	25
Figure 17	Les structures de différentes chaînes d'ubiquitine . . . . .	26
Figure 18	La dégradation protéasomale et lysosomale ubiquitine-dépendante . . . . .	28
Figure 19	Les multiples rôles des DUBs . . . . .	31
Figure 20	La structure du protéasome 26S . . . . .	32
Figure 21	Le mécanisme du protéasome . . . . .	34
Figure 22	La structure du domaine AZUL de Ube3a seul et en complexe avec le domaine RAZUL de Rpn10 . . . . .	37
Figure 23	L'organisation de l'oncoprotéine E6 de HPV16 . . . . .	39
Figure 24	La structure de l'interaction entre E6 et Ube3a . . . . .	40
Figure 25	La structure de l'interaction entre E6, Ube3a et p53 . . . . .	41
Figure 26	Les protéines contenant des domaines RLDs . . . . .	43

Figure 27	Les structures des domaines RLD1 et RLD3 de HERC2 . . . . .	43
Figure 28	Les "β-propeller" et leurs interactions protéine/protéine . . . . .	44
Figure 29	L'organisation de la région 15q11-q13 . . . . .	45
Figure 30	Un modèle hypothétique de la régulation de la méthylation de l'ADN dans la région PWS/AS par AS-IC et PWS-IC . . . . .	47
Figure 31	Les différentes causes génétiques du syndrome d'Angelman . . . . .	49
Figure 32	Les thérapies géniques et cellulaires . . . . .	51
Figure 33	L'activation de l'allèle paternel de Ube3a . . . . .	53
Figure 34	Crystal structure of the RLD2 domain of the HERC2 protein bound to the minimal HERC2-binding motif of Ube3a. (PDB : 7Q41) . . . . .	65
Figure 35	Interaction network between the DxDKDxD motif and the RLD2 do- main. . . . .	67
Figure 36	Conservation of HERC2 RLD2 domain across different species and of DxDKDxD motifs across different species and different proteins. . . . .	68
Figure 37	Crystallographic structure of the RLD2 domain in complex with dif- ferent peptide . . . . .	70
Figure 38	The interaction between DOCK10 and RLD2 is affected by mutations in their respective binding sites. . . . .	71
Figure 39	Conservation of HERC2 RLD2 domain across different species and of DxDKDxD motifs across different species and different proteins. . . . .	72
Figure 40	Expression of DOCK10mut, defective in RLD2 binding, leads to de- creased dendritic spine formation in primary hippocampal neurons. . . . .	73
Figure 41	Acrylamide gel of the different steps of the Ube3a protein purification	82
Figure 42	Measurement of the affinity of different peptides for the RLD2 domain	85
Figure 43	La prédiction des zones non repliées de la protéine Ube3a par Fol- dindex . . . . .	86
Figure 44	Les résultats du concours CASP . . . . .	88
Figure 45	Le fonctionnement du logiciel AlphaFold 2 . . . . .	89
Figure 46	Les différentes constructions de Ube3a . . . . .	90
Figure 47	Les profils d'élution des différentes constructions de Ube3a . . . . .	92
Figure 48	Les profils d'élution de la construction de Ube3a contenant trois sites TEV . . . . .	95
Figure 49	La structure AlphaFold 2 de Ube3a . . . . .	97
Figure 50	Le modèle AlphaFold 2 de l'interaction entre Ube3a et E6 . . . . .	99
Figure 51	Ubiquitination de P53 par Ube3a . . . . .	101

Figure 52	La prédiction Alphafold de l'interaction entre Ube3a et le domaine RLD2 . . . . .	102
Figure 53	La surface des différentes constructions de Ube3a . . . . .	103
Figure 54	La structure de Ube3a et sa conservation . . . . .	104
Figure 55	Un modèle de multimérisation de Ube3a . . . . .	107
Figure 56	La trimérisation du domaine HECT . . . . .	108
Figure 57	Les prédictions AlphaFold 2 des HERC 3, 4, 5 et 6 . . . . .	110
Figure 58	L'analyse phylogénétique du syndrome d'Angelman . . . . .	123

# Liste des tableaux

1	Les différentes UBLs et leurs enzymes E1 et E2 . . . . .	8
2	Les différentes enzymes E1 . . . . .	12
3	Les différents UBDs . . . . .	30
4	Les différentes isoformes de Ube3A . . . . .	36
5	Les traitements "Downstream" du syndrome d'Angelman . . . . .	54
6	Proteins containing DxDKDxD identified by SlimSearch . . . . .	83
7	Measurement of the affinity of different peptides for the RLD2 domain . . . . .	84
8	Les mesures d'affinités de 16E6 pour différents peptides . . . . .	100



Première partie

Introduction



# 1 | L'ubiquitination

## 1.1 Découverte de l'ubiquitination, historique

Chez les eucaryotes, il existe deux voies majeures de dégradation des protéines. La première est la protéolyse lysosomale. Cette voie consiste en la dégradation de protéines dans une structure cellulaire appelée lysosome. Cette structure contient de nombreuses enzymes, notamment des protéases.

L'autre principale voie de dégradation est la voie ubiquitine-protéasome, qui nécessite l'ubiquitination du substrat à dégrader.

L'ubiquitination est une modification post-traductionnelle consistant en la fixation d'une protéine, très conservée à travers les espèces et appelée ubiquitine ou d'une chaîne d'ubiquitines sur une lysine d'un substrat. Cette chaîne d'ubiquitine va, par la suite, jouer différents rôles selon sa nature. Le rôle principal et le plus connu de l'ubiquitination d'un substrat est son étiquetage pour sa dégradation par le protéasome.

Ce sont Irwin A. Rose, Avram Hershko et Aaron Ciechanover qui, les premiers, ont identifié l'ubiquitine comme signal de dégradation par le protéasome. Ils se sont vu décerner le prix Nobel de chimie en 2004 pour leurs travaux sur la dégradation ATP dépendante des protéines durant les années 1970-1980.

En effet, ils ont alors démontré que la dégradation des protéines était consommatrice d'énergie. Ceci indique que la dégradation des protéines implique un autre phénomène que l'hydrolyse qui, elle, n'est pas consommatrice d'ATP [1] [2]. Pour cela, ils ont étudié la protéolyse *in vitro* dans des réticulocytes de lapin. Les réticulocytes étant dépourvus de lysosome, cela leur a permis de mieux comprendre la dégradation des protéines sans interférence avec la dégradation lysosomale des protéines. Ainsi, ils ont révélé l'implication d'une autre protéine, l'ubiquitine, indispensable à la dégradation ATP-dépendante de protéines [3]. La nécessité de la liaison covalente de cette protéine au substrat a ensuite été démontrée [4].

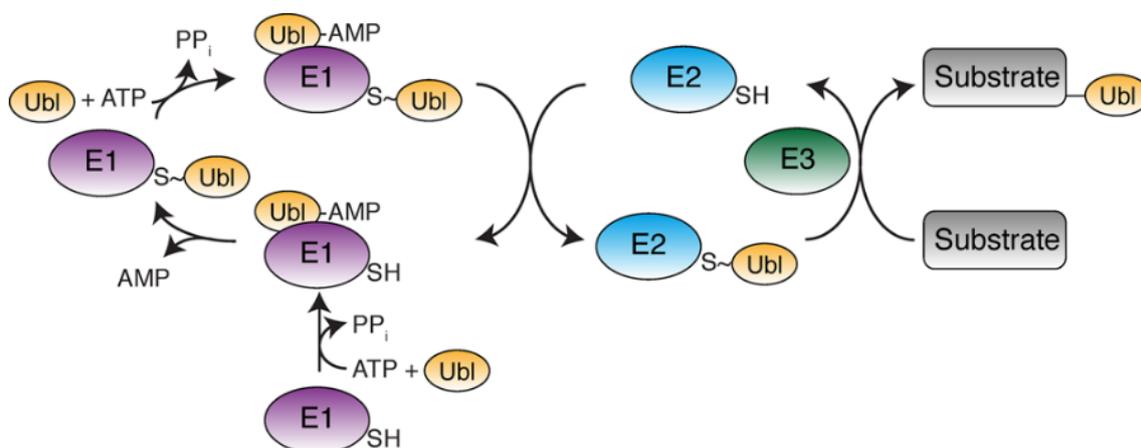
Le rôle de l'ubiquitine, alors appelée APF-1 pour facteur 1 de protéolyse ATP-dépendant,

a été identifié comme celui d'une étiquette de dégradation, indispensable à la dégradation de protéines. Ce processus d'étiquetage est consommateur d'ATP [5] [4]. APF-1 a ensuite été identifié comme étant l'ubiquitine [6] déjà connue notamment dans son rôle pour la différenciation des lymphocytes T et B [7] et est alors aussi appelée UBIP pour polypeptide hématopoïétique ubiquitaire.

Ce mécanisme de dégradation des protéines est dorénavant appelé UPS pour système ubiquitine protéasome.

## 1.2 L'ubiquitination : son mécanisme et ses différentes enzymes

Trois différents types d'enzymes interviennent dans le mécanisme d'ubiquitination d'un substrat : une enzyme d'activation de l'ubiquitine E1, une enzyme de conjugaison de l'ubiquitine E2 et enfin une enzyme ubiquitine ligase E3. Chacune de ces protéines joue successivement un rôle distinct jusqu'à l'ubiquitination du substrat. (Figure 1 page 4)



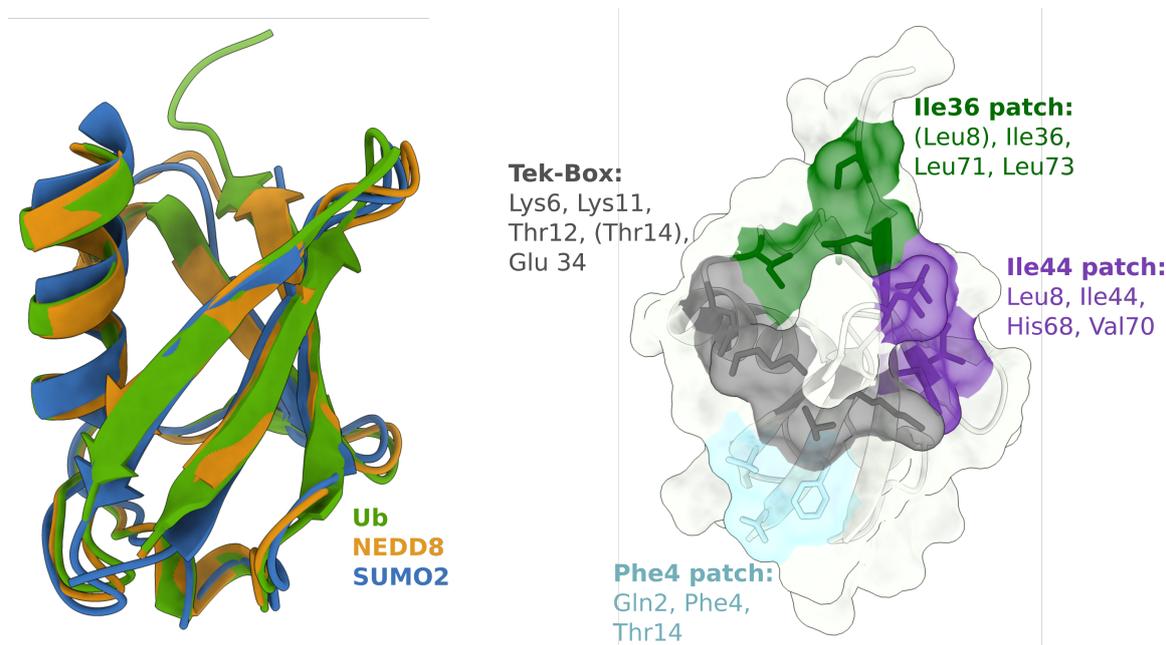
**FIGURE 1** – Figure représentant le mécanisme général de l'ubiquitination d'un substrat. Ce mécanisme fonctionne en cascade enzymatique : les enzymes E1 puis E2 et enfin E3. L'enzyme E1 se charge d'activer l'ubiquitine et de la transférer à l'enzyme E2 qui va, pour finir, transférer cette ubiquitine sur le substrat via l'intervention de l'enzyme E3. Figure de [8]

### 1.2.1 L'ubiquitine et les UBLs

#### 1.2.1.1 L'ubiquitine

L'ubiquitine est une protéine de 76 résidus dont la séquence est hautement conservée chez tous les organismes jusqu'aux plantes. Par exemple, seuls 3 acides aminés diffèrent entre l'ubiquitine humaine et l'ubiquitine de levure [9]. L'ubiquitine adopte un repliement dit

$\beta$ -grasp constitué d'un feuillet  $\beta$  composé de cinq brins  $\beta$  et d'une hélice  $\alpha$  ou plus [10]. L'ubiquitine possède différents patches appelés patch Ile 44, patch Ile 36, patch Phe 4 et TEK-box. [11]. (Figure 3 page 7)



**FIGURE 2** – À gauche : la superposition des structures de l'ubiquitine (en vert, PDB : 1UBQ), de NEDD8 (en orange, PDB : 1NND) et de SUMO2 (en bleu, PDB : 1WM3). À droite : la structure de l'ubiquitine (PDB : 1UBQ). Les différents patches sont affichés de couleurs différentes. Figure adaptée de [12].

Ces différents patches jouent un rôle dans l'interaction avec d'autres protéines. Ils sont impliqués dans l'interaction avec les désubiquitinasés (DUBs) qui sont des protéases coupant l'ubiquitine d'une protéine (détaillées dans la section 1.4.2 page 29), dans l'interaction avec des protéines possédant un domaine de liaison à l'ubiquitine (UBD) (Section 1.4.1 page 29) et dans l'interaction entre différentes ubiquitines d'une même chaîne (Section 1.3 page 23).

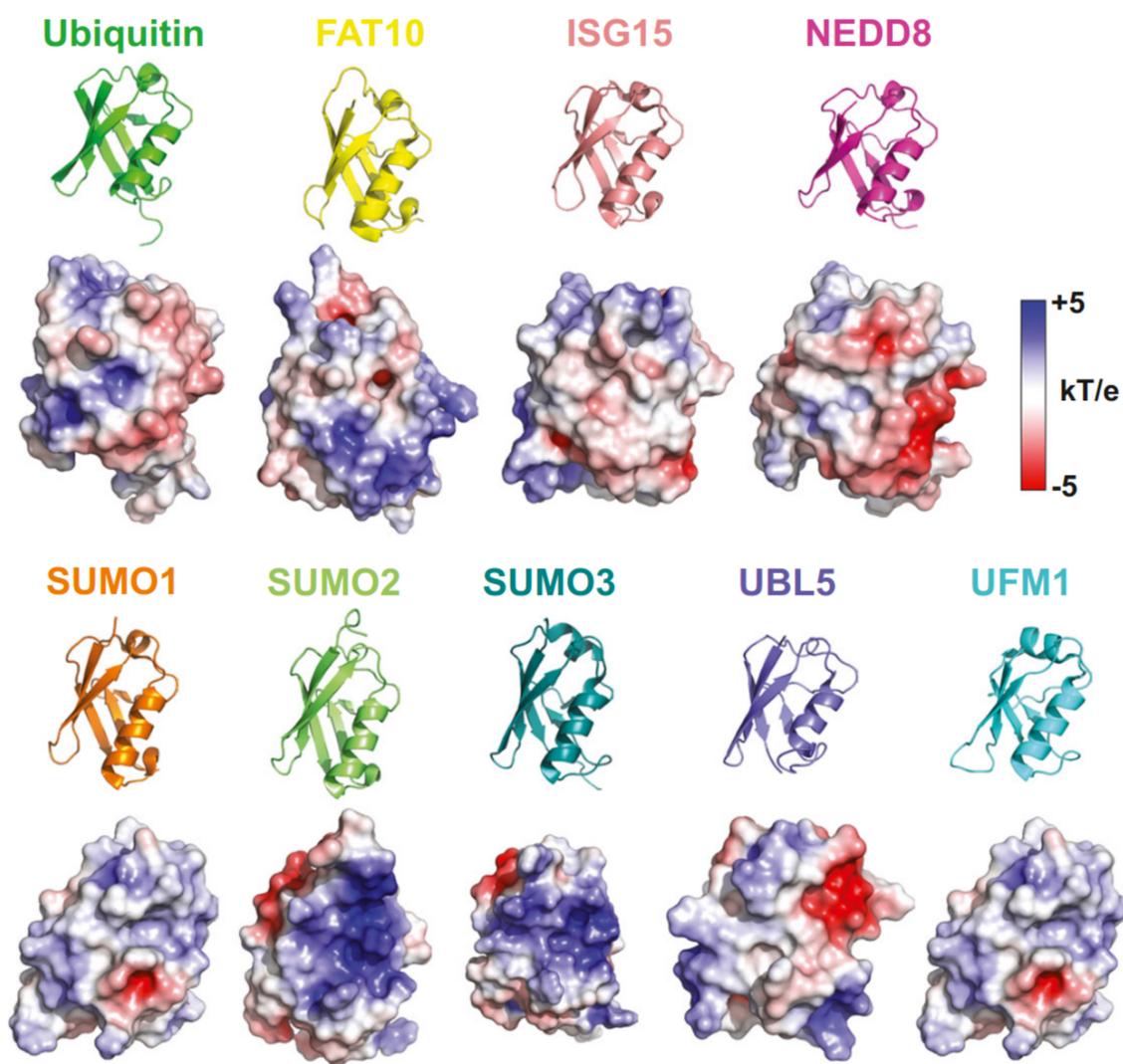
Le patch Ile 44, hydrophobe, interagit avec le protéasome et de nombreux UBDs (domaine de liaison à l'ubiquitine)[13] [14]. Le patch Ile 36, une autre surface hydrophobe, est reconnu par le domaine HECT des E3s [15] mais également par des UBDs [16] et des DUBs (désubiquitinasés)[17]. Cette surface est aussi impliquée dans l'interaction entre les ubiquitines d'une même chaîne. Le patch comprenant la Phe 4 interagit avec différentes DUBs, [17] et UBDs [18]. Le dernier patch appelé TEK-box est essentiel pour la formation des chaînes d'ubiquitines de type K11 que nous décrivons par la suite [19].

L'ubiquitine n'est pas produite comme une unité par un gène. Chez l'humain, l'ubiquitine est codée par différents gènes. Les gènes UbB [20] et UbC [21] codent respectivement

pour une répétition de trois et de neuf ubiquitines. L'ubiquitine est aussi produite liée à d'autres protéines. Elle est produite avec la protéine RPL40 (gène UBA52)[22] et la protéine RPS27A (gène UBA80) [23], deux protéines ribosomales. Les chaînes d'ubiquitines sont ensuite séparées en unité d'ubiquitine par les DUBs pour ensuite être utilisées [24].

#### 1.2.1.2 Les "ubiquitine-like"

Il existe d'autres protéines ayant un repliement similaire à l'ubiquitine, qui peuvent être liées à un substrat en tant que modification post-traductionnelle. Ces protéines sont appelées les UBLs ("Ubiquitin-like") et adoptent toutes un repliement  $\beta$ -grasp. Malgré une structure similaire, les UBLs diffèrent par leurs séquences. L'UBL la plus proche de l'ubiquitine, NEDD8, a une identité de séquence de 58% avec l'ubiquitine [25].(Figure 3 page 7)



**FIGURE 3** – Représentations de la structure secondaire de l'ubiquitine et des domaines analogues à l'ubiquitine (en haut), et de leur potentiel électrostatique (en bas). Les unités sont données en énergie par charge ( $k$ , constante de Boltzmann ;  $T$ , température en Kelvin ;  $e$ , charge d'un électron). (1UBQ [10], 5IA7[26], 1NDD [27], 6GF1 [28], 1Z2M [29], 1P0R [30], 4WJQ [31], 1WM2 [32], 1U4A [33]). Figure de [34].

Les surfaces importantes pour les interactions impliquant l'ubiquitine décrites précédemment ne sont pas conservées. Ainsi, les différentes protéines interagissant avec les différentes UBLs ne sont pas les mêmes. Elles diffèrent selon les caractéristiques physico-chimiques de leurs surfaces [35]. Le mécanisme qui entraîne la fixation d'ubiquitine ou d'UBLs sur un substrat, sera le même, mais utilisera différentes enzymes E1, E2 et E3. (Tableau 1 page 8).

Famille	Ubl	E1	E2
SUMO	SUMO1, SUMO2, SUMO3, SUMO4	UBA2/SAE1	UBC9
NEDD8	NEDD8	UBA3/NAE1	UBC12, UBE2F
ATG8	LC3A, LC3B, LC3B2, LC3C, GABARAP, GABARAPL1, GATE-16	ATG7	ATG3
ATG12	Atg12	ATG7	ATG10
URM1	URM1	UBA4	
UFM1	UFM1	UBA5	UFC1
FAT10	FAT10	UBA6	UBE2Z
ISG15	ISG15	UBA7	UBCH8

**TABLE 1** – Tableau des différentes familles d'ubiquitine-like, des membres de ces familles, de leurs enzymes E1 les activant et de leurs enzymes E2 impliquées dans l'ubiquitination d'un substrat. Tableau adapté de [8]

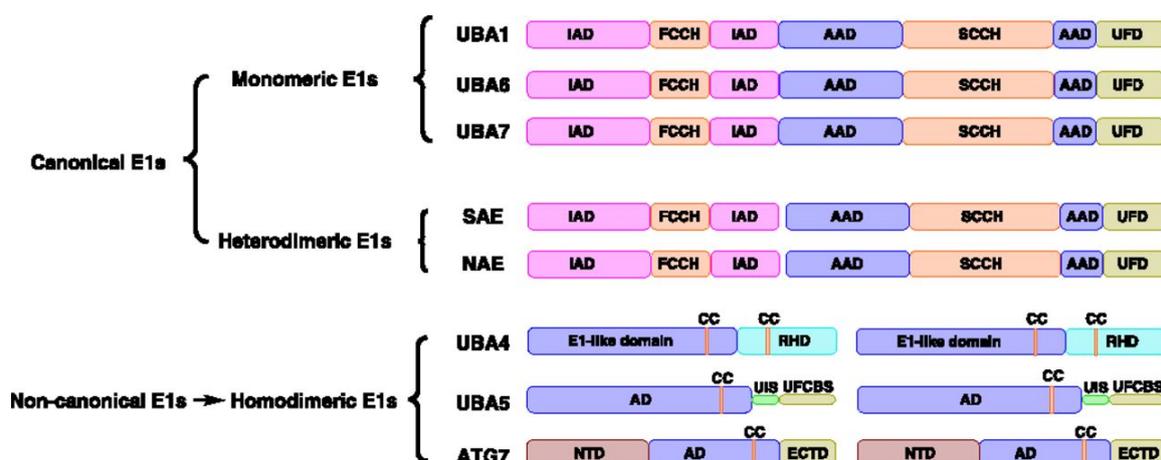
De plus, certaines protéines arborent des domaines ayant la même structure que l'ubiquitine/ubiquitine-like. Ces domaines sont appelés les domaines ubiquitine-like.

## 1.2.2 Activation de l'ubiquitine

La première étape dans le processus d'ubiquitination est menée par l'enzyme E1. Il s'agit de l'activation de l'ubiquitine.

### 1.2.2.1 Structures et classes

Il existe différentes protéines E1 qui sont, selon leurs propriétés structurales, subdivisées en deux classes. Les E1 dites canoniques, au nombre de 5, sont soit monomériques (UBA1, UBA6 et UBA7), soit hétérodimériques (SAE et NAE). Les E1 dites non canoniques (UBA4, UBA5 et ATG7) sont toutes homodimériques [36]. (Figure 4 page 9)



**FIGURE 4** – Figure représentant les enzymes E1, réparties en deux classes, les canoniques et les non canoniques. Les domaines des cinq enzymes canoniques et des trois enzymes non canoniques sont représentés par des blocs de couleurs. Notez que, l’AAD, les demi-domaines contenant la cystéine catalytique (FCCD et SCCD) et l’UFD sont conservés parmi les E1 canoniques. Les E1 non canoniques ne possèdent pas de CDD, la cystéine catalytique se situe au niveau de l’AD. Les E1 non canoniques comprennent également des domaines spécifiques aux E1 non canoniques. Figure de [37]

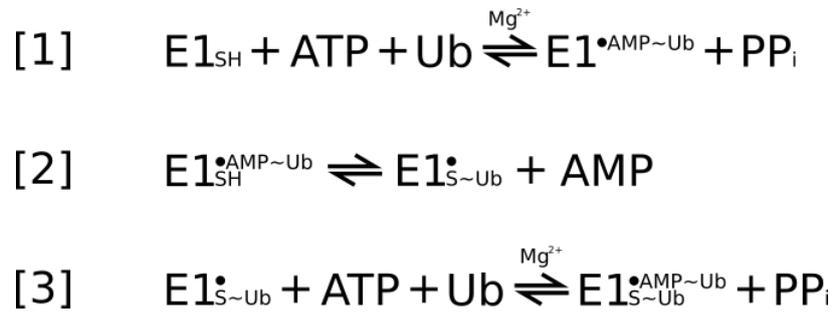
Toutes les enzymes E1 sont composées du domaine AD (Domaine d’adénylation), d’un domaine CCD (Domaine à Cystéine Catalytique) qui contient la cystéine catalytique ainsi que d’un domaine UFD (Domaine à repliement d’ubiquitine) [38].

Les E1 canoniques possèdent un domaine AD actif (AAD) et un domaine AD inactif (IAD). C’est le domaine AAD qui va reconnaître l’ubiquitine pour permettre son adénylation. Le domaine IAD jouerait un rôle dans la stabilité de la protéine. Le domaine CCD, quant à lui, est impliqué dans la formation de la liaison thioester avec l’ubiquitine [39]. Enfin, c’est le domaine UFD qui va permettre l’interaction avec l’enzyme E2 une fois l’ubiquitine activée.

Les E1 non canoniques sont toutes des homodimères. À la différence des E1 canoniques, elles ne possèdent pas de domaine CCD. En effet, les cystéines catalytiques de ces E1 se trouvent directement dans leurs domaines AD [8].

### 1.2.2.2 Mécanisme

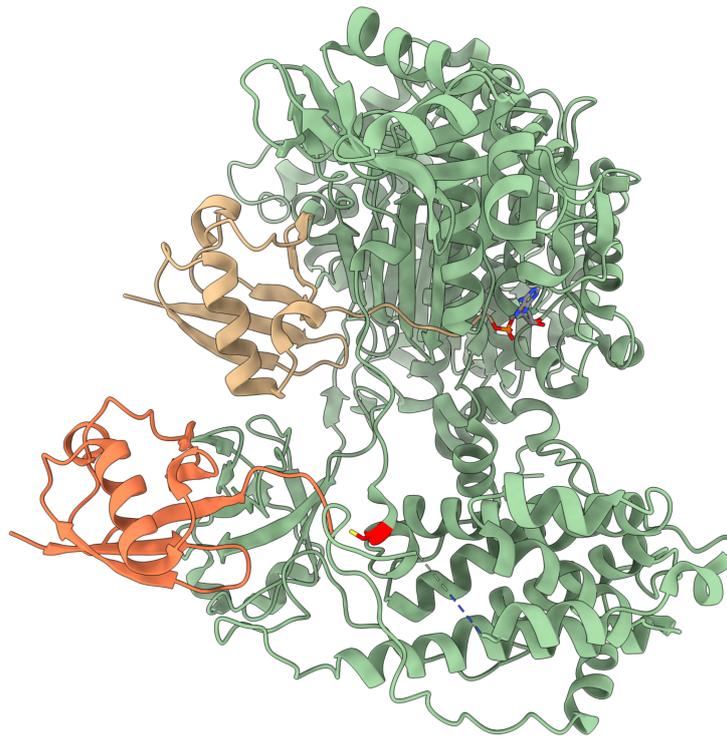
L’activation de l’ubiquitine peut se décomposer en trois étapes. La première étape consiste en l’adénylation de la partie carboxy-terminale de l’ubiquitine en présence d’une molécule d’ATP et d’un ion  $Mg^{2+}$ . Cette adénylation se fait grâce à la reconnaissance de l’ubiquitine par le domaine AAD de l’enzyme E1. De cette étape résulte une enzyme E1 liée de manière non covalente à l’ubiquitine adénylée libérant une molécule de  $PPi$  [40] (figure 5.1 page 10).



**FIGURE 5** – Cascade d'activation de l'ubiquitine par l'enzyme E1. 1) correspond à l'adénylation de la première ubiquitine (Ub). 2) correspond à la formation de la liaison thioester entre la cystéine catalytique de E1 et l'ubiquitine adénylée. 3) correspond à l'adénylation de la seconde ubiquitine. Figure adaptée de [40].

Ensuite, l'ubiquitine est transférée à la cystéine catalytique de l'enzyme E1. Cette étape consiste en une attaque nucléophile du groupement adénylate de l'ubiquitine par la cystéine catalytique de l'enzyme E1, libérant une molécule d'AMP. Ainsi, l'ubiquitine forme une liaison covalente avec la cystéine catalytique de E1 [40] (figure 5.2 page 10).

La troisième étape est l'adénylation d'une seconde molécule d'ubiquitine. En effet, il a été montré que les enzymes E1 sont capables de lier deux molécules d'ubiquitine en même temps [40]. Cette étape est la même que l'étape 1 décrite précédemment. De cette étape résulte une E1 liée de manière non covalente à une ubiquitine adénylée via son domaine ADD et de manière covalente à une molécule d'ubiquitine via la cystéine catalytique de son domaine CCD. (Figure 5.3 page 10). Cette E1 liant deux ubiquitines est plus favorable au transfert d'ubiquitine sur l'enzyme E2 qu'une E1 liant une seule ubiquitine [36] [8]. (figure 6 page 11)



**FIGURE 6** – Structure tridimensionnelle de l'enzyme E1 UBA1 (représentée en vert) liée de manière covalente à l'ubiquitine (en orange) via sa cystéine catalytique (colorée en rouge) et de manière non covalente à une AMP-Ubiquitine (en beige). (Code PDB : 4NNJ) [41]

Comme vu précédemment (section 1.2.1.2 page 6), l'ubiquitine fait partie de la famille des ubiquitine/ubiquitine-like. Ainsi les différentes E1 n'ont pas toutes l'ubiquitine comme substrat [42]. (Table 2 page 12)

E1	Structure	UBL	E2s	UBL Proteases
UBA1 (UBE1)	Monomer	Ubiquitin	Multiple	DUBs
UBA6 (UBE1L2)	Monomer	UBL FAT10	USE1 (UBE2Z)	DUBs Unknown
UBA7 (UBE1L)	Monomer	ISG15	UBCH8 UBCH6	USP18
SAE (SAE1-UBA2)	Heterodimer	SUMO 1/2/3	UBC9	SENPs
NAE (NAE1-UBA3)	Heterodimer	NEDD8	UBC12 UBE2Fs	CSN5 DEN1
UBA4 (MOCS3)	Heterodimer	URM1	—	—
UBA5 (UBE1DC1)	Homodimer	UFM1	UFC1	UFSP1 UFSP2
ATG7	Homodimer	ATG12 ATG8 family	ATG10 ATG3	—

**TABLE 2** – Les différentes enzymes E1 ainsi que leurs E2 et Ubiquitine/ubiquitine-like correspondantes. Toutes les enzymes E1 n'activent pas l'ubiquitine, certaines sont spécifiques à des ubiquitine-like. Tableau adapté de [37]

L'étape suivante de la cascade d'ubiquitination implique l'enzyme E2. Il s'agit du transfert de l'ubiquitine liée à la cystéine catalytique de E1 vers l'enzyme E2. L'interaction avec l'enzyme E2 se fait via le domaine UFD de E1. Une réaction de trans-thiolation va lier l'ubiquitine à la cystéine catalytique de l'enzyme E2 [40]. Une fois l'ubiquitine transférée à l'enzyme E2, l'enzyme E1 recommence un cycle d'activation de l'ubiquitine en liant de manière covalente l'AMP-Ubiquitine à laquelle elle est liée de manière non covalente.

### 1.2.3 Transport de l'ubiquitine

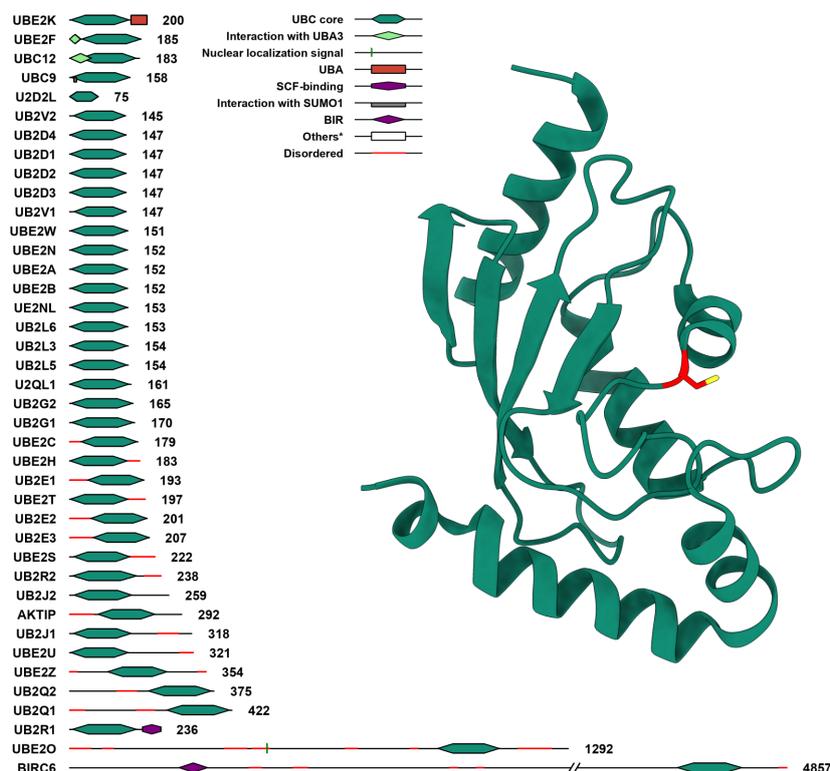
L'enzyme E2, aussi appelée protéine transporteuse d'ubiquitine, a été identifiée pour la première fois dans le lysat de réticulocyte de lapin [43]. L'ubiquitine alors liée à la protéine E1 va être transférée sur une cystéine de la protéine E2 via une liaison thioester.

#### 1.2.3.1 Structures et classes

Il y a plus d'une trentaine d'enzymes E2 humaines réparties en quatre classes. Les E2 de classe 1 ne possèdent que le domaine catalytique UBC, celles de classe 2 possèdent une extension en position carboxy-terminale de celui-ci, celles de classe 3 une extension en position amino-terminale. Enfin, les E2 de classe 4 possèdent à la fois une extension en position amino-terminale et une en position carboxy-terminale (Figure 7 page 13).

Ces extensions permettent des interactions avec d'autres protéines et jouent un rôle dans

la localisation cellulaire [44] [45] [46].



**FIGURE 7** – Figure représentant à gauche la liste des différentes enzymes E2 humaines. Le domaine UBC est représenté par un losange vert. Les domaines représentés sont ceux identifiés sur la base de données uniprot. Les protéines sont regroupées en fonction de leur similarité de compositions de domaines. Notez la présence d'un domaine UBA en position amino-terminale de la protéine UBE2K. Ce graphique a été réalisé en utilisant le logiciel ProFeatMap créé par Goran Bich [47]. Sur la droite de la figure est représentée la structure tridimensionnelle de l'enzyme E2H (PDB : 2Z5D) [48]. La cystéine catalytique du domaine UBC de E2H est représentée en rouge.

Le domaine UBC est un domaine d'environ 150 acides aminés composé de cinq hélices  $\alpha$  et d'un feuillet  $\beta$ . Ce domaine comporte la cystéine catalytique qui permet de former la liaison thioester avec l'ubiquitine [49] (Figure 7 page 13).

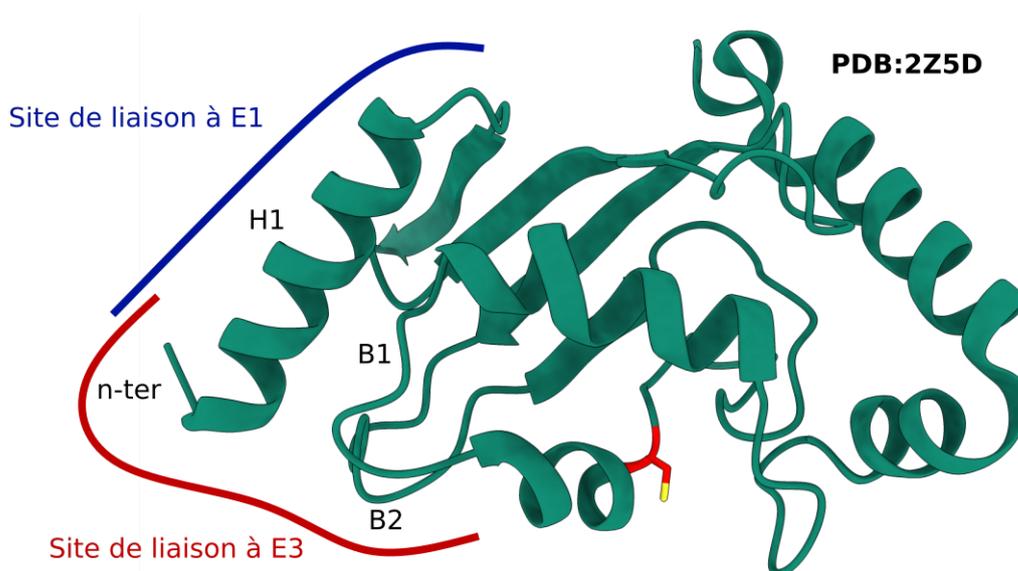
### 1.2.3.2 Mécanisme

Via son domaine UBC l'enzyme E2 est capable de lier l'enzyme E1 et ainsi permettre le transfert de l'ubiquitine de la cystéine catalytique de E1 vers la cystéine catalytique de E2. De plus, l'enzyme E2 possède une zone de liaison à E3.

L'interaction entre les protéines E1 et E2 entraîne un changement conformationnel de E1, rapprochant ainsi les deux sites catalytiques de E1 et E2 et permettant la réaction de transthiolation [50] [51]. Cette interaction se fait via le domaine UFD de E1. Ce domaine

via un sillon chargé négativement va interagir avec notamment une lysine de l'hélice 1 du domaine UBC de E2 [52].

Une fois lié à l'ubiquitine, l'enzyme E2 va pouvoir interagir avec l'enzyme E3 pour finalement entraîner l'ubiquitination d'un substrat. Cette interaction se fait via les boucles B1 et B2 ainsi que la partie amino-terminale de l'hélice 1 de E2 [52] (Figure 8 page 14).



**FIGURE 8** – Structure tridimensionnelle de l'enzyme E2H (PDB : 2Z5D) [48]. La cystéine catalytique du domaine UBC de E2H est représentée en rouge. Le site de liaison à E3 est représenté par une ligne rouge. Cette interaction se fait via les boucles B1 et B2 du domaine UBC et la partie amino-terminale de l'hélice H1. Le site de liaison à l'enzyme E1 (via le domaine UFD de E1) est représenté par une ligne bleue. Figure adaptée de [53].

De plus, les protéines E2 sont capables d'interagir avec d'autres protéines pouvant moduler leur activité [54] [55].

La plupart des E2 sont capables d'entraîner à la fois des réactions de transthiolation et d'aminolyse (transfert d'un groupement thioester à un groupement amine). D'autres E2 ne sont capables d'entraîner qu'un seul type de réaction [53].

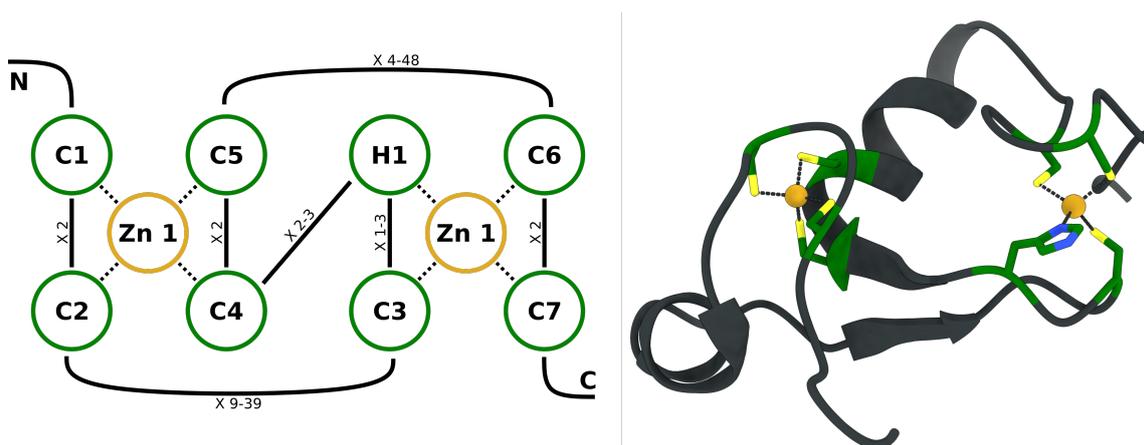
#### 1.2.4 Transfert de l'ubiquitine

La dernière étape de l'ubiquitination consiste au transfert de l'ubiquitine sur le substrat. Cette étape nécessite l'intervention d'une ubiquitine-ligase (enzyme E3). Les enzymes E3 sont classées selon leurs domaines catalytiques. Certaines ubiquitine-ligases vont jouer le rôle d'adaptateur entre la protéine E2 et le substrat. Dans ce cas, l'ubiquitine va être transférée directement de la protéine E2 au substrat. C'est le cas des E3 à domaine RING et des E3 à domaine U-Box. Pour les E3 à domaine HECT et à domaine RBR, l'ubiquitine

est transférée à une cystéine catalytique du domaine catalytique de E3, pour ensuite être transférée au substrat.

#### 1.2.4.1 Ubiquitine-ligases à domaine RING et à domaine U-BOX

Les ubiquitine-ligases de type RING (pour Really Interesting New Gene) comportent un domaine du même nom d'environ 50 acides aminés. Le domaine RING a été identifié pour la première fois en 1991 [56] et est caractérisé par le consensus suivant : Cys X2 Cys X(9-39) Cys X(1-3) His X(2-3) Cys X2 Cys X(4-48) Cys X2 Cys. Le domaine RING est caractérisé par la présence de ces 7 cystéines et d'une histidine qui permettent la liaison à deux atomes de zinc. Dans ce domaine de liaison au zinc, l'ordre linéaire dans la séquence des cystéines/histidine ne correspond pas à l'ordre de liaison aux atomes de zinc. En d'autres termes, les cystéines C1, C2, C4 et C5 interagissent avec un premier atome de zinc. Les cystéines C3, C6 et C7 ainsi que l'histidine interagissent avec un second atome de zinc. Cette caractéristique rend le domaine RING compact et globulaire [57] (Figure 9 page 15).



**FIGURE 9** – À gauche est représentée de manière schématique la structure du domaine RING, les cystéines sont numérotées selon l'ordre de la séquence. X représente le nombre d'acides aminés séparant les cystéines/histidine. À droite, est représentée la structure du domaine RING de la E3 RNF4 humaine (PDB : 4PPE). Les atomes de zinc sont représentés par des sphères oranges, les cystéines et l'histidine liant les atomes de zinc sont représentées en vert. Figure adaptée de [58]

Au fur et à mesure des découvertes de nouvelles protéines contenant un domaine RING, des variantes de celui-ci ont été identifiées. Ces variantes consistent en l'inversion d'une des cystéines impliquées dans la liaison de l'atome de zinc avec l'histidine [59].

D'autres domaines, tels que le domaine PHD et le domaine LIM, contiennent un schéma de cystéines et histidine similaire à celui du domaine RING. Cependant, il a été montré que les protéines à domaine PHD et LIM n'ont pas d'activité ubiquitine-ligase [60][61].

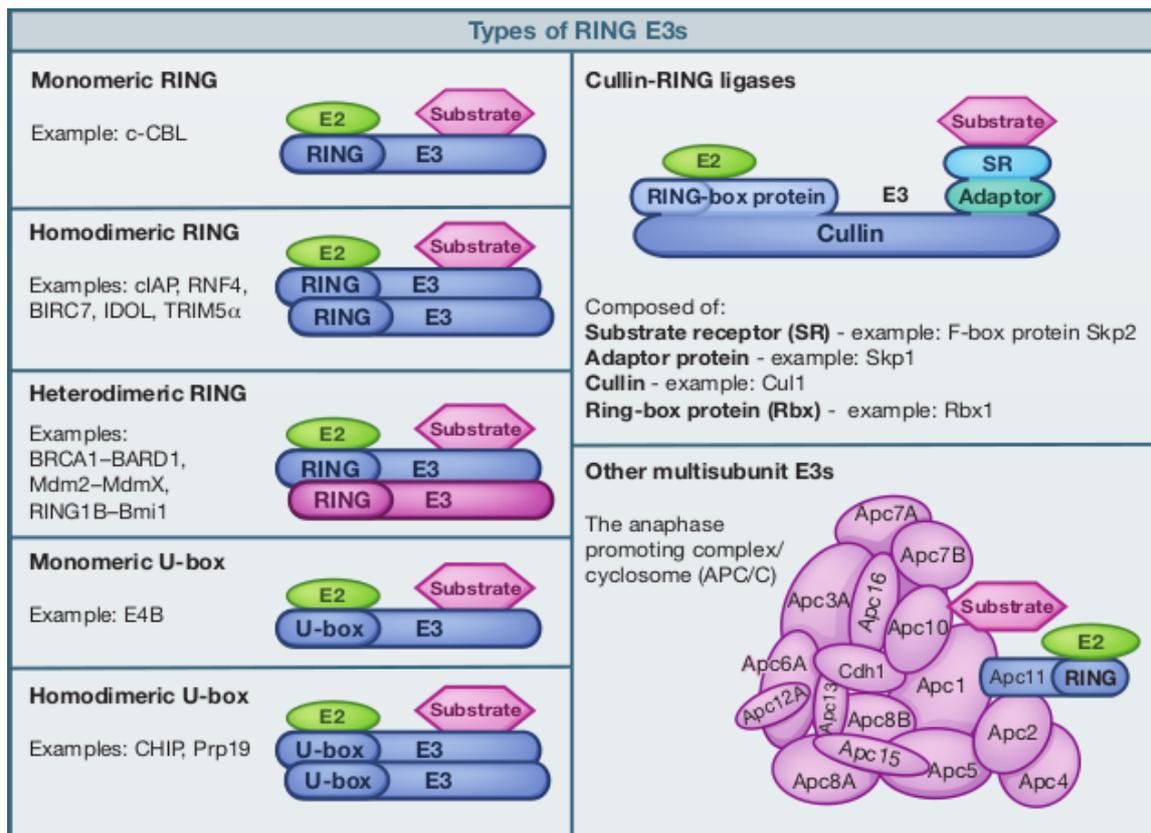
Contrairement à d'autres ubiquitine-ligases, les E3 à domaine RING servent d'interface entre la protéine E2 liée à l'ubiquitine et le substrat. L'ubiquitine ne passe pas par l'ubiquitine-ligase à domaine RING.

Les E3 à domaine RING peuvent être actives sous forme monomérique. C'est par exemple le cas de l'E3 c-CBL, dont la structure en complexe avec une E2 est connue [62]. De plus, sa dimérisation augmente son activité ubiquitine-ligase [63].

Certaines E3 à domaine RING forment des homodimères, c'est le cas, par exemple, des E3 cIAP [64], RNF4 [65] et BIRC7 [66]. Dans le cas des homodimères, les deux domaines RING possèdent la capacité de lier E2 Ub [65] [66].

Tous les domaines RING des E3 à domaine RING ne possèdent pas une activité ubiquitine-ligase intrinsèque. Ces domaines RING (sans activité ubiquitine-ligase) vont alors interagir et stimuler l'activité ubiquitine-ligase d'un domaine RING actif d'une autre protéine formant un hétérodimère. C'est notamment le cas pour les protéines Bard1, Bmil et MdmX qui vont respectivement s'associer aux domaines Brca1 [67] Ring1b [68] et Mdm2 [69].

Les E3 Cullin RING ligase (CRL) ont une structure différente. Elles sont composées d'une protéine Ring-box qui possède un domaine RING, une protéine cullin, un récepteur pour le substrat ainsi qu'une protéine adaptatrice[70]. (Figure 10 page 17)



**FIGURE 10** – Figure récapitulative des différentes organisations des ubiquitine-ligases à domaine RING et U-Box.[71]

Le domaine RING permet l'interaction à la protéine E2. La structure de la protéine Cbl en complexe avec son enzyme E2 Ubch7 résolue en 2000 [62] a permis la meilleure compréhension de cette interaction. Ainsi, comme vu précédemment (Figure 8 page 14), ce sont les boucles B1 et B2 de E2 qui interagissent avec le domaine RING de c-Cbl. (Figure 11 page 18)

À la surface du domaine RING, un sillon formé par les deux boucles liant le zinc et l'hélice entre ces boucles permet l'interaction avec les boucles B1 et B2 Ubch7. Cette interaction nécessite la Phe 63 de la boucle B1 ainsi que la Pro 97 et l'Ala 98 de la boucle B2 de Ubch7. Des interactions de type Van der Waals entre l'Ile 383 de la boucle 1 du domaine RING, le Trp 408, la Cys 404, la Ser 407 et la Ser 411 de l'hélice du domaine RING avec la surface hydrophobe de Ubch7 favorisent cette interaction [62] (Figure 11 page 18).

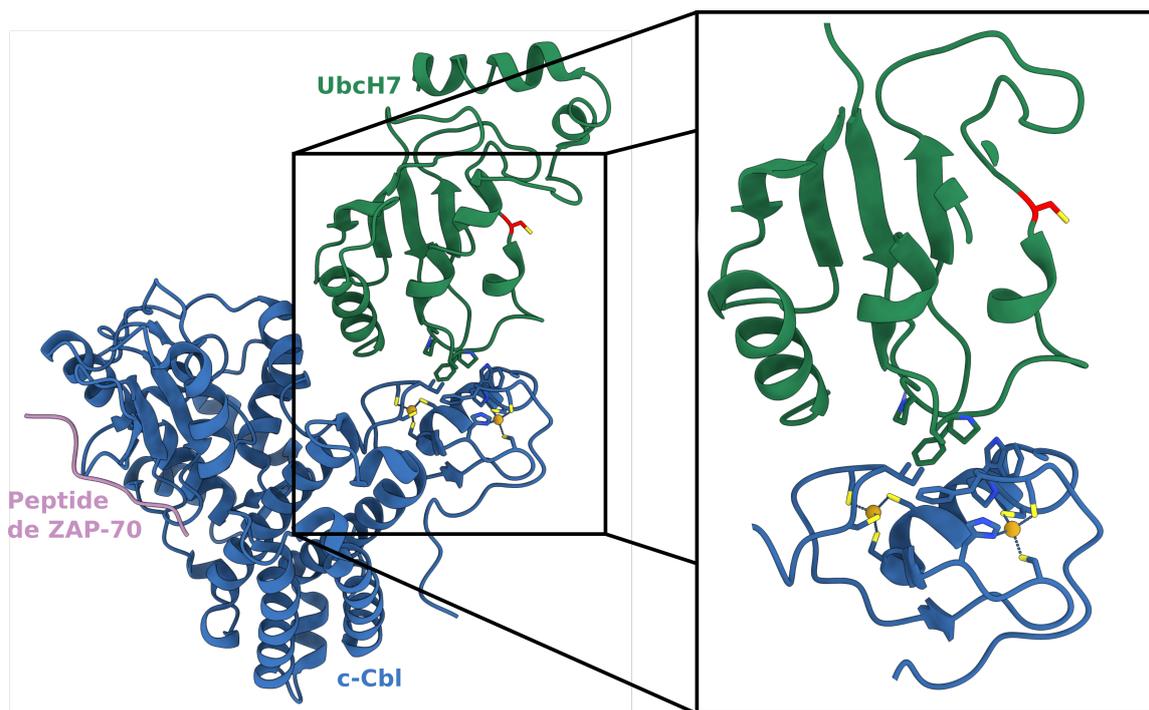


FIGURE 11 – Interaction entre le domaine RING de c-Cbl et l'enzyme E2 Ubch7. [58]

La surface de liaison de la protéine E2 au domaine RING est en partie superposée à la surface de liaison de E2 avec la protéine E1, qui sert au chargement en ubiquitine de la protéine E2 [72]. Ainsi, lorsque l'ubiquitine a été transférée au substrat, la protéine E2, pour être rechargée en ubiquitine, doit se dégager du domaine RING.

Les structures connues de l'interaction entre les E3 à domaine RING et les E2 indiquent que la cystéine catalytique de E2 se trouve éloignée du domaine RING [62]. Ceci s'explique par le rôle indirect des E3 à domaine RING dans l'ubiquitination du substrat. (Figure 11 page 18).

La structure de E2 en complexe avec le domaine RING n'est pas différente de la structure de E2 seule. Ces résultats indiquent que l'activation de E2 par le domaine RING ne serait pas due à un changement structural de E2 [62]. Cependant, une autre étude indique que le domaine RING activerait E2~Ub par une stabilisation de l'ubiquitine dans une conformation dite fermée [73].

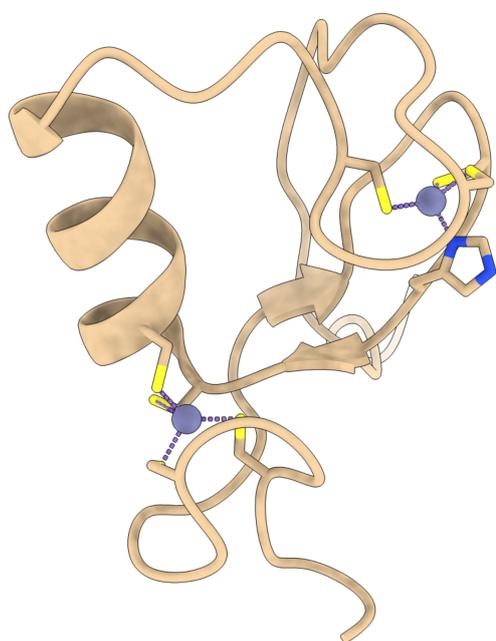
Ainsi les E3 à domaine RING permettent l'ubiquitination d'un substrat en liant le substrat et E2~Ub. Ceci se fait par la liaison du substrat par un domaine (autre que RING) et par la liaison de E2~Ub ce qui permet la stabilisation de l'ubiquitine.

Les E3 à sous unités multiples possèdent plusieurs domaines RING. C'est notamment le cas de la E3 GID [74], mais aussi des RING E3 composés d'une seule protéine comme la

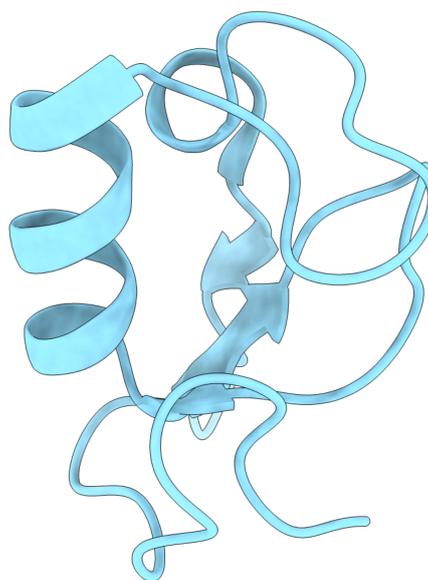
protéine mindbomb qui possède 3 domaines RING [75].

#### 1.2.4.2 E3 à domaine U-box

Un autre domaine d'environ 70 acides aminés, appelé U-BOX, possède un repliement similaire à celui du domaine RING. Cependant, à la différence du domaine RING le domaine U-Box ne possède pas de résidus liant le zinc. Ce sont des liaisons hydrogène qui permettent de conserver la même structure globulaire [76] (Figure 12 page 19).



RING domaine de CBL



U-Box domaine de Prp19

**FIGURE 12** – Comparaison du domaine RING de la E3 CBL (PDB : 1FBV) [62] à gauche et du domaine U-Box de Prp19 (PDB : 1N87) [77] à droite. Ces deux domaines ont un repliement similaire, mais seul le domaine RING lie deux atomes de zinc. La protéine U-Box se replie de la même façon via des liaisons hydrogène.

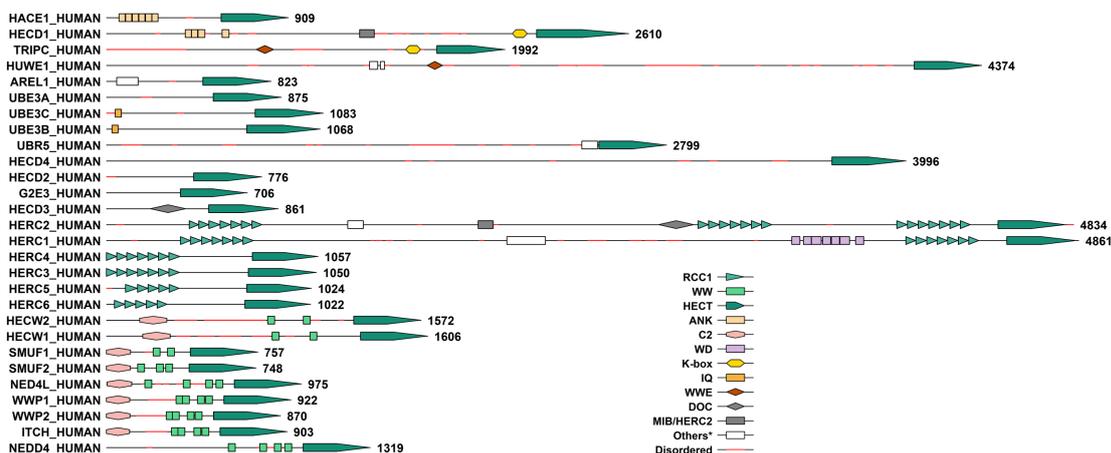
Comme les E3 à domaine RING, les E3 à domaine U-BOX peuvent être actives sous forme de monomères comme E4B mais peuvent aussi former des homodimères comme les protéines CHIP [78] et Prp19 [79]. Ce mécanisme des U-box est similaire à celui des E3 à domaine RING.

#### 1.2.4.3 E3 à domaine HECT

Un autre type d'ubiquitine-ligases est caractérisé par le domaine catalytique HECT (homologue à la partie C-terminale de E6AP)[80]. Contrairement aux ubiquitine-ligases à do-

maine RING, celles-ci ont une cystéine catalytique sur laquelle est transférée l'ubiquitine avant d'être transférée au substrat. C'est la structure du domaine HECT de l'E3 Ube3a (aussi appelé E6AP) qui a été résolue pour la première fois en 1999 [81]. Chez toutes les E3s à domaine HECT, le domaine HECT se trouve en position carboxy-terminale.

Il existe 27 ubiquitine-ligases à domaine HECT qui sont réparties en trois familles. La famille des HERC (ubiquitine-ligase à domaine HECT et RRC1-like) est caractérisée par la présence de un à trois domaines RRC1-like qui sont des domaines de types "β-propeller" à sept pales (détaillés dans la section 2.2.2 page 42). Cette famille est composée de six protéines HERCs. La famille des NEDD4 (aussi appelée C2-WW-HECT) est composée de 9 E3s (NEDD4, NEDD4L, HECW1, HECW2, ITCH, SMURF1, SMURF2, WWP1 et WWP2). Toutes possèdent un domaine C2, deux à quatre domaines WW et un domaine catalytique HECT. La dernière famille est constituée des autres ubiquitine-ligases à domaine HECT (Figure 13 page 20).



**FIGURE 13** – Représentation des différentes E3 à domaine HECT. Ce graphique a été réalisé en utilisant le logiciel ProFeatMap créé par Goran Bich [47]. Le domaine HECT, représenté par une flèche verte, se trouve toujours en position carboxy-terminale de la protéine. Ces E3 sont classées en fonction de leur organisation en domaine. Ainsi les HERC et Need8 sont regroupées. Les pales des domaines RLD (RCC1-like) sont représentées par les courtes flèches vertes (7 pales par domaine RLD).

Le domaine HECT est composé de deux lobes appelés respectivement le lobe carboxy-terminal (lobe-C) qui contient la cystéine catalytique et le lobe amino-terminal (lobe-N) qui sert à la liaison à la protéine E2. Ces deux lobes sont reliés par une boucle flexible [82] (figure 14 A, page 22).

Un aspect capital du domaine HECT est la zone flexible entre les deux lobes du domaine HECT agissant comme une charnière. Ainsi, la position relative des deux lobes varie lors des différentes étapes de l'ubiquitination. [81] [82] [83].

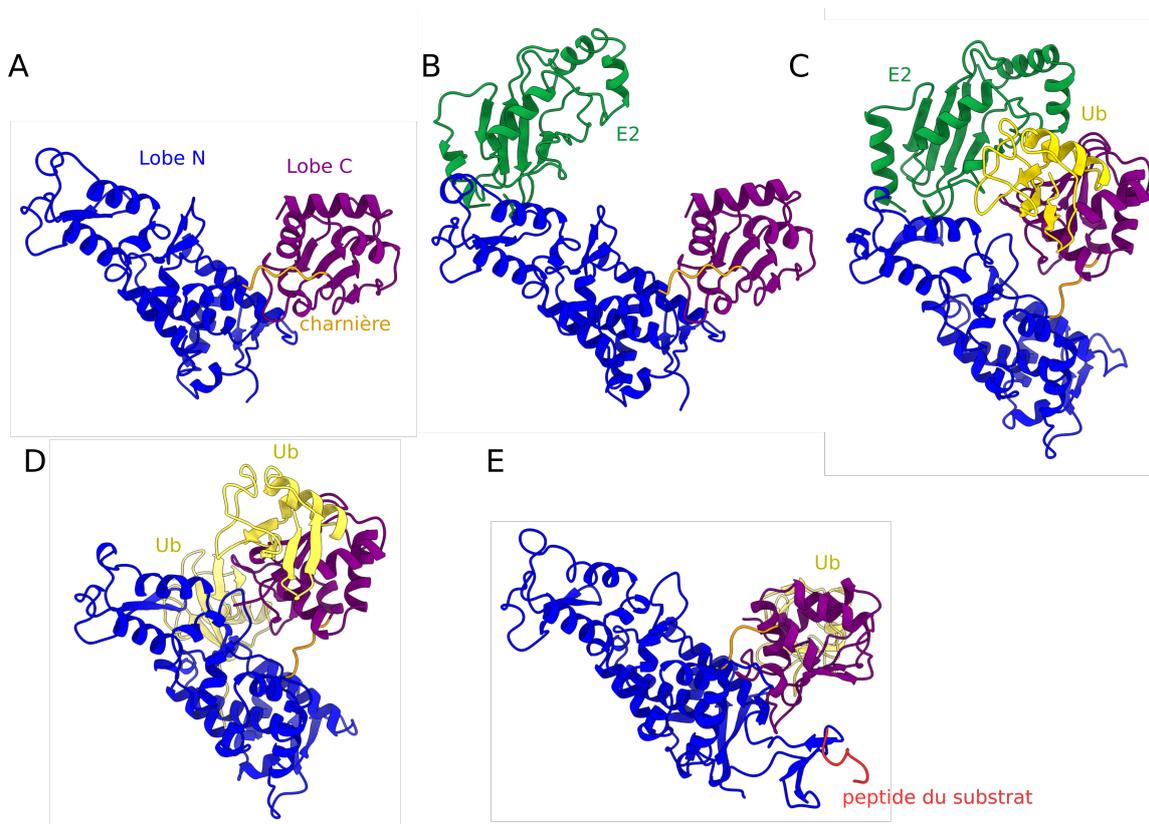
La première étape est la liaison du domaine HECT à la protéine E2, via le lobe-N. Cette interaction se fait via les boucles B1 et B2 et la partie amino-terminale de l'hélice  $\alpha_1$  de E2 [15] [84]. Les premières structures de complexe entre le domaine HECT et une E2 ont montré une grande distance entre la cystéine catalytique de E2 et la cystéine catalytique du domaine HECT, ce qui laisse supposer un changement structural pour la réaction de transthiolation [81]. (figure 14 B, page 22)

C'est la structure cristallographique de la protéine NEDD4L en complexe avec UBCH5-Ubiquitine (E2-Ub) qui a montré un changement de position relative des deux lobes, rendant le domaine HECT plus compact. Cette conformation du domaine HECT permet de rapprocher la liaison thioester E2-Ub de la cystéine catalytique du lobe-C. De plus, le lobe-C interagit avec l'ubiquitine via une interaction de type hydrophobe. Les résidus permettant cette interaction de type hydrophobe sont conservés au sein de la famille NEDD4 mais pas chez les autres E3 à domaine HECT. [15]. Ainsi le lobe-N interagit avec la protéine E2-Ub et la charnière initie une rotation du lobe-C qui permet un rapprochement de la cystéine catalytique avec la liaison thioester de E2 à l'ubiquitine. La distance entre les deux cystéines catalytiques reste de 8Å et une rotation supplémentaire serait donc nécessaire [15]. (figure 14C, page 22)

De plus, cette conformation entraîne une interaction entre le lobe-N et l'ubiquitine. Dans cette structure, une phénylalanine de la quatrième boucle de la protéine E2 interagit avec une surface hydrophobe du lobe-N du domaine HECT. Cette phénylalanine est conservée dans les E2 UBC4 et UBC5 [84].

La structure cristallographique de la protéine NEDD4-Ub a été résolue en 2013 [85]. Cette structure montre le domaine HECT après formation de la liaison thioester avec l'ubiquitine. Cette structure ne montre pas de différence avec la structure NEDD4L/E2-Ub au niveau de l'interaction entre le domaine HECT et l'ubiquitine. (figure 14 D, page 22)

Le transfert de l'ubiquitine liée à la cystéine catalytique du domaine HECT vers la lysine du substrat est favorisé par un rapprochement dans l'espace. La structure de RSP5-UB/Sna3 [86] montre une rotation de 130° via la charnière entre les deux lobes du domaine HECT par rapport aux structures de HECT/E2-Ub et HECT-Ub. Cette conformation est stabilisée par l'interaction entre les deux lobes. Dans cette structure, la phénylalanine 806 (aussi appelée phénylalanine -4) se situe dans une poche du lobe-N composée de la Phe 505 et la Leu 506. De plus, la mutation de cette phénylalanine empêche l'activité ubiquitine-ligase, indiquant que cette phénylalanine joue un rôle d'ancrage pour la position relative des deux lobes du domaine HECT [86]. Dans cette structure, la lysine du substrat n'est pas visible. [87] [88]. (figure 14E, page 22)



**FIGURE 14** – Structures connues des différentes étapes de l'ubiquitination d'un substrat par le domaine HECT. Le domaine HECT est représenté de 3 couleurs (Le lobe-N en bleu, le lobe-C en violet et la charnière en orange). E2 est représentée en vert, l'ubiquitine en jaune et le substrat en rouge. A) Structure historique du domaine HECT de Ube3a. B) Structure du domaine HECT de Ube3a en complexe avec l'enzyme E2 Ubch7. Notez que le domaine HECT reste dans la même position lorsqu'il se lie à la protéine E2 seul. C) Structure du domaine HECT de la protéine NEDD4L lié à l'ubiquitine (HECT/E2-Ub). Notez le mouvement du lobe-C du domaine HECT par rapport au lobe-N. Ce changement de position relative, via la charnière entre ces deux lobes, permet le rapprochement de la cystéine catalytique du domaine HECT et de celle de E2. D) Structure du domaine HECT lié de manière covalente à l'ubiquitine. Notez la capacité du domaine HECT à lier de manière non covalente une autre ubiquitine au niveau du lobe N. E) Structure du domaine HECT lié à l'ubiquitine en complexe avec le peptide de son substrat. Le lobe C effectue une rotation de 130° par rapport à la structure D).

#### 1.2.4.4 Domaine RBR

Dans le cas des ubiquitine-ligases à domaine RBR (domaine RING between-RING RING), l'ubiquitine est transférée sur l'E3 avant d'être transférée sur le substrat à la manière des E3 à domaine HECT.

Ce domaine RBR est composé d'un premier domaine RING (RING1) aussi appelé BRcat (Begnin RING catalytic) qui recrute l'enzyme E2 chargée en ubiquitine, d'un domaine IBR

(In Between RING) et d'un second domaine RING (appelé RING2 ou Rcat pour Ring Catalytic). Ce second domaine RING contient une cystéine catalytique sur laquelle l'ubiquitine va être transférée.

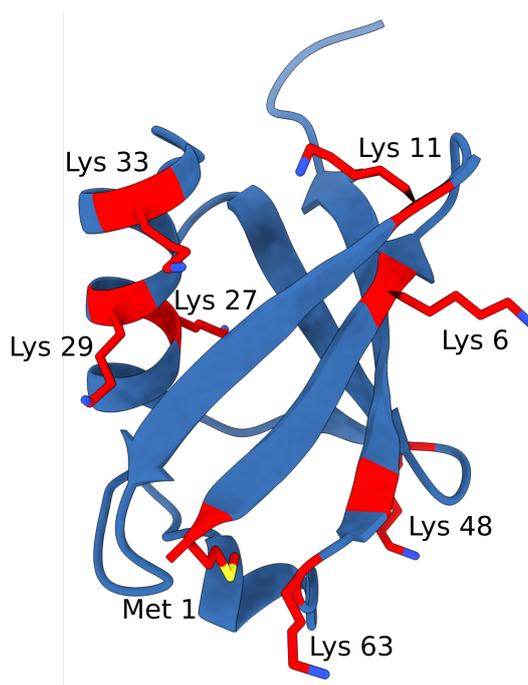
14 protéines à domaine RBR sont connues [89]. Le mécanisme des E3 à domaine RBR se situe à mi-chemin entre les E3 à domaine RING et les E3 à domaine HECT. Le domaine RING1 permet l'interaction avec la protéine E2-Ub, le domaine RING2 contient une cystéine catalytique sur laquelle l'ubiquitine va être transférée (à la manière des E3 à domaine HECT). Le domaine IBR est séparé des deux domaines RING par des zones flexibles. Ces zones flexibles permettent le changement relatif de ces 3 domaines les uns par rapport aux autres et jouent un rôle important dans le mécanisme d'ubiquitination d'un substrat.

L'interaction entre le domaine RING1 et E2 se fait de la même manière que pour les E3 à domaine RING classique [90]. Le transfert de l'ubiquitine sur le domaine RING2 se fait via un rapprochement des domaines RING1 et RING2 mais ce mécanisme n'est cependant pas encore parfaitement compris.

## 1.3 Variété des chaînes d'ubiquitines

### 1.3.1 Les différentes chaînes d'ubiquitines

Les ubiquitines peuvent former différents types de chaînes en se liant via différentes lysines. La liaison se fait entre l'extrémité carboxy-terminale d'une ubiquitine en bout de chaîne (dite ubiquitine distale) et une des lysines de l'ubiquitine proximale. Les résidus disponibles pour former ces chaînes sont les suivants : K6, K11, K27, K29, K33, K48, K63. De plus, les ubiquitines peuvent être ubiquitinées sur la méthionine M1. (Figure 15 page 24.

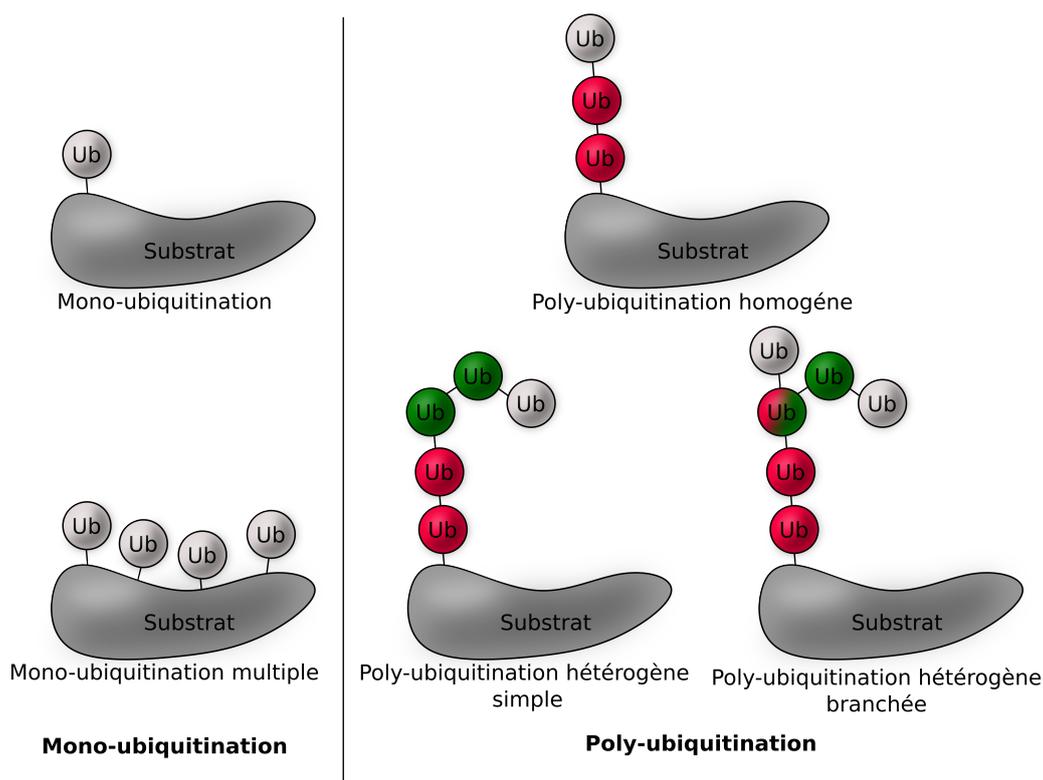


**FIGURE 15** – Structure de l'ubiquitine (PDB : 1UBQ) [10]. Les lysines et la méthionine pouvant former des chaînes d'ubiquitines sont colorées en rouge.

L'ubiquitination d'un substrat se fait sur une de ses lysines et peut prendre plusieurs formes. Un substrat peut être mono-ubiquitiné, dans ce cas, une seule molécule d'ubiquitine est liée au substrat. Dans le cas de la multi-mono-ubiquitination, plusieurs ubiquitines sont liées sur des lysines différentes du substrat.

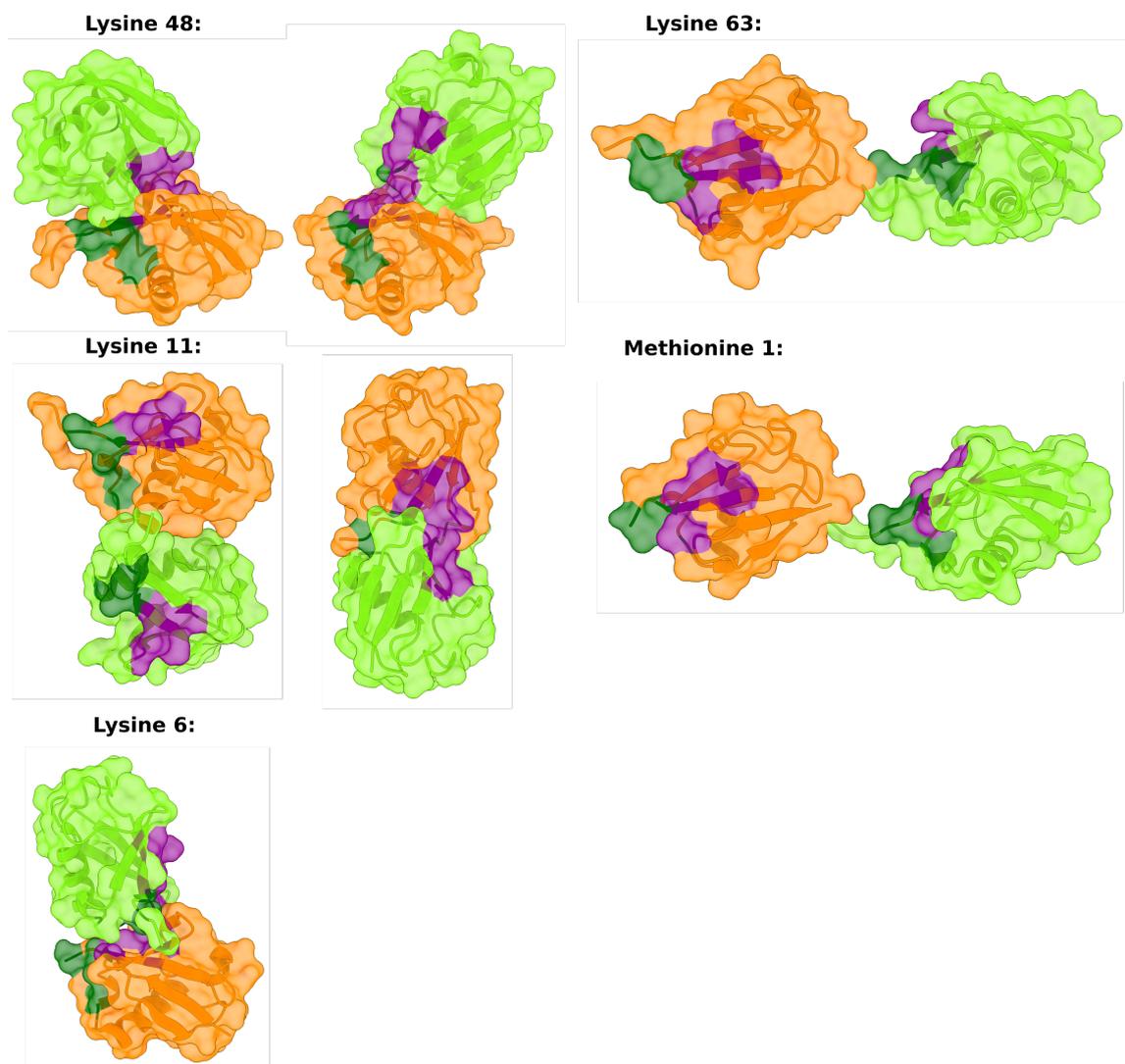
Un substrat poly-ubiquitiné a une chaîne de plusieurs ubiquitines liée sur une seule lysine. Ainsi, une poly-ubiquitination dite homogène consiste en la liaison d'une chaîne d'ubiquitines liées entre elles par la même lysine. Dans le cas des chaînes hétérogènes, les ubiquitines sont liées par différentes lysines (figure 16 page 25).

Les chaînes d'ubiquitines dites homogènes sont constituées d'ubiquitines liées par les mêmes lysines. Par opposition, les chaînes dites hétérogènes sont constituées d'ubiquitines liées via des liaisons mixtes. (figure 16 page 25)



**FIGURE 16** – Représentation schématique des différents types d'ubiquitination d'un substrat. À gauche sont représentées les mono-ubiquitinations. À droite sont représentées les poly-ubiquitinations. La couleur de l'ubiquitine correspond à la lysine impliquée dans la liaison avec l'ubiquitine suivante de la chaîne, par exemple rouge pour K63 et vert pour K11

Les variations de chaînes d'ubiquitines entraînent des différences structurales des chaînes elles-mêmes. En effet, certains types de liaison d'ubiquitine entraînent des interactions entre celles-ci via leurs patches hydrophobes (Ile 44 et Ile 36), ce qui amène à la formation d'une chaîne compacte d'ubiquitines. C'est le cas des chaînes de type Lys 6, Lys 11 et Lys 48. D'autres types de chaînes, comme les chaînes de type Met 1 et Lys 63, ne permettent pas ces interactions, ce qui rend la chaîne d'ubiquitines plus flexible. (Figure 17 page 26)



**FIGURE 17** – Structures de différentes chaînes de deux ubiquitines liées par différentes lysines. Les deux principaux patches, Ile 36 et Ile 44, sont respectivement représentés en vert et violet. L'ubiquitine proximale (attachée au substrat) est représentée en orange. L'ubiquitine distale est représentée en vert clair. Lysine 48 gauche PDB : 1AAR [91], droite PDB : 2PE9b [92]. Lysine 11 gauche PDB : 2XEW [93], droite PDB : 3NOB [94]. Lysine 6 PDB : 2XK5 [95]. Lysine 63 PDB : 2JF5 [96]. Méthionine 1 PDB : 2W9N [96].

Ces différents types de chaînes entraînent une diversité de structures globales et donc une diversité de motifs de répartition des patches sur celles-ci. Sur la figure 17 (Page 26), seules des chaînes de deux ubiquitines sont représentées, mais les chaînes d'ubiquitines peuvent être plus longues, ce qui enrichit encore la diversité de dispositions des patches.

Les modifications post-traductionnelles des ubiquitines elles-mêmes, comme des phosphorylations [97] et des acétylations [98], participent à la régulation de l'ubiquitination et

constituent une couche supplémentaire de diversités de ces chaînes [99].

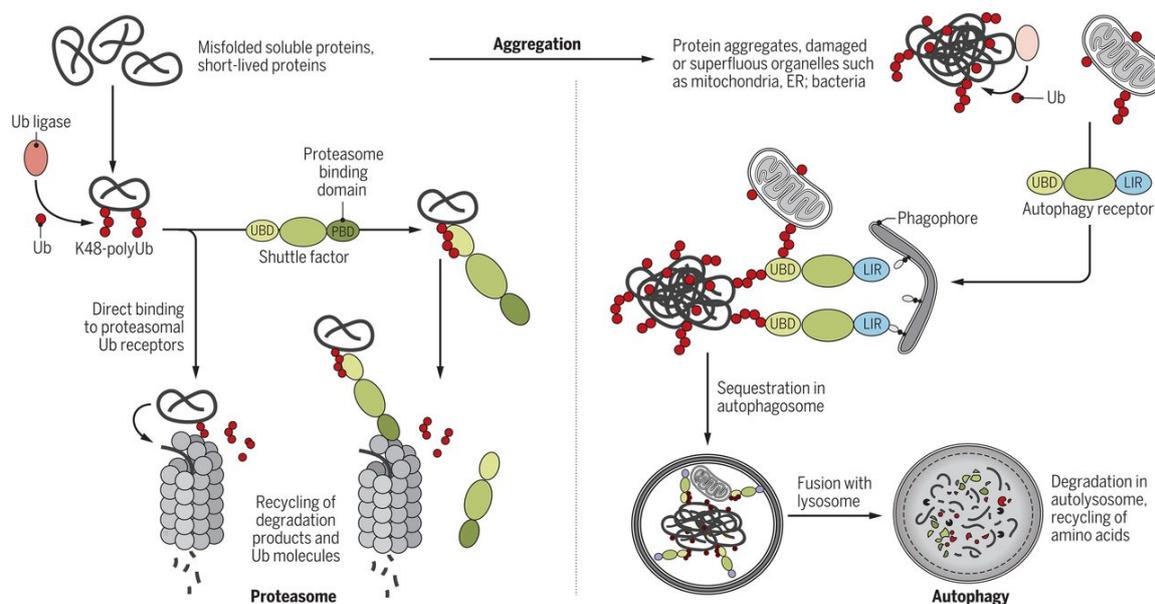
### 1.3.2 Rôles des différentes chaînes d'ubiquitines

Le premier rôle, le plus connu, des chaînes d'ubiquitination, est l'étiquetage du substrat pour entraîner sa dégradation par le protéasome. Cependant, toutes les chaînes d'ubiquitines ne sont pas impliquées dans cette voie de dégradation.

Les chaînes de type K48 sont les plus étudiées. Elles entraînent la dégradation du substrat par le protéasome. Ces chaînes peuvent être reconnues par des récepteurs d'ubiquitine du protéasome comme Rpn10 [100] et Rpn13 [101]. Elles peuvent aussi être reconnues par des protéines qui possèdent un site de liaison à cette chaîne et un domaine de liaison au protéasome (comme RAD23 [100]). Après reconnaissance, le substrat est déplié et dégradé par le protéasome. Les ubiquitines sont, quant à elles, relarguées afin d'être utilisées à nouveau (Figure 18 gauche, page 28).

Les chaînes d'ubiquitines peuvent aussi être un signal déclencheur de l'autophagie et ainsi dégrader des protéines, des agrégats de protéines et des organites. Ce phénomène implique des protéines ayant un domaine de liaison à l'ubiquitine et un domaine de liaison au récepteur d'autophagie. La protéine P62 est capable de reconnaître les chaînes d'ubiquitines K63 via son domaine UBD et déclenche le mécanisme d'autophagie via son domaine Lir [102] (Figure 18 droite, page 28).

Les chaînes K48, K63, K6 et K11 semblent avoir un rôle dans le phénomène de mitophagie. En effet, les mitochondries endommagées sont notamment ubiquitinées par l'ubiquitine-ligase E3 Parkin, capable de produire ces quatre types de chaînes. [103] [104]



**FIGURE 18** – À gauche : Les chaînes de type K48 sont accrochées par l'enzyme E3 sur le substrat. Ces chaînes sont reconnues par des protéines navettes qui sont équipées à la fois d'un domaine de liaison à l'ubiquitine et d'un domaine de liaison au protéasome (comme la protéine RAD23). Avant d'être dégradés par le protéasome, les substrats sont désubiquitinés et dépliés. À droite : L'autophagie sélective permet la dégradation des matériaux cytosoliques volumineux et hétérogènes, comme des protéines agrégées ou des organites. Un phagophore va reconnaître le motif LIR des récepteurs d'autophagie qui sont eux-mêmes liés à la chaîne d'ubiquitine par leurs domaines UBD. Finalement, le phagophore se referme pour donner naissance à l'autophagosome qui fusionne avec le lysosome. Figure de [105]

Les chaînes d'ubiquitines de type K6 ne sont pas impliquées dans la dégradation du substrat par le protéasome [106] [107]. Cependant, elles le sont dans la réponse cellulaire lors de l'exposition aux UV. L'exposition de cellules au rayonnement UV augmente la concentration cellulaire de chaînes d'ubiquitines de type K6 et de type K33 [108]. Cette augmentation est liée au rôle que jouent les chaînes de types K6 dans les mécanismes de réparation de l'ADN. En effet, l'ubiquitine-ligase BRAC1/BRAD1, connue pour son implication dans la régulation de la réparation de l'ADN, polyubiquitine ses substrats sous forme de chaînes contenant des liaisons K6 [109] [110].

La concentration cellulaire des chaînes de type K11 augmente lors de l'inhibition du protéasome, indiquant le rôle de ces chaînes dans la dégradation par le protéasome. De plus, les chaînes de type K11 sont importantes dans la régulation du cycle cellulaire. Le taux de ce type de chaînes augmente lors de la mitose. [94] Le complexe promoteur de l'anaphase (APC/C) est une ubiquitine-ligase à domaine HECT qui produit des chaînes de type K11, notamment lors de la mitose [19] [111]. Les chaînes de type K27 jouent un rôle dans la réparation de l'ADN [112] et dans la réponse immunitaire innée et adaptative [113]. La

fonction des chaînes de type K29 n'est pas encore bien définie, mais il semble qu'elles jouent un rôle dans la réponse cellulaire au stress protéotoxique et dans la régulation du cycle cellulaire [114]. Les chaînes d'ubiquitines de type K33 régulent la signalisation cellulaire lors des réponses immunitaires acquise et innée [115] [116].

Les chaînes de type M1 aussi appelées ubiquitinations linéaires semblent jouer, quant à elles, de multiples rôles de signalisations cellulaires [117] [118].

Les implications de ces différents types de chaînes d'ubiquitines ne sont pas encore totalement comprises et de nouvelles fonctions sont régulièrement découvertes [119].

## 1.4 Domaines et autres enzymes relatives aux ubiquitines

### 1.4.1 Les domaines de liaisons aux ubiquitines

L'ubiquitine et les chaînes d'ubiquitines peuvent être reconnues par différents domaines appelés UBDs (pour "ubiquitine Binding Domain"). Les différents UBDs sont capables de discriminer les différents types de chaînes d'ubiquitines.

Les UBDs sont regroupés en 5 familles. Les UBDs à hélice  $\alpha$ , ceux de type doigt de zinc (ZnF), ceux à domaine homologue à la pleckstrin et ceux à domaine Ubc-like. La dernière famille regroupe les autres UBDs. C'est cette diversité structurale des UBDs qui leur permet d'interagir avec différents types de chaînes d'ubiquitines [12] (Table 3 page 30).

### 1.4.2 Les désubiquitinases

Les désubiquitinases (DUBs) sont des enzymes qui jouent un rôle important dans le système ubiquitine/protéasome. 106 DUBs sont connus à ce jour et sont réparties en 5 familles ; la famille USP (Ubiquitine-spécifique Protéase), la famille UCH (Ubiquitine C-terminale Hydrolase), la famille OTU (Ovarian Tumor Protease), la famille MJD (Machado Joseph Disease), la famille des métallo-protéases à zinc et la famille JAMM (JAB1/MPN/MOV34) [96].

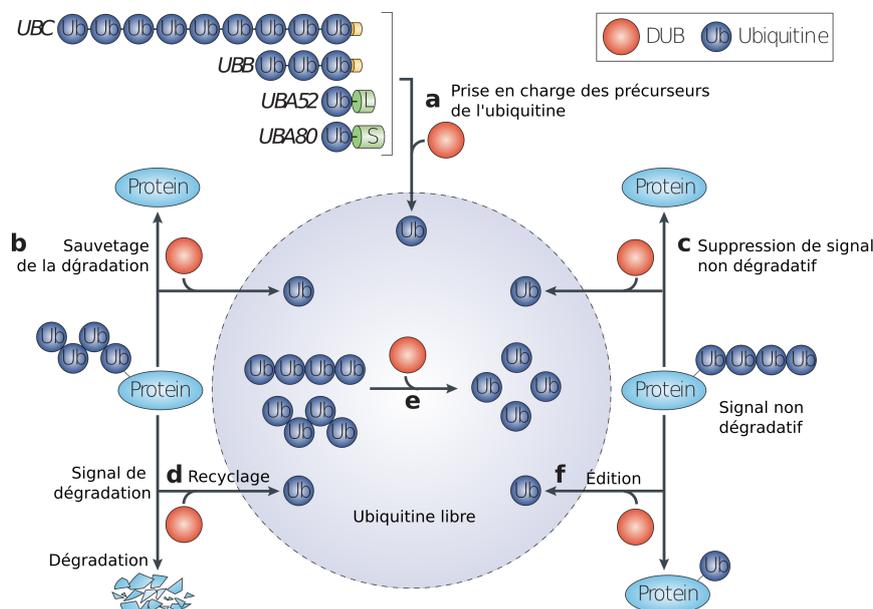
Les désubiquitinases sont des enzymes clivant la liaison entre une ubiquitine et une autre protéine (substrat ou ubiquitine). Ces enzymes peuvent être non spécifiques, coupant entre deux ubiquitines indépendamment du type de liaison, c'est le cas de la plupart des USPs. D'autres DUBs sont spécifiques, soit du substrat comme Ubp8, soit du type de chaîne d'ubiquitine. Les DUBs interviennent à de nombreux niveaux de la cascade d'ubiquitination. (Figure 19 Page 31)

Les DUBs peuvent être spécifiques à certains types de chaînes d'ubiquitines. Certaines coupent la chaîne au niveau du substrat (à la base de la chaîne) alors que d'autres sont des

Repliement	UBD	Protéine	Fonction
Helice $\alpha$	UIM	S5a (humain), Vps27, STAM, EPSINS, RAP80	Dégradation par le protéasome, Endocytose Biogenèse des MVB, réparation de l'ADN
	MIU or IUIM	RABEX-5	Endocytose
	DUIM	HRS	Biogenèse des MVB
	UBM	POL-i, REV-1	Réparation de l'ADN
	UBAN	NEMO, ABIN1-3, OPTINEURIN	Signalisation NF- $\kappa$ B
	UBA	Rad23/HR23A, Dsk2, NBR1	Régulation de kinase, autophagie, ciblage du protéasome
	GAT	GGA3, TOM1	Biogenèse des MVB
	CUE	Vps9, TAB2, TAB3	Endocytose, régulation de kinase
	VHS	STAM, GGA3	Biogenèse des MVB
	Doigt de zinc	UBZ	POL-h ; POL-k, Tax1BP1,
NZF		NPL4, Vps36, TAB2, TAB3	Biogenèse des MVB, régulation de kinase, régulation ERAD
ZnF A20		RABEX-5, A20	Endocytose, régulation de kinase
Domaine PH	Znf UBP	USP5/IsoT, HDAC6	Fonctionnement du protéasome, aggresome et autophagie
	PRU	RPN13	Fonctionnement du protéasome
Ubc-like	GLUE	EAP45	Biogenèse des MVB
	UEV	Uev1/Mms2	Réparation de l'ADN, Biogenèse des MVB, régulation de kinase
Autres	Ubc	UBCH5C	Transfert d'ubiquitine
	SH3	Sla1/CIN85	Endocytose
	PFU	Ufd3	ERAD
	Jab1/MPN	Prp8	Épissage

**TABLE 3** – Tableau des différents UBDs, leurs familles et leurs fonctions connues. Tableau adapté de [13]

endo-désubiQuitinases et coupent la chaîne entre 2 ubiquitines. Les exo-désubiQuitinase coupent l'ubiquitine distale d'une chaîne.



**FIGURE 19** – Figure représentant les multiples rôles des DUBs. a) L'ubiquitine est codée par 4 gènes (UBC, UBB, UBA52 et UBA80) et est transcrite et traduite sous la forme d'une fusion linéaire constituée de copies multiples d'ubiquitines ou d'ubiquitines fusionnées à l'extrémité amino-terminale de deux protéines ribosomales. La génération d'ubiquitine libre à partir de ces précurseurs est une fonction clé des désubiQuitinases. b) La désubiQuitination peut permettre aux protéines d'échapper à la dégradation. c) La désubiQuitination peut également supprimer les autres signaux des chaînes d'ubiquitines. d) Les DUBs jouent un rôle crucial dans le maintien de l'homéostasie de l'ubiquitine et dans la prévention de la dégradation de l'ubiquitine avec les substrats du protéasome 26S (recyclage de l'ubiquitine). e) Le désassemblage des chaînes libres d'ubiquitine permet les réutilisations des ubiquitines. f) Certaines DUBs pourraient avoir pour fonction d'éditer les chaînes d'ubiquitines et ainsi échanger un type de signal d'ubiquitine pour un autre. Figure adaptée de [120]

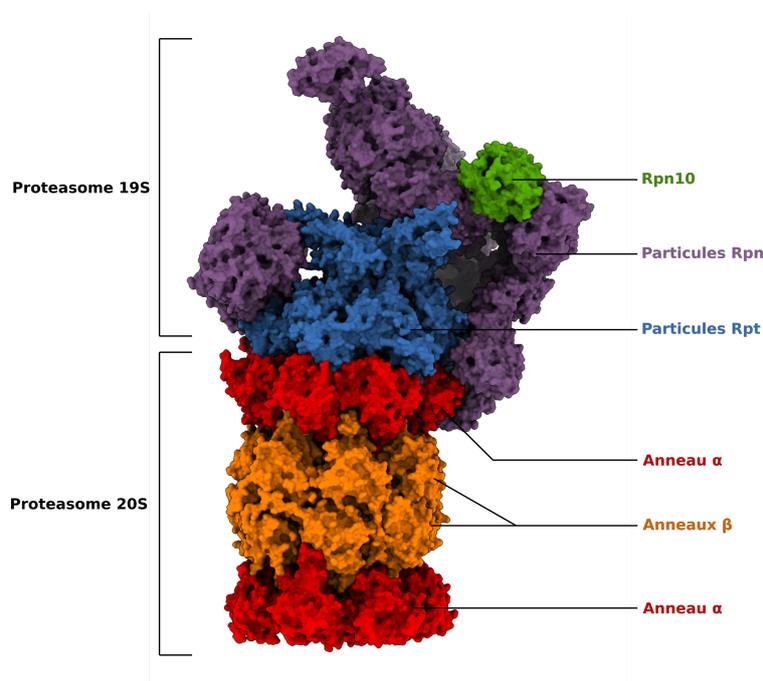
## 1.5 Le protéasome

Le protéasome est la machine qui dégrade les protéines qui sont marquées d'une chaîne d'ubiquitines (principalement K48). Le protéasome est localisé dans différents compartiments cellulaires : dans le cytoplasme et dans le noyau. La proportion de protéasome dans ces deux compartiments varie selon le type cellulaire [121] [122].

### 1.5.1 Organisation structurale

Les premières informations structurales sur le protéasome, révélant sa forme cylindrique, ont été obtenues en 1986 par microscopie électronique [123]. Les différents composants du protéasome sont nommés en fonction de leur coefficient de sédimentation respectif

(Svedberg). Ainsi le protéasome 26S est composé du cœur catalytique (20S) et d'un complexe régulateur (19S).



**FIGURE 20** – Structure du protéasome résolue par cryo-microscopie électronique (PDB : 5WVI)[124]. La sous-unité 20S (en bas sur la représentation) est composée de deux anneaux  $\beta$  représentés en orange et de deux anneaux  $\alpha$  représentés en rouge. La sous-unité 19S (en haut sur la représentation) est composée des particules régulatrices non-ATPases (Rpn) représentées en violet et des particules régulatrices AAA+ ATPases (Rpt) représentées en bleu. Rpn10, capable de se lier à la E3 Ube3a, une ubiquitine ligase qui nous intéresse dans ce travail (Section 2.1, page 35) est représentée en vert.

### 1.5.1.1 Le protéasome 20S

Le protéasome 20S est formé de structures heptamériques circulaires superposées. Les deux heptamères positionnés à l'extérieur sont composés des sous-unités  $\alpha$ . Les deux heptamères au centre sont composés de sous-unités  $\beta$ . Ce sont les heptamères  $\beta$  qui sont responsables de l'activité catalytique tandis que les heptamères  $\alpha$  ont une activité régulatrice.

Cette structure de superposition de 4 anneaux est conservée à travers le règne des eucaryotes [125] [126]. Les anneaux du protéasome 20S eucaryotique sont des hétéromères composés des sous-unités  $\alpha 1$  à  $\alpha 7$  pour les anneaux  $\alpha$  et  $\beta 1$  à  $\beta 7$  pour les anneaux  $\beta$  [127]. Le 20S procaryotique est composé d'homo-heptamères [128] [129].

Ce sont les sous-unités  $\beta 1$ ,  $\beta 2$  et  $\beta 5$  qui ont une activité catalytique, notamment grâce à leur thréonine en position amino-terminale (Protéase à thréonine) [130].

### 1.5.1.2 Le protéasome 19S

Le protéasome 19S est composé de 19 sous-unités regroupées en 2 classes, les particules régulatrices triphosphate ATPases (Rpt) et les particules régulatrices non-ATPases (Rpn) comme les Rpn10 et 13 décrits précédemment [131].

Le protéasome 19S est composé de 2 parties appelées base et couvercle. La base du protéasome 19S en contact avec le protéasome 20S, comporte les protéines Rpt 1 à 6, Rpn 1, 2, 10 et 13. Le couvercle est composé des protéines Rpn 3, 5, 6, 7, 8, 9, 11 et 12 [132].

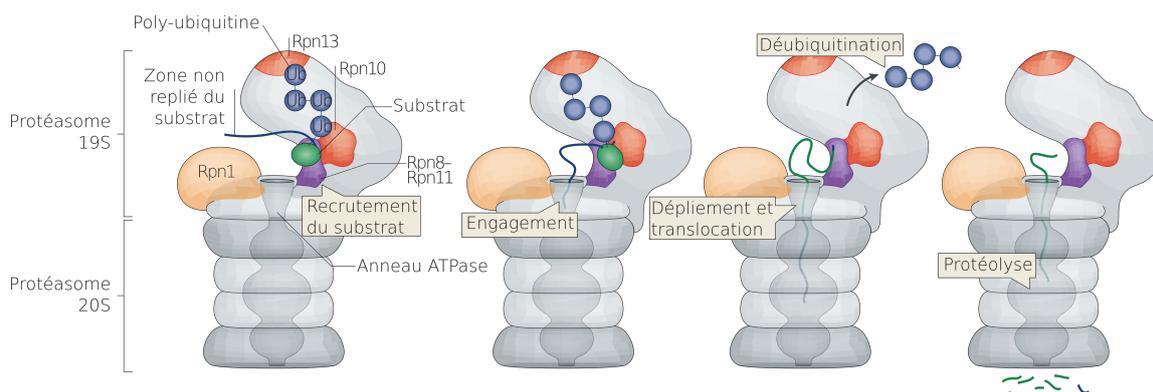
### 1.5.1.3 Les variations du protéasome

Différentes variantes du protéasome existent. En effet, les cellules immunitaires produisent des sous-unités  $\beta$  différentes. Ces sont les  $\beta 1i$ ,  $\beta 2i$  et  $\beta 5i$  qui remplacent les  $\beta 1$ ,  $\beta 2$  et  $\beta 5$  formant l'immuno-protéasome [133]. Le thymo-protéasome qui se retrouve dans les cellules épithéliales corticales du thymus est composé des sous-unités  $\beta 5t$ ,  $\beta 1i$  et  $\beta 2i$  [134].

De plus, différents complexes, autres que le protéasome 19S, peuvent se fixer sur le protéasome 20S : le protéasome 11S (aussi appelé PA28) et PA200. 3 différentes sous-unités PA28 existent :  $\alpha$ ,  $\beta$  et  $\gamma$ . Ces sous-unités forment 2 heptamères différents, un hétéroheptamère PA28 $\alpha\beta$  et un homoheptamère PA28 $\gamma$  [135] [136]. Leurs rôles sont toujours mal connus. PA200, quant à lui, semble être impliqué dans différents mécanismes, notamment dans la spermatogenèse [137] et la réparation de l'ADN [138].

## 1.5.2 Mécanisme de dégradation protéasomale

Le protéasome 19S permet l'initiation de la dégradation du substrat. La première étape pour la dégradation d'un substrat polyubiquitiné est sa reconnaissance par le protéasome 19S via les récepteurs d'ubiquitine Rpn1, Rpn10 et Rpn13 [139] [140]. Ensuite, le substrat est dénaturé par le module dit moteur, composé des 6 Rpt, ce sont les Rpt 1 à 6 [141] [142]. C'est cet hexamère qui entraîne le substrat dénaturé vers le corps catalytique (le protéasome 26S). La chaîne d'ubiquitines, quant à elle, est séparée du substrat par Rpn11 et Rpn8 qui sont des DUBs, ces ubiquitines pourront être utilisées à nouveau [143].



**FIGURE 21** – Étapes de la dégradation protéasomique. Le protéasome reconnaît les chaînes d'ubiquitines liées aux substrats par l'intermédiaire de ses récepteurs (ici représentées : les particules régulatrices non-ATPase 10 (Rpn10) et 13 (Rpn13)) et initie ensuite la dégradation d'une région non structurée du substrat. Au fur et à mesure que les moteurs ATPase tirent le substrat dans le canal de dégradation, la chaîne d'ubiquitines est clivée. Figure adaptée de [144]

Le passage du substrat dénaturé dans le canal du protéasome 20S est appelé translocation. La désubiquitination et la dénaturation d'une portion du substrat sont toutes deux nécessaires à la translocation de celui-ci [145].

La dégradation par le corps catalytique du protéasome se fait au niveau des anneaux  $\beta$ . Le substrat est poussé par le protéasome 19S dans le protéasome 20S. C'est ici que l'activité protéolytique des sous unités  $\beta 1$ ,  $\beta 2$  et  $\beta 5$  entraîne la dégradation du substrat. Les mécanismes d'action de ces trois sous unités diffèrent les uns des autres. Le site de type caspase-like de  $\beta 1$  coupe son substrat préférentiellement après un résidu acide. Le site trypsine de  $\beta 2$  coupe son substrat préférentiellement après un résidu basique. Enfin,  $\beta 5$  est une chymotrypsine qui coupe son substrat de préférence après un acide aminé hydrophobe [146] [147]. Dans le cas des immuno-protéases,  $\beta 1i$  coupe son substrat après un résidu hydrophobe. Ce changement est dû à une variation de séquence au niveau de son site de liaison au substrat [148]. Les peptides issus de la dégradation du substrat par le protéasome sont ensuite dégradés par des peptidases [149].

## 2 | Les ubiquitine-ligases Ube3a et HERC2

Dans ce travail, nous nous intéressons à 2 ubiquitine-ligases à domaine HECT (décrit précédemment dans la section 1.2.4.3 page 19) qui sont les protéines Ube3a et HERC2. Ces protéines sont toutes deux exprimées par des gènes situés dans la même région chromosomique du chromosome 15 appelée 15q11-q13 (décrite plus précisément dans la section 2.4 page 45).

### 2.1 L'ubiquitine-ligase Ube3a

La protéine Ube3a ("ubiquitin protein ligase E3A") aussi appelée E6AP ("E6-associated protein") est une ubiquitine-ligase à domaine HECT, proche parente des protéines HERC comme nous allons le voir dans la section 2 (page 86). C'est cette protéine qui a donné son nom au domaine HECT [80].

La protéine Ube3a existe sous trois isoformes différentes [150]. Ces isoformes diffèrent par une extension en position amino-terminale plus ou moins longue. L'isoforme 2, qui fait 872 acides aminés de long, possède une extension de 23 acides aminés par rapport à l'isoforme 1 et de 3 acides aminés comparé à l'isoforme 3. [151]. (Tableau 4 page 36)

Ces trois isoformes sont principalement localisées dans le cytoplasme. L'isoforme 1 est la plus représentée (environ 85%) dans les différents types de cellules étudiées (hESC et neurones) [151].

Caractéristiques de l'isoforme	isoforme 1	isoforme 2	isoforme 2
Numéro RefSeq	NM_130838.1	NM_000462.3	NM_130839.2
Taille de la protéine (aa)	852	875	872
Taille du transcrit (pb)	4491	5276	5211
Nombre d'exons	10	14	13

**TABLE 4** – Tableau représentant les différentes isoformes de l'ubiquitine-ligase Ube3A. Tableau adapté de [152]

## 2.1.1 Ube3a et le système ubiquitine-protéasome

### 2.1.1.1 Le domaine HECT et l'ubiquitination

Cette ubiquitine-ligase est connue pour lier et utiliser la protéine E2 Ubch7 pour l'ubiquitination de substrat. La structure du domaine HECT a notamment été résolue en complexe avec cette E2 [81]. En plus de pouvoir utiliser ubch7, il a été montré la capacité de Ube3a à lier d'autre E2 tel que Ubch5 et Ubch8 [88]. Ube3a est connue pour sa capacité à former des chaînes d'ubiquitines de type K48 [153].

Le domaine HECT de Ube3a est, comme les autres domaines HECT, composés de deux lobes séparés par une charnière. La charnière de Ube3a est composée de 3 résidus (738, 739 et 740 numérotations de l'isoforme 2). La cystéine catalytique Cys 820 se trouve dans le lobe C-terminal. Cette cystéine catalytique se trouve dans la boucle entre les brins  $\beta$  9 et  $\beta$  10. Cette boucle est impliquée dans l'interaction entre les deux lobes. La thréonine 820 et l'Asn 822 forment des liaisons hydrogène avec respectivement la Gln 553 et la lys 549 du lobe-N tandis que la phe 821 participe à une interaction de type WanderVals avec la Gly 546 du lobe-N. La cystéine catalytique est partiellement exposée au solvant.

La E2 Ubch7 interagit avec le lobe-N de Ube3a. Ce sont les boucles B1 et B2 ainsi que la partie amino-terminale de l'hélice 1 de Ubch7 qui interagissent avec un sillon formé des hélices H7 H8 et des brins S5 et S6 du lobe amino-terminal du domaine HECT. Un élément clef de cette interaction semble être la phénylalanine 63 de Ubch7 qui se positionne dans ce sillon hydrophobe (L642, L639, F690, I655, L659, L673). Cette phénylalanine est conservée dans le E2 spécifique aux E3 à domaines HECT [84].

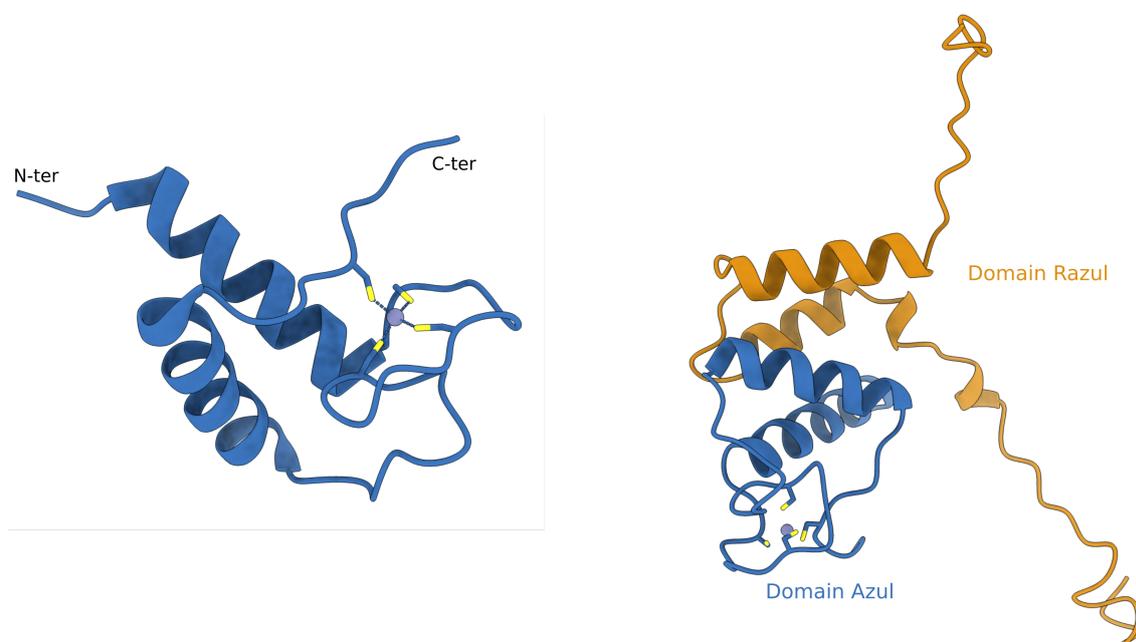
Différents substrats de Ube3a sont identifiés à ce jour. C'est notamment le cas des protéines RAD23A et B [154], alpha-Synuclein [155] [156], Ring1B[157] et BK channels [158].

### 2.1.1.2 Le domaine AZUL

Ube3a possède un domaine de liaison au zinc en position amino-terminale appelé AZUL ("Amino-terminal Zn-finger of Ube3a Ligase") [159]. Ce domaine adopte un repliement de

type hélice-boucle-hélice avec un atome de zinc coordonné par 4 cystéines positionnées de la manière suivante : Cys-X4-Cys-X4-Cys-X28-Cys.

Ce domaine est capable de lier le protéasome. Cette interaction avec le protéasome se fait via la protéine Rpn10 aussi appelée Psmd4 ("26S proteasome non-ATPase regulatory subunit 4"). Cette interaction se fait entre le domaine AZUL de Ube3a et le domaine RAZUL de Rpn10 [160]. (figure 22 page 37). Les structures de AZUL seul et du domaine AZUL en complexe avec le domaine RAZUL de Rpn10 sont connus. Cette interaction se fait via un empilement des deux hélices du domaine AZUL de Ube3a et des deux hélices du domaine RAZUL de Rpn10. L'hélice 1 du domaine RAZUL interagit avec les deux hélices du domaine AZUL. Cette interaction est principalement médiée par des contacts hydrophobes [161].



**FIGURE 22** – À gauche est représentée la structure du domaine AZUL de Ube3a résolue par RMN (PDB : 2KR1). Les cystéines interagissant avec l'atome de zinc se trouvent dans les boucles entre les hélices [162]. À droite, la structure du domaine AZUL de Ube3a en complexe avec le domaine RAZUL de la protéine Rpn10 (PDB : 6U19)[160].

Le domaine AZUL semble être capable de se lier à la protéine UBQLN1/2 (ubiquiline) via son domaine UBA ("ubiquitine associated domain") [163][164] [165]. Les ubiquilines ont un domaine Ubl et un domaine UBA. Elles sont ainsi capables d'interagir avec des protéines ubiquitinées via son domaine Uba et avec le protéasome via son domaine Ubl [166].

### 2.1.1.3 Interaction avec RAD23

D'autres protéines sont connues pour leur capacité à interagir avec Ube3a. RAD23 A et B sont des partenaires bien identifiés de Ube3a[165]. RAD23A serait un substrat de Ube3a[154]. Il est intéressant de noter que RAD23 est capable de se lier par ce même domaine Ubl à la protéine Rpn10 du protéasome [167]. RAD23 A et B sont des protéines intermédiaires entre le protéasome et des protéines ubiquitinées permettant ainsi leur colocalisation et leur dégradation par le protéasome [167].

### 2.1.2 Rôles de Ube3a

Plusieurs fonctions, liées ou non à son activité ubiquitine-ligase, ont été identifiées. Il a été démontré que cette protéine joue un rôle dans la régulation du cycle circadien via la régulation du facteur de transcription de BMAL1 [168]. Ube3a joue aussi un rôle dans la dégradation des agrésomes. Pour ce faire, Ube3a est capable de lier la protéine chaperonne Hsp70 elle-même liée à une protéine mal repliée, ce qui mènera à l'ubiquitination de cette dernière [169]. De plus, Ube3a activerait la transcription de récepteur de la progestérone [170]. Enfin, comme nous allons le voir par la suite, Ube3a joue un rôle capital dans le neurodéveloppement [171] (Section 3.1 page 48).

### 2.1.3 Structure globale de Ube3a

La protéine Ube3a est prédite comme entièrement repliée avec deux boucles, une boucle se trouverait entre le domaine AZUL et le reste de la protéine, une autre se trouverait entre deux blocs prédits comme repliés. La structure globale de la protéine Ube3a est discutée dans la seconde partie des résultats (Section 2 page 86).

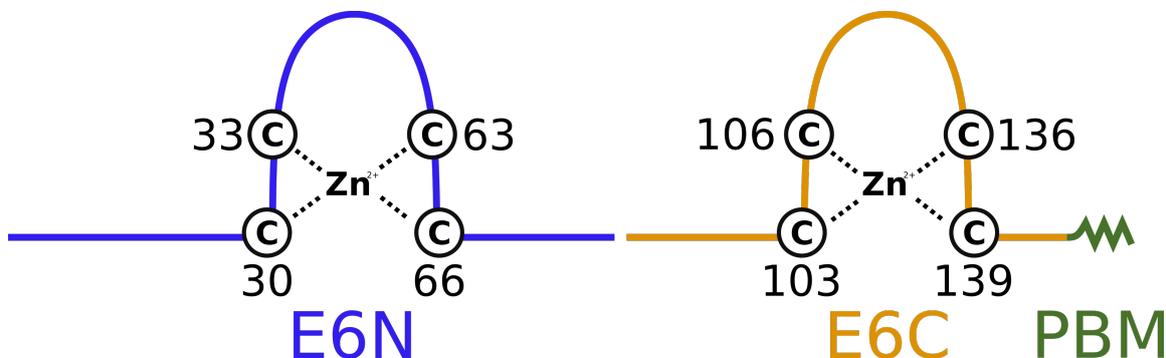
Une étude de 2014 tend à montrer que la trimérisation de Ube3a serait nécessaire à son activité ubiquitine-ligase [172]. Cette étude se base notamment sur la structure du domaine HECT de Ube3a résolue en 1999. Cette forme trimérique était alors décrite comme un artefact dû au "packing" cristallographique. Ce point est discuté dans la section résultats (Section 2 page 86).

### 2.1.4 Ube3a et les cancers liés aux papillomavirus

La protéine Ube3a est connue pour son rôle dans le développement de cancer en présence de papillomavirus (PV) [173]. Le papillomavirus est un virus à ADN qui est retrouvé dans des cancers tels que le cancer du col de l'utérus, mais aussi les cancers anogénitaux, de l'oropharynx et du larynx [174]. Les papillomavirus produisent une protéine, appelée E6, qui induit la formation de cancer.

E6 de HPV est une protéine de 158 acides aminés, cette protéine contient 2 domaines de

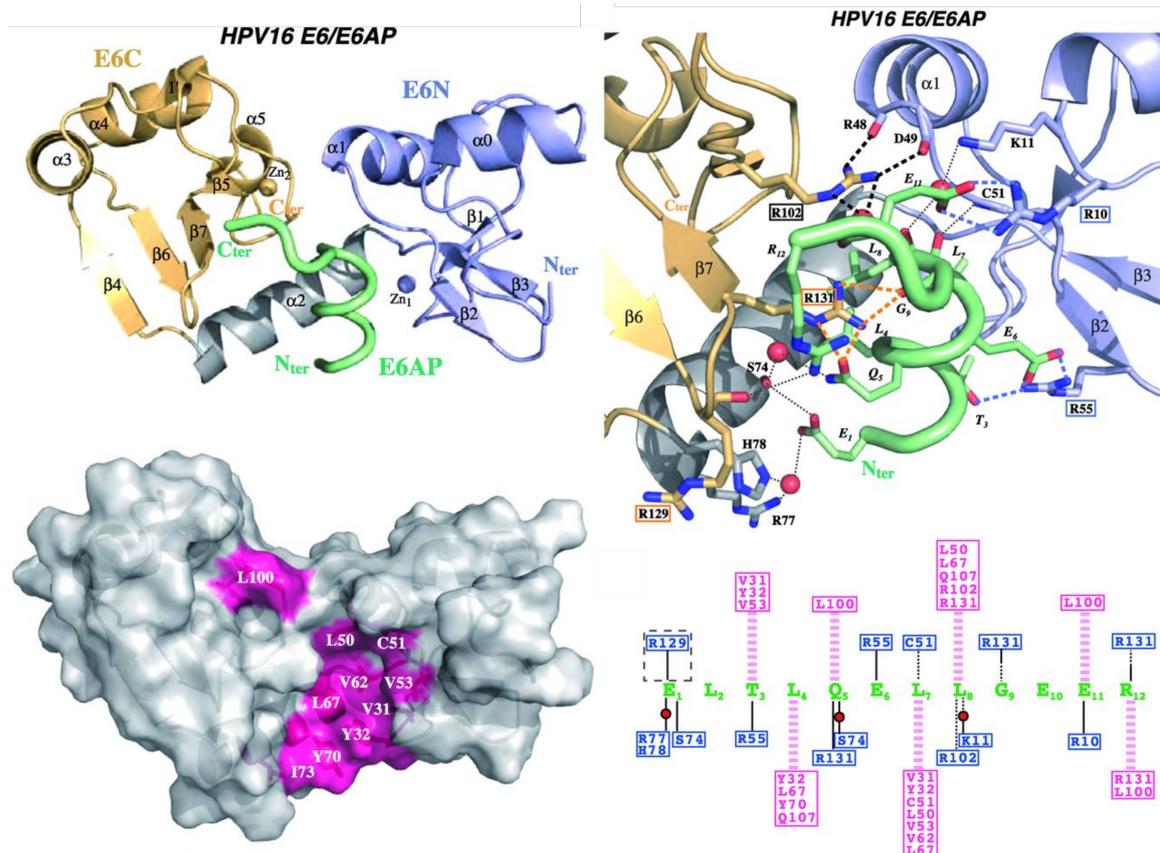
liaison au zinc appelé E6N (en amino-terminal) et E6C (en carboxy-terminal). La protéine E6 de certain HPV (comme HPV16) possède aussi un PBM ("PDZ Binding Motif") qui lui permet d'interagir avec le domaine PDZ d'autres protéines [175] (Figure 23, page 39).



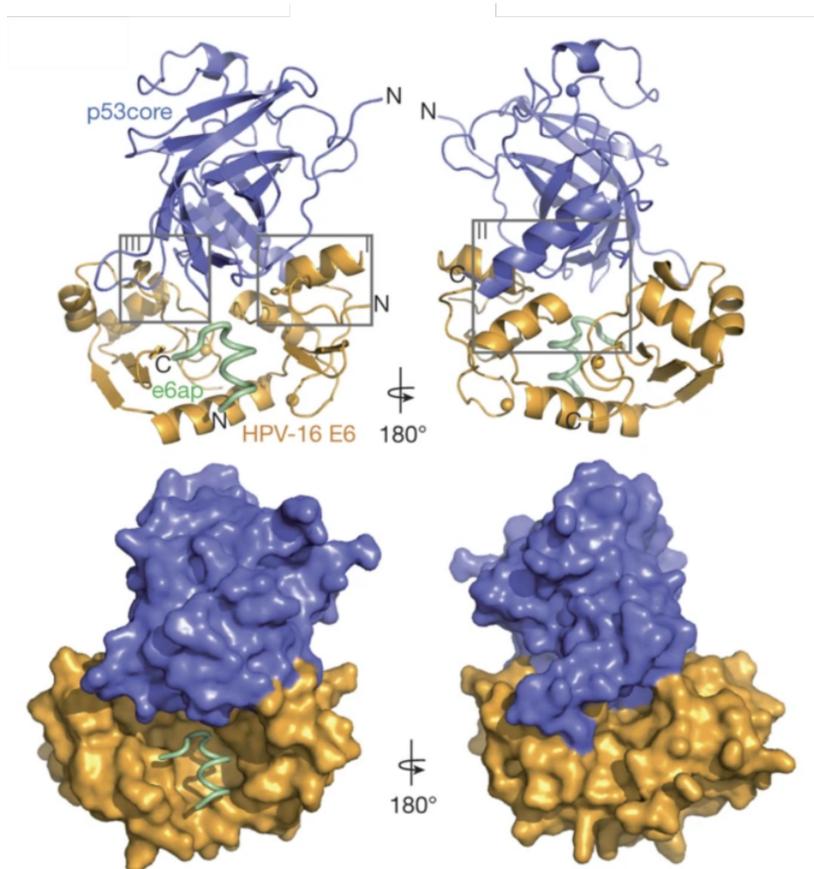
**FIGURE 23** – Organisation de l'oncoprotéine E6 de HPV16. L'oncoprotéine E6 est composée de 2 domaines de liaison au zinc appelés E6N représenté en bleu et E6C représenté en orange. Le motif de liaison au PDZ, positionné à l'extrémité carboxy-terminale, est coloré en vert. Figure adaptée de [176]

L'oncoprotéine E6 est capable d'interagir avec la protéine Ube3A en liant une séquence : le motif LxxLL. De plus, la protéine E6 interagit avec la protéine P53 et détourne l'activité ubiquitine-ligase de Ube3a pour entraîner l'ubiquitination et la dégradation de la protéine suppresseuse de tumeur P53 [177]. P53 est une protéine de 53 kDa, dite suppresseuse de tumeur. P53 est un acteur important dans la réponse cellulaire à différents stress. [178].

La résolution de la structure de l'interaction entre la protéine E6 et un peptide de Ube3A a permis de comprendre ce mécanisme. La liaison de E6 à Ube3a se fait via le motif LxxLL de Ube3a. Ce motif se trouve dans une région non repliée de Ube3a. La structure de la protéine E6 en complexe avec le motif LxxLL de Ube3a a été résolue en 2013 [179]. La structure du complexe Ube3a-LxxLL/16E6/P53 a été résolue en 2016 [180]. Cependant, le motif LxxLL ne semble pas être l'unique élément de Ube3a participant à l'interaction avec E6 [181]. Ce point sera discuté par la suite (Section 2, page 86).



**FIGURE 24** – Interaction entre l'oncoprotéine E6 du papillomavirus humain 16 (HPV 16) et le motif LxxLL de Ube3a (ici appelée E6AP). En haut est représentée la structure globale de l'interaction. Le domaine de liaison au zinc, E6C, est en jaune Le second domaine de liaison au zinc, E6N, est en violet. L'hélice liant ces deux domaines est représentée en gris. Le motif LxxLL de Ube3a est représenté en vert. Ce peptide de Ube3a adopte la forme d'une hélice dans cette interaction. En bas à gauche la surface de la protéine E6 de HPV est représentée, les acides aminés hydrophobes impliqués dans l'interaction sont en rose. En bas à droite, l'interaction est détaillée. Les interactions hydrophobes sont représentées par des pointillés roses, les liaisons hydrogènes par des lignes noires. Figure adaptée de [179]



**FIGURE 25** – Interaction entre l'oncoprotéine E6 du papillomavirus humain 16 (HPV 16), le motif LxxLL de Ube3a et le core domaine de P53. La protéine E6 de HPV16 est représentée en jaune, le core domaine de P53 en violet et le peptide de Ube3a (ici appelé E6AP) contenant le motif LxxLL en vert. Figure adaptée de [180]

En plus d'entraîner la dégradation de P53 par Ube3a, E6 entraîne une augmentation de l'activité ubiquitine-ligase de Ube3a [182]. Ube3a étant capable de s'ubiquitiner elle-même (en trans), E6 entraîne une augmentation de la dégradation de Ube3a [183] [184].

## 2.2 L'ubiquitine-ligase HERC2

Comme Ube3a, la protéine HERC2 est une ubiquitine-ligase à domaine HECT de la famille des HERC. HERC2 est longue de 4834 acides aminés. Elle est souvent qualifiée avec HERC1 de HERC géante.

Le gène HERC2 a tout d'abord été identifié comme le gène responsable du syndrome rjs ("Runty, Jerky and Sterile") chez les souris. Ce syndrome entraîne une petite taille, une espérance de vie réduite, la stérilité ainsi que des troubles neuromusculaires [185].

### 2.2.1 Différents rôles identifiés de HERC2

Les gènes HERC2 ainsi que OCA2 sont tous deux reliés à la pigmentation des yeux. En effet, beaucoup d'étude sur les gènes codant pour ces protéines montrent leur implication dans la pigmentation de l'iris, de la peau ainsi que des cheveux [186] [187] [188].

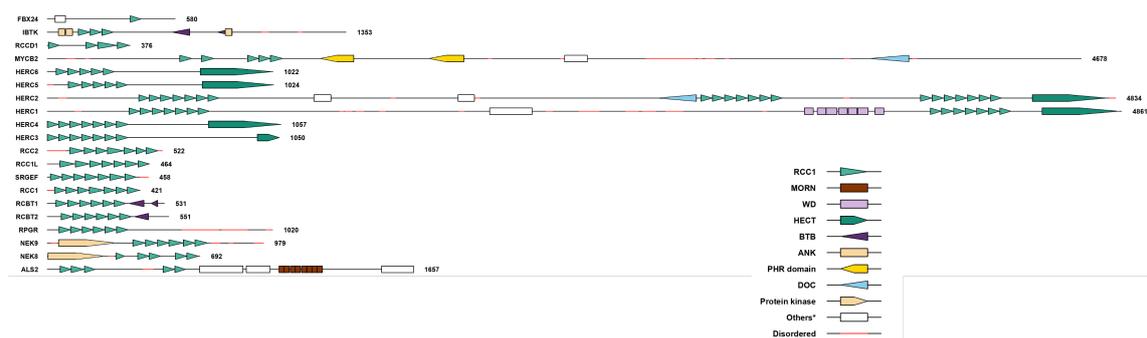
La protéine HERC2 est impliquée dans de multiples mécanismes de réponse aux dommages de l'ADN. HERC2 serait un facteur d'assemblage du complexe formé de l'E3 à domaine RING RNF8 et l'E2 Ubc13 [189]. RNF8 étant impliquée dans la régulation de voie de réparation de l'ADN. La protéine HERC2 joue un rôle dans la régulation de P53, dont nous avons parlé précédemment [190]. La protéine HERC2 interagirait avec P53 et ainsi régulerait son activité en participant à la tétramérisation de P53 [178]. De plus, la protéine HERC2 est capable d'interagir avec la protéine USP16, une désubiquitinase connue pour son rôle dans la désubiquitination de l'histone H2A. Cette régulation du taux d'ubiquitination de cette histone semble jouer un rôle dans la réponse aux dommages de l'ADN [191].

HERC2 est aussi impliquée dans la morphologie du centrosome. Les deux ubiquitine-ligases NEURL4 et HERC2 sont nécessaires à la bonne morphologie de celui-ci grâce à leur interaction avec la protéine centrosomale CP110 [192].

HERC2 serait impliquée dans la maladie de parkinson [193], la maladie de Crohn [194] et le diabète de type 1 [195]. Enfin, HERC2 est aussi impliquée dans le neurodéveloppement. Cette protéine régulerait la formation de boutons synaptiques dans les cellules de Purkinje [196]. Comme nous allons le voir, HERC2 est maintenant bien connue dans le cadre du syndrome affectant le neurodéveloppement appelé Angelman-like.

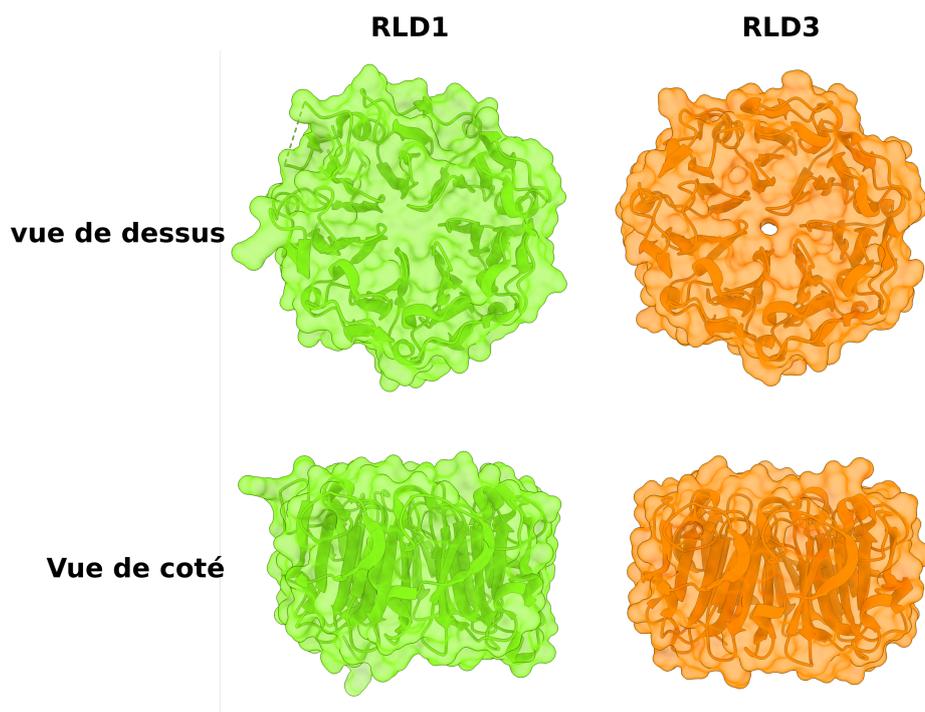
### 2.2.2 Les domaines RLDs

Les E3 de la famille des HERCs possèdent toutes au moins un domaine RLD (domaine RCC1-like). HERC2 possède 3 domaines RLDs appelés RLD1, 2 et 3. Les domaines RLDs sont retrouvés dans de nombreuses protéines et sont caractérisés par un repliement de " $\beta$ -propeller". Ce domaine a une forte similarité de séquence avec la protéine RCC1. Les domaines RLDs sont constitués de 7 répétitions de plus de 60 acides aminés. Chaque répétition forme un feuillet  $\beta$  et est appelé "pale". Les 7 pales forment un domaine ( $\beta$ -propeller) de forme torique. À ce jour, 19 protéines avec au moins un domaine RLD ont été identifiées. (Figure 26 page 43).



**FIGURE 26** – Protéines contenant des domaines RLDs. Chaque pale des domaines RLDs est représentée par un triangle vert. Notez que toutes les répétitions de pales de RLD ne sont pas identifiées sur Uniprot. Figure obtenue avec le logiciel ProFeatMap.

Les structures des domaines RLD1 et RLD3 ont été résolues par cristallographie (Figure 27 page 43). Ces structures ont été résolues par un consortium et n'ont pas plus été étudiées. Ces domaines prennent la forme classique des "β-propeller" à 7 pales.

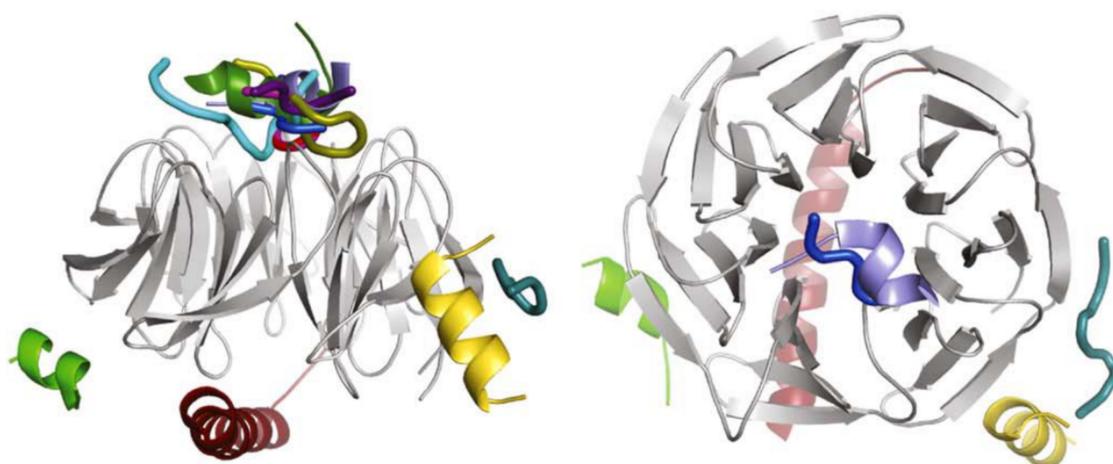


**FIGURE 27** – Structures des domaines RLD1 (PDB : 4L1M) et RLD3 (PDB : 3KCI) de HERC2. Ces deux domaines sont tous deux des "β-propeller" à 7 pales.

Les domaines de type "β-propeller" sont composés de feuillets appelés pales de 4 brins β antiparallèles. Les brins sont nommés A, B, C et D de l'intérieur vers l'extérieur. Par convention, le haut du domaine est le côté contenant les boucles reliant les brins D et A et les brins B et C [197].

Tous les "β-propeller" ne sont pas composés de 7 pales. Certains contiennent 4, 5, 6, 7, 8 ou 10 pales. Différentes familles de "β-propeller" existent en fonction de leur séquence et des motifs répétés qu'ils possèdent.

Les domaines adoptant des repliements de "β-propeller" sont connus pour leur implication dans des interactions domaine/motif. Selon les "β-propeller", l'interaction avec une autre protéine peut se faire sur le dessus, sur le dessous et sur le côté du "β-propeller". De nombreuses structures, notamment de domaine de type WD40, ont été résolues en complexe avec des peptides (Figure 29 page 45).



**FIGURE 28** – Interactions entre le domaine WD40 et différents peptides. Vue latérale et vue de haut du "β-propeller" avec 13 différents complexes superposés (seul un domaine WD40 est représenté). La majorité des peptides se regroupent autour du canal central supérieur (3IIW [198], violet ; 2QXV [199], rouge ; 2CE8 [200], bleu clair ; 2CE9 [200], bleu marine ; 2H13 [201], rouge ; 3EU7 [202], vert clair ; 3C9C [203], jaune, 1XHM [204], vert ; 2QNS (retracted), blue sarcelle ; 2OVQ [205], vert olive ; 1NEX [206], bleu ; 1P22 [207], cyan ; 3EMH [203], magenta. Figure de [208]

Les structures d'autres domaines de la protéine HERC2 ont été résolues. C'est notamment le cas pour celle de son domaine DOC (PDB : 7RGW) et son domaine Cytochrome-b5-like (PDB : 2KEO). Les structures du domaine RLD2, du doigt de zinc et du domaine MIB de HERC2 ne sont pas connus.

### 2.3 Interaction entre Ube3a et HERC2

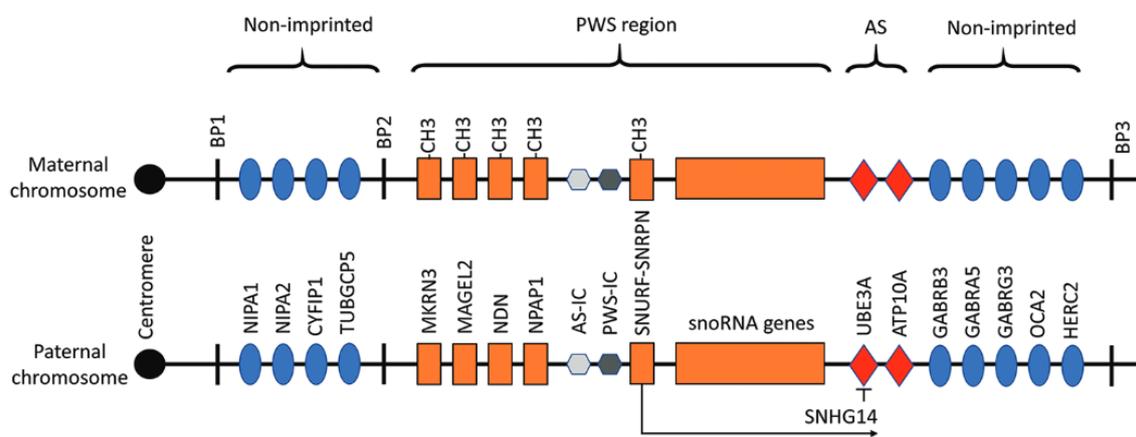
Il a été montré que les 2 ubiquitines ligases Ube3a et HERC2 sont capables d'interagir entre elles [209]. Cette interaction implique le domaine RLD2 de la protéine HERC2 et une région située entre le 150ème et le 200ème acide aminé de Ube3a [210].

De plus, il a été montré que la liaison au domaine RLD2 entraîne la stimulation de l'activité

ubiquitine-ligase de Ube3a [210]. Ici le test d'activité de Ube3a est effectué sur Ube3a elle-même (auto-ubiquitination). Cette étude ne permet pas de conclure sur le mécanisme de la hausse de l'activité ubiquitine-ligase de Ube3a par RLD2, ni sur l'existence de ce mécanisme *in vivo*.

La compréhension de l'interaction entre ces deux ubiquitine-ligases fait l'objet de la première partie des résultats (Section 1, page 1).

## 2.4 La région chromosomique 15q11-q13



**FIGURE 29** – Représentation de la région chromosomique 15q11-q13. Cette région chromosomique commence par 4 gènes non régulés par une empreinte parentale (ovales bleues). La région suivante est régulée par le centre d'empreinte IC-PW. Cette région comporte 5 gènes fonctionnels exprimés par l'allèle paternel (MKRN3, MAGEL2, NDN, NPAP1, SNURF-SNRPN) et des gènes codant pour snoRNA exprimés par l'allèle paternel (carrés oranges). Ces 5 gènes sont méthylés et donc non fonctionnels pour la copie maternelle. Les centres d'empreinte IC-PW et IC-AC sont inclus dans cette région. La région sous contrôle du IC-AS comporte 2 gènes exprimés par la mère, UBE3A et ATP10A (losanges rouges). Enfin, cette région contient 5 gènes qui ne sont pas sous contrôle d'une empreinte parentale : GABRB3, GABRA5, GABRG3, OCA2 et HERC2 (ovales bleues). BP : Point de rupture. Figure de [211]

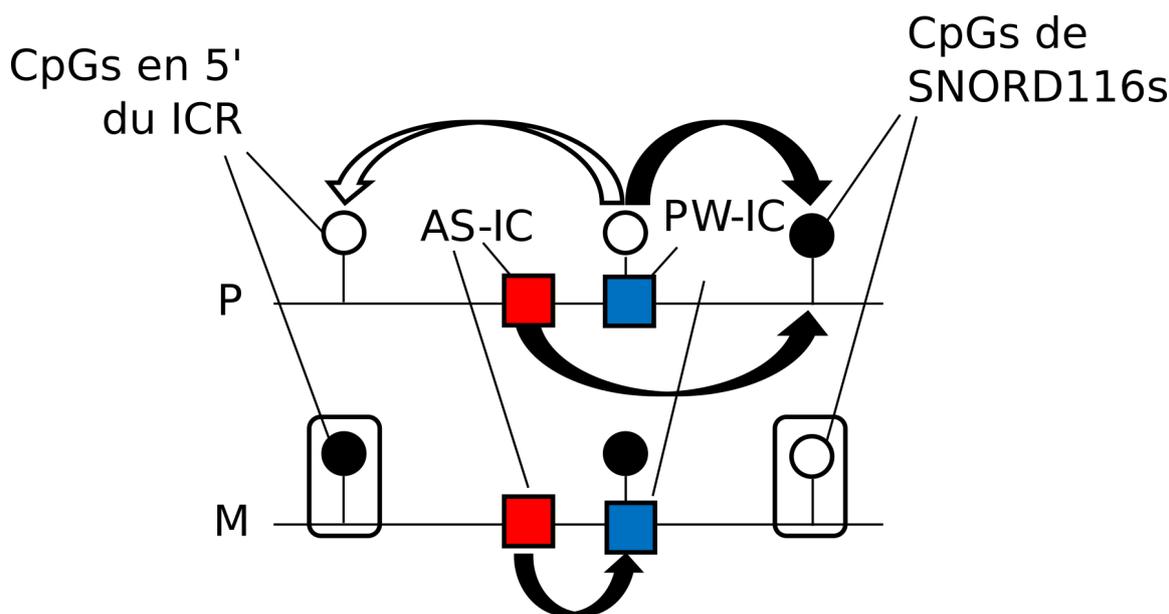
La région 15q11.2–q13.3 est une région de 6 Mb qui est impliquée dans des syndromes du neurodéveloppement. Différentes altérations de cette région chromosomique peuvent causer différents syndromes. Cette région contient les gènes codant pour les protéines Ube3a et HERC2 présentées précédemment (section 2 page 35).

Cette région a la particularité d'être soumise à une régulation très spécifique appelée empreinte parentale. Ceci signifie que les gènes maternels et paternels ne sont pas exprimés de la même manière. Cette empreinte parentale est contrôlée par 2 centres d'empreinte (IC pour imprinting center) appelés PW-IC (pour centre d'empreinte du syndrome prader Willi) et AS-IC (pour centre d'empreinte du syndrome d'Angelman) [212] [213]. Ces 2 IC

sont regroupés dans une même région appelée IPR ("imprinting control region"). Le mécanisme de l'empreinte parentale de cette région n'est pas encore entièrement compris.

L'IPR a été définie comme étant la plus petite région chevauchante identifiée dans les cas de PW et AS ayant des micro-délétions [214]. Ce centre de l'empreinte parental est dit bipartite, étant régulé par le AS-IC et le PW-IC [215]. Le PW-IC est une région de 4.1 Kb qui s'étend sur le promoteur et le premier exon du gène SNURF/SNRPN. Ce PW-IC inclut une zone de méthylation différenciée appelée iDMR ("imprinted differentially methylated region"). Cette région est méthylée pour l'allèle maternel et non méthylée pour l'allèle paternel. Le AS-IC est une région de 880 bp qui se trouve à environ 35 Kb du PW-IC. À la différence du PW-IC le AS-IC ne comporte pas de zone de iDMR. Ces 2 ICs jouent un rôle capital dans la régulation épigénétique et donc dans la régulation des gènes de ce locus [216]. D'autres zones de cette région chromosomique ont été identifiées comme étant des iDMRs. C'est notamment le cas d'une région proche du gène SNURF/SNRPN et de sites CpG dans le cluster SNORD116. Ces 2 régions ne sont pas méthylées de la même manière pour l'allèle maternel et l'allèle paternel [217]. Cette méthylation différenciée des iDMRs n'est pas encore totalement comprise.

Une étude récente met en évidence le rôle des 2 ICs dans la méthylation de ces iDMRs (Figure 30 page 47).



**FIGURE 30** – Modèle hypothétique de la régulation de la méthylation de l'ADN dans la région PWS/AS par AS-IC et PWS-IC. Les cercles indiquent les sites CpGs présentant des schémas de méthylation différents en fonction de l'origine parentale. Les cercles pleins représentent les sites CpGs méthylés, et les cercles vides les sites CpGs non méthylés. P désigne l'allèle hérité paternellement, et M l'allèle hérité maternellement. Sur l'allèle P, les CpGs dans le PWS-IC et le promoteur des gènes exprimés par la mère dans la région 5 en amont de l'ICR sont non méthylés, et les CpGs situés de manière éparse dans le cluster SNORD116s sont méthylés. On pense que le PWS-IC non méthylé sur l'allèle P conduit à des CpGs avec un schéma de méthylation paternelle au niveau des CpGs à la fois dans la région en amont (non méthylés, représentés par des flèches blanches) et dans le cluster SNORD116s (méthylés, représentés par des flèches noires). L'AS-IC sur l'allèle P a méthylé les CpGs dans le cluster SNORD116s. Sur l'allèle M, le PWS-IC et les CpG dans les iDMR des gènes imprimés en amont sont méthylés, et les CpG dans le cluster SNORD116s sont non méthylés. L'AS-IC de l'allèle M a méthylé le PWS-IC. Figure adaptée de [218]

Cette empreinte parentale entraîne l'expression d'un ARN non codant par l'allèle paternel. Dans les neurones, cet ARN non codant s'étend jusqu'à se superposer avec la région codant pour Ube3a comme un transcrit antisens (Ube3a-ATS) ce qui entraîne le "silencing" du gène Ube3a paternel [219]. Ce "silencing" n'est donc pas dû à une méthylation du gène Ube3a. Ce mécanisme n'est, lui non plus, pas encore totalement compris.

Ainsi, dans les neurones, seule le gène Ube3a maternel est exprimé. Si celui-ci est non fonctionnel, est délété ou code pour une protéine Ube3a non fonctionnelle/incomplète alors, aucun gène Ube3a n'est fonctionnel dans les neurones.

## 3 | Syndromes impliquant la région 15q11–q13

### 3.1 Le syndrome d'Angelman

#### 3.1.1 Historique

Le syndrome dit d'Angelman a été identifié pour la première fois en 1965 par le médecin pédiatre Anglais : Harry Angelman. Il donne alors le nom de "puppet children" au syndrome porté par ces enfants. Le syndrome est alors caractérisé par une liste de 7 symptômes, particularités physiques et comportements [220]. Par la suite, ce syndrome portera le nom de "Happy Puppet" [221] puis syndrome d'Angelman. La région chromosomique (15q11-q13) a été identifiée comme responsable du syndrome d'Angelman en 1987 [222], puis plus précisément le gène Ube3a codant pour l'ubiquitine ligase Ube3A en 1997 [223].

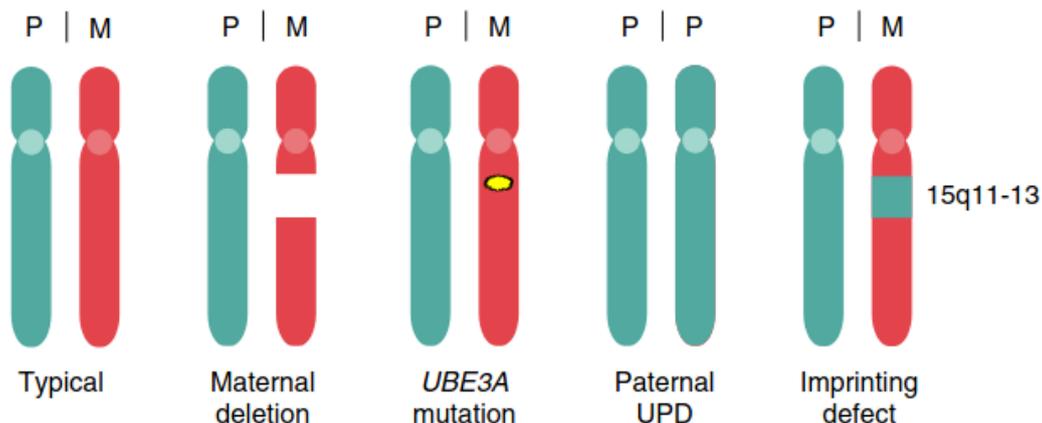
#### 3.1.2 Les différents types du syndrome et les symptômes associés

Le syndrome d'Angelman touche environ une personne sur 15000 et est caractérisé par différents symptômes. Ce syndrome implique une déficience intellectuelle. Celle-ci est associée à des difficultés d'apprentissage et notamment des troubles du langage pouvant aller jusqu'à l'absence de langage verbal. Les troubles moteurs (ataxie), une tendance accrue au rire et au sourire et un comportement décrit comme jovial sont des autres symptômes caractéristiques du syndrome d'Angelman. De plus, peuvent s'ajouter des troubles du sommeil, une microcéphalie, des crises d'épilepsie, des crises atoniques (relâchement musculaire total) et myocloniques (tremblements musculaires) [224].

En raison de l'empreinte parentale qui entraîne uniquement l'expression du gène Ube3a maternel dans les neurones, il suffit que seul l'allèle maternel soit affecté pour entraîner ce syndrome. Ceci explique pourquoi ce syndrome est plus fréquent que d'autre maladie génétique.

Les études génétiques sur des patientes atteintes du syndrome d'Angelman ont révélé

cinq mécanismes différents à l'origine de ce syndrome. La délétion *de novo* du chromosome 15q11-q13 maternel représente entre 70 et 80 pour cent des cas du syndrome. Les mutations du gène Ube3a représente 10 à 20 pour cent des cas. 3 à 5 pour cent des cas sont dus à une disomie uniparentale paternelle du chromosome 15. Enfin, 3 à 5 pour cent des cas sont expliqués par un défaut du mécanisme de l'empreinte parentale affectant l'expression du gène Ube3a maternel [223].



**FIGURE 31** – Causes génétiques du syndrome d'Angelman. Représentation des chromosomes paternels (P) et maternels (M) chez les individus atteints du syndrome d'Angelman par rapport aux individus neurotypiques. Notez que dans tous les cas, le gène Ube3a hérité par la mère est affecté, soit par une délétion du locus 15q11-13 maternel, soit par une mutation intragénique dans le gène Ube3a, soit par son absence due à l'héritage de deux copies du chromosome paternel (disomie uniparentale [UPD]), soit par des défauts d'empreinte qui répriment l'expression de l'Ube3a maternelle. Figure de [225]

Les différents types du syndrome d'Angelman causent des symptômes qui peuvent varier dans leur intensité. La délétion semble entraîner des symptômes plus prononcés [226]. A contrario, le syndrome dû à une monodisomie paternelle semble entraîner des symptômes moins prononcés [227] [228] [229] [230].

Avant l'identification des causes du syndrome d'Angelman, il regroupait d'autres syndromes ayant les mêmes symptômes. Par la suite, la meilleure compréhension du syndrome a amené à la mise en place de dépistage en cas de suspicions permettant l'identification ou non du dysfonctionnement de l'allèle maternel de Ube3a.

### 3.1.3 Pistes de traitements

À ce jour, seuls des traitements non spécifiques au syndrome d'Angelman existent. Ce sont des traitements des symptômes qui s'adaptent aux besoins du/de la patient-e. Des antiépileptiques peuvent être prescrits pour les crises d'épilepsie. L'intervention d'un-e psy-

chomotricien·ne peut permettre d'améliorer la qualité de vie dans le cas des troubles du comportement comme l'hyperactivité et les troubles de l'attention. Il existe aussi des traitements du trouble du sommeil, comme l'utilisation de mélatonine.

De plus, différentes stratégies pour des traitements spécifiques au syndrome d'Angelman existent. Chacune de ces stratégies de traitement visent à rétablir l'expression d'un gène Ube3a dans les neurones, ou à rétablir les fonctions affectées par l'absence d'expression de ce gène dans les neurones.

Il est important de noter que le syndrome d'Angelman est un syndrome du neurodéveloppement, cela implique la nécessité d'un traitement au plus tôt pour espérer récupérer le plus possible de fonctions affectées par ce syndrome [231].

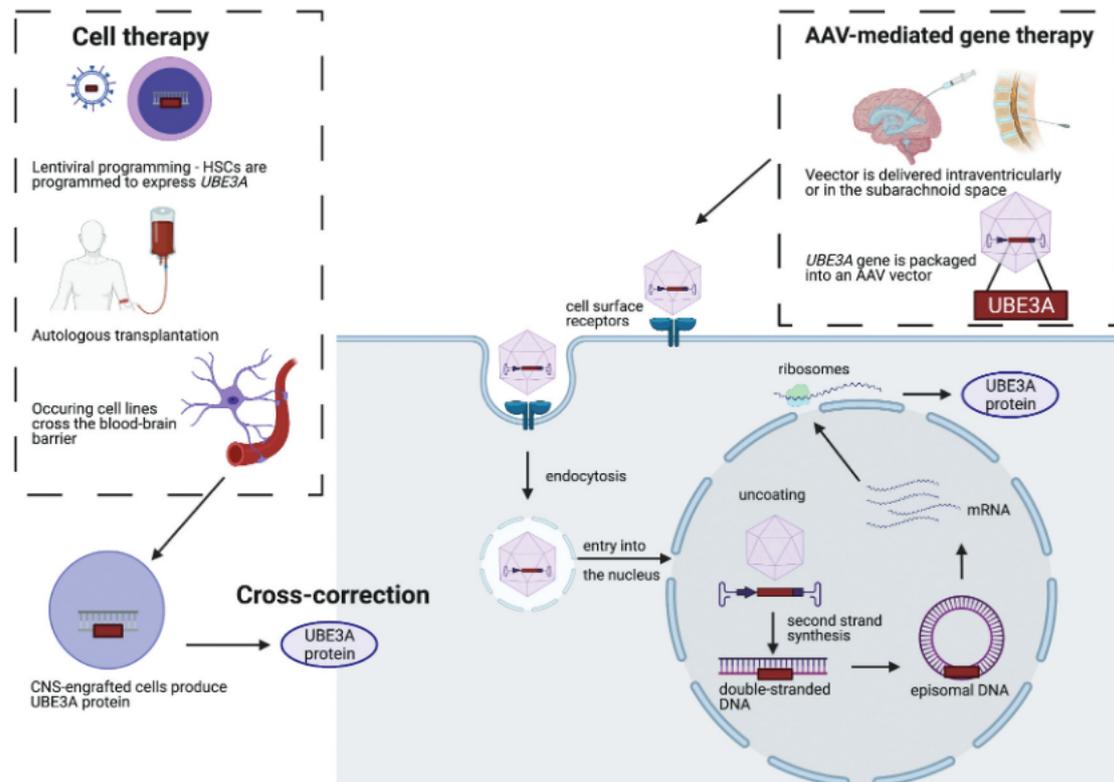
#### 3.1.3.1 Thérapies géniques, cellulaires et enzymatiques

Ces méthodes ont pour but l'apport d'un gène fonctionnel ou de protéines fonctionnelles dans les neurones.

La première méthode est la thérapie génique, Le but est d'apporter un gène additionnel et fonctionnel de Ube3a dans les cellules nerveuses. Pour cela, cette stratégie consiste en l'utilisation d'un adénovirus modifié (AAV) qui va permettre, après son entrée dans les cellules et les noyaux des cellules, la production de ce gène [232].

Une autre stratégie est la thérapie cellulaire. Cette méthode nécessite d'utiliser les HSC (cellules souches hématopoïétiques) du·de la patient·e et d'y insérer une copie fonctionnelle du gène Ube3a au moyen d'un lentivirus. Ces HSC modifiées sont ensuite injectées en IV (intraveineux) chez le·la patient·e. Les cellules immunitaires, qui sont capables de passer la barrière hémato-encéphalique, issues de ces HSC modifiées, vont pouvoir sécréter la protéine Ube3a dans le SNC (Système Nerveux Central) [233].

Enfin, la méthode du remplacement protéique (ERT pour enzyme replacement therapies) consiste en l'injection de protéine Ube3a purifiée dans les neurones [234].



**FIGURE 32** – Thérapie génique *in vivo* et thérapie cellulaire *ex vivo*. Dans le cas de la thérapie génique médiée par les AAV, les vecteurs AAV transformés pour porter une copie du gène à remplacer (*Ube3a*), sont reconnus par les récepteurs de la surface cellulaire des cellules cibles et sont internalisés par endocytose. Ils sont ensuite transportés à l'intérieur des cellules dans des vésicules endosomales. Après être entrés, ils libèrent le gène *Ube3a*. Le génome est généralement stabilisé sous forme d'épisomes circulaires, qui peuvent ensuite être transcrits en ARNm et traduits en protéines. Le génome des AAV peut également s'intégrer dans l'hôte à une faible fréquence. En thérapie cellulaire, les CSH du patient sont isolées. Après une transplantation autologue, les CSH portant la copie normale du gène se différencient en cellules capables de traverser la barrière hémato-encéphalique. Une fois dans le SNC, les cellules produisent la protéine *Ube3a*. Figure de [235]

### 3.1.3.2 Activation de l'allèle paternel

Comme seul l'allèle maternel de *Ube3a* est exprimé dans les neurones, une stratégie consiste en l'empêchement du mécanisme qui entraîne la non-production de l'allèle paternel.

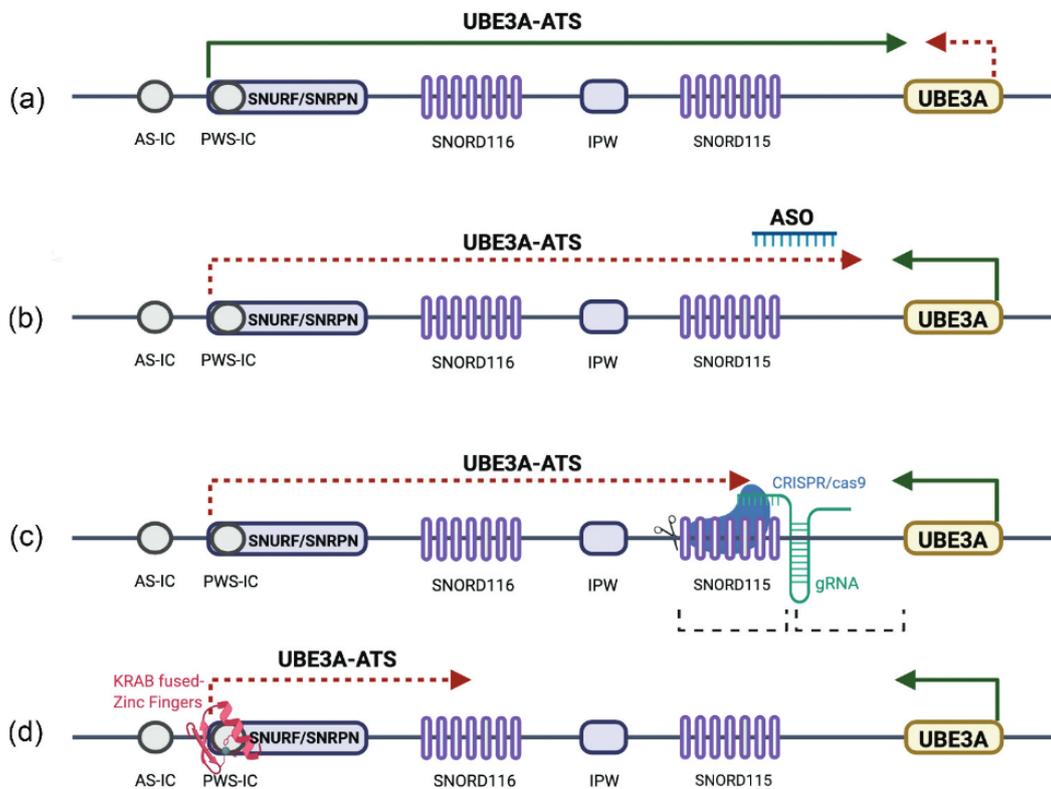
Une stratégie pour activer l'allèle paternel est l'utilisation d'un oligonucléotide antisens (AOS pour Antisense oligonucleotides). Il s'agit de l'utilisation d'un AOS complémentaire à la partie distale de *Ube3a*-ATS [236].

L'utilisation de la stratégie CRISPR/cas9 est une autre piste prometteuse pour activer le

gène Ube3a paternel. Ici cette technique est utilisée pour muter la région Ube3a-ATS et ainsi bloquer son expression pour permettre l'expression du gène Ube3a paternel [237].

La méthode du facteur de transcription artificiel (ATF) utilise un système composé d'une région se liant à l'ADN et d'un effecteur permettant de réguler l'expression du gène cible. Dans le cas du syndrome d'Angelman les ATF à doigt de zinc constitue une piste étudiée, car il sont capables de traverser la barrière hémato-encéphalique. Ici, l'ATF est composé d'un doigt de zinc fusionné au répresseur de transcription KRAB (Krüppel associated box) et se lie au promoteur de Snurf/Dnrpn pour empêcher le blocage par l'Ube3a-ATS [238].

L'utilisation d'inhibiteurs de topoisomerase est aussi une piste pour activer l'allèle paternel de Ube3a. En effet, des "screenings" de molécules permettant d'activer l'allèle paternel ont permis d'identifier des inhibiteurs de protéase. C'est notamment le cas de la molécule Topotecan. Cependant, les tests effectués sur ce type de molécules entraînent la dérégulation de l'expression d'autres gènes [239].



**FIGURE 33** – Stratégies thérapeutiques pour l'activation du gène Ube3a paternel. A) La copie paternelle normale du gène Ube3a sur le chromosome 15 est "réduite au silence" en raison de l'empreinte parentale. La transcription de l'ARN non codant, l'Ube3a-ATS est régulée à partir de la zone située au niveau ou en amont du PWS-IC. L'Ube3a-ATS traverse le Snurf/Snrpn, Snord116, Ipw, Snord115 et la région codante de l'Ube3a antisens. Sur le chromosome paternel 15, la transcription de l'Ube3a-ATS entraîne le "silencing" du gène Ube3a. B) Les ASO ("Antisense Oligonucleotide" complémentaires à la partie distale de l'Ube3a-ATS peuvent conduire par l'hybridation ASO/ARN à la terminaison prématurée de la transcription de l'Ube3a-ATS. En l'absence de transcription de l'Ube3a-ATS, l'allèle paternel est exprimé. C) La mutagenèse, permise par CRISPR/Cas9, dans les zones génomiques qui codent pour l'Ube3a-ATS pourrait conduire à l'activation de l'allèle paternel par l'arrêt précoce de la transcription de l'Ube3a-ATS. D) Les protéines à doigt de zinc fusionnées à KRAB qui se lient au promoteur de l'Ube3a-ATS peuvent empêcher la transcription de l'Ube3a-ATS et activer l'allèle paternel. Cette approche est encore en développement préclinique. Figure de [235]

### 3.1.3.3 Les traitements "Downstream"

La dernière méthode ne consiste pas au rétablissement de la production de la protéine Ube3a fonctionnelle dans les neurones, mais vise à rétablir les fonctions qui sont affectées par l'absence de Ube3a. Différents tests sont effectués. (Tableau 5 page 54)

### 3. SYNDROMES IMPLIQUANT LA RÉGION 15Q11-Q13 et syndromes dits Angelman-like

Composé	Institut/Industriel	Mécanisme d'action	référence
Gaboxadol (OV101)	Ovid Therapeutics	Restauration de l'inhibition tonic de GABA	[240]
IGF-2 R ligands	NYU	Améliore la croissance et le maintien des synapses	[241]
Analogue de cyclique glycine-proline (NNZ-2591)	Neuren Therapeutics	Améliore la croissance et le maintien des synapses	[235]
PP2A inhibitor (LB-100)	UC Davis, Lixte	Inhibition de PP2A : améliore la fonction synaptique, renforce la plasticité synaptique.	[242]
NSI-189 phosphate	Seneca Biopharma	Améliore la fonction synaptique, renforce la plasticité synaptique	[243]
SAGE-324	Sage Therapeutics	Améliore la transmission GABAergique	[235]
Ketone esters	University of Colorado Trumacro Nutrition	Crises d'épilepsie	[244]

**TABLE 5** – Les traitements "Downstream" du syndrome d'Angelman. Abbreviations : IGF-2 R : insulin-like growth factor-2 receptor, NYU : New York University, PP2A : protein phosphatase 2A, UC Davis : University of California Davis, GABA : gamma-aminobutyric acid. Tableau adapté de [235]

### 3.2 Les syndromes dits Angelman-like

Lorsqu'un·e patient·e présente des symptômes s'approchant de ceux du syndrome d'Angelman mais qu'aucun défaut, ni d'empreinte parentale dans la région 15q11-q13, ni de délétion ou de mutation au niveau du gène Ube3a, n'est détecté, ce syndrome est appelé Angelman-like.

Différents gènes ont été identifiés comme responsable de syndromes dits Angelman-like. De manière notable, le gène HERC2 codant pour la protéine du même nom a été identifié à plusieurs reprises comme responsable de ce type de syndrome.

Le gène HERC2 n'étant pas soumis à une empreinte parentale, la délétion de ce gène ou des mutations affectant le bon fonctionnement de cette protéine sur les deux copies du gène (maternel et paternel) sont nécessaires au développement de ce syndrome. Ceci explique le faible nombre de syndromes de type Angelman-like dus au gène HERC2. Différentes études ont été faites sur les différents cas de syndromes Angelman-like liés à des mutations de la protéine HERC2.

Les premiers cas étudiés sont issus d'une famille de la communauté Ammische. 7 individus de cette famille souffrant de symptômes de neurodéveloppement, pouvant être asso-

ciés à des troubles du spectre autistique, possèdent une mutation faux-sens (Pro594Leu). Cette mutation se situe au niveau du domaine RLD1 de HERC2 [245]. Une seconde étude montre que cette mutation sur le domaine RLD1 affecte l'expression de HERC2 et sa stabilité [246]. Bien que les symptômes soient proches du syndrome d'Angelman, il est important de noter que dans ces cas le syndrome d'Angelman n'as pas été suspecté.

Une seconde étude porte sur un patient ayant une délétion homozygote contenant les exons 57 à 93 du gène HERC2 et les exons 1 à 19 du gène OCA2. Cette délétion entraîne la production d'une HERC2 dépourvue de sa partie carboxy-terminale (absence notamment des domaines RLD2, RLD3 et du domaine catalytique HECT). Les symptômes observés chez ce patient comprennent une hypotonie néonatale, un retard de développement global et une démarche instable (ataxie). Pour ce patient, les symptômes semblent être plus sévères que pour les 7 patient·e·s de l'étude précédemment évoquée. Cela s'expliquerait par une absence de production de la protéine fonctionnelle [247]. Une autre étude de 2021 s'intéresse à une autre mutation homozygote (Arg1542His) [248].

De manière notable, HERC1, une autre ubiquitine-ligase de la famille HERC a été identifiée comme étant impliquée dans des troubles du neurodéveloppement [249]. D'autres syndromes se rapprochant du syndrome d'Angelman impliquent d'autres gènes. Ces syndromes, rares, sont souvent appelés Angelman-like [250].

### 3.3 Le syndrome de Prader-Willi

Le pendant opposé du syndrome d'Angelman est le syndrome de Prader-Willi. Ce syndrome touche environ 1 personne sur 15000. Il a été identifié en 1956 par trois médecins qui donneront leurs noms à ce syndrome [251]. Ce syndrome est caractérisé par une hypotonie, des troubles de croissance et une hyperphagie due au manque de sensation de satiété [252].

Ce syndrome est dû à l'absence de l'expression des gènes paternels de la région 15q11q13. Ce manque d'expression peut avoir différentes origines (comme pour le syndrome d'Angelman). La délétion de cette région est impliquée dans 65 à 75 pour cent des cas, la disomie uniparentale maternelle du chromosome 15 dans 20 à 30 pour cent des cas. Enfin, 1 à 3 pour cent des cas sont dus à un problème d'empreinte parentale pouvant impliquer des micros délétion dans le IC [253].

### 3.4 Le syndrome dup15q

Un autre syndrome du spectre autistique, le syndrome dup15q, touche environ 1 personne sur 30000. Ce syndrome implique cette même région chromosomique et a pour cause la

multiplication de la région 15q11q13 [254].

Les symptômes les plus décrits sont : l'hypotonie, le retard de développement, la déficience intellectuelle, l'épilepsie et des troubles du spectre autistiques. Ce syndrome est donc proche du syndrome d'Angelman [255] [256].

Cette multiplication de cette région peut se faire par deux mécanismes. Le premier est appelé duplication interstitielle. Dans ce cas, la duplication se fait au sein du chromosome 15. Dans l'autre cas, la duplication dite isodicentrique implique l'apparition d'un chromosome supplémentaire.

La multiplication de cette région spécifique est due à la présence de LCRs ("Long Copy Repeats") aussi appelés BPs ("BreakingPoints") qui favorisent les réarrangements génomiques [257].

L'origine de la duplication, maternelle ou paternelle, a un effet sur les symptômes [258]. La multiplication de la région maternelle semble être associée à des symptômes du spectre autistique proche du syndrome d'Angelman tandis qu'une multiplication d'origine paternelle semble engendrer des symptômes plus proches du syndrome Pradder-Willi. Cependant, cette corrélation ne semble pas totalement claire.

Ainsi, le dosage de la protéine Ube3a dans les neurones et dans le reste des cellules semble être capital. Un surdosage de cette protéine (Dup15Q) tout comme un sous dosage (Angelman) ont un impact sur le neurodéveloppement.

## 4 | Objectifs de la thèse

Dans cette introduction, nous avons vu le mécanisme d'ubiquitination d'un substrat et la diversité des rôles joués par l'ubiquitination. Les principaux protagonistes de ce mécanisme sont les enzymes E1, E2 et E3 qui interviennent à tour de rôle.

De plus, deux E3 à domaine HECT, HERC2 et Ube3a, sont impliquées dans des syndromes du neurodéveloppement. De manière notable, ces deux gènes sont situés dans une même région chromosomique. La protéine Ube3a est impliquée dans le syndrome d'Angelman mais aussi dans les syndromes de Prader-Willi et dup15q. Le dosage de cette protéine semble être capital pour le bon fonctionnement du neurodéveloppement. Un sous dosage dans les neurones cause le syndrome d'Angelman alors qu'un surdosage cause le syndrome Dup15Q. L'absence ou la défaillance de la protéine HERC2 est reliée, quant à elle, à un syndrome proche du syndrome d'Angelman appelé Angelman-like.

De plus, ces deux protéines, HERC2 et Ube3a sont capables de se lier l'une à l'autre. La région de Ube3a impliquée dans cette interaction a été identifiée comme un fragment non replié de la protéine Ube3a (résidus 150-200). Cette région interagit avec le second domaine RCC1-like de la protéine HERC2. Cependant, le mécanisme moléculaire de cette interaction, sa structure et ses implications ne sont pas connues.

De plus, la protéine Ube3a (aussi appelée E6AP dans les études traitant de HPV) peut être liée par l'oncoprotéine E6 de papillomavirus. Cette interaction entraîne l'augmentation de l'activité ubiquitine-ligase de Ube3a. De plus, cette interaction entraîne aussi la dégradation de la protéine P53 qui, en absence de E6, n'est pas un substrat de Ube3a.

Ainsi, ce travail se sépare en deux axes qui ont pour but de mieux comprendre les mécanismes mis en jeu dans ces syndromes.

Le premier axe de ce projet vise à participer à une meilleure compréhension de l'interaction entre les deux ubiquitine-ligases HERC2 et Ube3a. Nous avons d'abord précisé cette interaction, c'est-à-dire délimité la région minimale d'Ube3a permettant l'interaction avec le domaine RLD2 de HERC2. Ensuite, nous avons étudié la structure du domaine RLD2 seul et la structure du domaine RLD2 lié à cette région d'Ube3a. Ceci nous a permis de

#### 4. OBJECTIFS DE LA THÈSE

---

déterminer une séquence d'acides aminés, appelée motif, requise pour cette interaction. Enfin, nous nous sommes interrogés sur la présence de ce motif dans d'autres protéines leur permettant d'interagir avec le domaine RLD2.

Le second axe de ce projet consiste à déterminer la structure de la protéine Ube3a. En effet, seuls les structures des domaines HECT et AZUL d'Ube3a sont connues.

Dans ce projet, nous visons à mieux connaître la structure entière d'Ube3a ainsi que la structure de ses différents domaines. Pour cela, nous avons utilisé différentes constructions de cette protéine dans le but de les purifier et de tenter de les cristalliser. Nous avons associé cette méthode aux structures prédites par le logiciel AlphaFold (un logiciel de prédiction de structures de protéine).

La connaissance de la structure de Ube3a pourra permettre la meilleure compréhension de son mécanisme d'ubiquitination et de ses rôles cellulaires. En effet, les mécanismes d'ubiquitination par Ube3a, notamment son activation par la protéine E6, ne sont pas encore totalement compris. L'étude de la structure et des mécanismes de Ube3a est capitale pour l'étude des syndromes présentés précédemment.

Deuxième partie

Résultats & Discussions



# 1 | Le domaine RLD2 et son interactome

## 1.1 Avant-propos

### 1.1.1 Introduction

Cette section présente les travaux effectués sur le domaine RLD2 et sa capacité à lier un motif présent dans différentes protéines, dont Ube3a.

La capacité de la protéine HERC2 à interagir via son domaine RLD2 avec la protéine Ube3a est connue. Cependant, le détail de cette interaction n'a pas été étudié. Ici, nous nous intéressons à cette interaction d'un point de vue structural puis d'un point de vue interactomique.

Nous avons décrit structurellement la nature de l'interaction entre ces deux protéines et identifié les éléments clés de celle-ci. En effet, la structure de l'interaction minimale entre Ube3a et HERC2 nous a permis de déterminer la courte séquence de Ube3a, appelée motif, permettant cette interaction. Nous avons par la suite identifié d'autres protéines humaines possédant ce motif déterminant.

De plus, ces deux protéines étant impliquées dans des syndromes du neurodéveloppement, nous avons pu replacer ce réseau d'interactions dans son contexte biologique. Nous avons identifié le domaine RLD2 de HERC2 comme jouant un rôle important dans le développement des neurones via sa capacité à lier d'autres protéines.

Cette partie résultats est présentée sous la forme d'un papier qui sera soumis prochainement.

### 1.1.2 Contributions

Dans le travail qui suit, j'ai participé à l'optimisation du protocole de purification avec Eduardo Howard. J'ai réalisé la partie structurale (purification, cristallisation, et résolu-

tion des structures) du domaine RLD2 et du domaine RLD2 en complexe avec les différents peptides ainsi que l'analyse de ces structures. Avec Alastair Mc Ewen, Alexandra Cousido-Siah et André Mitschler, j'ai réalisé les mesures de diffraction des rayons-X au synchrotron. J'ai aussi réalisé les mesures d'affinités (mesures par HoldUp non présentées ici et mesures par polarisation de fluorescence). Les premières mesures d'affinité des différents peptides de Ube3a pour RLD2, réalisées par HoldUp ont été effectuées par Anna Bonheure. Les peptides utilisés ont été produits par Pascal Eberling. J'ai effectué les analyses phylogénétiques de séquences. L'étude sur les effets de l'interaction entre RLD2 et DOCK10 a été menée par Susanne Schmidt et Anne Debant.

## 1.2 Abstract

Ube3A and HERC2 are two mutually-interacting HECT E3 Ub-ligases whose genes, both encoded by the unstable 15q 11-13 chromosomal region, are altered in Autism Spectrum Disorder Dup15q, Angelman Syndrome (AS) and other phenotypically-related mental retardation syndromes. We crystallized the RCC1-Like Domain 2 (RLD2) domain of HERC2 complexed to a short region of Ube3A comprising the consensus motif "DxDKDxD". Ube3A conserves a DxDKDxD motif only in animals possessing a central nervous system (CNS) whereas HERC2 conserves its DxDKDxD-binding surface across most animal species, suggesting that this motif may be present in proteins other than Ube3A. Accordingly, we identified and validated by bioinformatic, crystallographic, quantitative interactomic and cellular approaches RLD2-binding DxDKDxD motifs in a handful of proteins (Ube3A, DOCK10, PCM1, USP35, BAZ2B, ARID4A, ARIP4, RERE and MYT1), several of which are notoriously relevant for brain development. In particular, mutagenesis and/or RLD2 binding of the DxDKDxD motif of the RhoGEF DOCK10 disrupted the GEF activity of DOCK10 and its capacity to regulate spine density in hippocampal neurons. Our findings indicate that the RLD2 domain of HERC2 establishes with its DxDKDxD-containing partners a tightly balanced interaction network, whose alteration at the genetic and/or protein levels can produce neurodevelopmental defects.

## 1.3 Introduction

Ube3A (also called E6AP) is a E3 ubiquitin (Ub)-ligase whose maternal allele is deleted, mutated or inactivated by de novo non parental-inherited genetic alterations in the severe neurodevelopmental disease Angelman syndrome (AS) [259]. The maternal Ube3A allele is also duplicated or triplicated in an autism spectrum disorder (ASD), called Dup15q syndrome, covering approximately 1p of ASD cases [171]. The specific involvement of the maternal Ube3A allele in neurodevelopment syndromes is due to the fact that the paternal

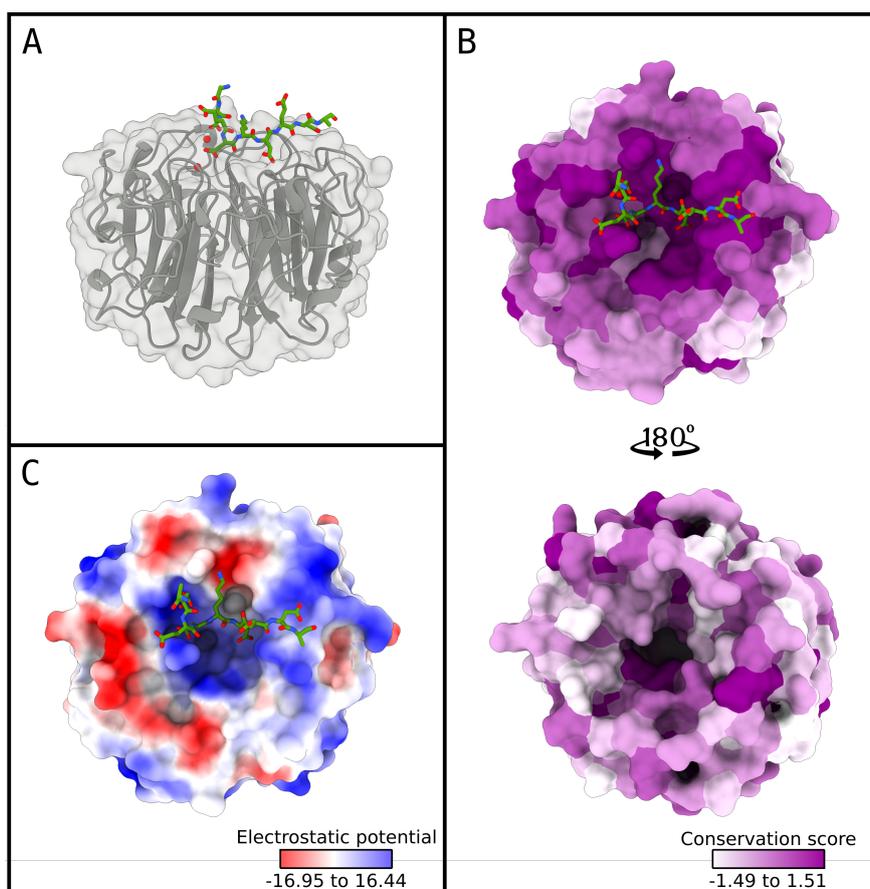
allele of Ube3A gene is epigenetically silenced in most brain regions, while both alleles are expressed elsewhere, allowing the paternal allele to compensate the defective maternal allele in the rest of the body. This epigenetic mechanism is directly linked to the presence, proximal to the Ube3A-encoding gene, of the "Angelman Syndrome - Prader Willi Syndrome Imprinting center"(AS-PWS-IC) [152] [260].

HERC2 is another E3 ubiquitin ligase, whose gene is also altered in rare neurodevelopmental syndromes, presenting several symptoms reminiscent of AS [245] [246] [247][250]. Remarkably, Ube3A and HERC2 genes are chromosomal neighbours, both belonging to the 15q 11-13 region. Some (yet not all) of the systematic Ube3A maternal allele deletions observed in AS patients also involve a deletion of the neighbour HERC2 gene. Furthermore, at the protein level, Ube3A (852 residues) and HERC2 (4834 residues) belong to the same E3 ubiquitin ligase family. They both contain a "HECT" C-terminal domain responsible for the ligase activity. The high sequence similarity of both domains suggest that the two genes originated from a duplication event during evolution, which may explain their proximal chromosomal location. Moreover, both proteins physically interact with each other [261] [210] [165] [209] and binding of HERC2 to Ube3A enhances Ube3A ligase activity [210]. A recent transcriptional study suggested Ube3A and HERC2 proteins may cooperate in Dup15q syndrome [262]. The Ube3A-HERC2 complex thus emerges as a critical actor of brain formation, whose perturbation leads to neurodevelopmental diseases.

Here, we define the minimal interface of the HERC2-Ube3A complex, involving the RCC1-Like Domain 2 (RLD2) of HERC2 and a "DxDKDxD" sequence motif of Ube3A. The HERC2-Ube3A interface is conserved in most animals possessing a centralized nervous system, pointing to the importance of this complex for brain development. We identify a set of HERC2-binding proteins containing the DxDKDE motif, including Ube3A, DOCK10, PCM1, USP35, BAZ2B, ARID4A, ARIP4, RERE and MYT1. We validate most of these complexes by X-ray crystallography, affinity measurements and unbiased AP-MS interactomics. We further investigate the interaction of HERC2 with DOCK10, a Rho-GEF protein involved in dendritic spine morphogenesis. Mutating the DxDKDxD motif of DOCK10 disrupts both the DOCK10-HERC2 interaction and the Rho-GEF activity of DOCK10. We also find that PCM1, a key pericentriolar component already known to bind HERC2, can be displaced from the pericentriole region by overexpression of the RLD2 domain, suggesting that binding to HERC2 is essential to PCM1 pericentriolar localization. Altogether, these data point to the existence of a network of DxDKDxD-containing proteins, all capable of binding to HERC2. At critical steps of brain development, the absence or disruption of any member of this network, such as Ube3A, DOCK10, BAZ2B, PCM1 or HERC2 itself, may create an imbalance in the network, leading to altered neurodevelopment.

## 1.4 Results

The HERC2-Ube3A interaction was previously found to involve the RCC1-like domain 2 (RLD2) of HERC2 and Ube3A residues 173 –223 (numbering of the 875-residue Ube3A human isoform II) [263] [209]. By probing the binding of recombinant RLD2 to an array of Ube3A-derived peptides, we further defined the minimal HERC2-binding region of Ube3A (Supplemental table 7, page 84) and subsequently crystallized a complex composed of HERC2 RLD2 (residues 2941-3342) bound to Ube3A residues 183-197 (Figure 34). RLD2 displays a donut-shaped, seven-bladed  $\beta$ -propeller fold typical of the RCC1 family, with a central funnel connecting the “bottom” and “top” sides of the donut [264]. The Ube3A peptide lies on a highly conserved groove situated on the top face of RLD2 domain (Figure 34.B). The seven HERC2-binding residues of Ube3A visible in the crystal are essentially charged, with a central lysine flanked by aspartic and glutamic acids on both sides (sequence : D1E2D3K4D5E6D7). Accordingly, the charge distribution of the motif-binding surface of RLD2 is perfectly complementary to that of the peptide (Figure 34.C).



**FIGURE 34** – Crystal structure of the RLD2 domain of the HERC2 protein bound to the minimal HERC2-binding motif of Ube3a. (PDB : 7Q41). A. RLD2 adopts the canonical donut-shaped, 7-bladed beta-propeller fold of RCC1-like domains. The peptide interacts with the "top" side of the RCC1-like fold according to its conventional representation [264]. The peptide is oriented from N-terminus (left) to C-terminus (right) in this representation and the following ones. B. AL2CO conservation score [265], calculated using the Chimera program from the HERC2 sequences of 7 representative animal organisms (see sequences in Figure 36A, page 69) and mapped on the surface of the RLD2 domain. Note the particularly high conservation ("hot" purple color) of the motif-binding groove. C. Surface electrostatic charges of the surface of the RLD2 domain, calculated using the Chimera program. A positively charged groove hosts the two first aspartic acids, D1 and D3, of the D1E2D3K4D5E6D7 motif. The central lysine K4 is embedded in a negatively charged pocket. The two last aspartic acids D5 and D7 interact with positively charged areas of the RLD2 domain. Further details of the complex are shown in Figure 35, page 67.

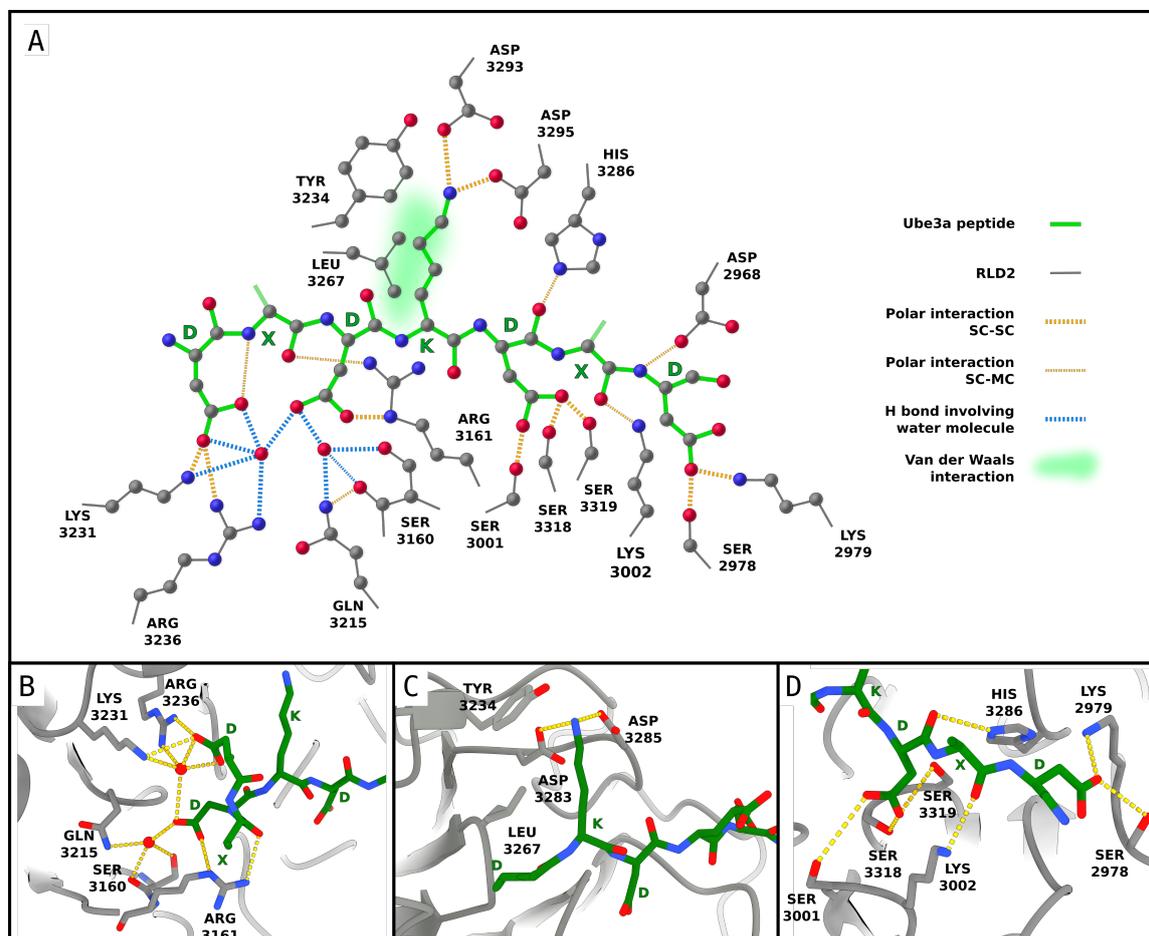
By combining visual inspection of the complex (Figure 35, page 67) and quantitative assessment of its buried interface (Figure 36A page 69), we found that the Lys residue and the surrounding aspartic residues of Ube3A establish critical charged and hydrophobic interactions with the RLD2 domain (Figure 35A C, page 67). By contrast, the two glutamic residues (positions 2 and 6) point towards the solvent and are therefore not involved in the

interface. The key HERC2-binding residues of Ube3A therefore boil down to the following sequence motif : D1x2D3K4D5x6D7, where x stand for non-determining residues.

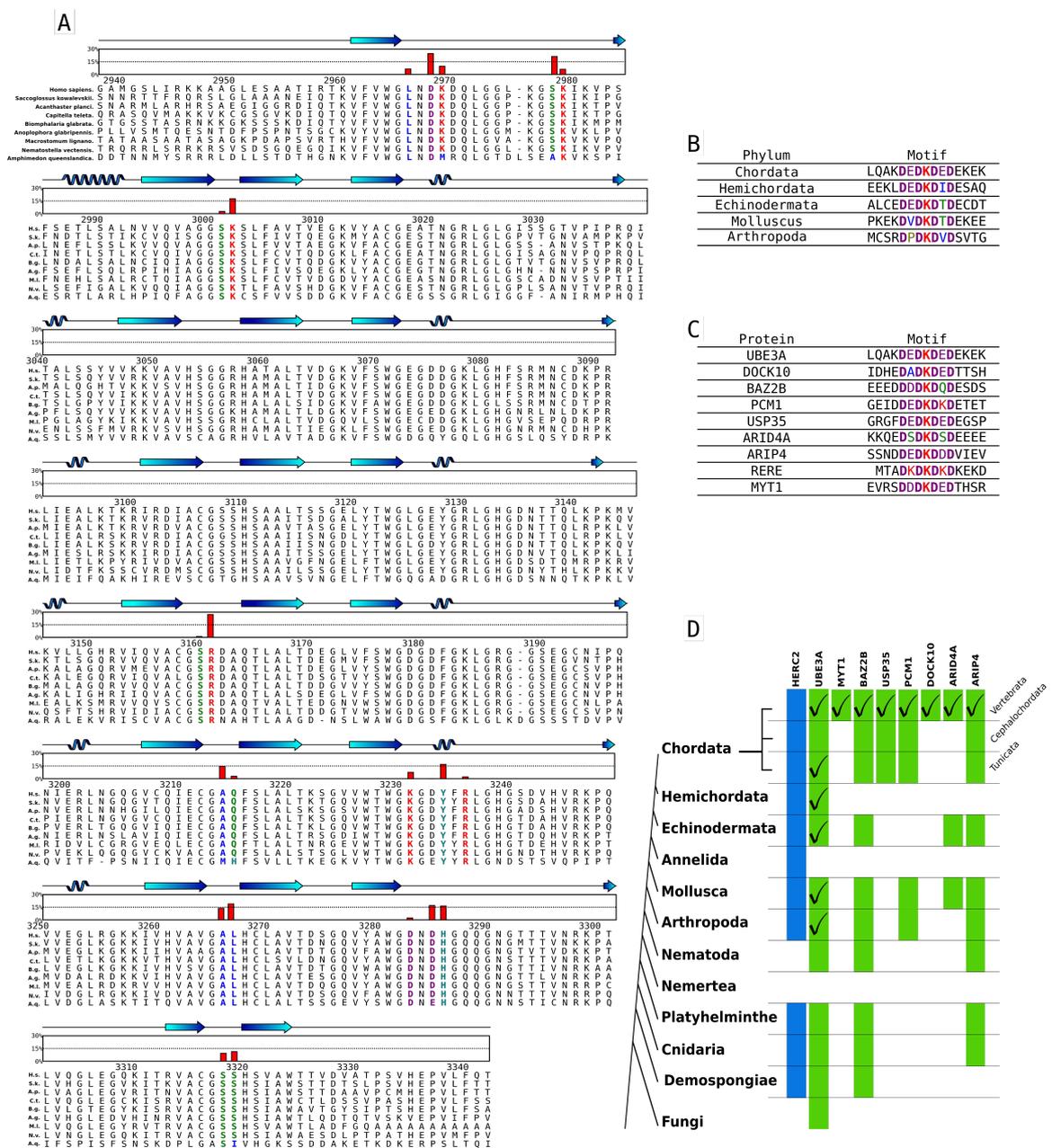
HERC2 proteins, characterized by their large size ( 5000 residues in human), their three RLD domains, and a C-terminal E3 Ubiquitin-ligase HECT domain, are found in most animal species, yet not in fungi nor more distant organisms. By aligning HERC2 sequences of a diversity of organisms (Multiple alignment, Figure 36.A), we found that the DxDKDxD-binding surface of HERC2 is strongly conserved across most animal species (Figure 36.A). Ube3A-like proteins, characterized by their 850-residue size, their proteasome-binding N-terminal AZUL domain and their C-terminal E3 Ub-ligase HECT domain, are found in most animals as well as in a number of fungi (Figure 36.BD). However, we found that Ube3A-like proteins display the HERC2-binding DxDKDxD motif only in animals that possess a brain, or centralized nervous system (CNS), such as vertebrates, molluscs, or arthropods (Figure 36 B D page 69). This further indicates that the HERC2-Ube3A interaction participates in the development and/or functioning of animal brain.

This phylogenetic sequence analysis also revealed that HERC2 proteins conserve a DxDKDxD-binding surface in animals devoid of CNS, where Ube3A does not display anymore a DxDKDxD motif. This indicated, that the RLD2 domain of HERC2 may bind to other candidate DxDKDxD motifs, found within proteins distinct from Ube3A. A bioinformatic search across the full human proteome identified a perfect DxDKDxD motif within only 14 proteins including Ube3A (Supplemental Table 6, page 83). In 9 of them (Figure 36C, page 69), the motif was situated in a predicted disordered region favoring its availability for partner binding, was conserved across most vertebrate species (Figure 36D) and turned out to display significant RLD2-binding in subsequent assays (Supplemental table 7, page 84).

We moved on to solve the structures of RLD2 bound to five of the above-mentioned DxDKDxD motifs, derived from Dock10, Pcm1, Usp35, Baz2b and Myt1 (Figure 37, page 70). The presence of side chains of variable bulkiness at the non-conserved "x" positions, and the higher resolution of some of the structures (culminating at 1.95 Å for the RLD2-BAZ2B complex) allowed us to ascertain the orientation and position of the peptide in the pocket, which was strictly conserved in all the RLD2-DxDKDxD complexes, including the initial RLD2-Ube3A structure.



**FIGURE 35** – Interaction network between the DxDKDxD motif and the RLD2 domain. A. Schematized network of contacts between RLD2 domain and DxDKDxD motif, conserved in all six RLD2-motif structures solved in this work. Note the importance of ionic and polar contacts. Two water molecules buried at the interface mediate a number of critical H-bonds between several RLD2 residues and aspartates D1 and D3 of the motif. Van der Waals contacts are also observed, with Tyr 3234 and Leu 3267 facing the hydrophobic side-chain of lysine K4 of the motif (C). Note that the two "x" amino acids are exposed to solvent and do not contribute to the interaction (B and D). B. Detail of the RLD2-DxDKDxD complex, showing contacts of the N-terminal part of the motif, involving the two buried water molecules. This view and the two next ones were prepared from the RLD2/BAZ2B structure (PDB :7Q42), which displayed the best resolution (1.95 Å). The binding mechanism is the same for all crystallised complexes. C. Detail of contacts involving the central lysine of the motif. D. Detail of contacts involving the C-terminal part of the motif.



**FIGURE 36** – Conservation of HERC2 RLD2 domain across different species and of DxDKxD motifs across different species and different proteins. A. Sequence alignment of HERC2 RLD2 domain of species from 7 different animal phyla (accession numbers of these and all other sequences shown in the figure are provided in supplemental data). Blades are structurally aligned based on their secondary structure elements, displayed on top of the sequences. The parts of the beta strands oriented towards the "top" side of the propeller supporting the bound peptide motif are colored in dark blue. For each amino acid of the domain, the percentage of accessible surface area buried by the bound peptide is plotted on top of the sequence alignment. Aminoacids buried by the peptide are colored according to physico-chemical properties. RLD2 amino acids are numbered according to human HERC2 Uniprot sequence. B. Conservation of DxDKxD motif across UBE3A proteins of different

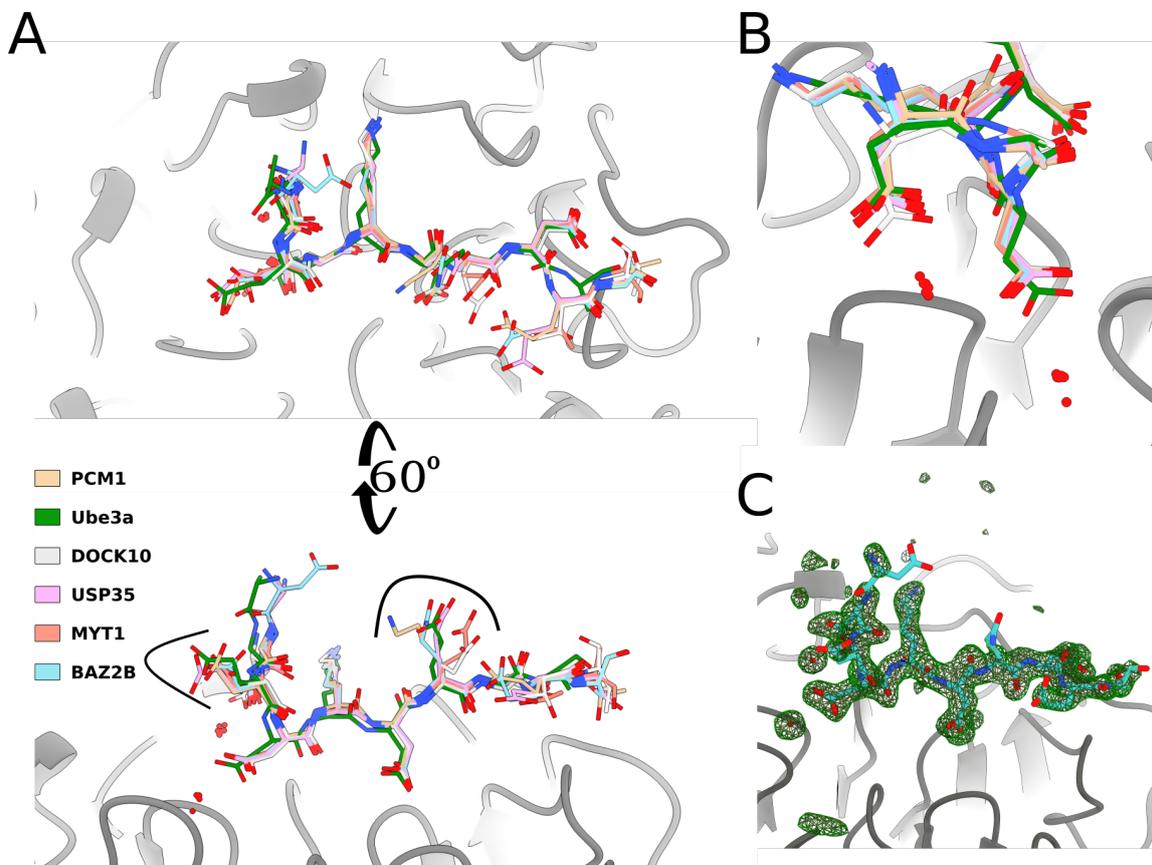
phyla. C. Presence of DxDKDxD motif in other human proteins. D. Conservation of RLD2 domain and DxDKDxD motifs across proteomes of different phyla. Blue squares indicates that HERC2 (and RLD2) was found in the proteome. The green squares indicate that the indicated protein was found. The check marks indicate the motif was conserved in the indicated protein.

We used a competitive fluorescence polarization assay (cFP) (see methods) to measure the affinity of the HERC2 RLD2 domain for DxDKDxD motif-containing peptides derived from Ube3A, Dock10, Usp35, Pcm1, Baz2b, Myt1, Arid4a, Arip4, RERE, Smc6 and MED1. Most peptides bound in a comparable affinity range ( $K_D$  3 to 30  $\mu$ M) (Supplemental table 7, page 84), except those from Smc6 and MED1, which exhibited very weak or undetectable binding.

The affinity of some peptides could vary depending on their length and/or post-translational modification status. For instance, Arid4a contains a DSDKDSD sequence with S538 and S542 being annotated on the phosphosite database ([www.phosphosite.org](http://www.phosphosite.org)) as phosphorylatable positions. The affinity of Arid4a-derived peptides varied from  $K_D$  100  $\mu$ M for a short 15-mer centered on the DxDKDxD motif to  $K_D$  10  $\mu$ M for a 25-mer bearing a single phosphorylation on S538 (Supplemental table 7, page 84). As concerns Ube3A, we tested four different polypeptides, all comprising the DxDKDxD motif : a minimal 15-mer synthetic peptide, a 25-mer, a 51-mer and the full-length protein (Uniprot isoform 1, 852 residues) that we produced and purified as a monomeric sample (Supplemental, Figure 41, page 82). The 15-mer, 51-mer and the full-length Ube3A all bound RLD2 in the micromolar affinity range ( $K_D$  = 6  $\mu$ M, 5  $\mu$ M and 1,5  $\mu$ M, respectively). By contrast, the 25-mer peptide bound to RLD2 at least 100-fold stronger, in the nanomolar range ( $K_D$  < 50 nM) (Supplemental table 7, page 84). In the Ube3A 25-mer peptide, we noticed the occurrence of two other K residues surrounded by acidic residues, which may serve as secondary HERC2-binding sites that potentially reinforce binding affinity via avidity effects. In support of that hypothesis, RLD2 readily crystallized in the presence of a 15-meric Ube3A peptide but never crystallized with the higher affinity 25-meric Ube3A peptide. The presence of multiple alternative binding sites may have hindered crystallization. We hypothesize, that in full-length Ube3A as well as in the 51-meric peptide, these secondary binding sites engage long-range contacts with sequentially distal regions of Ube3A, so that they can no longer participate in RLD2 binding.

All the identified DxDKDxD motifs are situated in presumably disordered regions. This suggests that, in the context of the full-length proteins they belong to, these motifs are freely available for interaction with HERC2. One can reasonably infer that the differences of RLD2-binding affinity measured for the various peptides are representative of the differences of affinity of their full-length proteins of origin (Ube3A, Dock10, Pcm1, Usp35,

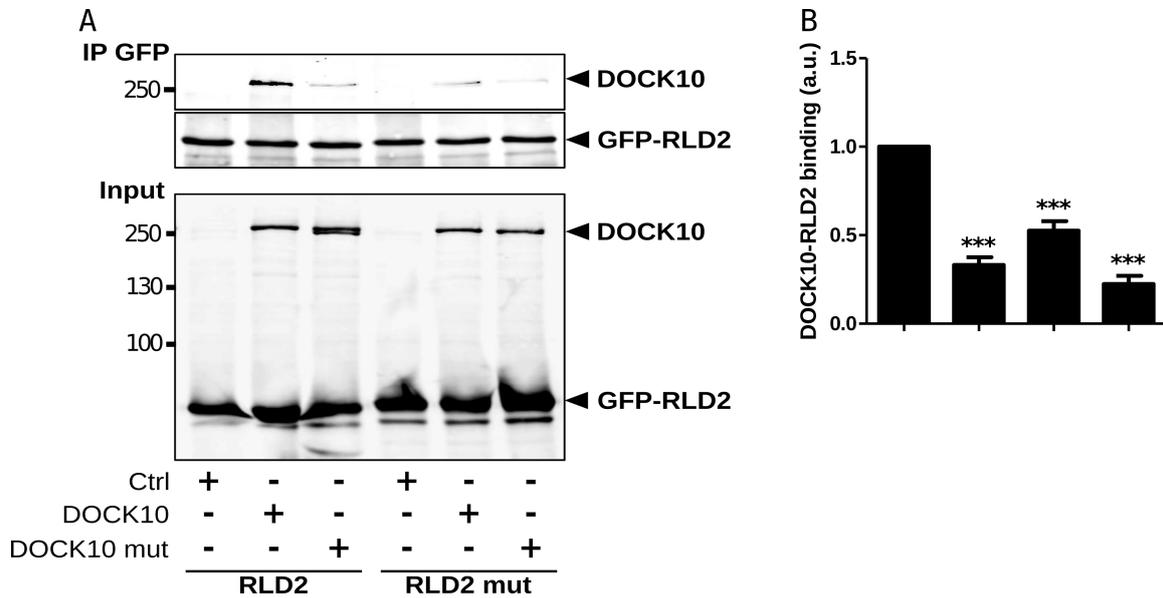
Baz2b, Myt1, and so forth) towards full-length HERC2.



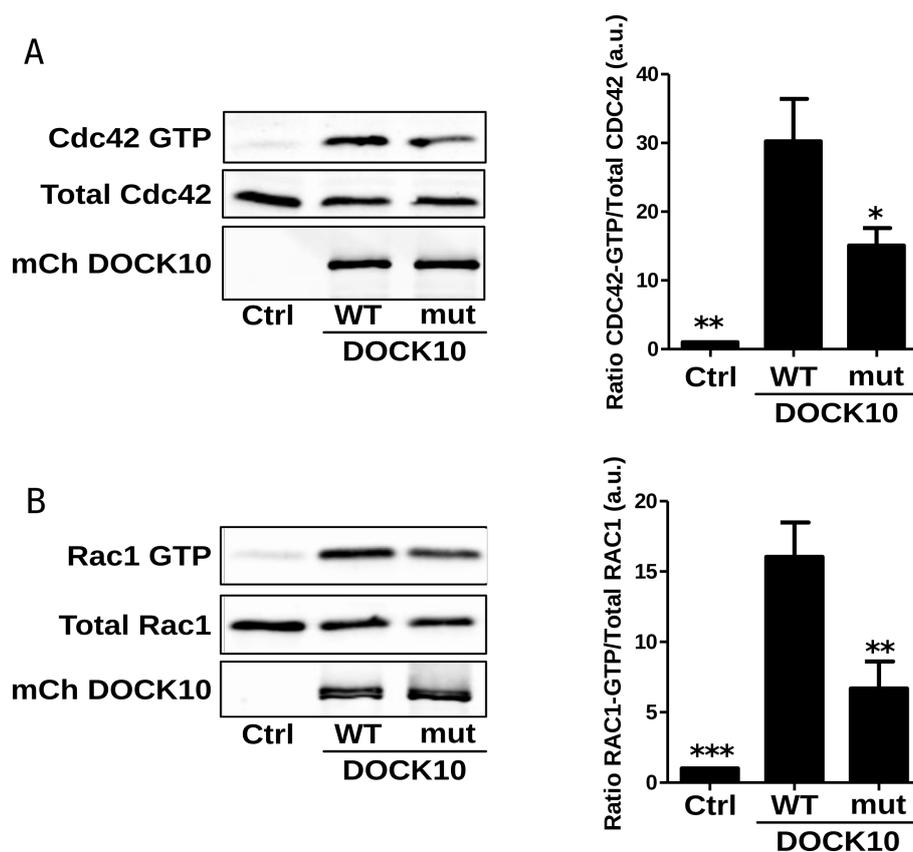
**FIGURE 37** – A) Crystallographic structure of the RLD2 domain in complex with different peptide containing the DxDKDxD motif. B) The 2 water molecules involved in the interaction with the first two aspartic acids are present in all structures. D) The omit map for BAZ2B.  $|(F_o)-(F_c)|$  difference density contoured at  $3\sigma$  (green)

We further investigated the interaction of Ube3A with DOCK10, a Rho-GEF known to regulate dendritic spine growth at critical steps of post-embryonary neurodevelopment [266]. Co-precipitation and mutagenesis experiments confirmed that the HERC2 RLD2 domain binds to DOCK10, and that this interaction requires both an intact DxDKDxD-binding groove in HERC2 and a preserved DxDKDxD motif in DOCK10 (Figure 38, page 71). Furthermore, mutagenic inactivation of the DxDKDxD motif in the N-terminal region of DOCK10 significantly disrupted the Rho-GEF activity of DOCK10, despite the fact that this activity lies in the distal DHR2 domain of DOCK10, situated at its C-terminus (Figure 39, page 72). Finally, mutagenic inactivation of the DxDKDxD motif of DOCK10 also disrupted the capacity of DOCK10 to enhance dendritic spine formation in transfected cultured hippocampal neurons (Figure 40, page 74). These findings point to the DxDKDxD motif of DOCK10 as a critical modulator of DOCK10 GEF activity and subsequent regulation of dendritic spine formation. They suggest the intriguing possibility that binding of HERC2 to

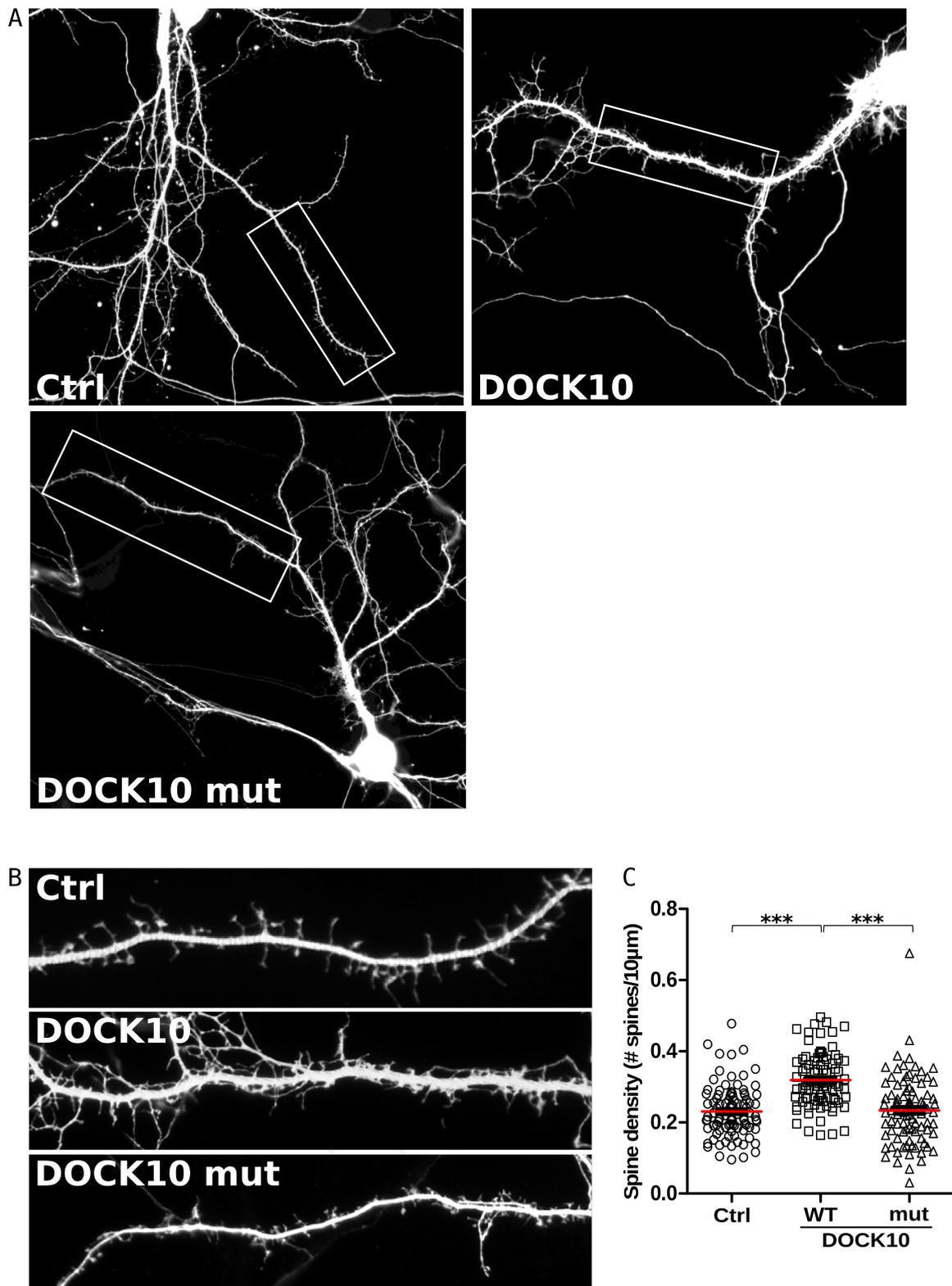
the DxDKDxD motif of DOCK10 could negatively modulate its GEF activity and subsequent control of dendritic spine growth.



**FIGURE 38** – The interaction between DOCK10 and RLD2 is affected by mutations in their respective binding sites. A. Western Blot analysis of a co-immunoprecipitation assay of GFP-RLD2 and mCherry-DOCK10. HEK293T cells were transfected with GFP-RLD2 (or GFP-RLD2mut, mutated on the DxDKDxD-binding interface as described in methods) and mCherry-DOCK10 (or mCherry-DOCK10mut, mutated on the DxDKDxD motif as described in methods) or a control empty vector. RLD2 was immunoprecipitated using an anti-GFP antibody, and the co-precipitating DOCK10 was detected using a mCherry antibody. B. Histogramm of the quantification of the experiment presented in A. Data are presented as the mean  $\pm$  SEM of 5 independent experiments. Statistical analyses were made using one-way ANOVA followed by Dunnett's test. Asterisks indicate datasets significantly different to the wt interaction between RLD2 and DOCK10 (\*\*\*) $p < 0.001$ ).



**FIGURE 39** – The activation of CDC42 and of RAC1 by DOCK10 is decreased for the DOCK10mut construct impaired for binding to RLD2. A. CDC42-GTP pull down assay. HEK293T cells (transfected with mCherry DOCK10, mCherry DOCK10mut or a control vector) were lysed and active GTP-CDC42 was affinity purified using the CDC42/RAC1-interactive binding (CRIB) domain of PAK1, immobilised on Glutathione- Sepharose beads. Purified GTP-bound and total CDC42 were detected by Western blot, using an anti-CDC42 antibody. Protein expression in the cell lysates was verified by immunoblotting with an anti-mCherry antibody. One representative experiment is shown. Right hand panel : histogram of the quantification of the experiment presented in the left-hand panel. CDC42 activation mediated by wt DOCK10 was arbitrarily set to 100%, in order to be able to compare the individual experiments. The % of CDC42 activation was calculated from at least 4 independent experiments (mean  $\pm$ SEM). Statistical analyses were made using one-way ANOVA followed by Dunnett's test (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ ). B. RAC1-GTP pull down assay. Experiments were performed as described in A, except that an anti-RAC1 antibody was used for immunoblotting.



**FIGURE 40** – Expression of DOCK10mut, defective in RLD2 binding, leads to decreased dendritic spine formation in primary hippocampal neurons. A. Representative micrographs of cultured murine hippocampal neurons transfected with mCherry DOCK10, mCherry

DOCK10mut or a control vector. Neurons were transfected at 8 DIV and fixed at 11 DIV (experimental conditions to reveal spine formation). Scale bar, to be added. B. Higher magnification images of the regions selected in A, illustrating the dendritic spine morphology and density in the neurons transfected as described in A. C. Quantification of dendritic spine density of hippocampal neurons transfected as described in A, measured by the number of spines/10  $\mu\text{m}$ . Error bars indicate SEM ( $n = x$  independent experiments).

PCM1 [267] (Pericentriolar Component Protein 1) also contains a DxDKDxD motif, which ranks among the top RLD2-binders in our affinity-based classification (Supplemental table 7, page 84). PCM1 has been previously and independently shown to interact with HERC2 by two other groups [REFs], and the binding site to PCM1 had been restricted to a HERC2 fragment that encompassed the RLD2 domain [Ref], confirming that the DxDKDxD motif is a hallmark for HERC2 RLD2-binding .

The DNA-binding transcriptional regulator ARID4A is another DxDKDxD motif-containing protein. Quite remarkably, ARID4 is known to directly regulate the epigenetic modification of the AS/PWD IC (Angelman Syndrome / Prader Willi Disorder imprinting center) [268]. That ARID4A is both (i) one of the few (14) human proteins containing a perfect DxDKDxD sequence able to competitively bind to the Ube3A-binding site of HERC2, and (ii) a regulator of the very imprinting center that rules the expression of Ube3A-encoding alleles, is remarkable and unlikely to be fortuitous. Phylogenetic sequence studies have previously shown that, during mammalian evolution, the AS/PWD imprinting center must have emerged in a common ancestor of all eutherian mammals, after the separation from the marsupial branch [260]. Interestingly, sequence analysis reveals that ARID4a proteins of all eutherian mammals contain the HERC2-binding DxDKDxD motif, while in marsupials this motif is present in some species, and absent in others. This suggests that the DxDKDxD motif of ARID4A was already present in the common ancestor of all mammals, but remains indispensable only in eutherian mammals, which possess the AS/PWD IC. One intriguing hypothesis could be that variations of Ube3A protein expression would operate a retro-feedback control on Ube3A gene transcriptional dosage, by inducing variations of the HERC2-bound fraction of ARID4A, that would modulate the action of ARID4A on the AS/PWD IC.

We retrieved from the PaxDb proteomic database the relative abundances of HERC2 and its diverse DxDKDxD motif-containing partners in various tissues and cell types. Using a recently published approach [269], we combined these relative abundance data with our RLD2-peptide affinity data to predict the relative distribution of the complexes engaged by HERC2 with these different targets. These predictions indicate that, among this list, Ube3A should be the main DxDKDxD-containing protein bound to HERC2 in the brain of a normal individual, since Ube3A is the highest expressed DxDKDxD protein in brain, and it appears

to be roughly 50-fold more abundant than HERC2.

This distribution is expected to change dramatically if Ube3A is no more expressed in brain. In such a situation, HERC2 is expected to be redirected to form complexes predominantly with DOCK10 and PCM1, and in minor amounts with USP35, ARID4A, SMC6 and MYT1. Our cellular data indicate that the functioning of both DOCK10 and PCM1 may be altered upon binding to the RLD2 domain of HERC2. Subsequently, one can expect that the activity of at least those two proteins may be significantly altered in brains displaying a lower expression of Ube3A.

## 1.5 Discussion

Ube3A and HERC2 genes are altered in neurodevelopmental syndromes presenting overlapping clinical pictures, and the Ube3A and HERC2 proteins interact with each other. This promoted us to investigate the molecular details of the Ube3A-HERC2 complex. We found that this complex boils down to a domain-motif interaction involving a short DxDKDxD motif of Ube3A and the RLD2 domain of HERC2. Remarkably, sequence conservation analysis supports the existence of such complex in most animals possessing a centralized nervous system. Furthermore, we identified and validated a handful of additional HERC2-binding proteins that also contain a bona fide RLD2-binding DxDKDxD motif.

We were particularly intrigued by the DOCKD GEF family members DOCK9, 10 and 11, which all contain a variant of the DxDKDxD motif within a small presumably disordered stretch of their N-terminal region, sequentially distal from the C-terminal region responsible for GEF activity. We focused on DOCK10, which contains a perfect version of the motif and was formerly found to be involved in dendritic spine formation at the early steps of mouse brain post-embryonic development. Mutagenesis of the DxDKDxD motif of DOCK10 or co-expression with RLD2 domain altered the ability of DOCK10 to promote dendritic spine formation. Furthermore, the GEF activity of DOCK10 was compromised by mutation of its DxDKDxD motif. This suggests that a proper dosage of HERC2-DOCK10 complexes and conformational changes potentially undergone by DOCK10 upon binding to HERC2 might be important for correct neurodevelopment, including dendritic spine formation.

We also found that endogeneous PCM1, a key pericentriolar component already well-known to bind HERC2 [209] [165] [192] [270] was displaced from the pericentriole region by overexpression of the RLD2 domain, suggesting that PCM1 binding to HERC2 is essential to PCM1 pericentriolar localization. PCM1 is proposed to participate in microtubule organization in early neuronal development [271] [271], mutations of PCM1 in human have been linked to schizophrenia, possibly through the induction of ciliopathic defects [272], and PCM1-KO mice display progressive ciliary, anatomical, psychomotor, and cognitive

abnormalities [273] [272]. The PCM1-HERC2 interaction is also remarkable by the fact that HERC2 and Ube3A have both been mentioned to localize at the centrosome at particular steps of the cell cycle. Again, it is tempting to speculate that an altered dosage of HERC2-PCM1 complexes might alter some of the potential functions of PCM1 in neurodevelopment.

The DNA-binding transcriptional regulator ARID4A/RbBP1 is another DxDKDxD motif-containing protein. Quite remarkably, ARID4 is a chromatin remodeler [274] involved in spermatogenesis [275], also known to directly regulate the epigenetic modification of the AS/PWD IC (Angelman Syndrome / Prader Willi Disorder imprinting center)[268]. That ARID4A is both (i) one of the few (14) human proteins containing a perfect DxDKDxD sequence able to competitively bind to the Ube3A-binding site of HERC2, and (ii) a regulator of the very imprinting center that rules the expression of Ube3A-encoding alleles, is remarkable and unlikely to be fortuitous. Phylogenetic sequence studies have previously shown that, during mammalian evolution, the AS/PWD imprinting center must have emerged in a common ancestor of all eutherian mammals, after the separation from the marsupial branch [260]. Interestingly, sequence analysis reveals that ARID4A proteins of all eutherian mammals contain the HERC2-binding DxDKDxD motif, while in marsupials this motif is present in some species, and absent in others. This suggests that the DxDKDxD motif of ARID4A was already present in the common ancestor of all mammals, but remains indispensable only in eutherian mammals, which possess the AS/PWD IC. One intriguing hypothesis could be that variations of Ube3A protein expression would operate a retro-feedback control on Ube3A gene transcriptional dosage, by inducing variations of the HERC2-bound fraction of ARID4A, that would modulate the action of ARID4A on the AS/PWD IC.

Among the other DxDKDxD-containing proteins, BAZ2B (bromodomain adjacent to zinc finger 2) is a component of chromatin remodeling complexes whose gene haploinsufficiency is strongly associated to developmental delay, intellectual disability, and autism spectrum disorder [276] [277] [278]. ARIP4/RAD54L2 is a steroid hormone receptor co-factor involved in chromatin remodeling and Androgen Receptor regulation in sex differentiation and spermatogenesis [279], as well as a major binder of the DYRK1A kinase implicated in neuronal differentiation defects in ASD and Down Syndrome [280] [281] [282]. Myt1 (Myelin Transcription Factor 1) is cell cycle kinase [283] and proneuronal transcription factor [284] that binds to CoREST protein [282] to serve as a neuron-specific subunit of the LSD1/CoREST chromatin remodeling complex [285] [286], and Myt1 gene deletion promotes mental retardation<sup>31</sup>. In total, 4 out of 10 DxDKDxD motif proteins (ARID4A, BAZ2B, ARIP4, and MYT1) are involved in chromatin remodeling. This is remarkable, as HERC2 itself is a strong candidate for chromatin remodeling activities as it contains a

zz domain that specifically interacts with Histone H3 [287], and HERC2 gene mutations were early recognized to induce sperm defects in mice [288]. Finally, our last validated DxDKxD protein, USP35, is a de-ubiquitinase involved in various cancers [289][290] [291] [292] [293] that was recently suggested to play an active role in the development of ciliopathies [294].

The ensemble of our observations thus point to a network of proteins sharing the ability to bind to a defined domain of the giant protein HERC2. Remarkably, most of these proteins, including HERC2 itself, show documented implications in normal and/or aberrant neurodevelopment, through cellular functions that include ubiquitinylation pathways, cell cycle control, centrosome regulation, ciliogenesis, transcription control and chromatin remodeling.

In homeostatic conditions, one can expect that HERC2 will form complexes in different proportions with the different DxDKxD partners, depending on the tissue and cell types, the amounts of HERC2 and of each partner, their localization, and their conformational states which may be regulated by binding to third party proteins or molecules, or post-translational modifications. For instance, our affinity measurements have shown that Ube3A may display at least two alternative binding modes to RLD2 : one in the 50 nM range, and one in the 10  $\mu$ M range. Ube3A may switch between those two affinity modes depending on post-translational modifications or binding to other proteins. Similar affinity-modulatory switches may also occur for the other DxDKxD members of the network, as well as for HERC2.

In pathologic conditions, where a member of the network (HERC2, Ube3A or another DxDKxD motif protein) is missing or defective due to genetic alterations, one can expect an imbalanced network, where complexes will be redistributed in anomalous proportions. In brain tissues where Ube3A is missing or misfolded (AS), HERC2 is likely to form more abundant complexes with its other potential partners, thereby altering their conformation, their localization, their availability for other interactions, and so forth. In brain tissues where Ube3A is expressed in higher amounts (Dup15q), an anomalous redistribution of the network should be also expected. Finally, in non-brain tissues where Ube3A is expressed in lower amounts, we should also expect a redistribution of HERC2 complexes engaged with the other DxDKxD proteins. In this regard, we find it remarkable that three of the identified HERC2-binding DxDKxD proteins, PCM1, BAZ2B, and USP35, showed a genetic or functional link to ciliopathies.

We advocate for a generalization of the interactome imbalance approach, including quantitative affinity measurements, protein abundance data and estimation of resulting complexes, in the study of complex neurodevelopmental syndromes.

More generally : exceptional cognitive and social performances of humans have been op-

timized by means of an ultra-rapid evolutionary process. This practically excludes the potential emergence of novel molecular actors. Rather, optimized cooperation of pre-existing actors : optimization of networking. Actors that are also used for other body functions are optimally organized in the brain. Small failures of the optimized cooperation between such actors can lead to more or less strong cognitive or behavioural divergences, which would remain un-noticed in other animals but are immediately perceived in human communities. Hence the multi-genetic origins of autism and also of even more pronounced syndromes such as those presenting “Angelman-like” features.

## 1.6 Material and methods

### Protein expression and purification

The RLD2 constructs is expressed in *Escherichia coli* BL21. Transformed bacteria are selected by adding 1mM kanamycin. The culture is performed with 0.2% glucose and shaking at 190 rpm. After induction by addition of 0.5mM IPTG (final concentration) the temperature is lowered to 16°C overnight.

Bacterial pellets are suspended in lysis buffer containing 50 mM tris-HCl pH 8, 400 mM NaCl, 5 mM glycerol, 5 mM beta-mercaptoethanol (BME) and protease inhibitors. Cell lysis is performed by sonication for 1 minute 45 seconds in 2 second on/2 second off steps. Then the lysate is centrifuged at 125000g for 45 minutes to obtain the clear lysate. The construct is purified by nickel affinity chromatography. The sample is loaded on a nickel column in a buffer containing 20 mM HEPES pH8, 300mM NaCl, 3mM BME and washed with the same buffer to which 40 mM imidazole is added. The protein is eluted by a buffer containing 250 mM imidazole. The next step is an SP Sepharose ion exchange chromatography. The elution is done by a NaCl gradient (from 100mM to 1000mM NaCl), containing 20 mM HEPES pH8 and 3mM BME. After cleavage by TEV ("Tobacco Etch Virus") protease to remove MBP, the RLD2 domain is further purified by ion exchange chromatography (SP sepharose) and a steric exclusion chromatography step (S75 16/600) in a buffer containing 20 mM HEPES pH8, 300 mM NaCl and 3mM BME. Protein concentration is performed using centricon (10K cutoff) to reach the concentration of 4mg/mL.

### Crystalization and structure determination

RLD2 apo and RLD2 in complex with different peptides were crystallized using vapor diffusion technique in hanging drop. The precipitant solution used contain 20% poly ethylene glycol 3350 and 0.18M tris ammonium citrate. The ratio protein/precipitant solution was 1/1. the cryo protectant solution used contain 20% ethylene glycole 20% poly ethylene glycol 3350 and 0.18M tris ammonium citrate. For the crystallization of RLD2 in complex with

the different peptides. RLD2 (4mg/ml) was mixed with the peptide in a ration of 1molar for 5 peptide (molarity) Data were collected at the Soleil synchrotron (proxima 1 and proxima 2A beam lines) and at the Swiss Light Source (XD6DA). XDS and ccp4 were used for integrate the Data. Model building and refinement was done using Coot and phenix.

### peptide selection and production

Peptide selection was done by searching the motif (DxDKDxD) or degraded motif in the SlimSearch software. This software allow to find the protein containing the motif in unfolded region. Unfolded region is favorable for motif-domain interaction. Peptide for crystallization and fluorescence polarization measurements were produced and purified by Pascal Eberling at the peptide synthesis platform of IGBMC.

### Fluorescence polarization

FP directs measurements were measured diluting RLD2 in buffer containing 25mM HEPES pH 7.5, 150mM NaCl, 1mM TCEP, 0.005% Tween20 and 50nM fluo peptide (Ube3a 10 aa linked to fluoreceine).

FP competitive measurements of the peptides were measured in competition with the same fluo peptide. The non fluo peptide was diluted in mix containing 25mM HEPES pH7.5, 150mM NaCl, 1mM TCEP 0.005% Tween20, 50nM fluo peptide and 3.33mM RLD2. The measurement was performed with plate reader Pherastar. The different dilution for direct and competitive measurement was transferred from 96 well plate to 384 black wells plate in triplicate of 10 ul. Fitting of the FP data was performed using Profit software.

### Conservation

To find the different orthologs, we used different strategy : The first on was to use the NCBI tool to search similar proteins via the global architecture. This strategy works well for UBE3A. In fact, ube3a is the only protein with AZUL domain and HECT domain but this strategy misses the sequences not well annotated. The other strategy used was to do a systematic blast-p of the proteins versus the proteins of the different phylums one by one.

For the sequences conservation of the RLD2 domain of different species, we used the conservation tool of chimera X. This tool calculates the AL2CO score and map it on the structure. This tool use MSA that was performed using clustal omega. Human (O95714), *Saccoglossus kowalevskii* (XP\_006821411.1), *Acanthaster planci* (XP\_022098915.1), *Capitella teleta* (ELT92472.1), *Biomphalaria glabrata* (XP\_013080917.1), *Anoplophora glabripennis* (XP\_018562376.1), *Macrostomum lignano* (PAA75977.1), *Nematostella vectensis* (XP\_001629259.1), *Amphimedon queenslandica* (XP\_019851474.1).

### **amino acids exposition**

the online software Accessible surface area and accessible calculation protein version 1.2 was used to identify the surface of amino acids of RLD2 in contact with the peptide UBE3A. The percentage of the surface of each amino acid of RLD2 domain in contact with the solvent is calculated in presence and in absence of the peptide of BAZ2B (best resolution). By doing the difference we know the percentage of surface of each amino acid of RLD2 in contact with the peptides.

### **Co-immunoprecipitation assays**

HEK293T cells were co-transfected as indicated with the plasmids pEGFP-RLD2 domain of HERC2 and pmCherry-DOCK10 (full length), in their wildtype or mutant version, or control plasmids. 36h after transfection, cells were lysed in lysis buffer [50 mM Tris HCl pH 7.5, 150mM NaCl, 1 mM EGTA, 1 mM EDTA, 0.5% (w/v) Triton X-100, 10% sucrose, 1 mM dithiothreitol (DTT), 1mM Benzamidine, 0.5 mM AEBSF (4-(2-aminoethyl) benzenesulfonyl fluoride hydrochloride)]. For each condition, GFP-RLD2 was immunoprecipitated from 1mg of lysate by addition of 2 $\mu$ g anti-GFP antibody (Torrey Pines Biolabs, TP401), coupled to 40  $\mu$ l of Protein G-coated magnetic beads (Dynabeads Protein G, Invitrogen), and the co-precipitating mCherry-DOCK10 was detected by immunoblotting with a mCherry antibody (Thermo, PA5-34974). Total cell lysates (input) were analyzed with the relevant anti-mCherry and anti-GFP antibodies. Immunoblot detections and band-intensity quantifications were made with the Odyssey system from Li-COR Biosciences.

### **RAC1 and CDC42 activation assays**

The RAC1-GTP and CDC42-GTP pull down assays were performed as described (Jaudon et al, 2015). Briefly, HEK293T cells were transfected with pmCherry-DOCK10 plasmids, either wildtype or mutated, or with a control plasmid. 36h post-transfection, endogenous GTP-RAC1 (or GTP-CDC42) was pulled down by incubating the protein lysates for 1h at 4°C with the CDC42/RAC1-interactive binding (CRIB) domain of PAK1 immobilized on GST-Sepharose beads (GE-Healthcare). Total HEK293T lysates and the corresponding pull-downs were processed for Western blotting using anti-Rac1 (BD, 610650) or anti-CDC42 (BD, 610929) antibodies and anti-GFP (Torrey Pines Biolabs, TP401). The level of GTP-bound (active) RAC1 (or CDC42) protein was assessed by band-intensity quantification using the Odyssey system from Li-COR Biosciences, and normalized to the total amount of RAC1 or CDC42 detected in the cell lysate.

### Primary culture and transfection of hippocampal neurons

Wild-type RjOrl :Swiss pregnant female mice (Janvier Labs, St Berthevin, France) were housed at the animal house facility of the Institut de Génétique Moléculaire de Montpellier (Montpellier, France). Animals had a libitum access to food and water, with 12h-light/dark cycle. The mouse facility has been approved by the Département des pratiques de recherche réglementées “Animaux à des fins scientifiques” – AFiS, under the approval number F3417216.

Neuronal hippocampal cultures were prepared from embryonic day 17.5 RjOrl :Swiss embryos, as described (Bonnet et al, 2023). Neurons were transfected at 8 DIV (Days In Vitro) with cDNA plasmid constructs of pmCherry-DOCK10, wildtype or mutated, or a control plasmid, using the Lipofectamine 2000 reagent (ThermoFisher), according to the manufacturer’s instructions. At DIV11, neurons were fixed with 4% paraformaldehyde/4% sucrose in PBS for 10 min and processed for microscopic examination.

### Immunolabelling and dendritic spine quantification

Fixed neurons were permeabilized in 0.15% Triton-X-100/PBS for 3 min. Immunostaining was performed with GFP antibody (Aves Labs, GFP-1020) and mCherry (Thermo, PA5-34974) and the corresponding Alexa Fluor 488 and 546-conjugated secondary antibodies (ThermoFisher Scientific). All coverslips were mounted in Mowiol reagent (Sigma Aldrich). Images were acquired using a fluorescence microscope (Zeiss Axioimager Z2) and a 63x objective (NA 0.8). Morphometric analyses were performed in different fields from at least three different cultures, using the ImageJ software (National Institutes of Health, Bethesda, MD). Spines were defined as dendritic protrusions with a neck and a head. Dendrites were randomly selected and spines manually counted over a 50- $\mu$ M length of dendrite. Data were then expressed as density of spines/10- $\mu$ m length of dendrite. All experiments were conducted in a blinded manner, and dendritic spines were analyzed from at least 20 neurons from at least three independent cultures.

### Sequences

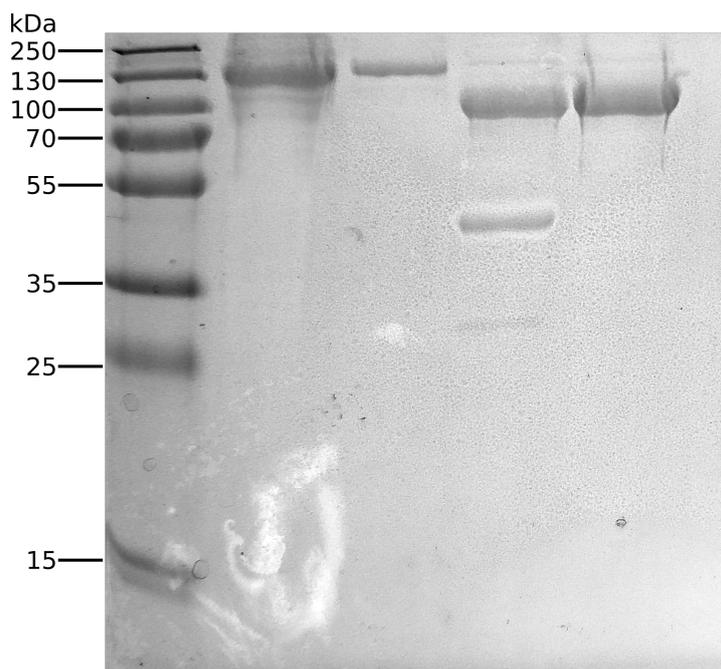
#### 6H-MBP-TEV-RLD2 :

```
MKHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFW
AHDRFGGYAQSGLLAEITPKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSILYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDNAGAKAGLTFLVDLIKHKHMNADT
DYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFL
ENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAAS
GRQTVDEALKDAQTNSSNNNNNNNNNNNNPMSENLYFQGGAMGSLIRKKAAGLESAATIRTKVFWVWGLNDKD
QLGGLKGSKIKVPSFSETLSALNVVQVAGGSKSLFAVTVEGKVVYACGEATNGRLGLGISSGTVPPIRQITALS
SYVVKKVAVHSGGRHATALTVDGKVFVSWGEGDDGKLGHF SRMNC DKPRLIEALKTKRIRDIACGSSHSAAAL
TSSGELYTWGLGEYGRLGHDNTTQLKPKMVKVLGHRVIQVACGSRDAQTLALTDEGLVFSWGDGDFGK
LGRGGSEGCNIPQNIERLNGQGVQCIECGAQFSLALTKSGVVWWTWGKGDYFRLGHGSDVHVRKPQVVEG
LRGKKIVHVAVGALHCLAVTDSGQVYAWGDNHDGQQGNGTTTVNRKPTLVQGLEGGKITRVACGSSHSVA
WTTVDVATPSVHEPVLFTQ
```

#### RLD2 mut :

GSLIRKKAAGLESAATIRTKVFVWGINaaDQLGGLKGseIKVPSFSETLSALNVVQVAGGeeSLFAVTVEGKVY  
ACGEATNGRLGLGISSGTVPIPRQITALSSYVVKVAVHSGGRHATALTVDGKVFWSWEGDDGKLGHF SRM  
NCDKPRLEALKTKRIRDIACGSSHAALTSSGELYTWGLGEYGRGLGHGDNTTQLKPKMVKVLGHRVIQVA  
CGseDAQTLALTDEGLVFSWGDGDFGKLRGGSEGCNIPQNIERLNGQGVCQIECGaqFSLALTKSGVVWT  
WGeGdtFqLGHGSDVHVRKPVVEGLRGKKIVHVAVGaIHCLAVTDSGQVYAWGrNaaGQQGNGTTTTVNRK  
PTLVQGLEGQKITRVACGssHSAVATTVDVATPSVHEPVLFT

## 1.7 Supplemental



**FIGURE 41** – acrylamide gel of the different steps of the Ube3a protein purification. From left to right : after nickel and ion exchange column, after selection of monomeric protein by SEC, after TEV protease cleavage, and after SEC

Protein	Sequence	Start	End
PCM1	ageidDEDKDKDetetv	1737	1743
RCN2	drfvnDYDKDNDgrldp	238	244
ARID4A	ikkqeDSKDSDeeek	537	543
USP35	wgrgfDEDKDEdegspg	990	996
CALR	edkedDEDKDEDeede	388	394
MYT1	vevrsDDDKDEDthsrk	349	355
DOCK10	eidheDADKDEDttshs	297	303
ARIP4	fssndDEDKDDDvietv	1454	1460
MIGA1	nsladDIDKDTDitmkg	297	303
Ube3a	slqakDEDKDEDekeka	185	191
BAZ2B	seeedDDDKDQDesdsd	651	657
SMC6	qeeleDFDKDGDedeck	25	31
MED1	kdrdrDRDKDRDkkksh	1515	1521
RERE	mtaDKDKDKDkekdr	4	10

**TABLE 6** – Proteins containing DxDKDxD identified by SlimSearch

Protéine	Peptide name	Sequence	kd in $\mu$ M
	Ube3a 173-187	HTKEELKSLQAKDED	23.69
	Ube3a 178-192	LKSLQAKDEDEKDEDE	107.03
	Ube3a 183-197	AKDEDEKDEDEKEKAA	6.36
	Ube3a 188-202	KDEDEKEKAACSAAA	226.41
	Ube3a 193-207	KEKAACSAAA MEEDS	no binding
	Ube3a 198-212	CSAAAMEEDSEASSS	no binding
	Ube3a 203-217	MEEDSEASSSRIGDS	no binding
	Ube3a 208-222	EASSRIGDSSQGDNN	no binding
	Ube3a 150-200	HTKEELKSLQAKDEDEKEKAACSAAA MEEDSEASSSRIGDSSQGDNN	4.59
	Ube3a 157-174	HTKEELKSLQAKDEDEKDEDEKEKAA	<0.5
	ube3a full length	Uniprot : Q05086	1.52
DOCK10	Dock10 156-170	HEDADKDEDTTSHSS	3.59
	Dock10 146-170	KLPSHSFEIDHEDEADKDEDTTSHSS	5.29
USP35	Usp35 988-1002	GFDEDEKDEDEGSPGG	9.73
PCM1	Pcm1 1737-1751	DEDKDKDETETVKQT	8.22
BAZ2b	Baz2b 649-663	EDDDDKDQDESDDT	12.18
Myt1	Myt1 347-360	RSDDDKDEDTHSRK	31.88
	ARID4A 535-549	QEDSDKDSDEEEES	97.19
	ARID4A 525-549	PKQKEKKIKKQEDSKDSDEEEES	21.07
ARID4A	ARID4A 525-549 pi1 pi2	PKQKEKKIKKQEDpSDKDPSEEEES	14.44
	ARID4A 525-549 p1	PKQKEKKIKKQEDpSDKDSDEEEES	10.78
	ARID4A 525-549 pi2	PKQKEKKIKKQEDSDKpSDEEEES	30.83
	ARID4 638-332	KKKQKKAKNKEDSEKDEKRDEERQ	27.21
ARIM4	Arip4 152-167	NDDEDKDDVIEVTGK	26.94
REER	REER 0-15	MTADKDKDKKDKDR	32.21
Erbin	Erbin 462-476	FESDEDKDEREAPPR	311.64
ARMC9	Armc9 598-612	EADLDKDELIQPQLG	156.16
SMC6	Smc6 23-37	LEDFDKGDEDECKG	138.9
Rfx1	Rfx1 918-932	PLDPDKDEEEEEEE	103.43
phf201	Ph201 198-213	SLQAKDEDEKDEKEKAA	no binding
Med 1	Med1 1513-1528	DRDDRDKDRDKKSHS	no binding

TABLE 7 – Measurement of the affinity of different peptides for the RLD2 domain

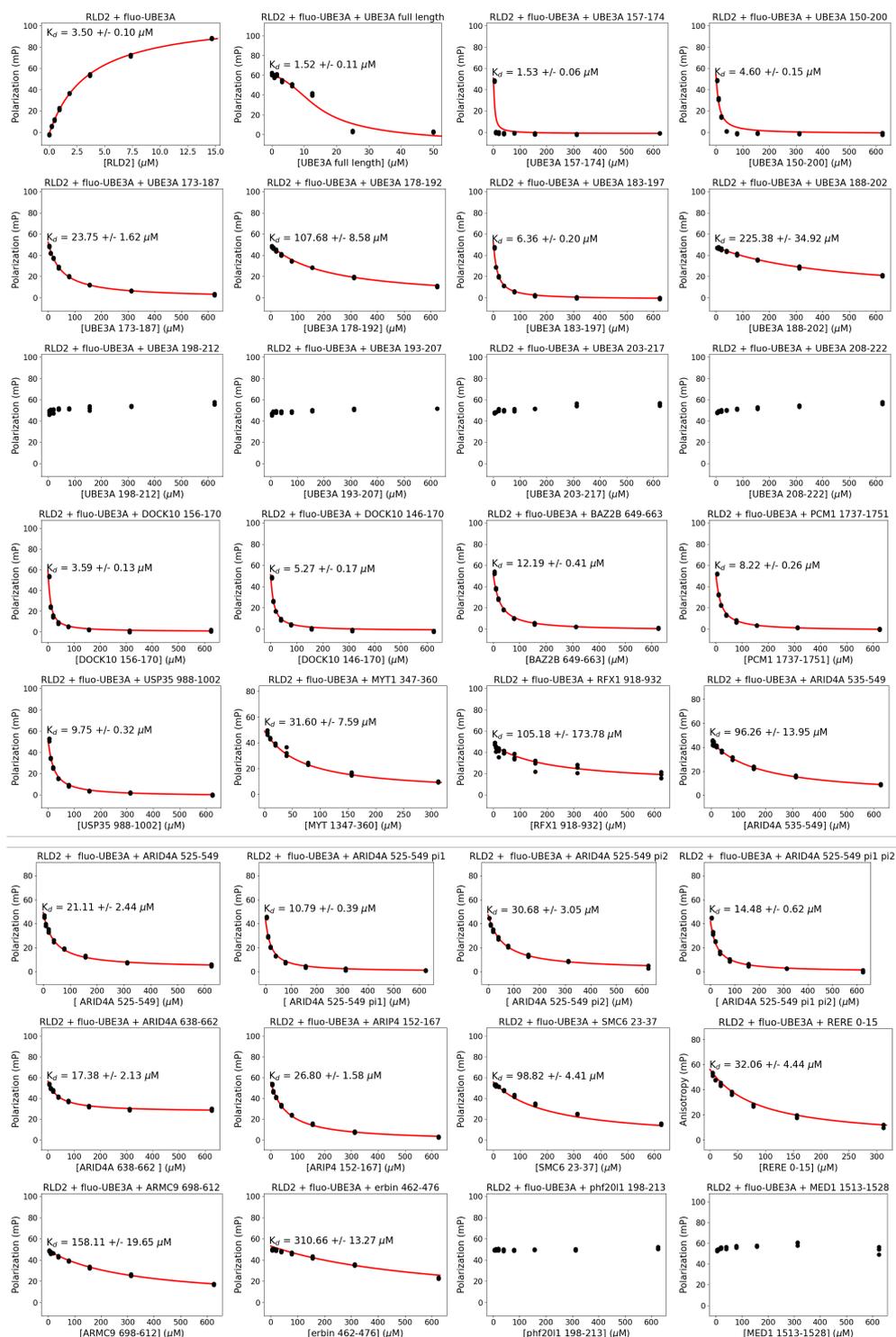


FIGURE 42 – Measurement of the affinity of different peptides for the RLD2 domain

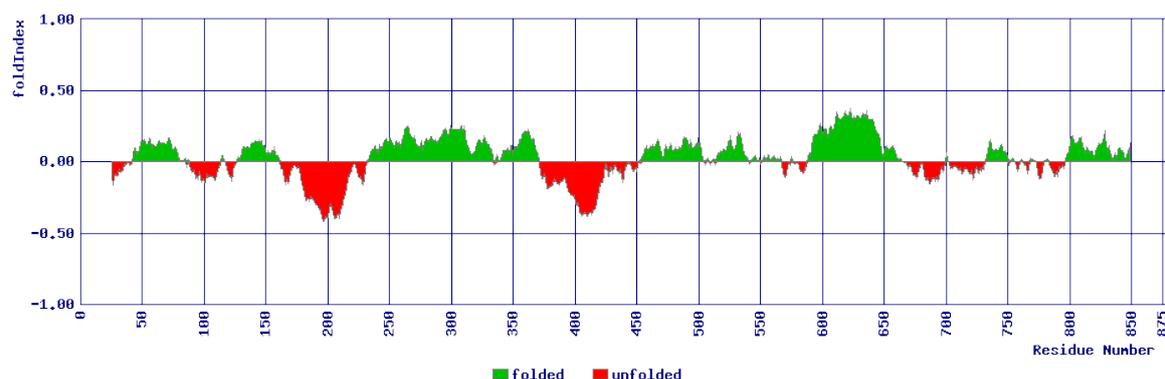
## 2 | Étude structurale de Ube3a

### 2.1 Introduction

#### 2.1.1 Stratégie

Le second aspect de ce projet de thèse porte sur l'étude structurale de la protéine Ube3a. Comme vu précédemment, seules les structures de deux de ses domaines, HECT et AZUL, sont connus. La stratégie mise en œuvre pour tenter d'obtenir des informations structurales sur la protéine Ube3a est l'expression et la purification de différentes constructions de la protéine Ube3a. Une fois capable de purifier ces différentes constructions, nous avons effectué des tests de cristallisation.

La conception des différentes constructions a été faite en utilisant des prédictions de zones désordonnées de la protéine Ube3a. Pour cela, nous avons utilisé des logiciels tels que MeDor ou foldindex (Figure 43, page 86).



**FIGURE 43** – Graphique représentant les zones prédites comme non repliées de la protéine Ube3a. Ce graphique, produit en utilisant le logiciel Foldindex [295], représente en abscisses la numérotation des acides aminés de Ube3a et en ordonnées le score de repliement. Une zone ayant un score supérieur à zéro (en vert) est prédite comme repliée, Une zone ayant un score inférieur à zéro est prédite comme désordonnée (en rouge). Notez que les zones contenant les motifs permettant l'interaction avec RLD2 et avec E6 sont prédites comme désordonnées.

Ces prédictions ont permis de déterminer des zones non repliées. C'est notamment le cas de la zone d'interaction avec E6 contenant le motif LxxLL, de la zone d'interaction avec le domaine RLD2 de HERC2 ainsi que d'une zone en carboxy-terminal du domaine AZUL. Ainsi, nous avons identifié trois domaines que nous avons appelés DUS-1, DUS-2 et DUS-3 ("Domain of Unknown Structure"). Différentes constructions de Ube3a ont donc été produites grâce à ces données (Figure 46 page 90). Nous avons élaboré des constructions des domaines DUS seuls, mais aussi avec des fragments plus longs de Ube3a contenant le domaine catalytique (DUS2-DUS3-HECT et DUS3-HECT).

Au cours de ce projet, nous nous sommes de plus en plus intéressés aux prédictions de structure. En effet, au-delà des prédicteurs de zones désordonnées que nous avons utilisés au début de ce projet, d'autres outils permettant d'obtenir des données structurales de manière non expérimentale ont émergé. C'est notamment le cas de RosettaFold et de AlphaFold puis AlphaFold 2.

### 2.1.2 AlphaFold 2, un tournant ?

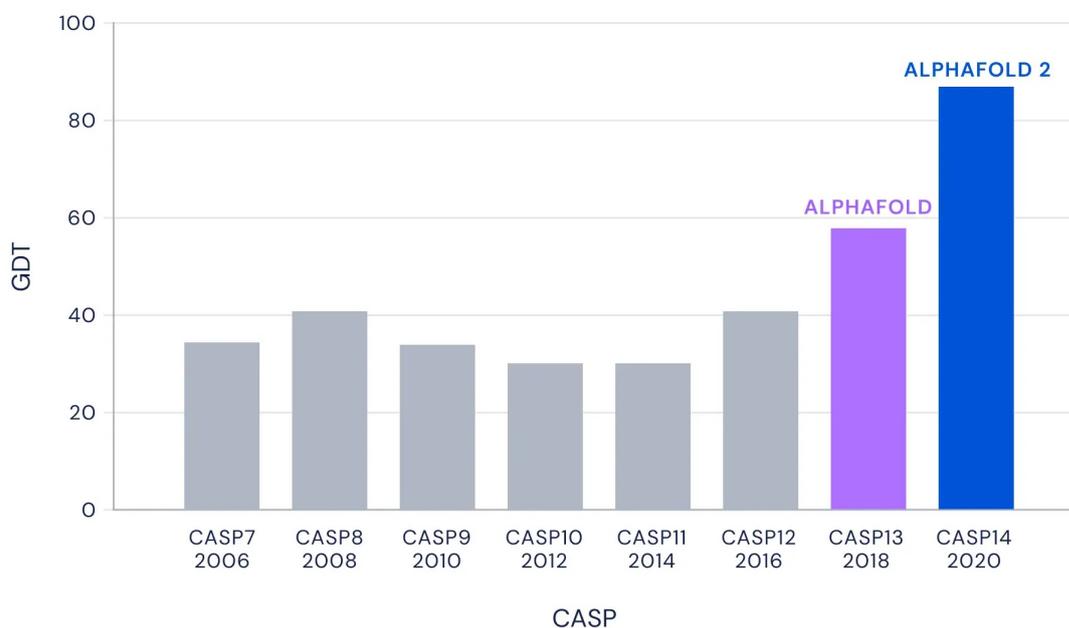
L'arrivée du logiciel AlphaFold 2 a été une petite révolution dans le domaine de la biologie structurale. Ce logiciel prend en entrée une séquence protéique et produit plusieurs prédictions de structure sous la forme de fichiers PDB. En plus de cette structure, AlphaFold génère un score de confiance global, un score de confiance acide aminé par acide aminé et une prédiction de l'erreur de position relative pour chaque paire d'acides aminés (en ångström).

C'est en 2021 qu'a été publié le papier sur AlphaFold 2 de Deepmind (Google) [296]. Ce logiciel a notamment gagné le concours CASP ("Critical Assessment of Structure Prediction") en 2018 avec AlphaFold 1 puis en 2020 avec AlphaFold 2. C'est un concours où s'affrontent des équipes travaillant sur les prédictions de structures (Figure 44, page 88).

Ce logiciel utilise une approche d'apprentissage profond ("deeplearning"). Il utilise deux méthodes différentes de manière simultanée, et les informations résultantes de ces deux méthodes sont utilisées de nouveau dans chacune de ces deux méthodes. La première méthode est basée sur des alignements multiples de séquences similaires à celle entrée dans le logiciel. La seconde est basée sur l'appariement ("pairing") des acides aminés de la séquence d'entrée via les structures connues déposées sur la PDB ("Protein Data Bank").

La première étape du logiciel consiste d'un côté en la sélection des séquences similaires à la séquence d'entrée (protéines homologues, paralogues et orthologues) et en la génération d'un alignement multiple de séquences (MSA). D'un autre côté, le logiciel AlphaFold 2 sélectionne les structures déjà connues ayant une séquence proche de la séquence

### Median Free-Modelling Accuracy



**FIGURE 44** – Amélioration de la précision médiane des prédictions dans la catégorie de la modélisation libre pour la meilleure équipe de chaque CASP, mesurée en GDT. Le GDT ("Global Distance Test") est un score qui permet de déterminer la similarité de structure de deux modèles ayant la même séquence, un GDT de 100 correspond à deux structures identiques [297]. Figure de [298].

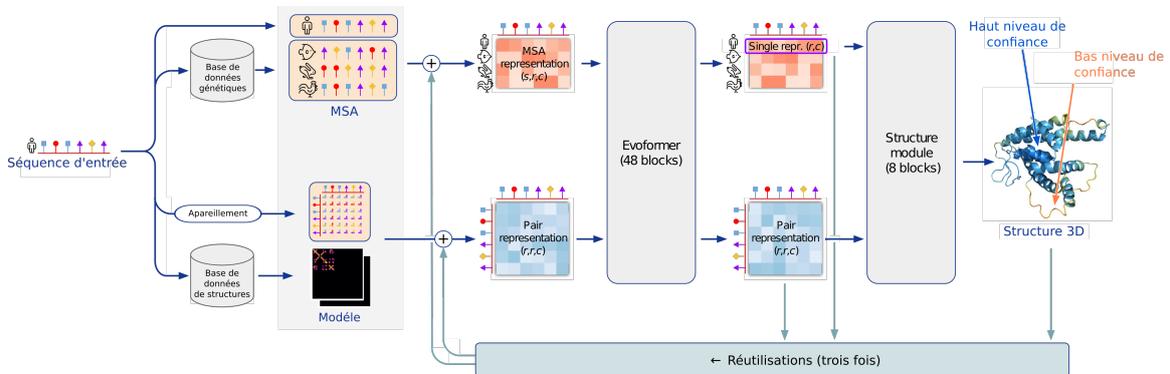
d'entrée et va ainsi en retirer une matrice de données d'appariement pour chaque paire d'acides aminés.

Ensuite, ces données vont être utilisées dans deux modules. Le premier module est appelé EvoFormer. Il améliore les données de MSA et d'appariement. L'amélioration des données de MSA est notamment effectuée par l'analyse des séquences par ligne (en simplifiant, le logiciel identifie si deux acides aminés ont co-évolué et pourraient être mis en jeu dans une même interaction) et colonne par colonne (en simplifiant, certains acides aminés pourraient être conservés dans toutes les séquences et joueraient un rôle capital dans le repliement).

Enfin, AlphaFold utilise le module de structure. Celui-ci utilise les données sortant du module EvoFormer pour donner des coordonnées dans l'espace à chaque acide aminé. Ce module ne prend pas en compte de restrictions structurales comme les angles, les distances ou encore les "clashes" stériques.

Ces deux modules sont utilisés de manière itérative. À la fin des itérations, une étape de

relaxation utilisant le champ de forces AMBER permet d'éliminer les potentiels "clashes" stériques ainsi que les angles et distances non autorisés.

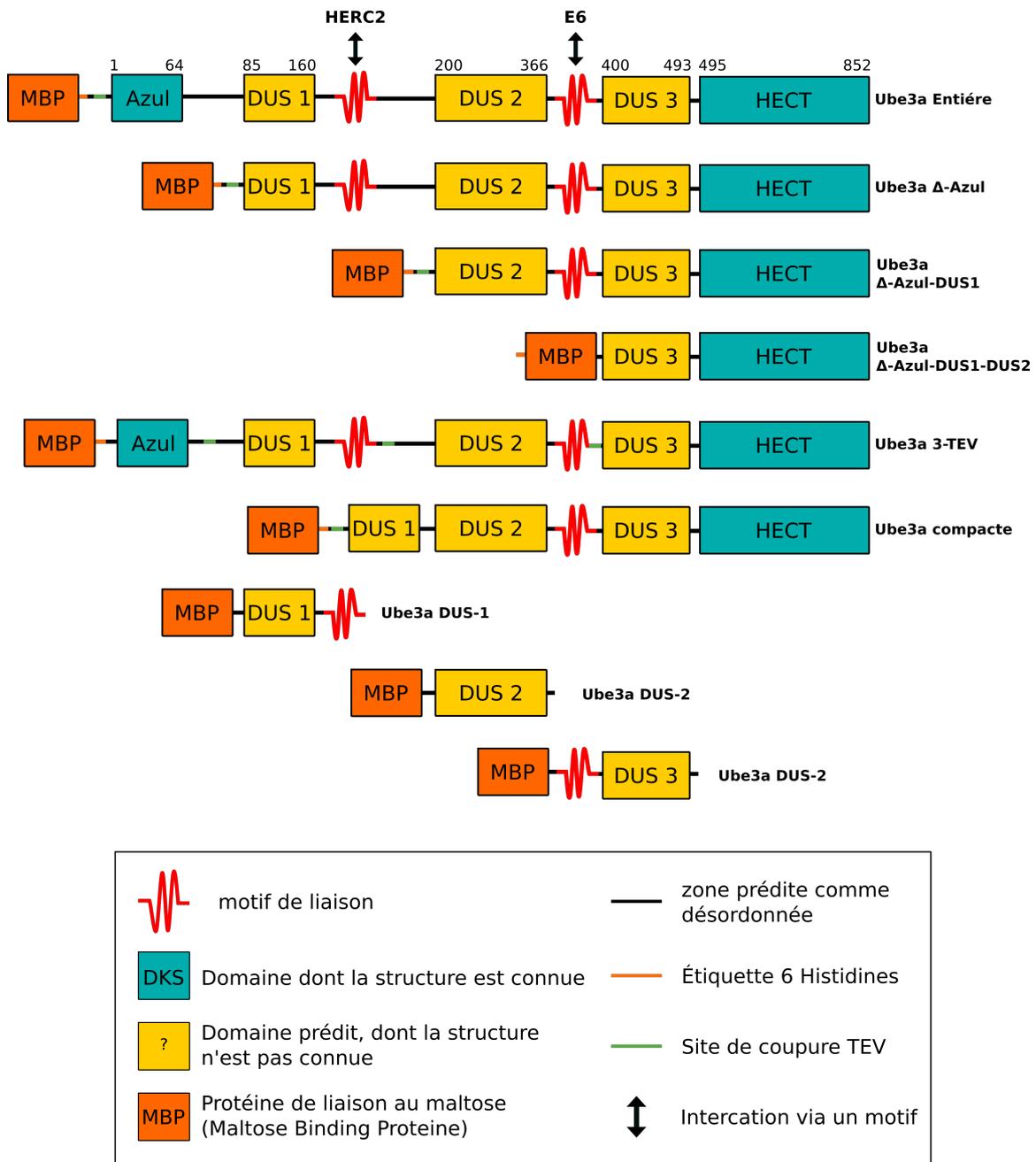


**FIGURE 45** – Architecture du logiciel AlphaFold 2. Les flèches montrent le flux d'informations entre les différents composants d'AlphaFold.  $s$  : nombre de séquences ;  $r$  : nombre de résidus ;  $c$  : nombre de canaux.

## 2.2 Contributions

Dans les résultats présentés, j'ai optimisé avec Eduardo Howard et Irina Paula Suarez le protocole de purification de Ube3a et ses différentes constructions. J'ai purifié les différentes constructions de Ube3a et effectué les différents tests de cristallisation. Les différentes constructions ont été clonées par Camille Kostmann Soufflet. J'ai utilisé AlphaFold pour produire les différentes prédictions de structures ainsi que l'analyse de ces prédictions. Des tests de microscopie électronique (non présentés ici) ont été réalisés avec Irina Paula Suarez. Les purifications de la protéine 16E6 et du domaine Ubl de RAD23A ont été réalisées par Alexandra Cousido-Siah. Les peptides de 16E6 ont été produits par Pascal Eberling.

## 2.3 Résultats



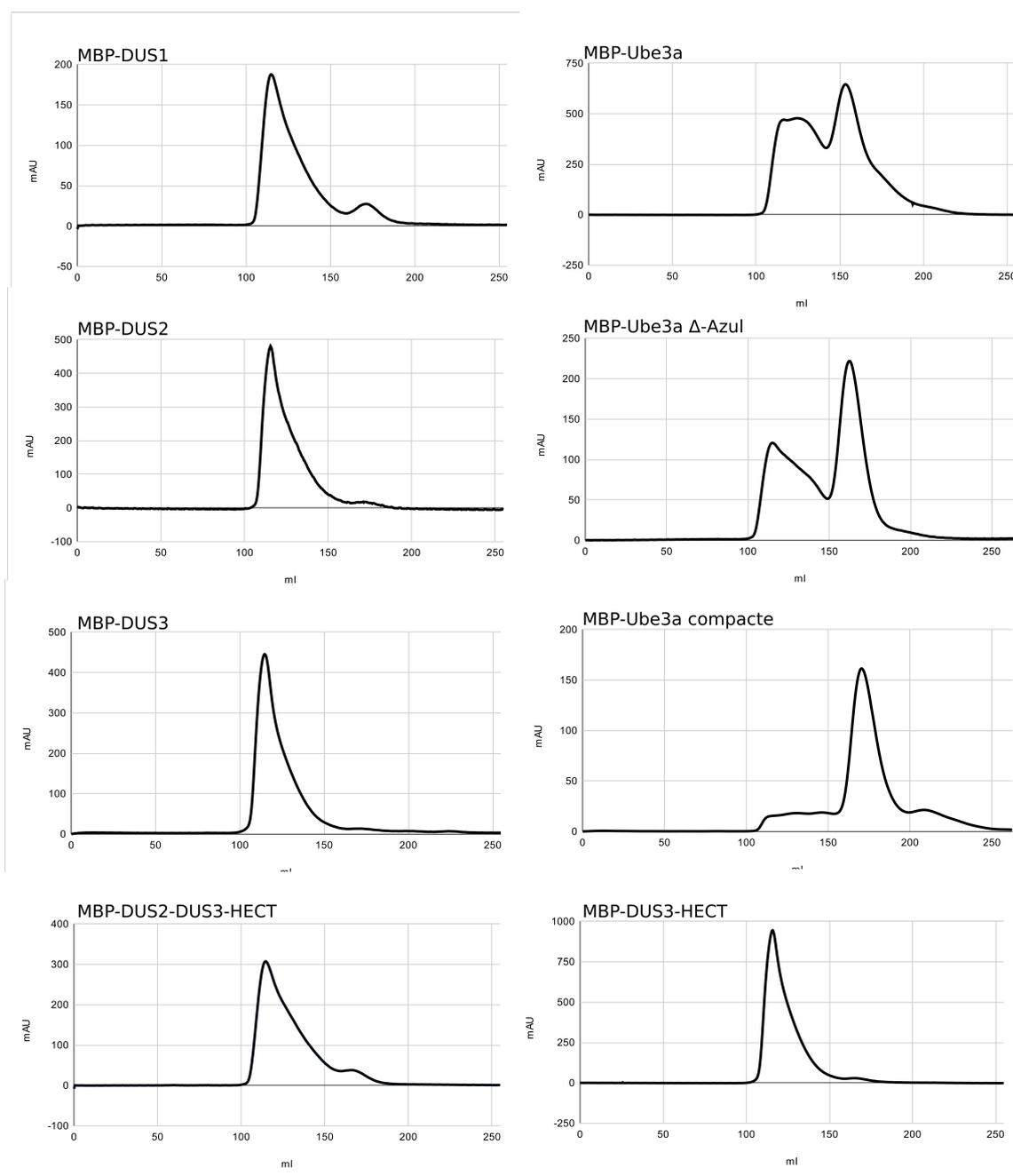
**FIGURE 46** – Schéma représentant les différentes constructions de Ube3a ainsi que leurs domaines connus et supposés.

Les différentes constructions de Ube3a (Figure 46, page 90) ont été purifiées en utilisant la méthode décrite dans la partie matériel et méthodes (section 1.2.2, page 114).

L'étape de SEC ("Size Exclusion Chromatography") avant la coupure à la TEV est capi-

tale dans la purification. En effet, en plus d'être une étape de purification qui permet de séparer les particules selon leur taille, cette étape permet de contrôler la qualité de la purification et le comportement de la construction. Elle permet d'identifier si la construction se purifie sous la forme monomérique ou sous la forme multimérique/d'agrégats solubles. Ici, la SEC utilisée, une Superdex (S200 26/60), permet une séparation des particules de 10 kDa à 600 kDa (kilo dalton). La SEC sépare en fonction du volume hydrodynamique. Ne connaissant pas le volume hydrodynamique de ces particules, nous approximons en utilisant la masse en kDa qui, elle, est connue. Ainsi, les particules de plus de 600 kDa ne sont pas résolues et sont éluées dans le volume mort de la colonne.

Tout d'abord, les constructions des domaines DUS-1, DUS-2 et DUS-3 ont été purifiées pour tenter de les cristalliser et ainsi obtenir des données structurales. Les purifications de ces trois domaines non connus MBP-DUS1, MBP-DUS2 et MBP-DUS3 ont été effectuées par colonne d'affinité amylose et colonne échangeuse d'ions. Après cela, dans le cas de ces trois différentes constructions, les profils d'élutions de SEC sont similaires. Un premier pic se trouve au volume mort de la colonne (26/60 S200). Ce qui est élué dans le volume mort de la colonne n'a pas pu être séparé. Ce volume correspond à une taille de protéine supérieure à 600 kDa. Les trois constructions, MBP-DUS1, MBP-DUS2 et MBP-DUS3 ont des tailles respectives de 56 kDa, de 62 kDa et de 54 kDa. Ainsi, ces différentes constructions ne se purifient pas sous forme monomérique, et semblent former des multimères ou agrégats solubles. (Figure 47, page 92).



**FIGURE 47** – Profils d'éluion des différentes constructions de Ube3a. Ces profils d'éluion représentent en ordonnées l'absorbance à 280 nm et en abscisses le volume d'éluion. Ces profils sont tous issus de l'éluion de la SEC 26/60 S200 (Cytiva) après une purification par colonne d'affinité et échangeuse d'ions, Q sepharose.

La construction entière de Ube3a ne se purifie pas non plus uniquement sous la forme monomérique (Figure 47, page 92). En effet, l'étape de SEC avant coupure à la TEV produit deux pics. Ici, on remarque un pic au volume mort correspondant à la limite séparative de la colonne. Le second pic, à 155 ml, correspond, quant à lui, à la forme monomérique

de la protéine Ube3a.

Le profil d'élution de la construction Ube3a sans son domaine AZUL (que l'on appellera  $\Delta$ -AZUL) est comparable à celui de Ube3a entière. Un premier pic d'élution au volume mort et un second à un volume de 160 mL correspondant à la taille de la construction monomérique (Figure 47, page 92).

Le second pic de l'élution, correspondant à la protéine monomérique, a été utilisé pour la suite de la purification. Elle consiste en la séparation de la MBP ("Maltose Binding Protein") par la coupure au niveau du site TEV. Les constructions de Ube3a et  $\Delta$ -AZUL ont été purifiées sous forme monomérique, et de nombreux tests de cristallisation ont été effectués. Ces tests de cristallisation ont été faits avec et sans la MBP, mais aussi avec et sans E6 et RLD2. Ces différents tests de cristallisation n'ont jamais permis l'obtention de cristaux.

Il est intéressant de noter que les monomères de Ube3a sont stables. En effet, après la coupure TEV à température ambiante, une seconde étape de SEC ne laisse apparaître qu'un pic correspondant à une forme monomérique. Ceci indiquerait la stabilité des monomères de Ube3a dans les conditions de la purification (contenant 3mM de BME, un antioxydant).

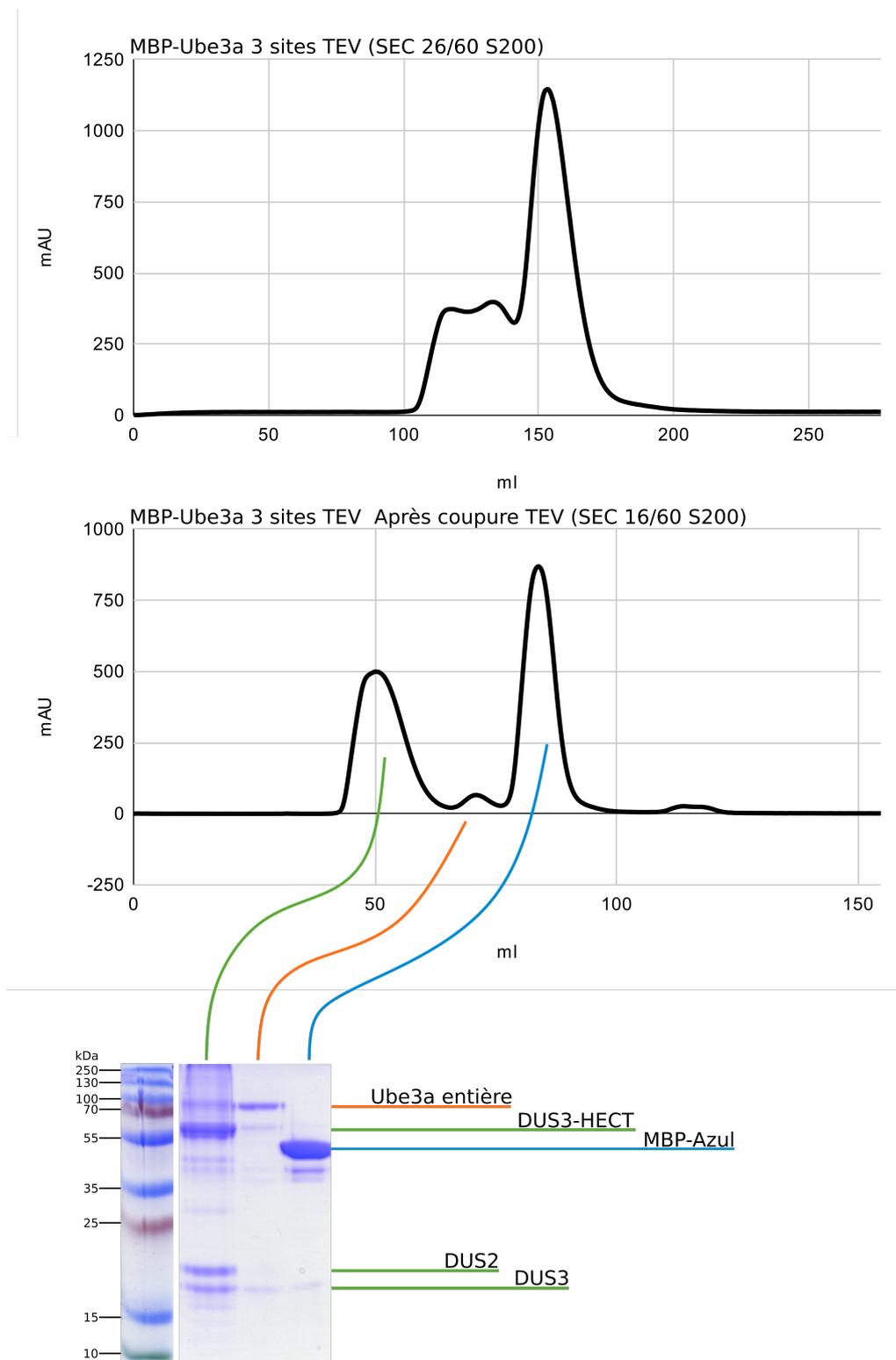
Ne parvenant pas à purifier les différents domaines DUS seuls sous forme monomérique et les constructions de Ube3a entière et Ube3a  $\Delta$ -AZUL ne cristallisant pas, nous avons purifié la construction contenant le domaine HECT et le domaine DUS-3 ainsi que la construction contenant le domaine HECT, le domaine DUS-3 et DUS-2.

Les purifications de ces deux constructions donnent des profils d'élution (avant coupure TEV) similaires aux domaines DUS-1, 2 et 3 seuls. Ainsi, les constructions dépourvues du domaine DUS-1 et dépourvues des domaines DUS1 et DUS2, ne se purifient pas sous forme monomérique. (Figure 47, page 92) Ceci indiquerait que ces différents domaines font partie d'un repliement global de la protéine et que l'absence d'un ou plusieurs de ces domaines déstabilise ce repliement et entraînerait la formation de multimères voir d'agrégats. Des tests de SEC ont tout de même été effectués avec les fractions multimériques de ces constructions, et ont permis l'obtention de cristaux pour deux constructions. Cependant, La diffraction trop faible n'a pas permis l'obtention de structure.

Pour valider la supposition d'un repliement global impliquant les domaines DUS-1, DUS-2, DUS-3 et HECT, nous avons purifié une autre construction de la protéine Ube3a. Cette construction consiste en une Ube3a entière, fusionnée en position amino-terminale à la MBP et possédant trois sites de coupure à la TEV. Un premier site de coupure se trouve entre le domaine AZUL et le domaine DUS-1. Le second site est positionné entre le domaine DUS-1 et le domaine DUS-2. Le dernier site TEV est positionné entre le do-

maine DUS-2 et le domaine DUS-3. Après sélection de la forme monomérique permise par l'étape de SEC et la coupure aux niveaux des sites TEV, cette construction produit quatre fragments : la fusion de la MBP et du domaine AZUL, le domaine DUS-1 seul, le domaine DUS-2 et un fragment composé du domaine DUS-4 et du domaine HECT. (Figure 47, page 92)

Après la coupure TEV, une étape de SEC (16/60 S200) permet de voir si les différents fragments de Ube3a sont élués ensemble. Cette élution révèle deux pics. Un gel acrylamide permettant de séparer les différents fragments de chaque pic ainsi qu'une analyse par spectrométrie de masse ont permis d'identifier les trois fragments DUS-1, DUS-2 et DUS-3/HECT dans les mêmes fractions d'élution. Ces trois fragments ont des tailles différentes et sont élués dans un volume d'élution correspondant à leur tout. Un second pic correspond à l'élution du fragment MBP-AZUL seul (Figure 48, page 95). Ainsi les trois domaines DUS-1, 2 et 3 ainsi que le domaine HECT semblent être un même domaine séparé par des boucles non structurées.

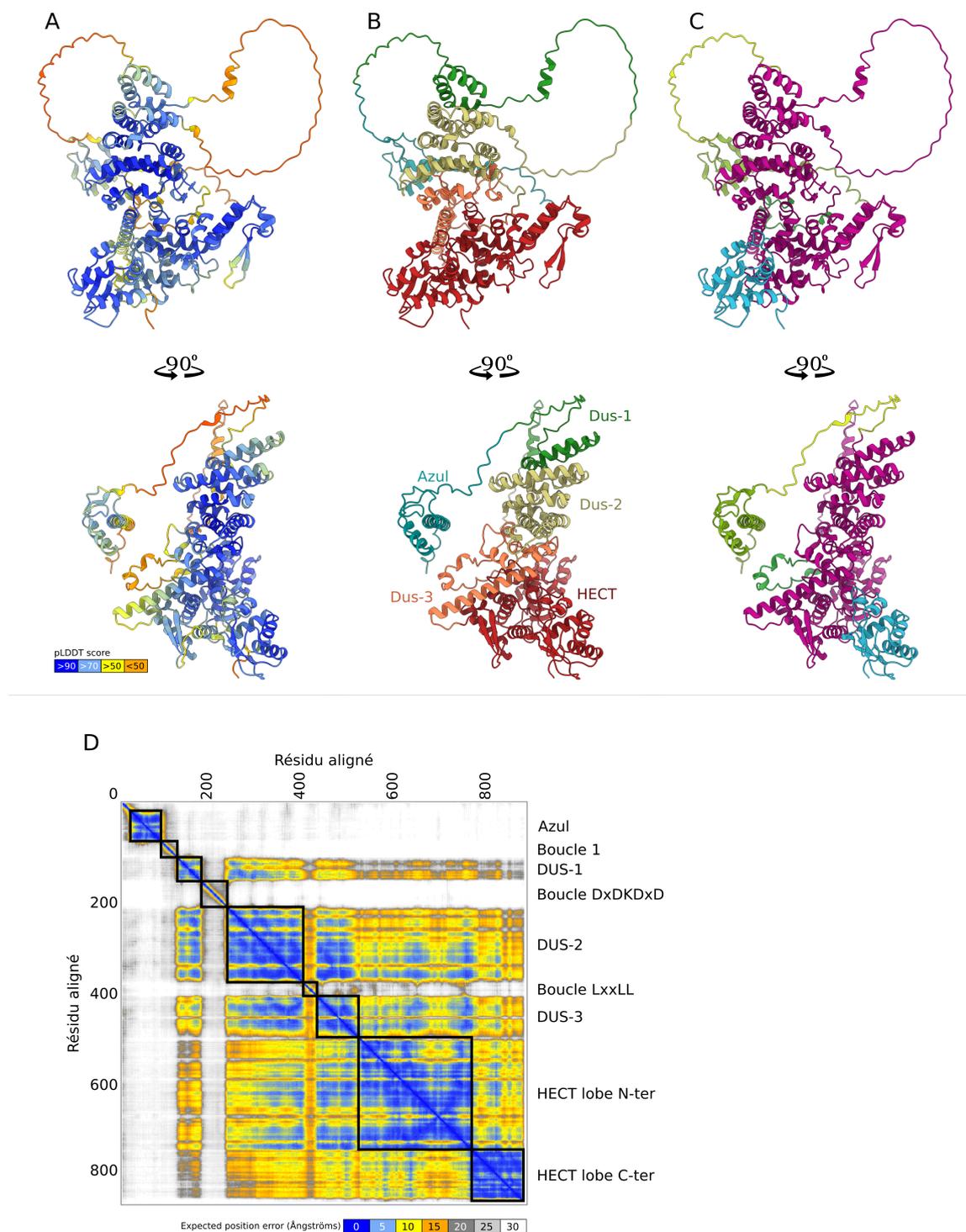


**FIGURE 48** – Profils d'éluion de la construction Ube3a contenant 3 sites TEV avant et après coupure à la TEV. En haut, le pic d'éluion monomérique 155 ml avant coupure (SEC 26/60 S2000) a été sélectionné pour la coupure TEV. Sur le second profil d'éluion, en bas, après coupure TEV (SEC 16/60 S200) les domaines DUS-1 DUS-2 et DUS-3-HECT sont élués ensemble à 50 ml. Le domaine AZUL fusionné à la MBP est élué seul à un volume de 80ml. Les différents fragments sont identifiés par gel acrylamide.

Cette idée d'un domaine formé de DUS 1, 2, 3 et du domaine HECT a été validée par la prédiction de la structure de Ube3a par le logiciel AlphaFold 2. En effet, l'arrivée du logiciel AlphaFold 2 nous a permis de générer différentes prédictions de structures de Ube3a. La prédiction de la structure de Ube3a entière (Figure 49 A page 98) est en accord avec les différentes observations et informations que nous avons déjà sur Ube3a. Cette structure prédite par AlphaFold 2 est formée d'un bloc composé du domaine DUS-1, DUS-2, DUS-3 et du domaine HECT. Le domaine AZUL semble être un domaine n'interagissant pas avec le reste de la protéine. (Figure 49 B page 98).

Le score de qualité de la prédiction (score pLDDT) des acides aminés du domaine AZUL et des domaines DUS-1, 2, 3 et HECT est élevé. Ce score est plus bas pour les zones prédites comme non structurées. C'est le cas de la longue boucle reliant le domaine AZUL à DUS-1, de la longue boucle possédant le motif DxDKDD entre les domaines DUS-1 et DUS-2 ainsi que de la boucle contenant le motif LxxLL. Il est normal que les zones non structurées aient un score plus faible (Figure 49 A page 98). Ainsi, la structure des différents domaines semble de bonne qualité.

La matrice de score PAE ("Predicted Aligned Error") (Figure 49 D page 98) montre l'estimation de l'erreur relative en distance d'une paire d'acides aminés. Pour une paire d'acides aminés (x,y) il s'agit de la prédiction de l'erreur de position de l'acide aminé y en ångström si l'acide aminé x de structure prédite était aligné à ce même acide aminé x de la structure réelle (non connue). Ici, on remarque les différents blocs de faibles erreurs correspondant aux différents domaines identifiés (le domaine AZUL, les domaines DUS ainsi que le domaine HECT). De plus, l'erreur prédite est très faible pour les acides aminés impliqués dans l'interaction entre les différents domaines DUS-1, 2, 3 et HECT. Ceci indique que la prédiction de la position relative de ces différents domaines semble sûre. A contrario, la prédiction de l'erreur de position des acides aminés du domaine AZUL par rapport aux autres domaines est élevée, indiquant que le logiciel AlphaFold ne semble pas savoir où placer le domaine AZUL relativement aux autres domaines, ce qui concorde avec une non-interaction entre AZUL et les autres domaines. Cette matrice peut être utilisée pour déterminer de nouveaux domaines. Cette détermination de domaines ne se base pas sur la fonction des domaines, mais sur la prédiction d'erreur et donc des zones rigides. Ainsi, le logiciel ChimeraX, en utilisant ces données, reconnaît trois domaines et deux boucles : le domaine AZUL, un domaine composé des domaines DUS-1, 2, 3 ainsi que le lobe-N du domaine HECT et un dernier domaine composé du lobe-C du domaine HECT (Figure 49 C page 98). Le fait que les deux lobes du domaine HECT soient identifiés comme faisant partie de deux domaines distincts correspond bien à la mobilité connue des deux lobes l'un par rapport à l'autre lors des différentes étapes de l'ubiquitination d'un substrat.



**FIGURE 49** – Structure prédite par AlphaFold 2 de la protéine Ube3a (PDB : AF\_AFQ05086F1). A) représente la structure de Ube3a dont les acides aminés sont colorés en fonction du score de confiance par résidu pLDDT ("per-residue confidence score"). B) représente la même structure où sont affichés de couleurs différentes les différents domaines que nous avons identifiés au début du projet (AZUL, DUS-1, 2, 3 et le domaine HECT). C)

représente la même structure avec les domaines prédits par ChimeraX en utilisant la matrice PAE (Predicted aligned error). D) Sur cette matrice la couleur aux coordonnées x,y indique l'erreur de position attendue par AlphaFold au niveau du résidu x si la structure prédite et la structure réelle étaient alignées sur le résidu y. Les différents domaines que nous avons identifiés au début du projet sont encadrés. Notez que le domaine AZUL est bien défini mais l'erreur de sa position par rapport au reste de la protéine est de plus de 30Å.

Les connaissances obtenues par la purification des différentes constructions de Ube3a ainsi que cette prédiction de structure nous ont permis de concevoir une nouvelle construction de Ube3a. Le but de cette nouvelle construction est d'obtenir une protéine Ube3a compacte, ce qui pourrait favoriser sa cristallisation. Cette construction est débarrassée du domaine AZUL, et de la Boucle entre DUS-1 et DUS-2 (contenant le motif DxDKxD). Une prédiction de structure de cette construction a été faite grâce à AlphaFold (Figure 57, page 110).

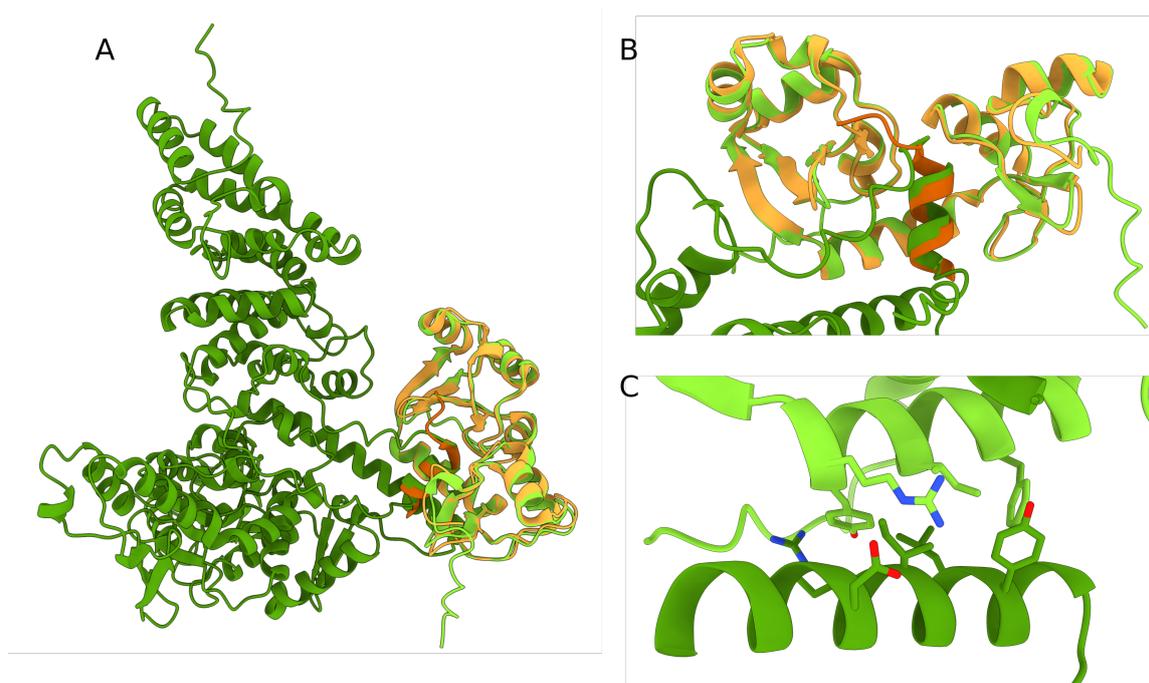
Ainsi, la purification de cette Ube3a compacte nous a permis d'obtenir un profil d'élution différent de ceux obtenus jusqu'alors. En effet, le profil d'élution de la première SEC, avant la coupure TEV, ne contient pas de pic au niveau du volume mort, ce qui indique que cette construction ne produit pas de multimères. (Figure 48, page 95)

Après coupure TEV pour se débarrasser de la MBP, la construction Ube3a compacte a été utilisée dans des tests de cristallisation seule, avec le domaine Ubl (ubiquitine like) de la protéine RAD23A, et avec E6 de HPV16. Des cristaux ont été obtenus avec Ube3a compacte seule. Malgré des optimisations pour obtenir des cristaux de plus grande taille et de meilleure qualité, les cristaux sont encore petits et non monocristallins. Des tests préliminaires de diffractions aux rayons X n'ont pas permis de déterminer si ces cristaux sont des cristaux de sel ou de la protéine. Des cristaux ont aussi été obtenus avec le domaine Ubl de RAD23A mais n'ont, à ce jour, pas encore été testés aux rayons X.

AlphaFold 2 a été utilisé pour générer différentes prédictions de structure du complexe entre Ube3a et la protéine 16E6 (Figure 50 A et B page 99). Les différentes prédictions de Ube3a entière avec 16E6 ne sont pas concluantes. La position relative entre E6 et Ube3a est prédite avec une forte erreur et ne correspond pas aux structures connues de l'interaction.

La prédiction de structure de Ube3a compacte avec 16E6 semble beaucoup plus fiable et concorde avec la structure obtenue par cristallographie [179] [180]. AlphaFold 2 prédit correctement la conformation en hélice  $\alpha$  prise par le peptide contenant le motif LxxLL de Ube3a. En plus de la liaison de E6 au motif LxxLL de Ube3a, E6 semble interagir avec une hélice du domaine DUS-3 (Figure 50 A et C page 99). Cette interaction semble impliquer les résidus suivants : Glu 504, Arg 506, Ile 507, Leu 510 et Tyr 511 de Ube3a

et Phe 76, Ile 80, Tyr 83 et Arg 84 de E6. Cette région de Ube3a additionnelle au motif LxxLL, impliquée dans cette interaction, avait déjà été identifiée en 2020 [181]. L'estimation de l'erreur de distance entre les acides aminés impliqués dans l'interaction entre l'hélice LxxLL et E6 est très faible. Cette estimation d'erreur de distance entre les acides impliqués dans cette nouvelle surface d'interaction est bien plus importante. Ceci pourrait s'expliquer par l'absence de structure connue de cette interaction dans la PDB.



**FIGURE 50** – Figure représentant la structure cristallographique de E6 de HPV (orange foncé) en complexe avec le motif LxxLL de Ube3a (orange clair) superposé à la prédiction de AlphaFold 2 de l'interaction entre la construction compacte de Ube3a (vert foncé) et la protéine E6 de HPV16 (vert clair). AlphaFold 2 ne parvient pas à générer un modèle concordant à la structure cristallographique en présence du domaine AZUL et de sa boucle. A) Structure globale de ce complexe. B) Zoom de la région résolue par cristallographie. C) Surface d'interaction additionnelle entre Ube3a et E6.

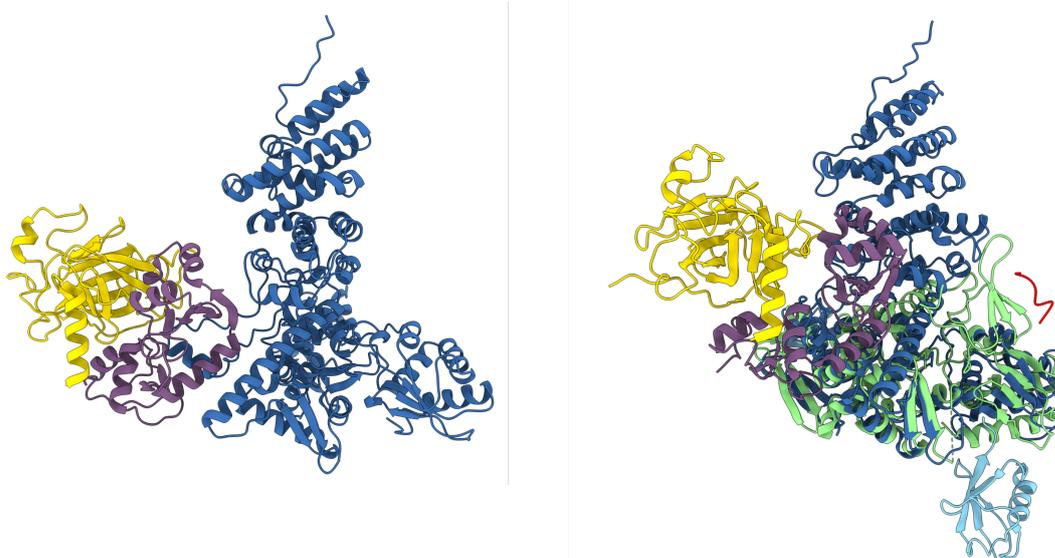
Des mutants de Ube3a et de E6 de HPV16 permettant d'empêcher cette nouvelle interaction ont été produits pour être étudiés prochainement. De plus, nous avons élaboré des peptides combinant l'hélice LxxLL ainsi que cette hélice de DUS-3. L'ajout de cette hélice supplémentaire pourrait augmenter l'affinité de E6 au peptide par rapport à l'hélice LxxLL seule. Des mesures préliminaires d'affinité de ces peptides pour MBP-16E6 ont été faites par polarisation de fluorescence (Tableau 8, page 8). Ici trois différents peptides ont été utilisés. Un peptide contenant le motif LxxLL et trois peptides contenant l'hélice de DUS-3 liée au motif LxxLL. Ces trois peptides diffèrent par la séquence qui lie les deux hélices, il s'agit de 3 GS (glycine sérine), 4 GS et de la séquence en amino-terminal de l'hélice

Peptide	Séquence	Affinité
Fluo-LxxLL	PESELTLQELLGEER	2.88 $\mu$ M
h2-3GS-LxxLL	RMYSEERRITVLYSLVQQQGS GSGSESELTLQELLGEERR	2.64 $\mu$ M
h2-4GS-LxxLL	RMYSEERRITVLYSLVQQQGS GSGSGSESELTLQELLGEERR	2.30 $\mu$ M
h2-LxxLL	RMYSEERRITVLYSLVQQQED DEEAAAESSELTLQELLGEERR	0.5 $\mu$ M

**TABLE 8** – Mesures d'affinités des différents peptides de Ube3a pour la protéine MBP-E6 de HPV16. Ces mesures ont été effectuées par polarisation de fluorescence.

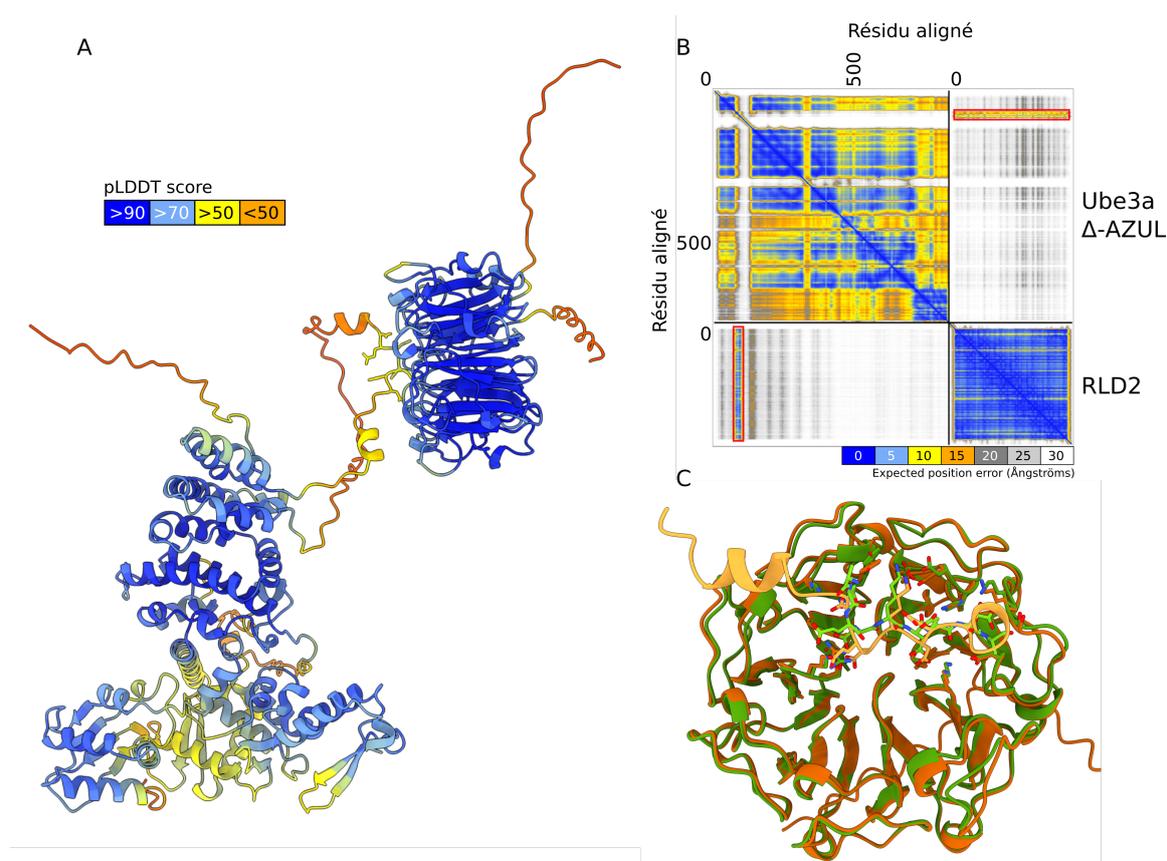
de DUS-3 et en carboxy-terminal du motif LxxLL. Les peptides LxxLL ainsi que les deux peptides avec les GS ont une affinité similaire (autour de 2.5  $\mu$ M). Le peptide possédant les deux hélices séparées par la séquence de Ube3a a une affinité de 0.5  $\mu$ M pour E6. Ainsi l'hélice de DUS-3 semble bien jouer un rôle dans l'interaction entre Ube3a et E6.

Cette prédiction de structure a permis d'explorer le mécanisme de l'ubiquitination de P53 par Ube3a en présence de E6. Ce modèle utilise la prédiction AlphaFold 2 ainsi que différentes structures résolues par cristallographie. AlphaFold 2 n'est pas capable de prédire l'interaction entre Ube3a (entière ou compacte), E6 et P53. La structure de Ube3a, seule et en complexe avec E6, semble être la même (selon les prédictions). De plus, la protéine P53 se situe éloignée du domaine HECT et ne pourrait pas être atteinte par le lobe-C pour son ubiquitination. Cependant, l'extension située en carboxy-terminal du core domaine de P53, connue pour être la zone ubiquitinée de P53 [299] [300], peut atteindre cette zone.



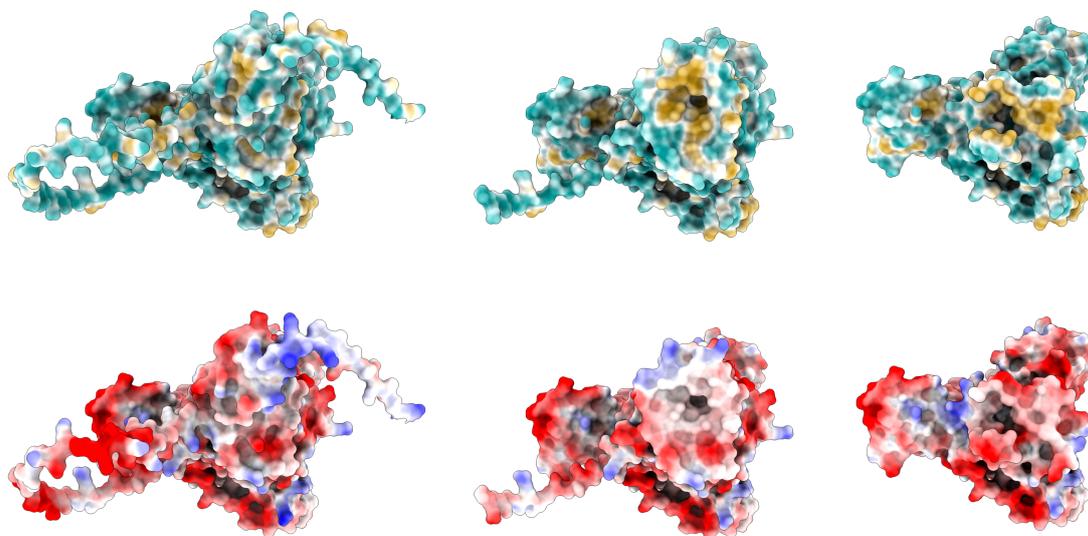
**FIGURE 51** – Modèle du complexe entre E6, P53, et Ube3a. À gauche, la structure de 16E6 (en violet) et de P53 (en jaune) en interaction avec le motif LxxLL (non affiché) est superposée à la structure prédite par AlphaFold de la construction compacte de Ube3a (en bleu) et 16E6 (non affiché). À droite, la structure du domaine HECT de la E3 Rsp5 (vert) en complexe avec un peptide de son substrat (rouge) et l'ubiquitine (bleue) (PDB : 4LCD).

La prédiction de structure faite de Ube3a avec le domaine RLD2 de HERC2 semble être très proche de la structure que nous avons obtenue. À noter que cette prédiction a été faite après la déposition des différentes structures sur la PDB (Section 1, page 61). Avant la déposition de ces structures sur la PDB, les générations de prédiction par une ancienne version de AlphaFold du domaine RLD2 de HERC2 avec un peptide de Ube3a n'avaient pas permis de prédire l'interaction.



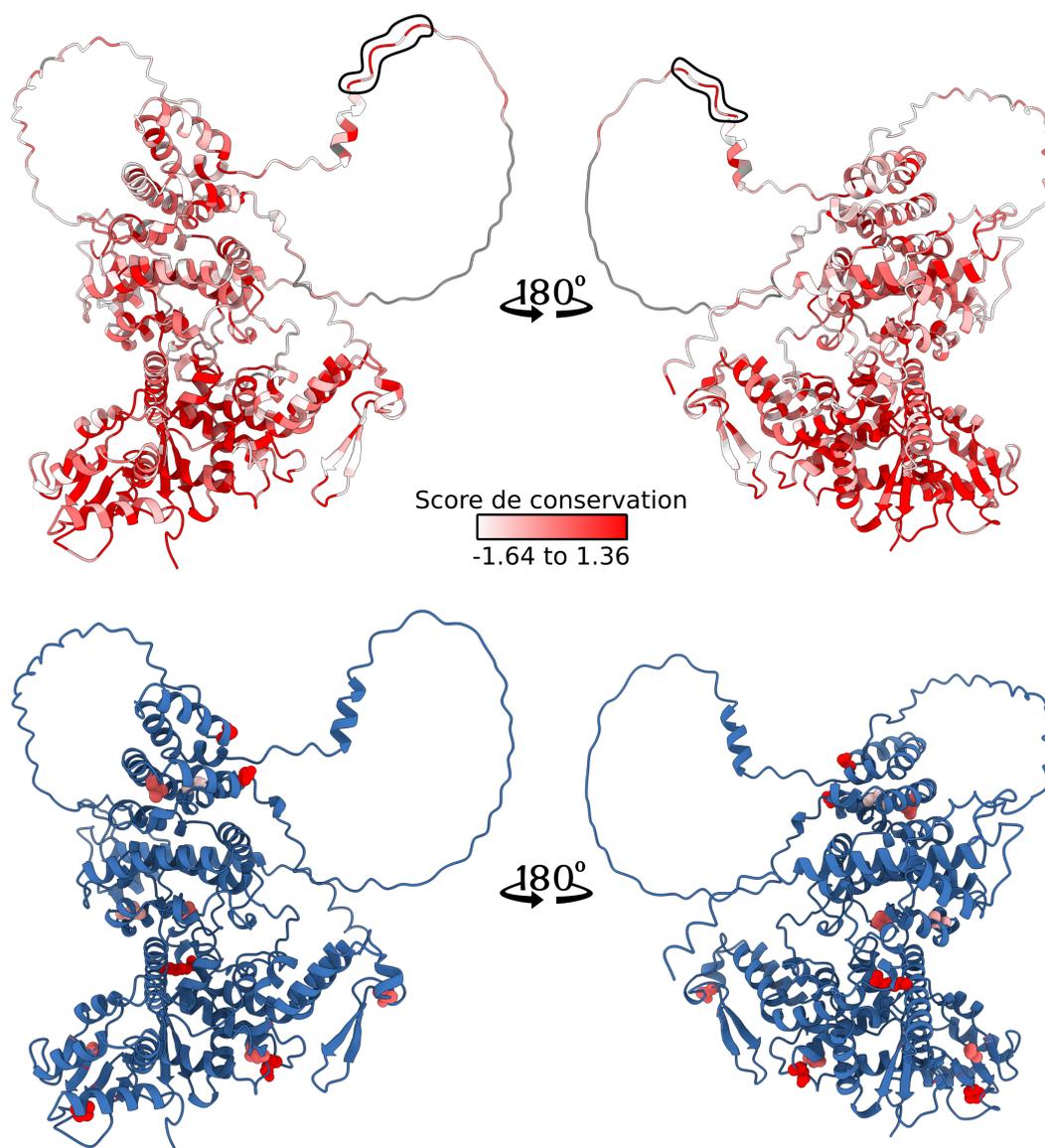
**FIGURE 52** – A) Structure prédite par AlphaFold 2 de l’interaction entre Ube3a  $\Delta$ -AZUL et du domaine RLD2 de HERC2. Cette prédiction a été faite après le dépôt des structures sur la PDB. Les acides aminés sont colorés en fonction du score pLDDT. B) matrice de PAE. Notez que la prédiction d’erreur de position entre le domaine RLD2 et le motif LxxLL est faible (encadrée en rouge). C) superposition de la prédiction (orange foncé et clair) avec la structure déposée sur la PDB (vert foncé et clair).

La prédiction AlphaFold 2 permet aussi d’expliquer la formation de multimères de la construction ne contenant pas le domaine DUS-1 et de celle ne contenant ni le domaine DUS-1 ni le domaine DUS-2. En effet, la délétion du domaine DUS-1 expose une surface hydrophobe de DUS-2 au solvant. De même, la délétion de DUS-1 et DUS-2 découvre une zone hydrophobe de DUS-3. (Figure 53, page 103)



**FIGURE 53** – Structures prédites de différentes constructions de Ube3a vues du dessus (domaine HECT en dessous). De gauche à droite, les constructions sans AZUL, sans AZUL ni DUS-1 et sans AZUL ni DUS-1 ni DUS-2. En haut, la représentation de la surface en fonction de l'hydrophobicité (beige : hydrophobe, Cyan : hydrophile). En bas, la représentation de la surface en fonction de la charge électrostatique. On remarque l'exposition au solvant d'une zone hydrophobe lors de la délétion du domaine DUS-1 et du domaine DUS-2. Ceci pourrait expliquer l'absence de monomère lors de la purification de ces constructions.

Une étude de la conservation de séquence à travers différentes espèces permet de mettre en évidence la haute conservation du domaine HECT. Les domaines DUS-1, 2 et 3 sont moins conservés. De plus, on remarque la haute conservation du motif DxDKDxD. (Figure 54, page 104). On remarque aussi que les mutations de Ube3a impliquées dans le syndrome d'Angelman répertorié dans Uniprot sont situées sur le domaine HECT et dans les domaines DUS-1, 2 et 3.



**FIGURE 54** – Figure représentant la surface de la prédiction de structure de Ube3a par AlphaFold 2. Le score de conservation de séquence à travers différentes espèces (voir matériel et méthodes) est affiché sur la structure. Le motif DxDKDxD permettant l'interaction avec le domaine RLD2 de HERC2 est entouré. Ici, toutes les séquences sélectionnées possèdent le motif DxDKDxD. Les acides aminés en gris sont des extensions de Ube3a humaine. En bas est représenté la même structure avec, affichés sous forme de sphères, les acides aminés mutés dans des cas d'Angelman (répertorié dans Uniprot). Ces acides aminés sont colorés selon le score de conservation

## 2.4 Discussion

### 2.4.1 Structure de Ube3a

Les différentes constructions de Ube3a et leurs purifications nous ont permis d'obtenir beaucoup d'informations qui sont en adéquation avec la prédiction de structure de Ube3a faite par AlphaFold 2. Ainsi, nous avons maintenant une idée de plus en plus précise de l'organisation structurale de Ube3a.

Cette protéine semble être composée d'une partie globulaire formée du domaine HECT et des trois domaines que nous appelons DUS-1, 2 et 3. L'utilisation de ce modèle, en adéquation avec la purification de la construction Ube3a multi TEV, nous a permis de concevoir une construction compacte dénuée du domaine AZUL, de la boucle entre le domaine AZUL et le domaine DUS-1 ainsi que de la boucle liant RLD2 entre le domaine DUS-1 et le domaine DUS-2. Cette construction ne forme pas de multimère et semble cristalliser.

Cependant, ces résultats sont en désaccord avec des modèles de Ube3a proposés. En effet, une étude de spectrométrie de masse associée à du "crosslinking" [301] identifie le domaine AZUL comme interagissant avec le reste de la protéine Ube3a et plus particulièrement le domaine HECT. Cette étude montre que l'interaction de E6 avec Ube3a renforce ce rapprochement entre le domaine AZUL et le domaine HECT. Ainsi, elle suggère un réarrangement structural de Ube3e lors de la liaison de E6 à Ube3a.

La purification de la protéine Ube3a n'est pas effectuée de la même manière dans cette étude et dans nos travaux. Cela pourrait expliquer cette divergence de résultats. En effet, dans cet article, comme dans notre protocole, la purification consiste en une colonne d'affinité Ni (Ube3a contenant une étiquette 6-His) et une étape de purification par une colonne échangeuse d'ions (Q sepharose). Cependant, contrairement à notre protocole, elle ne comprend pas d'étape de SEC, cruciale pour sélectionner Ube3a sous forme monomérique. Ainsi, cette interaction entre le domaine AZUL et le domaine HECT de cette étude pourrait être due à la multimérisation de Ube3a. Il est important de noter qu'à la différence de cette étude, la construction multi TEV que nous utilisons est fusionnée à la MBP.

Les prédictions de la structure de Ube3a ainsi que l'estimation d'erreur de ces modèles, semblent indiquer une non-interaction du domaine AZUL avec le reste de la protéine.

### 2.4.2 Ube3a et son interaction avec E6 et RLD2

La prédiction de structure de Ube3a en complexe avec E6 semble indiquer une zone d'interaction supplémentaire à l'hélice contenant le motif LxxLL. Cette zone d'interaction est

constituée par la dernière hélice du domaine DUS-3 (détaillé ultérieurement section 2.4.4, page 108). Cette hélice avait déjà été identifiée comme faisant partie du domaine HECT. En effet, une étude de séquence des différents domaines HECT avait fait apparaître des conservations de séquences en n-ter du domaine HECT [302]. Cette hélice, aussi appelée  $\alpha 1$  (première hélice du domaine HECT), n'est pas présente dans la structure résolue en 1999 du domaine HECT de Ube3a mais apparaît dans la structure de domaine HECT d'autres E3. En effet, cette hélice ne faisant pas partie de la définition du domaine HECT, elle n'était pas présente dans la construction qui a cristallisé en 1999. Dans les prédictions AlphaFold, l'hélice  $\alpha 1$  semble être plus longue chez Ube3a que chez d'autres E3 à domaine HECT (notamment de la famille NEDD4).

La structure prédite de Ube3a et son interaction avec le domaine RLD2 ne permettent pas d'indiquer ou d'expliquer une stimulation de l'activité ubiquitine-ligase de Ube3a par RLD2. En effet, l'interaction se fait dans une boucle de Ube3a et ne semble pas changer la structure ou la conformation de Ube3a. Cette liaison entre les deux E3, Ube3a et HERC2, pourrait jouer un rôle de localisation cellulaire. Cette capacité à se lier entraînerait la colocalisation de Ube3a et de HERC2. De même, le domaine AZUL de Ube3a semble jouer le rôle de colocalisation de Ube3a avec le protéasome via son interaction avec le domaine RAZUL de Rpn10. Ainsi ceci entraînerait la colocalisation de HERC2, de Ube3a et du protéasome.

### 2.4.3 Multimérisation de Ube3a

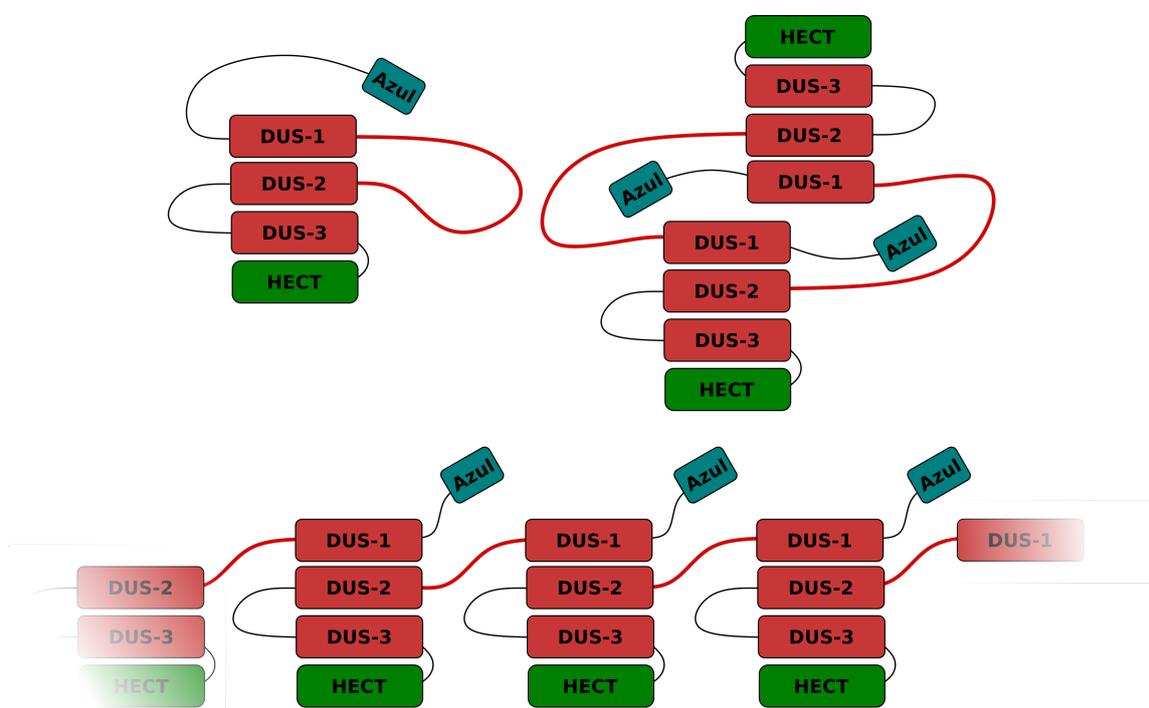
La purification des différentes constructions de Ube3a ainsi que la prédiction de structure nous permettent de mieux comprendre la nature des multimères lors de la purification des constructions Ube3a entière et de Ube3a  $\Delta$ -AZUL.

La construction Ube3a entière et la construction Ube3a  $\Delta$ -AZUL, qui possèdent la longue boucle entre le domaine DUS-1 et DUS-2, forment des multimères. A contrario, la construction compacte, similaire à Ube3a  $\Delta$ -AZUL, mais sans cette boucle, ne forme pas de multimères. La délétion de la boucle de liaison au domaine RLD2 de HERC2 (entre DUS-1 et DUS-2) empêche donc la formation de ces multimères.

Cependant, les constructions DUS3-HECT ainsi que DUS2-DUS-3-HECT, qui ne possèdent pas cette boucle, ne se purifie que sous forme de multimères. Ainsi, on peut supposer que les multimères des constructions entières/ $\Delta$ -AZUL et des autres constructions ne sont pas de même nature. Dans le cas des constructions entières et  $\Delta$ -AZUL de Ube3a, il s'agit de multimérisations permises par la boucle de liaison à RLD2. Nous supposons que cette multimérisation se fait par l'interaction du domaine DUS-1 avec le domaine DUS-2 en trans, un phénomène aussi appelé échange de domaine [303]. Ceci permettrait la formation de dimères, trimères et d'autres multimères (Figure 55, page 107). Le fait que nous

surexprimons la protéine Ube3a pour la purifier pourrait favoriser la formation de ces multimères, nous ne pouvons donc pas dire si ces multimères de Ube3a se forment dans des conditions physiologiques.

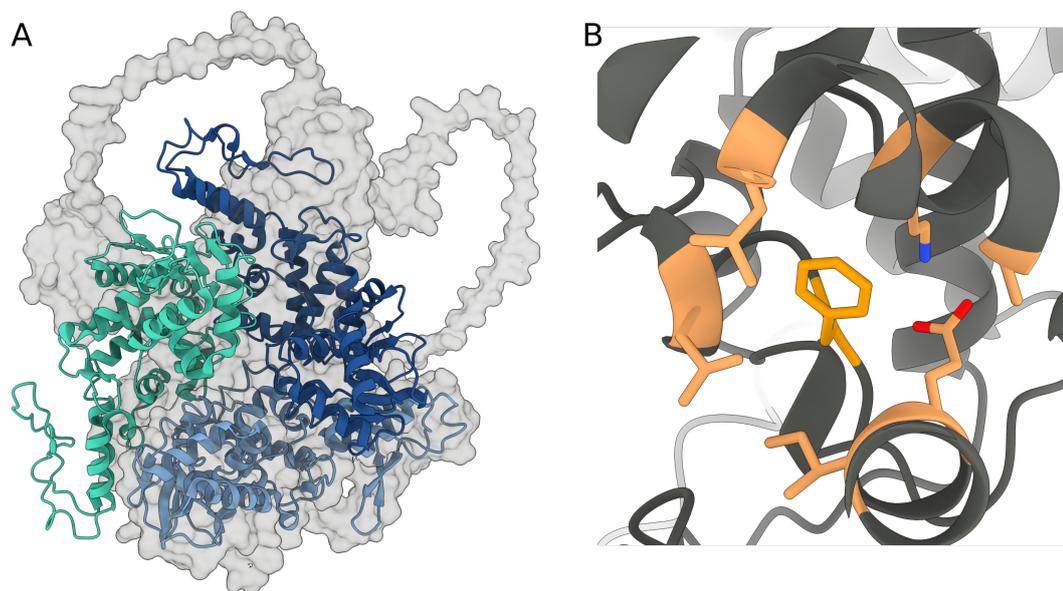
La multimérisation des autres constructions semblerait être d'une autre nature. Comme nous l'avons vu avec la prédiction AlphaFold (Figure 53, page 103), la délétion des domaines DUS-1 et la délétion des domaines DUS-1 et DUS-2, exposent une surface hydrophobe de DUS-2 et de DUS-3. L'exposition au solvant de ces surfaces hydrophobes entrainerait l'agrégation de ces différentes constructions et expliquerait l'absence de monomère lors de leur purification.



**FIGURE 55** – Représentation schématique du modèle d'oligomérisation de Ube3a permise par la boucle de liaison au domaine RLD2 (représentée en rouge). En haut à gauche est représentée Ube3a sous sa forme monomérique. En haut à droite est représentée Ube3a sous sa forme dimérique. Cette interaction supposée du domaine DUS-1 d'une molécule de Ube3a avec le domaine DUS-2 d'une autre permettrait de former des multimères de diverses tailles, représentés en bas.

La formation de multimères par la protéine Ube3a avait déjà été étudiée. Un modèle de trimérisation nécessaire à son activité ubiquitine-ligase avait été proposé (Section 2.1.3, page 38). En effet, cette étude tendait à montrer que la trimérisation de Ube3a par son domaine HECT était nécessaire à son activité ubiquitine-ligase [172]. Cependant cette trimérisation pourrait être due à l'absence de l'hélice  $\alpha 1$  du domaine HECT (Figure 56, page 108).

De plus, en superposant la prédiction de structure d'AlphaFold 2 au trimère de la structure de 1999, on remarque que les domaines DUS-1, DUS-2 et DUS-3 de Ube3a ne permettraient pas cette trimérisation. En effet, les domaines DUS-1, 2 et 3 se superposent aux deux autres domaines HECT du trimère (Figure 56, page 108).



**FIGURE 56** – A) structure du trimère du domaine HECT de Ube3a résolue en 1999 en bleu clair, bleu marine et cyan. En gris, la surface de la structure prédite d'AlphaFold 2 dont le domaine HECT est superposé à celui du trimère (bleu clair). On remarque que les domaines DUS-1 2 et 3 prédits par AlphaFold 2 se superposent aux autres domaines HECT de la structure trimérique. B) Dans la structure prédite par AlphaFold, la phénylalanine (orange), décrite comme importante dans cette trimérisation du domaine HECT, se trouve dans une poche hydrophobe composée des acides aminés suivants : leucine 723, leucine 470, lysine 466, alanine 463, acide glutamique 731 et isoleucine 732 de Ube3a

#### 2.4.4 Activité ubiquitine-ligase de Ube3a et sa modulation

E6 est connue pour sa capacité à stimuler l'activité ubiquitine-ligase de Ube3a [182]. La structure globale prédite par AlphaFold semble correcte. Cependant, la structure prédite de Ube3a seule et de Ube3a en complexe avec E6 sont similaires et ne montrent pas de changement de conformation de Ube3a.

L'interaction entre E6 et la dernière hélice du domaine DUS-3 (aussi appelée hélice  $\alpha 1$  du domaine HECT) pourrait cependant expliquer cette stimulation de l'activité enzymatique de Ube3a. En effet, cette hélice semble jouer un rôle important, mais pas encore définie, dans la modulation de l'activité ubiquitine-ligase des domaines HECT. Une étude indique que la délétion de l'hélice  $\alpha 1$  du domaine HECT de l'E3 HUWE1 entraîne une meilleure activité ubiquitine-ligase comparée à celle de ce même domaine HECT contenant cette hélice

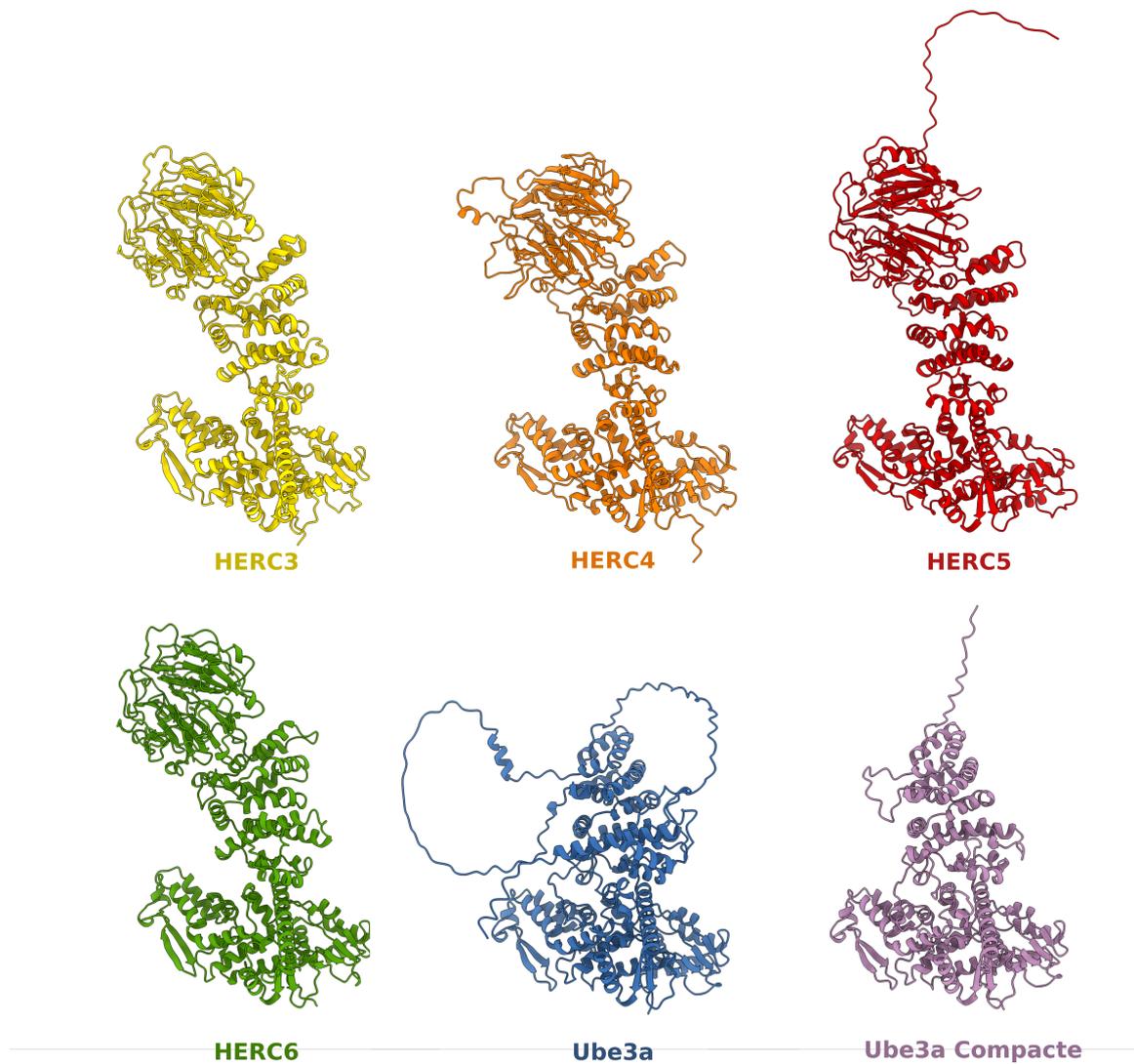
[304]. L'hypothèse du rôle de stabilisation du domaine HECT par cette hélice avait alors été émise. L'absence de cette hélice permettrait au domaine HECT d'adopter différentes conformations et favoriserait le transfert de l'ubiquitine de E2 sur sa cystéine catalytique et de la cystéine catalytique à son substrat.

De plus, la structure Rsp5 (PDB : 4LCD) indique qu'un feuillet  $\beta$  (qui n'existe pas dans les domaines HECT des autres E3) permet l'interaction avec son substrat. Ce feuillet  $\beta$  est situé à proximité de l'hélice  $\alpha 1$ . En supposant que le mouvement du lobe-C du domaine HECT pour l'ubiquitination du substrat est le même pour les différents domaines HECT, on peut imaginer que cette zone (proche de  $\alpha 1$ ) est la zone de liaison du substrat. Ainsi la stabilisation de l'hélice  $\alpha 1$  du domaine HECT pourrait libérer une zone de liaison au substrat, ce qui entrainerait l'activité ubiquitine-ligase de Ube3a.

Ici, nous pouvons proposer un modèle de la stimulation de l'activité ubiquitine-ligase de Ube3a par E6. Dans le cas des prédictions de structures de Ube3a, la présence de la protéine E6 semble stabiliser l'hélice  $\alpha 1$  du domaine HECT de Ube3a. La stabilisation de cette hélice pourrait permettre la libération d'un site de liaison au substrat. Une autre hypothèse pourrait être la mobilité du lobe-C du domaine HECT permise par la stabilisation de cette hélice, rendant possible le transfert de l'ubiquitine au substrat.

#### 2.4.5 Ube3a une proche parente des HERCs

La génération des prédictions AlphaFold des protéines HERC-3 4, 5 et 6 révèle une structure similaire à la prédiction de Ube3a. En effet, la protéine Ube3a semble être composée des mêmes domaines que les HERCs. Elle semble être une HERC dénuée de domaine RLD avec un domaine AZUL et une boucle supplémentaire. Les structures DUS-1, 2 et 3 sont similaires chez les HERCs. Tout comme les HERCs, Ube3a semble avoir une hélice  $\alpha 1$  plus longue que chez les autres E3 [302]. La ressemblance des structures prédites de Ube3a et des HERCs pourrait indiquer des mécanismes similaires de régulation de leur activité ubiquitine-ligase. (Figure 57, page 110)



**FIGURE 57** – Prédiction des structures des protéines HERC3, HERC4, HERC5, HERC6, Ube3a et Ube3a compacte. Notez que toutes les structures prédites des HERCs sont similaires. La structure de la construction Ube3a compacte est extrêmement proche de celles des HERCs.

Troisième partie

Matériel & méthodes



# 1 | Matériel & méthodes

## 1.1 Expressions

Les différentes constructions sont exprimées en *Escherichia coli* BL21. Les bactéries transformées sont sélectionnées en ajoutant 1mM de kanamycine. Pour les constructions contenant un domaine de liaison au zinc (E6 et les constructions contenant un domaine Azul), 0.01 mM de zinc est ajouté. La culture est réalisée avec 0,2% de glucose et une agitation de 190 rpm. Après induction par l'ajout de 0,5mM d'IPTG (concentration finale) la température est abaissée à 16°C sur la nuit.

## 1.2 Purifications

### 1.2.1 Purification de RLD2

Les culots de bactéries sont suspendus dans un tampon de lyse contenant 50 mM de tris-HCl pH 8, 400 mM de NaCl, 5% de glycérol, 5 mM de beta-mercaptoéthanol (BME) et des inhibiteurs de protéases. La lyse de la cellule est réalisée par sonication pendant 1 minute 45 secondes, par étapes de 2 secondes on/2 secondes off. Ensuite, le lysat est centrifugé à 125000g pendant 45 minutes pour obtenir le lysat clair. La construction est purifiée par chromatographie d'affinité nickel. L'échantillon est chargé sur une colonne nickel dans un tampon contenant 20 mM d'HEPES pH8, 300mM de NaCl, 3mM de BME puis est lavé avec le même tampon auquel est ajouté 40 mM d'imidazole. La protéine est éluée par un tampon contenant 250 mM d'imidazole. L'étape suivante est une chromatographie échangeuse d'ions SP Sepharose. L'éluion se fait par un gradient de NaCl (de 100mM à 1000mM de NaCl), contenant 20 mM d'HEPES pH8 et 3mM de BME. Après un clivage par la protéase TEV ("Tobacco Etch Virus") pour enlever la MBP, le domaine RLD2 est de nouveau purifié par chromatographie d'échange d'ions (SP sépharose) et par une étape de chromatographie d'exclusion stérique (S75 16/600) dans un tampon contenant 20 mM d'HEPES pH8, 300 mM de NaCl et 3mM BME. La concentration de la protéine est effectuée à l'aide d'un "centricon" (seuil de 10K).

### 1.2.2 Purification des constructions de Ube3a

Les culots de bactéries sont suspendus dans un tampon de lyse contenant 50 mM de tris-HLC pH 8, 400 mM de NaCl, 5% de glycérol, 5 mM de beta-mercaptoéthanol (BME) et des inhibiteurs de protéases. La lyse de la cellule est réalisée par sonication pendant 1 minute 45 secondes, par étapes de 2 secondes on/2 secondes off. Ensuite, le lysat est centrifugé à 125000g pendant 45 minutes pour obtenir le lysat clair.

Les différentes constructions sont d'abord purifiées par une colonne d'affinité, soit par colonne d'affinité amylose (liant la MBP) pour les constructions sans étiquettes 6HIS, soit par une colonne d'affinité nickel (liant l'étiquette 6HIS). Le lysat est chargé sur une colonne nickel dans un tampon contenant 50 mM de tris-HCl pH8, 300mM de NaCl et 3mM de BME puis est lavé avec le même tampon auquel est ajouté 35 mM d'imidazole. La protéine est éluée par un tampon contenant 80 mM d'imidazole. Dans le cas de la purification par colonne d'affinité amylose, un tampon contenant 50 mM de tris HCl pH8, 300mM de NaCl, 3mM de BME et 10mM de maltose permet l'éluion.

La suite de la purification est commune à toutes les constructions de Ube3a. La seconde étape est une purification par une colonne échangeuse d'ions (Q-sepharose). L'échantillon est chargé sur colonne dans un tampon contenant 100 mM de NaCl, 50 mM de tris HCl pH 8 et 3mM de BME. L'éluion est faite par un gradient de NaCl (100 à 600 mM de NaCl). Une étape de SEC (Chromatographie d'exclusion stérique) utilisant une colonne 26/60 S200 (Cytiva) avec une solution tampon contenant 20 mM de Tris HCl pH 8, 200 mM de NaCl et 3mM de BME permet de séparer les monomères des multimères.

Les monomères sont sélectionnés pour retirer la MBP via une coupure par la protéase TEV. Cette coupure est faite sur la nuit à température ambiante. La rétention de la MBP (couplée à l'étiquette 6HIS) est faite via une colonne d'affinité nickel (tampon contenant 35 mM d'imidazole). Ensuite, une étape de purification par une colonne échangeuse d'ions Q-sepharose (similaire à la précédente) est effectuée. Enfin, une étape de SEC (16/60 S200) dans un tampon contenant 20 mM de Tris HCl pH 8, 200 mM de NaCl et 3mM de BME permet l'obtention de la protéine pure et monomérique. La concentration de la protéine a été effectuée à l'aide d'un "centricon" (seuil de 10K).

### 1.2.3 Purification de MBP-16E6

Les bactéries sont lysées dans un tampon contenant 50mM de tris Hcl pH 6.8, 400 mM de NaCl et 5% de glycérol en utilisant le microfluidizer. Ensuite, le lysat est centrifugé à 125000g pendant 45 minutes pour obtenir le lysat clair.

La première étape de purification est effectuée via une colonne d'affinité amylose. Après la charge du lysat clair sur la colonne et un lavage avec un tampon contenant 50mM de

Tris HCl pH 6.8, 400mM de NaCl et 2mM de DTT, la protéine est éluée grâce à un tampon contenant 50mM de Tris-hCl pH 6,8, 400mM de NaCl, 2mM de DTT et 10mM de maltose. L'étape suivante consiste en l'élimination des agrégats solubles par une centrifugation de 15 heures à 350000 RPM. Enfin, la protéine pure et monomérique est obtenue par une SEC (26/60 S200) dans un tampon contenant 50mM de Tris-HCl pH6.8, 200mM de NaCl et 2mM de DTT.

### 1.3 Cristallisation

Les criblages de conditions de cristallisation sont faites en gouttes assises avec la technique de la diffusion de vapeur. Ces criblages sont faits via le robot de cristallisation (mosquito) et des kits de criblages commerciaux contenant 96 conditions par kits. Différents ratios précipitant/protéine sont testés. Les boîtes de cristallisations sont stockées dans un formulatrix ce qui permet la prise de photo régulière des gouttes de cristallisation. Les optimisations de conditions de cristallisation sont faites à la main, en gouttes suspendues. Les différents essais de cristallisation sont faits à 20 et 4 °C. Les tests de cristallisation de Ube3a avec RLD2, 16E6 et le domaine Ubl de RAD23A sont faits de manière équimolaire.

RLD2 seul et RLD2 en complexe avec différents peptides sont cristallisés à une concentration de 4mg/mL (avant mélange avec la solution précipitante) en utilisant la technique de diffusion de vapeur dans une goutte suspendue. La solution de précipitant utilisée contient 20% de polyéthylène glycol 3350 et 0,18 M de citrate de tris ammonium. Le rapport protéine solution précipitante est de 1 volume pour 1 volume. Pour la cristallisation de RLD2 en complexe avec les différents peptides, RLD2 (4mg/ml) est mélangé avec le peptide dans un rapport de 1 M de RLD2 pour 5 M de peptides. La solution cryoprotectrice utilisée contient 20% de polyéthylène glycol 3350, 0,18 M de citrate de tris ammonium et 20% de glycérol. Les données sont collectées au synchrotron Soleil (lignes proxima 1 et proxima 2A) et à la Swiss Light Source (XD6DA). XDS et ccp4 sont utilisés pour intégrer les données. La construction et l'affinement des modèles sont réalisés à l'aide de Coot et Phenix.

### 1.4 Mesures d'affinités par polarisation de fluorescence

Pour les mesures d'affinités à RLD2, les mesures directes de polarisation de fluorescence sont effectuées en diluant RLD2 dans un tampon contenant 25mM d'HEPES pH 7,5, 150mM de NaCl, 1mM de TCEP, 0,005% de Tween20 et 50nM de peptide fluorescent (Ube3a 10 aa lié à la fluorescéine). Les mesures des peptides non fluorescents sont effectuées en compétition avec ce même peptide fluorescent (Ube3a 10 aa lié à la fluorescéine). Le peptide non fluorescent est dilué dans un mélange contenant 25mM d'HEPES

pH 7.5, 150mM de NaCl, 1mM de TCEP, 0.005% de Tween20, 50nM de peptide fluorescent et 3.33mM de RLD2.

Pour les mesures d'affinités à MBP-16E6, les mesures directes de polarisation de fluorescence sont effectuées en diluant MBP-16E6 dans un tampon contenant 25mM d'HEPES pH 7.5, 150 mM de NaCl, 1 mM de TCEP, 0.005% de Tween20 et 50 nM de peptide fluorescent (peptide contenant le motif LxxLL de Ube3a). Les mesures des peptides non fluorescents sont effectuées en compétition avec ce même peptide fluorescent. Le peptide non fluorescent est dilué dans un mélange contenant 25 mM d'HEPES pH 7.5, 150 mM de NaCl, 1 mM de TCEP, 0.005% de Tween20, 50 nM de peptide fluorescent et 3.33 mM de MBP-16E6.

Les mesures sont effectuées avec le lecteur de plaques Pherastar. Les différentes dilutions pour les mesures directes et compétitives sont transférées d'une plaque de 96 puits à une plaque de 384 puits noirs en triplicat de 10  $\mu$ l. Le traitement des données de polarisation de fluorescence est effectué à l'aide du logiciel Profit.

#### 1.4.1 Génération des modèles par AlphaFold 2

La génération des différentes prédictions de structure de Ube3a sont faites par AlphaFold 2 [296] via le notebook [305] exécuté sur google colab. La structure de Ube3a entière utilisée est celle déposée sur AlphaFold Protein Structure Database (<https://alphafold.ebi.ac.uk/>)

#### 1.4.2 Analyses phylogénétiques de séquences

Les séquences de HERC2 de différentes espèces sont trouvées par des BLASTs systématiques de la séquence humaine dans différents phylums : Humain (O95714), *Saccoglossus kowalevskii* (XP\_006821411.1), *Acanthaster planci* (XP\_022098915.1), *Capitella teleta* (ELT92472.1), *Biomphalaria glabrata* (XP\_013080917.1), *Anoplophora glabripennis* (XP\_018562376.1), *Macrostomum lignano* (PAA75977.1), *Nematostella vectensis* (XP\_001629259.1), *Amphimedon queenslandica* (XP\_019851474.1). L'alignement des séquences est fait par Clustal Omega [306]. La coloration des acides aminés par le score de conservation est issue de ChimeraX calculant le score AL2CO.

Les séquences de UBE3a de différentes espèces sont trouvées via l'outil NCBI de recherche d'architecture globale de protéines [307] (protéines contenant un domaine Azul et un domaine HECT) : Humain (Q05086), hémichordés (XP\_006813929.1), Echinodermata (XP\_033118421.1), Mollusca (XP\_033742467.1), Arthropoda (XP\_038213425.1). L'alignement de ces séquences est fait avec TCOFFEE [308]. La coloration des acides aminés par le score de conservation est issue de ChimeraX calculant le score AL2CO.

### 1.4.3 Mesure de surface accessible au solvant

Le logiciel en ligne Accessible surface area and accessible calculation protein version 1.2 [309] est utilisé pour identifier la surface des acides aminés du domaine RLD2 en contact avec le peptide. Le pourcentage de la surface de chaque acide aminé du domaine RLD2 en contact avec le solvant est calculé en présence et en absence du peptide de BAZ2B (structure avec la meilleure résolution). En faisant la différence de ces pourcentages de surface, nous connaissons le pourcentage de surface de chaque acide aminé de RLD2 en contact avec le peptide.

## 1.5 Séquences

RLD2 :

```
MKHHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFW
AHDRFGGGAQSGLLAEITPKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDNAGAKAGL TFLVDLIKHKHMNADT
DYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFFKGQPSKPFVGVLSAGINAASPNKELAKEFL
ENYLLTDEGLEAVNKDKPLGAVALKS YEEELAKDPRIAATMENAQKGEIMPNIPOMSAFWYAVRTAVINAAS
GRQTVDEALKDAQTNSSSNNNNNNNNNNNPMSENLYFQGAMGSLIRKKAAGLESAATIRTKVFWGLNDKD
QLGGLKGSKIKVPSFSETLSALNVVQVAGGSKSLFAVTVEGKVYACGEATNGRLGLGISSGTVPIPRQITALS
SYVVKKVAVHSGGRHATAALTVDGKVFWSWEGDDGKLGHF SRMNC DKPRLIEALKTKRIRDIACGSSSHAAL
TSSGELYTWGLGEYGRLGHDNTTQLKPKMVKVLGHRVIQVACGSRDAQTLALTDEGLVFSWGDGDFGK
LGRGGSEGCNIPQNIERLNGQVCQIECGAQFSLALTKSGVVWWTWGKGDYFRLGHGSDVHVRKPVVVEG
LRGKKIVHVAVGALHCLAVTDSGQVYAWGDNHDGQQGNGTTTTVNRKPTLVQGLEGQKITRVACGSSHSVA
WTTVDVATPSVHEPVLFTQ
```

Ube3a entière :

```
MKHHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFW
AHDRFGGGAQSGLLAEITPKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDNAGAKAGL TFLVDLIKHKHMNADT
DYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFFKGQPSKPFVGVLSAGINAASPNKELAKEFL
ENYLLTDEGLEAVNKDKPLGAVALKS YEEELAKDPRIAATMENAQKGEIMPNIPOMSAFWYAVRTAVINAAS
GRQTVDEALKDAQTNSSSNNNNNNNNNNNPMSENLYFQGAMAMSEIKMNKKGARIDFKDVTYLTTEEKVY EIL
ELCREREDYSPLIRVIGRVFSSAEALVQSFRKVKQHTKEELKSLQAKDEKDEDEKEKAACSAAMEEDSEA
SSSRIGDSSQGDNNLQKLGPDVSDIDAIRRVYTRLLSNEKIETAFLNALVYLSPNVECDLTYHNVYSRDPN
LYNLFIIVMENRNLHSP EYLEMALPLFCAMSKLPLAAQGGKLR LWSKYNADQIRMMETFQQLLTYKVISNEF
NSRNLVNDDDAIVAASKCLKMVYANVVGGVDTNHNEEDDEEPIPESELTLQELLGEEERNNKKGPRVDP
LETELGVKTLDCRKLIPFEFFINEPLNEVLEMDKDYTFKVFETENKFSFMTCPFILNAVTKNLGLYDNRIRM
YSERRITVLYSLVQGGQLNPLRLKVR RDHIIDALVRLEMIAMENPADLKKQLYVFEFEGEQGVDEGGVSK
FFQLVVEEIFNPDIGMFTYDESKLFWFNPSSFETEGQFTLIGIVLGLAIYNNCILDVHFPMVYVYRKL
MGKKGTFRDLDGSDHPVLYQSLKDLLEYEGNVEDDMMITFQISQTDLFGNPMMYDLKENGDKIPITNE
YILNKSVEKQKAFRRGFHMVTNESPLKYLFRPEEIEILLICGSRNLDFQALEETTEYDGGYT
RDSVLIREFWEIVHSFTDEQKRLFLQFTTGTDRA PVGGLGKLMIIAKNGP DTERLPTSHTCFNVLLPEYSS
KEKLERLLKAITYAKGFGML
```

Ube3a Δ-Azul :

```
MKHHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFW
AHDRFGGGAQSGLLAEITPKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALD
KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDNAGAKAGL TFLVDLIKHKHMNADT
DYSIAEAAFNKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFFKGQPSKPFVGVLSAGINAASPNKELAKEFL
ENYLLTDEGLEAVNKDKPLGAVALKS YEEELAKDPRIAATMENAQKGEIMPNIPOMSAFWYAVRTAVINAAS
GRQTVDEALKDAQTNSSSNNNNNNNNNNNPMSENLYFQGAMAMSEIKMNKKGARIDFKDVTYLTTEEKVY EIL
ELCREREDYSPLIRVIGRVFSSAEALVQSFRKVKQHTKEELKSLQAKDEKDEDEKEKAACSAAMEEDSEA
SSSRIGDSSQGDNNLQKLGPDVSDIDAIRRVYTRLLSNEKIETAFLNALVYLSPNVECDLTYHNVYSRDPN
LYNLFIIVMENRNLHSP EYLEMALPLFCAMSKLPLAAQGGKLR LWSKYNADQIRMMETFQQLLTYKVISNEF
NSRNLVNDDDAIVAASKCLKMVYANVVGGVDTNHNEEDDEEPIPESELTLQELLGEEERNNKKGPRVDP
LETELGVKTLDCRKLIPFEFFINEPLNEVLEMDKDYTFKVFETENKFSFMTCPFILNAVTKNLGLYDNRIRM
YSERRITVLYSLVQGGQLNPLRLKVR RDHIIDALVRLEMIAMENPADLKKQLYVFEFEGEQGVDEGGVSK
FFQLVVEEIFNPDIGMFTYDESKLFWFNPSSFETEGQFTLIGIVLGLAIYNNCILDVHFPMVYVYRKL
MGKKGTFRDLDGSDHPVLYQSLKDLLEYEGNVEDDMMITFQISQTDLFGNPMMYDLKENGDKIPITNE
YILNKSVEKQKAFRRGFHMVTNESPLKYLFRPEEIEILLICGSRNLDFQALEETTEYDGGYT
RDSVLIREFWEIVHSFTDEQKRLFLQFTTGTDRA PVGGLGKLMIIAKNGP DTERLPTSHTCFNVLLPEYSS
KEKLERLLKAITYAKGFGML
```

Ube3a DUS2-DUS3-HECT :

MKHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFW  
 AHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALD  
 KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGL TFLVDLIKXKHMNADT  
 DYSIAEAAFNKGETAMTINGPWAWNSNIDTSKVNYGVTVLPFTFKGQPSKPFVGVLSAGINAASPNKELAKEFL  
 ENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAAS  
 GRQTVDEALKDAQTNSSSNNNNNNNNNPMSENLYFQGAMAGDNNLQKLGPDVSDIDAIRRVYTRLLS  
 NEKIETAFNLALVYLSPNVECDLTYHNVYSRDPNYLNLFIIVMENRNLHSPYELEMALPLFCCKAMSKLPLAAQ  
 GKLIIRLWSKYADQIRRMETTFQQLITYKVISNEFNSRNLVNDDDAIVAASKCLKMVYYANVVGGEVDTNHN  
 EEDDEEPIPELSELTLQELLGEERRNKKGPRVDPLETELGVKTLDCRKLIPFEFINEPLNEVLEMDKDYTF  
 FFKVETENKFSFMTCPFILNAVTKNLGLYYDNRIRMYSEERRITVLYSLVQGGQLNPYLRLKVRRDHIIDALV  
 RLEMIAMENPADLKKQLYVEFEQGVDEGGVSKEFFQLVVEEIFNPDIGMFTYDESTKLFWFNPSSFETEGQFT  
 LIGIVLGLAIYNNCILDVHFPVYRKLGMKGGTFRDLGDSHPVLYQSLKDLLEYEGNVEDDMMITFQISQTD  
 LFGNPMYDLKENGDKIPITNENRKEFVNLYSYILNKSVEKQKAFRRGFHMTNESPLKYLFRPEEIEILLIC  
 GSRNLDFQALEETTEYDGGYTRDSVLIREFWEIVHSFTDEQKRLFLQFTTGDRAPVGGGLGKLMIIAKNGP  
 DTERLPTSHTCFNVLLLPEYSSKEKLERLLKAITYAKGFGML

### Ube3a DUS3-HECT :

MKHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFW  
 AHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALD  
 KELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGL TFLVDLIKXKHMNADT  
 DYSIAEAAFNKGETAMTINGPWAWNSNIDTSKVNYGVTVLPFTFKGQPSKPFVGVLSAGINAASPNKELAKEFL  
 ENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAAS  
 GRQTVDEALKDAQTNSSSNNNNNNNNNPMSENLYFQGAMESSELTLQELLGEERRNKKGPRVDPLETEL  
 GVKTLDCRKLIPFEFINEPLNEVLEMDKDYTFKKVETENKFSFMTCPFILNAVTKNLGLYYDNRIRMYSEERR  
 ITVLYSLVQGGQLNPYLRLKVRRDHIIDALVREMIAMENPADLKKQLYVEFEQGVDEGGVSKEFFQLV  
 VEEIFNPDIGMFTYDESTKLFWFNPSSFETEGQFTLIGIVLGLAIYNNCILDVHFPVYRKLGMKGGTFRDLG  
 DSHVPVLYQSLKDLLEYEGNVEDDMMITFQISQTDLFGNPMYDLKENGDKIPITNENRKEFVNLYSYILNKS  
 VEKQKAFRRGFHMTNESPLKYLFRPEEIEILLICGSRNLDFQALEETTEYDGGYTRDSVLIREFWEIVHSFT  
 DEQKRLFLQFTTGDRAPVGGGLGKLMIIAKNGP DTERLPTSHTCFNVLLLPEYSSKEKLERLLKAITYAKGFG  
 ML

### Ube3a 3 sites TEVs :

MKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDRFGGYA  
 QSGLLAEITPAAAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALDKELKAKGKS  
 ALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGL TFLVDLIKXKHMNADTDYSIAEAAF  
 NKGETAMTINGPWAWNSNIDTSKVNYGVTVLPFTFKGQPSKPFVGVLSAGINAASPNKELAKEFL ENYLLTDEG  
 LEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDAAL  
 AAAQTNAAMAMKRAAAKHLIERYHQLTEGCGNEACTNEFCASCPTFLRMDNNAAAIKALELYKINAKLCD  
 PPHSKGASSAYLENSKGGSGSENLYFQGGSGSAPNNSCSEIKMNKKGARIDFKDVTYLTEEKVYEILELCR  
 EREDYSPLIRVIGRVFSSAEALVQSFRKVKQHTKEELKSLQAKDEKDEDEKEKAACSAAMEEDSGSGSE  
 NLYFQGGSGSEASSRIGDSSQGDNNLQKLGPDVSDIDAIRRVYTRLLSNEKIETAFNLALVYLSPNVECD  
 LTYHNVYSRDPNYLNLFIIVMENRNLHSPYELEMALPLFCCKAMSKLPLAAQGKLIIRLWSKYADQIRRMET  
 FQQLITYKVISNEFNSRNLVNDDDAIVAASKCLKMVYYANVVGGEVDTNHNHSGSENLYFQGGSGSEEDDEE  
 PIPESSELTLQELLGEERRNKKGPRVDPLETELGVKTLDCRKLIPFEFINEPLNEVLEMDKDYTFKKVETEN  
 KFSFMTCPFILNAVTKNLGLYYDNRIRMYSEERRITVLYSLVQGGQLNPYLRLKVRRDHIIDALVREMIAMEN  
 PADLKKQLYVEFEQGVDEGGVSKEFFQLVVEEIFNPDIGMFTYDESTKLFWFNPSSFETEGQFTLIGIVL  
 GLAIYNNCILDVHFPVYRKLGMKGGTFRDLGDSHPVLYQSLKDLLEYEGNVEDDMMITFQISQTDLFGNPM  
 MYDLKENGDKIPITNENRKEFVNLYSYILNKSVEKQKAFRRGFHMTNESPLKYLFRPEEIEILLICGSRNL  
 FQALEETTEYDGGYTRDSVLIREFWEIVHSFTDEQKRLFLQFTTGDRAPVGGGLGKLMIIAKNGP DTERLPT  
 SHTCFNVLLLPEYSSKEKLERLLKAITYAKGFGML

### Ube3a compacte :

MKGLNDIFEAQKIEWHEKHHHHHPMKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEK  
 FPQVAATGDGPDIIIFWAHDRFGGYAQSGLLAEITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNK  
 DLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAG  
 L TFLVDLIKXKHMNADTDYSIAEAAFNKGETAMTINGPWAWNSNIDTSKVNYGVTVLPFTFKGQPSKPFVGVLS  
 AGINAASPNKELAKEFL ENYLLTDEGLEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMS  
 SAFWYAVRTAVINAASGRQTVDEALKDAQTNSSSNNNNNNNNNPMSENLYFQGAMSEIKMNKKGARIDF  
 KDVTYLTEEKVYEILELCREREDYSPLIRVIGRVFSSAEALVQSFRKVKDNNLQKLGPDVSDIDAIRRVYTR  
 LLSNEKIETAFNLALVYLSPNVECDLTYHNVYSRDPNYLNLFIIVMENRNLHSPYELEMALPLFCCKAMSKLPLA  
 AQGKLIIRLWSKYADQIRRMETTFQQLITYKVISNEFNSRNLVNDDDAIVAASKCLKMVYYANVVGGEVDTN  
 HNEEDDEEPIPELSELTLQELLGEERRNKKGPRVDPLETELGVKTLDCRKLIPFEFINEPLNEVLEMDKDY  
 TFFKVETENKFSFMTCPFILNAVTKNLGLYYDNRIRMYSEERRITVLYSLVQGGQLNPYLRLKVRRDHIIDALV  
 RLEMIAMENPADLKKQLYVEFEQGVDEGGVSKEFFQLVVEEIFNPDIGMFTYDESTKLFWFNPSSFETEG  
 QFTLIGIVLGLAIYNNCILDVHFPVYRKLGMKGGTFRDLGDSHPVLYQSLKDLLEYEGNVEDDMMITFQIS  
 QTDLFGNPMYDLKENGDKIPITNENRKEFVNLYSYILNKSVEKQKAFRRGFHMTNESPLKYLFRPEEIE  
 LLICGSRNLDFQALEETTEYDGGYTRDSVLIREFWEIVHSFTDEQKRLFLQFTTGDRAPVGGGLGKLMIIAK  
 NGP DTERLPTSHTCFNVLLLPEYSSKEKLERLLKAITYAKGFGML

### Ube3a DUS-1 :

MKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGIKVTVEHPDKLEEKFPQVAATGDGPDIIIFWAHDRFGGYA  
 QSGLLAEITPAAAFQDKLYPFTWDAVRYNGKLIAYPIAVEALS LIYNKDLLPNPPKTWEEIPALDKELKAKGKS  
 ALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGL TFLVDLIKXKHMNADTDYSIAEAAF  
 NKGETAMTINGPWAWNSNIDTSKVNYGVTVLPFTFKGQPSKPFVGVLSAGINAASPNKELAKEFL ENYLLTDEG  
 LEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDAAL  
 AAAQTNAAMADPHPSKKGASSAYLENSKGGSGSENLYFQGGSGSAPNNSCSEIKMNKKGARIDFKDVTYL  
 TEEKVYEILELCRERE  
 DYSPLIRVIGRVFSSAEALVQSFRKVKQHTKEELKSLQAKDEKDEDEKEKAA

## Ube3a DUS-2 :

MKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGKIVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYA  
QSGLLAEITPAAAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKS  
ALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAF  
NKGETAMTINGPWAWSNIDTSAVNYGTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEG  
LEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDAAL  
AAAQTNAAAMAGDNNLQKLGPDVSDIDAIRRVYTRLLSNEKIETAFNLALVYLSNVVYSRD  
PNYLNLFIIVMENRNLHSPYELMALPLFKAMSKLPLAAQGLIRLWSKYNADQIRMMETFQQLITYKVISN  
EFNSRNLVNDDDAIVAASKCLKMVYYANVVGGEVDTNHNEEDDEEPIPESELTLQE

## Ub3a DUS-3 :

MKIEEGKLVWINGDKGYNGLAEVGGKFEKDTGKIVTVEHPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYA  
QSGLLAEITPAAAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSIIYNKDLLPNPPKTWEEIPALDKELKAKGKS  
ALMFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGLTFLVDLIKNKHMNADTDYSIAEAAF  
NKGETAMTINGPWAWSNIDTSAVNYGTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEG  
LEAVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDAAL  
AAAQTNAAAMESSELTLQELLGEERRNKKGPRVDPLETELGVKTLDCRKLIPFEFFINEPLNEVLEMDKDYT  
FFKVETENKFSFMTCPFILNAVTKNLGLYYDNIRMYSEERRITVLYSLVQGGQLN

## 16E6 :

MFNLQEPYFTWPLIAADGGYAFKYENGGYDIKDVGVNDAGAKAGLTFLVDLIKTKHMNADTDYSIAEAAFNK  
GETAMTINGPWAWSNIDTSKVNYGTVLPTFKGQPSKPFVGVLSAGINAASPNKELAKEFLENYLLTDEGLE  
AVNKDKPLGAVALKSYEEELAKDPRIAATMENAQKGEIMPNIQMSAFWYAVRTAVINAASGRQTVDEALKD  
AQTNSSSNNNNNNNNNNNPMSENLYFQGAMFQDPQERPRKLPQLCTELQTTIHDIILECVYCKQQLLRREYV  
DFARRDLCIVYRDGNPYAVCDKCLKFYISKISEYRHYSYSLYGTTLQYQYNKPLSDLLIRCINCQKPLSPEEKQ  
RHLDKKQRFHNIRGRWTGRCMCSRSRSTRRETQL



Quatrième partie

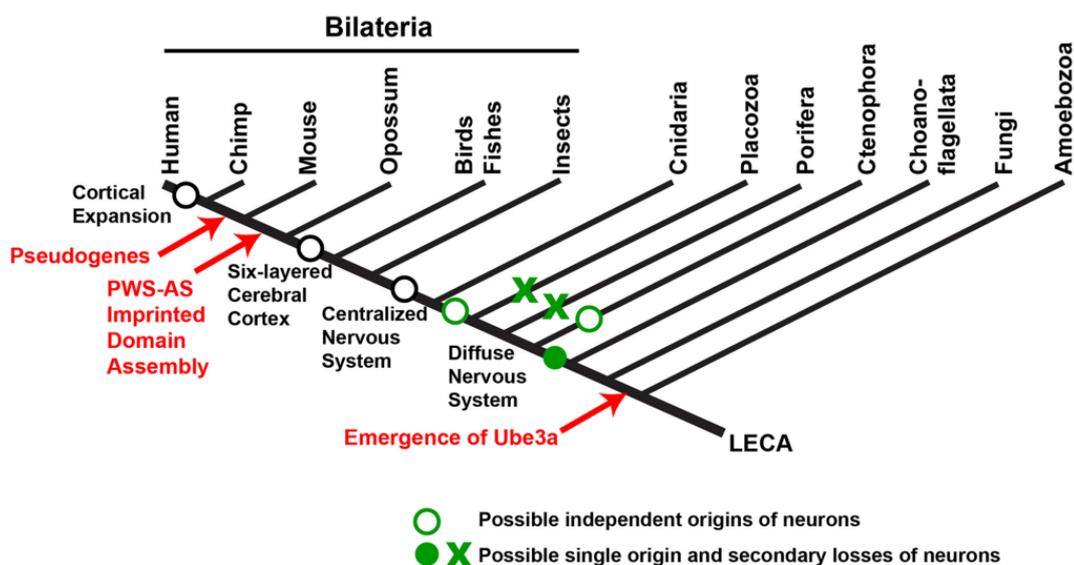
Conclusion & Perspectives



# 1 | Conclusion

La résolution des structures i) du domaine RLD2 de HERC2 et ii) du complexe formé par le domaine RLD2 et le peptide de Ube3a participe à une meilleure compréhension de l'interaction entre ces deux protéines. Cette interaction se fait entre une surface hautement conservée du domaine RLD2 de HERC2 et un motif que nous avons défini comme étant DxDKDxD.

L'analyse phylogénétique des séquences de HERC2 et de Ube3a a permis d'identifier la conservation du domaine RLD2 et de sa capacité à lier ce motif jusqu'au démosponge. Le motif DxDKDxD de Ube3a est, quant à lui, conservé jusqu'aux arthropodes. Ainsi les interactions entre ces deux E3 semblent être possibles chez les animaux, des mammifères jusqu'aux arthropodes. Cette capacité d'interaction entre Ube3a et HERC2 semble concorder avec l'apparition d'un système nerveux central chez les animaux [260].



**FIGURE 58** – Évolution de Ube3a en relation avec l'évolution du système nerveux. LECA pour "last eukaryotic common ancestor". Figure de [260]

De plus, une recherche de ce motif dans le protéome humain nous a permis d'identifier dif-

## 1. CONCLUSION

---

férentes protéines susceptibles de pouvoir interagir avec le domaine RLD2 de HERC2. La mesure de l'affinité des peptides de ces différentes protéines humaines contenant ce motif pour le domaine RLD2 ainsi que la résolution de structures du domaine RLD2 en complexe avec certains de ces peptides ont permis de mettre en évidence un réseau d'interactions (inconnu jusqu'alors) centré sur le domaine RLD2 de HERC2. Ainsi, nous avons identifié 8 protéines humaines, dont Ube3A, la RhoGEF DOCK10, la protéine péricentriolaire PCM1, la dé-ubiquitinase USP35 et quatre protéines du remodelage de la chromatine, BAZ2B, ARID4A, ARIP4, et MYT1 capables d'interagir avec le domaine RLD2 de HERC2.

Par la suite, nous avons étudié le rôle de l'interaction entre le domaine RLD2 de HERC2 et le motif DxDKxD de la protéine DOCK10. Cette étude révèle son rôle dans la formation des épines dendritiques des neurones. Ainsi, ce réseau d'interactions pourrait expliquer certains symptômes des syndromes d'Angelman et Angelman-like, mais aussi d'autres syndromes ou troubles du spectre autistique. La disparition ou le dysfonctionnement de la protéine HERC2 entraînerait la disparition totale de ce réseau d'interactions. C'est notamment le cas dans des syndromes Angelman-like. De plus, la disparition d'une des protéines de ce réseau (autre que HERC2), comme c'est le cas dans le syndrome d'Angelman où la protéine Ube3a n'est plus exprimée dans les neurones, entraîne une refonte de ce réseau. En effet, les différentes protéines possédant le motif DxDKxD sont en compétition pour interagir avec le domaine RLD2. Ainsi, en plus de la disparition de ses fonctions propres, l'absence de Ube3a entraînerait un changement de la hiérarchie des interactions avec RLD2. Ceci pourrait expliquer certains symptômes du syndrome d'Angelman n'étant pas directement liés aux fonctions de Ube3a.

La purification de différentes constructions de Ube3a associée à des prédictions de structures de celle-ci a permis de mieux comprendre la structure de Ube3a. Ici, nous avons identifié expérimentalement et par prédiction de structures l'interaction de différents domaines de Ube3a entre eux formant un domaine globulaire. Cette structure globulaire est composée des domaines DUS-1, DUS-2, DUS-3 et HECT. Des boucles entre ces domaines permettent l'interaction de Ube3a avec d'autres protéines. La boucle entre le domaine DUS-1 et DUS-2 contient le motif DxDKxD permettant l'interaction avec le domaine RLD2 de HERC2. Celle entre le domaine DUS-2 et DUS-3 contient le motif LxxLL permettant l'interaction avec E6.

Ces différentes observations ont permis la purification d'une construction compacte de Ube3a. Cette construction est dépourvue du domaine Azul, de la boucle entre celui-ci et le domaine DUS-1 ainsi que de la longue boucle entre le domaine DUS-1 et DUS-2. Cette construction se purifie uniquement sous forme monomérique et pourrait permettre la cristallisation et la résolution expérimentale de la structure de la protéine Ube3a.

Bien que le logiciel AlphaFold permette de produire des prédictions de bonne qualité, il

## 1. CONCLUSION

---

demeure important de travailler à la résolution de structures de la protéine Ube3a. En effet, ce logiciel ne permet pas, pour l'instant, de produire des prédictions de protéines avec des ligands et ne permet pas de prédire avec une grande précision la structure de ces complexes.

## 2 | Perspectives

Les différents résultats présentés ici ouvrent de nouvelles pistes de recherche pour, d'une part, tenter de mieux comprendre les syndromes Angelman et associés et, d'autre part, étudier les mécanismes de l'activité ubiquitine-ligase de Ube3a et sa régulation.

Tout d'abord, l'étude structurale de la protéine Ube3a seule et avec ses différents interactants est toujours en cours. La mise en place d'une purification permettant l'obtention d'une protéine Ube3a pure et monomérique dans des quantités suffisantes permet la continuation des tests de cristallisation de cette protéine.

De plus, les améliorations des techniques de microscopie électronique, depuis la préparation de grilles jusqu'aux méthodes de traitement de données, permettent l'étude structurale d'objets de tailles de plus en plus petites. Ainsi, des études de la structure de différentes protéines pourraient permettre de mieux comprendre leur rôle dans ces syndromes et dans le neurodéveloppement. L'étude structurale des protéines HERC2 et Ube3a par microscopie électronique est importante. C'est aussi le cas pour les différentes protéines de la famille DOCK-D (9, 10 et 11). Cette étude pourra être associée à des tests de cristallographie pour déterminer la structure des différents domaines de ces protéines.

L'utilisation des dernières méthodes développées dans l'équipe pourrait permettre la meilleure compréhension du rôle du réseau d'interactions centré sur le domaine RLD2. La technique de "holdup assay" [310] [311] ainsi que la technique "native HoldUp" récemment développée dans l'équipe [184] seront appliquées à ce projet. Cette dernière technique permettra d'identifier les partenaires d'un motif, d'un domaine ou d'une protéine dans des extraits cellulaires de mammifères. Cette identification de partenaires se fera par spectrométrie de masse et permettra la détermination de constantes d'affinités pour ces protéines entières dans des conditions proches du contexte *in vivo*.

Enfin, l'exploration du mécanisme d'ubiquitination de la protéine Ube3a et de la stimulation de son activité ubiquitine-ligase par la protéine E6 est en projet. En effet, la liaison de la protéine E6 à l'hélice  $\alpha 1$  du domaine HECT pourrait jouer un rôle dans cette activité. Ainsi, différentes constructions de la protéine Ube3a (entière,  $\Delta$ -Azul et compacte) mais

## 2. PERSPECTIVES

---

aussi des constructions de Ube3a mutée au niveau de cette hélice seront utilisées pour effectuer des tests d'ubiquitination avec et en absence de la protéine E6.

Pour conclure, les travaux présentés ici, ainsi que les travaux réalisés par d'autres laboratoires sur les syndromes de type Angelman offrent des clés de compréhension de ceux-ci et des mécanismes du neurodéveloppement. À terme, ces différentes études pourraient permettre une meilleure prise en charge des patient·e·s.



## Cinquième partie

### Annexes



# 1 | Annexe 1

Le papier suivant porte sur le développement d'une stratégie de purification de protéines de fusion monodisperses et actives. Cette méthode est appliquée à la purification de la protéine E6 de HPV.

La purification de la protéine E6 de HPV est périlleuse. En effet, cette protéine a une forte tendance à oligomériser. La purification de cette protéine, notamment pour des études cristallographiques, nécessite quelques précautions. La purification de cette protéine est faite en présence d'un antioxydant. Différentes stratégies ont été mises en place pour purifier les protéines E6. Celles-ci consistent en la mutation des cystéines exposées au solvant et non impliquées dans le repliement de la protéine. De plus, les tampons utilisés pour la purification sont dégazés et argonés pour empêcher la formation de ponts disulfures. En effet, la protéine E6 contient 8 cystéines conservées permettant sa liaison au zinc. La fusion à la MBP et une étape d'ultracentrifugation permettent la purification de E6 sous forme monomérique.

Une méthode de purification de E6 en une seule étape est présentée ici. Il s'agit d'une purification utilisant une résine nickel pour purifier E6 associée à une étiquette 6HIS. Différents types de résines et différents temps d'incubation sont testés pour obtenir le meilleur taux de E6 monomérique. Ensuite, des tests ont été effectués pour vérifier que cette E6 purifiée est biologiquement active, via sa capacité à lier le motif LxxLL.

J'ai participé à la réalisation des expériences lors de mon stage de master sous l'encadrement d'Anna Bonheure.

RESEARCH

Open Access



# One-step affinity purification of fusion proteins with optimal monodispersity and biological activity: application to aggregation-prone HPV E6 proteins

Anna Bonhoure<sup>1</sup>, Auguste Demenge<sup>1</sup>, Camille Kostmann<sup>1</sup>, Leticia San José<sup>2</sup>, Eva De la Cal<sup>2</sup>, Pilar Armisen<sup>2</sup>, Yves Nominé<sup>1</sup> and Gilles Trave<sup>1\*</sup>

## Abstract

**Background:** Bacterial expression and purification of recombinant proteins under homogeneous active form is often challenging. Fusion to highly soluble carrier proteins such as Maltose Binding Protein (MBP) often improves their folding and solubility, but self-association may still occur. For instance, HPV E6 oncoproteins, when produced as MBP-E6 fusions, are expressed as mixtures of biologically inactive oligomers and active monomers. While a protocol was previously developed to isolate MBP-E6 monomers for structural studies, it allows the purification of only one MBP-E6 construct at the time. Here, we explored a parallelizable strategy more adapted for biophysical assays aiming at comparing different E6 proteins.

**Results:** In this study, we took advantage of the distinct size and diffusion properties of MBP-E6 monomers and oligomers to separate these two species using a rapid batch preparation protocol on affinity resins. We optimized resin reticulation, contact time and elution method in order to maximize the proportion of monomeric MBP-E6 in the final sample. Analytical size-exclusion chromatography was used to quantify the different protein species after purification. Thus, we developed a rapid, single-step protocol for the parallel purification of highly monomeric MBP-E6 samples. MBP-fused HPV16 E6 samples obtained by this approach were validated by testing the binding to their prototypical peptide targets (the LXXLL motif from ubiquitin ligase E6AP) by BIAcore-SPR assay.

**Conclusions:** We have designed a rapid single-step batch affinity purification approach to isolate biologically active monomers of MBP-fused E6 proteins. This protocol should be generalizable to isolate the monomer (or the minimal biologically active oligomer) of other proteins prone to self-association.

**Keywords:** Protein aggregation, Protein solubility, MBP fusion, IMAC, One-step purification

\*Correspondence: gilles.trave@igbmc.fr

<sup>1</sup> Équipe Labellisée Ligue 2015, Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258/CNRS UMR 7104/Université de Strasbourg, 1 rue Laurent Fries, BP 10142, 67404 Illkirch, France

Full list of author information is available at the end of the article



## Background

Protein solubility is a major issue in recombinant protein purification from *Escherichia coli*. It is influenced by different parameters, such as proper folding and/or aggregation. Under the stress induced by high rates of heterologous protein expression, inactive misfolded polypeptides accumulate as inclusion bodies. They can be solubilized by using chaotropic agents (urea or guanidium hydrochloride) and then refolded in vitro (by dilution or dialysis). However, such an approach can be time-consuming since it requires optimization adapted to every protein, and the yield of final soluble product can be low. Another strategy consists of fusing the protein of interest (also called “passenger protein”) to a solubility-enhancing protein (also called “carrier protein”), such as Thioredoxin, glutathione-S-transferase (GST) or maltose-binding protein (MBP). MBP was reported to be particularly efficient in improving the solubility of its fusion partners [1, 2]. In addition to its ability to stabilize and solubilize the passenger protein, MBP can be used as an affinity tag for purification on amylose resin [3, 4]. However, an MBP-fused protein can sometimes be solubilized in the form of a mixture of properly folded monomers (or minimal biologically active oligomers) and of large oligomers in which the passenger protein is self-associated and its folding and/or biochemical activity may be altered [5–13]. In such situations, the challenge consists in exploring conditions of expression and purification favoring the biologically active monomeric -or minimally oligomeric-MBP-fused samples [9, 11, 14, 15].

E6 proteins produced by oncogenic human papillomaviruses (HPV) are a prototypical case of proteins that display unstable biochemical behavior when produced recombinantly [16–18]. While unfused E6 proteins are expressed in insoluble form, MBP-fused E6 proteins are soluble yet prone to self-association, leading to soluble oligomers, which in turn are unable to specifically interact with protein partners [9, 10, 19]. Furthermore, E6 proteins are generally cysteine-rich. E6 proteins contain two zinc-binding domains, each involving four cysteine residues to coordinate one zinc ion. In addition to these eight cysteine residues highly conserved for structural purposes, E6 proteins contain additional non-conserved cysteine amino acids. For instance, wild-type HPV16 E6 and HPV8 E6 have a total of 14 and 16 cysteine residues, respectively. Oxidation promotes the formation of intermolecular disulfide bridges, and thus increases the aggregative propensity of E6 proteins.

Papillomaviruses E6 oncoproteins establish numerous interactions with host proteins [16–18, 20]. For instance, E6 from “high-risk” mucosal HPV was reported to hijack the ubiquitin ligase E6AP (E6-associated protein), resulting in the proteasome-mediated degradation of the

tumor suppressor p53 [21]. Crystal structures of the HPV16 E6 oncoprotein from high-risk mucosal HPV16 in complex with a minimal target fragment from E6AP and the core domain of p53 were solved by X-ray crystallography, using solubility-enhanced mutants of E6 fused to MBP [22, 23].

There are more than 200 HPV types, with different tropisms (mucosal, cutaneous) causing a large variety of phenotypes ranging from warts and condylomas for low-risk HPVs to malignant tumors for high-risk HPVs [24–26]. The present work will focus on E6 oncoproteins from HPV16 and HPV8. HPV16 is the highest-risk mucosal HPV, responsible for more than 50% of cervical cancers and more than 90% of HPV-positive oropharyngeal cancers [27, 28]; whereas HPV8 is a cutaneous HPV type that can generate skin cancers, in particular in individuals with genetic or immunological diseases [29].

Since E6 proteins from different HPV types were reported to have distinct protein interaction preferences by proteomic studies [20], it would be interesting to further decipher their binding properties by quantitative biophysical and structural approaches. The protocol used in our former publications [9–12] allows the isolation of soluble, monomeric E6 proteins by amylose affinity chromatography followed by overnight ultracentrifugation and size-exclusion chromatography (SEC) for the elimination of protein oligomers. Such a strategy is appropriate for structural studies that require relatively large amounts of high quality protein material. However, it allows the purification of only one protein at a time. To decipher the protein–protein interactions of a panel of E6 oncoproteins, parallel purifications would be more appropriate. In addition, a faster purification protocol would reduce the risk of protein aggregation over time.

In the present work, we exploited the distinct size and diffusion properties of monomers and oligomers of MBP-E6 to separate these two species using a rapid batch affinity preparation approach. This led us to obtain a fast, single-step protocol for the preparation of two test E6 oncoproteins under monomeric and biologically active form: HPV16 F47R 4C/4S E6 (a solubility-enhanced mutant of HPV16 E6) and HPV8 E6 (prone to rapid self-association). The long-lasting and non-parallelizable steps of the previous protocol (ultracentrifugation and preparative gel-filtration) were substituted by a fast batch affinity chromatography on reticulated nickel resins. The reticulation state of the resin, the elution protocol and the contact time were adjusted for optimal separation of monomeric and oligomeric MBP-fused E6 proteins. We show that this new and fast protocol is effective enough to obtain soluble and active E6 protein samples suitable for protein–peptide interaction assay as revealed by surface plasmon resonance (SPR). This protocol is amenable

to robotization, offering the ability to prepare in a parallel fashion multiple protein samples displaying optimal monodispersity.

## Results

### Protein constructs used for this study

To develop and evaluate a new E6 purification protocol, we used two distinct HPV E6 oncoproteins. On the one hand, HPV16 E6 F47R 4C/4S (hereafter named 16E6mut) is a solubility-optimized mutant of HPV16 E6. The structure of this mutant has been solved by X-ray crystallography and several of its interactions have been precisely characterized by surface plasmon resonance (SPR) and Isothermal Titration Calorimetry [22]. On the other hand, HPV8 E6 (hereafter named 8E6) has a strong tendency to self-associate and is more challenging to purify as a monomer (unpublished observations).

As depicted in Fig. 1, HPV E6 constructs were overexpressed as fusions to the C-terminus of bacterial MBP. MBP is known to favor solubilization of recalcitrant recombinant proteins [1, 2], and can be used for affinity purification on amylose resin. In addition, the constructs include a 6-His tag for immobilized-metal affinity chromatography (IMAC) on Nickel resin and a TEV (Tobacco Etch Virus) protease cleavage site allowing elimination of N-ter purification tags.

### Impact of the reticulation level of affinity resins on the aggregation level of purified MBP-E6 samples

In the protocol published previously for the purification of soluble MBP-E6 protein [9, 10], the initial affinity step was performed on amylose resin. In such conditions, a fraction of the MBP-E6 sample formed oligomers that were detected by static light scattering in a fluorimeter. These soluble oligomers can be eliminated by means of overnight ultracentrifugation followed by preparative gel-filtration [19]. However, this approach requires substantial amounts of protein in order to be efficient and is not suitable for parallel purification of HPV E6 proteins. We hypothesized that testing resins of different reticulation degree might be an alternative strategy to reduce the binding of the oligomeric proteins to the affinity resin since large particles cannot interact with a highly

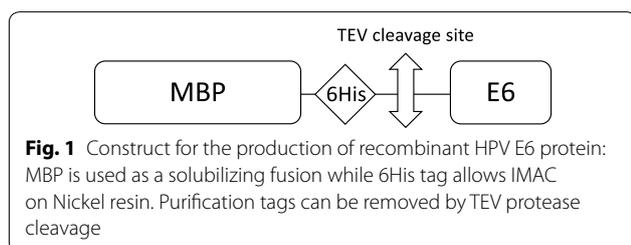
reticulated resin, thus promoting the preferential interaction of small molecules with the affinity resin.

Here, we compared the aggregation levels of purified MBP-E6 protein obtained either with the same amylose resin than the one used in the initial protocol (hereafter called A), or with six different reticulated nickel resins (numbered from N1 to N6, as detailed in Table 1). In order to test the preferential binding of the monomer to the different resins, elution was performed by centrifugal filtration on a filter plate to maximize the recovered liquid fraction and the detection of all protein species eluted from the resin. The purity of the final fractions was satisfactory, as checked by microfluidics capillary electrophoresis in denaturing conditions (Fig. 2). The bands corresponding to the protein of interest are visible between the 44 and 72 kDa markers for MBP-16E6mut (expected MW = 61.4 kDa) and between the 48 and 69 kDa markers for MBP-8E6 (expected MW = 60.9 kDa).

The oligomeric state of the purified protein samples was assessed by analytical SEC, based on absorbance at 280 nm. Figure 3 shows chromatograms of MBP-16E6mut (A) and MBP-8E6 (B) purified using the initial protocol in which the bacterial lysate of overexpressed E6 protein was incubated 2 h with amylose resin. After washing the resin, protein was eluted by maltose-containing buffer and recovered by centrifugal filtration.

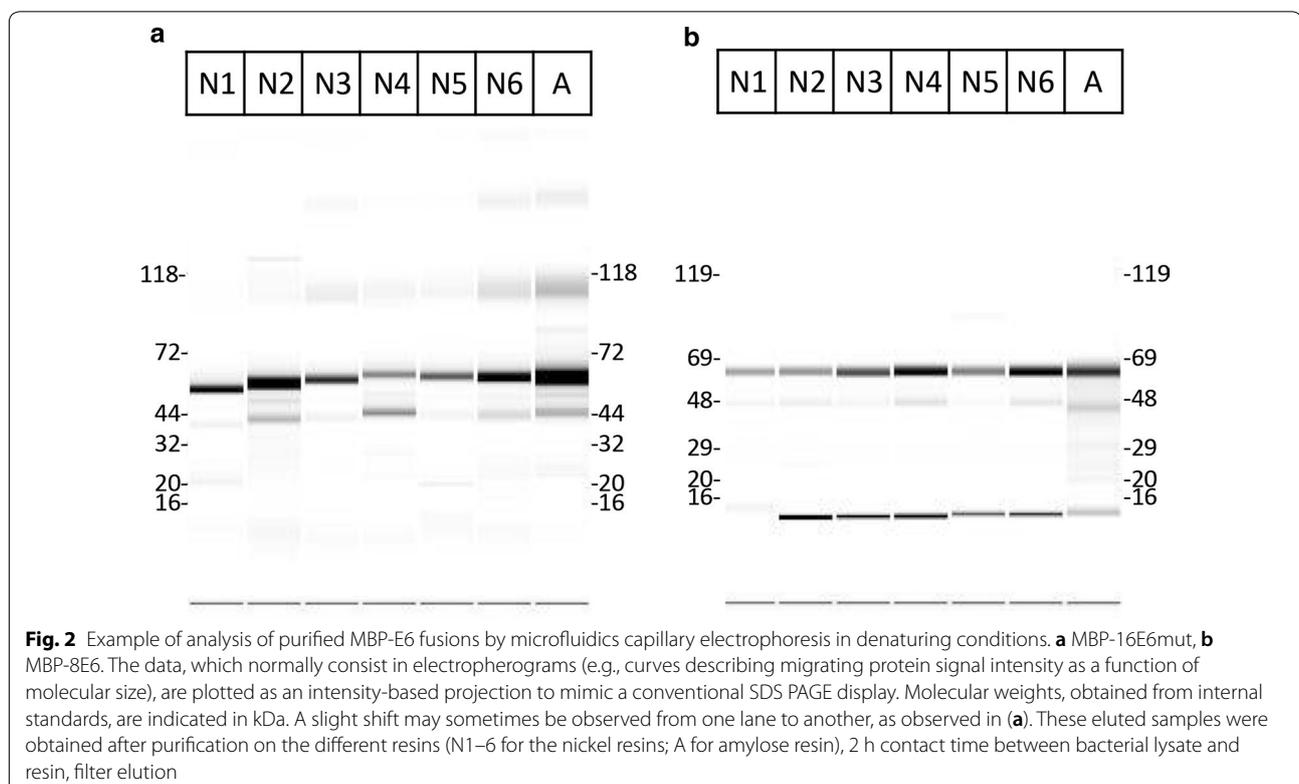
According to the column calibration based on (Eqs. 1 and 2), the elution volume for MBP-16E6mut purified on amylose resin ( $V_e = 1.90$  mL) corresponds to a molecular weight MW of 75 kDa and a hydrodynamic radius  $R_h$  of 37 Å. For MBP-8E6, a similar peak appears at an elution volume  $V_e$  of 1.91 mL, which would correspond to a molecular weight MW = 72 kDa and a hydrodynamic radius  $R_h = 37$  Å. Considering the accuracy of the method and the fact that calibration is based on the diffusion properties of globular proteins, we assume that both peaks refer to monomeric MBP-E6 (theoretical molecular weights: MW (MBP-16E6mut) = 61.4 kDa, MW (MBP-8E6) = 60.9 kDa). Since MBP-E6 fusions are not globular [22], the experimental elution volumes are expected to be slightly lower than the calculated elution volume obtained from the calibration ( $V_{e\text{theo}} = 1.95$  mL).

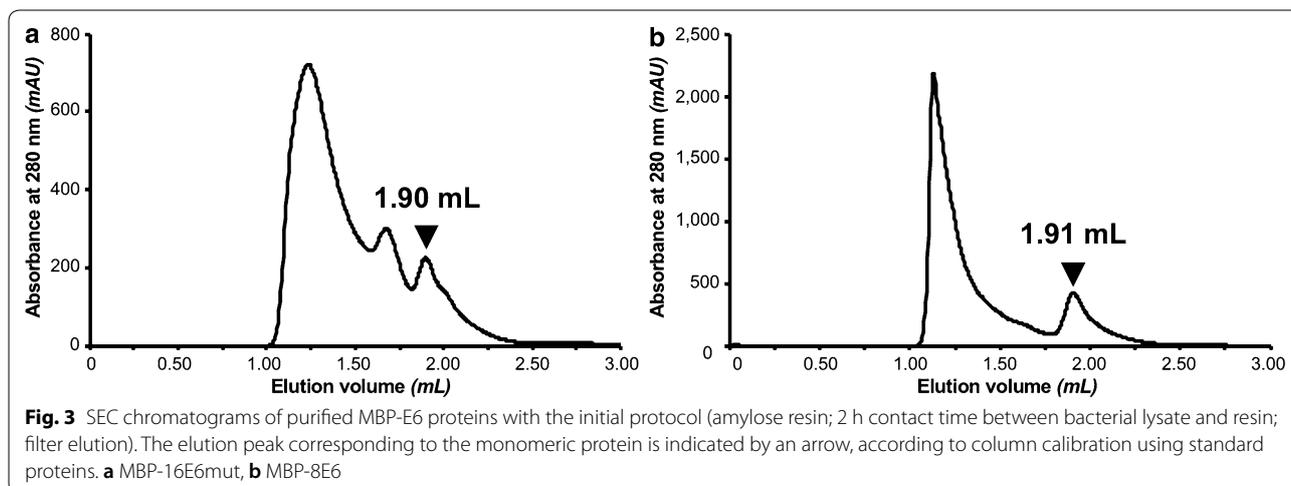
For each tested resin, analytical SEC was performed. Based on absorbance at 280 nm, the monomeric and oligomeric protein were detected at a given elution volume. We measured the following elution volumes for monomeric MBP-16E6mut and MBP-8E6 at  $V_e = 1.92 \pm 0.01$  mL and MBP-8E6:  $V_e = 1.99 \pm 0.04$  mL, respectively. All peaks with  $V_e < 1.90$  mL correspond to particles larger than monomer, that were therefore considered as oligomeric proteins (Fig. 3, Additional file 1). Elution peaks were integrated based on the signal of absorbance at 280 nm. Their areas were utilized



**Table 1** Panel of resins used for affinity purification, gathering technical data provided by the manufacturers

	Manufacturer	Catalogue ref	Reference	Affinity	Nickel ligand	Features	Exclusion limit for globular proteins	Metal ion capacity/binding capacity	Max. pressure
N 1	GE Healthcare	17-5318-06	Ni Sepharose 6 Fast Flow	Nickel	NTA	6% cross-linked agarose	4000 kDa	Approx 15 $\mu\text{mol Ni}^{2+}$ /mL gel	0.1 MPa
N 2	ABT Agarose Beads Bio-technologies	10BCL-QHNI-5	High Density Nickel 10 BCL-QHNI-5	Nickel	IDA	10% cross-linked agarose	500 kDa	20–40 $\mu\text{mol Ni}^{2+}$ /mL gel	Low pressure
N 3	ABT Agarose Beads Bio-technologies	6BCL-QHNI-2	High Density Nickel 6 BCL-QHNI-2	Nickel	IDA	6% cross-linked agarose	4000 kDa	20–40 $\mu\text{mol Ni}^{2+}$ /mL gel	Low pressure
N 4	ABT Agarose Beads Bio-technologies	6BCL-NTANI-2	Nickel NTA agarose resin 6 BCL-NTANI-2	Nickel	NTA	6% cross-linked agarose	4000 kDa	5–19 $\mu\text{mol Ni}^{2+}$ /mL gel	Low pressure
N 5	ABT Agarose Beads Bio-technologies	6NiRR-2	Nickel Rapid Run 6 NiRR-2	Nickel	IDA	6% agarose	4000 kDa	5–19 $\mu\text{mol Ni}^{2+}$ /mL gel	0.3 MPa
N 6	ABT Agarose Beads Bio-technologies	6NTANI-2	Nickel NTA rapid run 6 RR-NTANI-2	Nickel	NTA	6% agarose	4000 kDa	5–19 $\mu\text{mol Ni}^{2+}$ /mL gel	0.3 MPa
A	NEB New England Biolabs	E8022S	Amylose resin high flow	Maltose		Cross-linked agarose	Unspecified	7 mg MBP-protein/mL gel	0.5 MPa





to quantify monomeric and oligomeric species. As summarized in Table 2, the fraction of monomer in the same purification conditions is generally higher for MBP-16E6mut than for MBP-8E6. This result confirms that MBP-8E6 is highly prone to self-association, consistent with our previous observations.

The difference between the MBP-16E6mut reference protein and the challenging MBP-8E6 protein is particularly visible in Fig. 4, since the proportion of oligomers over total purified protein is much higher for MBP-8E6 (values comprised between 70 and 93% for the different tested nickel resins, Table 2) than for MBP-16E6mut (between 5 and 70%). It also clearly appears that amylose resin retains a higher proportion (90%) of MBP-16E6mut oligomers than nickel resins, confirming the hypothesis of oligomer exclusion by reticulated nickel resins. A standard deviation of 15% unit was inferred from the duplication of several experiments, warranting significance. Three nickel resins allow to purify 85% or more of monomeric MBP-16E6mut, namely N1 (95%), N2 (87%) and N4 (85%). In addition, quantification of the total protein amount (Table 2, see “Total A280 GF Peak Area”) shows that N4 resin has the highest recovery capacity of total purified protein. Regarding the purified MBP-8E6 sample, the same three resins N1, N2 and N4 show the maximal ratio of monomeric state (19%, 30% and 22%, respectively), although the monomer fraction obtained with the amylose resin (16%) is within the same range as that obtained with nickel resins (7 to 30%).

The present assays showed the usefulness of using a reticulated nickel resin for the purification of MBP-16E6mut, since purification trials allowed to reach up to 95% monomer in a single purification step. However, MBP-8E6, being more prone to oligomerization, could

not yet be purified as a predominantly monomeric state at this step.

#### Optimizing the elution protocol to maximize the proportion of monomeric protein

At the previous stage, resins were thoroughly dried by centrifugal filtration in order to recover maximal liquid fraction and to quantify all oligomer species that were bound (filter elution protocol). After probing the efficiency of nickel reticulated resins, we implemented another elution method aimed at reducing the fraction of oligomers in the final purified fraction. Instead of being filtrated, the resin/elution buffer mixture was centrifuged at low speed (5 min at 500×g) and the supernatant was collected immediately after centrifugation. The oligomers are expected to diffuse out of the resin slower than the monomers, thereby increasing the proportion of monomers in the final pipetted material.

Using this adapted elution protocol, the proportion of monomeric MBP-8E6 was significantly improved and reached up to 71% for resin N1 (Fig. 5). Resins N2 and N4 allowed to purify up to 56% of monomeric protein, but N1 remains the only resin allowing to purify the highest total protein concentration and the largest monomer ratio (as defined by the proportion of monomer over the total amount of protein in the sample) for both MBP-8E6 and MBP-16E6mut samples. For MBP-16E6mut, the two resins reaching highest monomer ratio with the new elution method are N1 and N2, with approximately 86% and 88%, respectively. Although the monomer fraction was significantly increased with this “decant and take up” elution procedure, it is noticeable that the total protein amount was reduced by two-fold for most of the resins. By adapting the elution method, we have shown that it is possible to significantly reduce the diffusion of

**Table 2 Summary table for purification results: Monomer are defined based on their elution volume, as it follows: MBP-16E6mut:  $V_e = 1.92 \pm 0.01$  mL; MBP-8E6:  $V_e = 1.99 \pm 0.04$  mL Oligomers are defined as all protein species eluted between void volume ( $V_0$ ) (included) and monomer elution volume (excluded)**

	Monomer (%)	Oligomer (%)	Total A280 GF peak area (mAU × mL)
A: MBP-16E6mut			
(1) Filter elution			
N1	95	5	256
N2	87	13	243
N3	52	48	167
N4	85	15	480
N5	30	70	116
N6	68	32	345
A	10	90	263
(2) Pipetting elution			
N1	86	14	366
N2	88	12	125
N3	60	40	125
N4	72	28	128
N5	39	61	39
N6	69	31	155
B: MBP-8E6			
(1) Filter elution			
N1	19	81	467
N2	30	70	113
N3	12	88	230
N4	22	78	256
N5	7	93	194
N6	9	91	404
A	16	84	486
(2) Pipetting elution			
N1	71	29	301
N2	56	44	117
N3	28	72	227
N4	56	44	121
N5	16	84	156
N6	74	26	89
(3) Optimization of lysate incubation time			
N1			
5 min	74	26	185
30 min	77	23	160
2 h	56	44	220
N2			
5 min	92	8	25
30 min	71	29	38
2 h	77	23	138

Quantification of monomeric and oligomeric species is based on the integration of elution peak (absorbance at 280 nm), with estimated uncertainty of  $\pm 15\%$  for monomer/oligomer ratio and  $\pm 57$  mAU × mL for total A80 peak area based on duplicated experiments

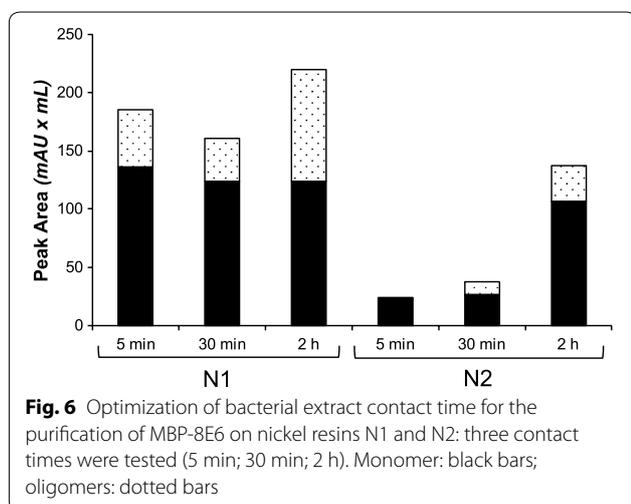
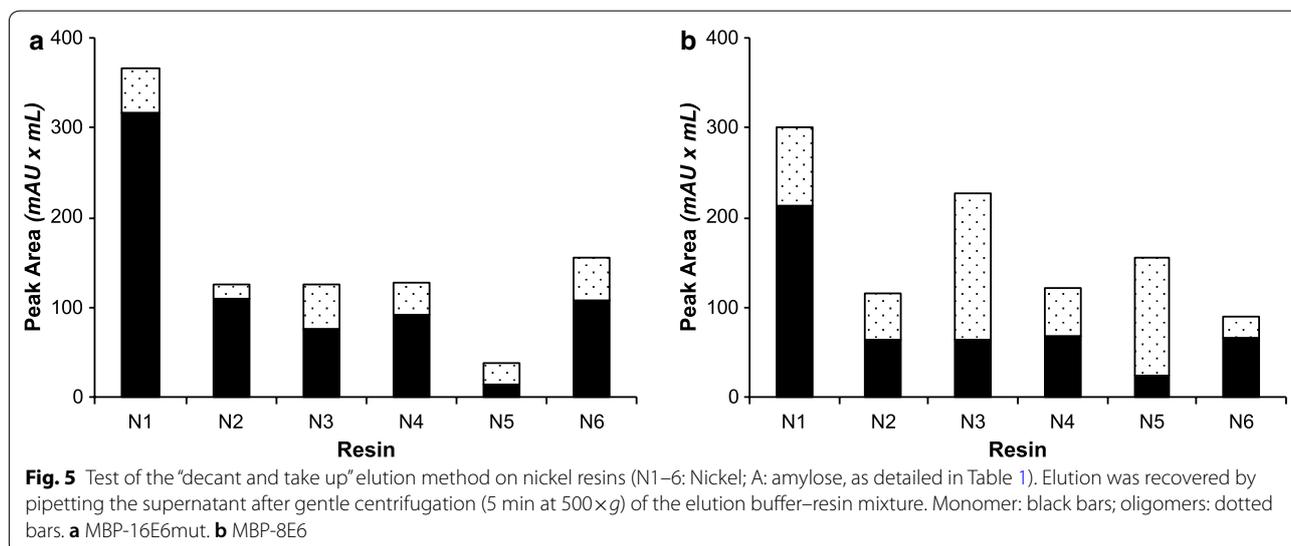
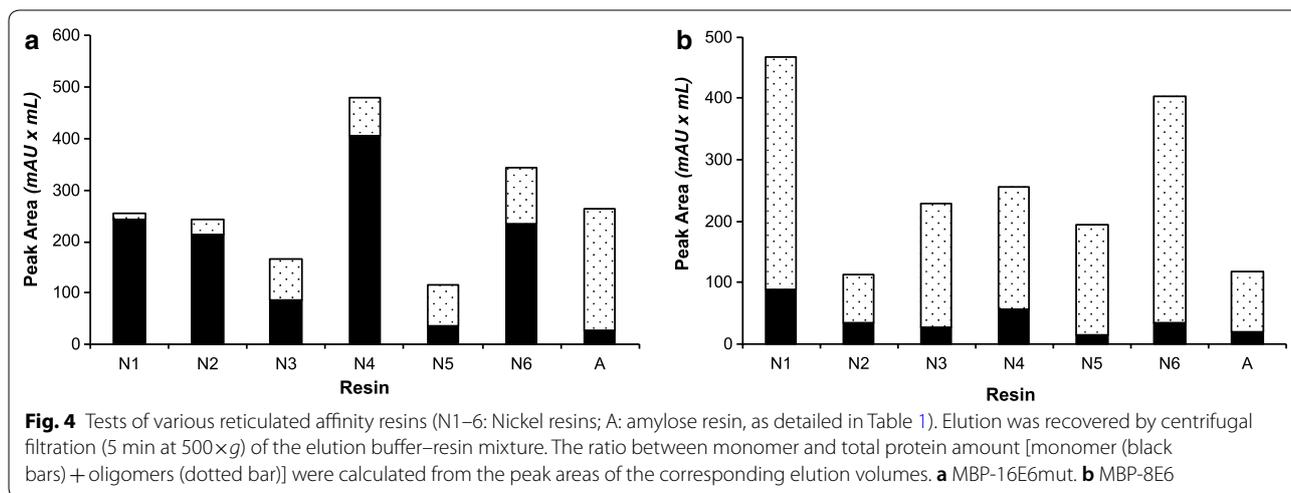
oligomeric species in solution, and subsequently to maximize the ratio of monomeric purified protein.

#### Reducing the contact time of bacterial extract with resin during the initial binding step limits protein oligomer binding

Contact time between affinity resin and cleared bacterial lysate during the initial binding step is another critical parameter for protein purification since it determines the amount and the nature of protein species bound to the resin. On the one hand, a long incubation time should increase the total amount of protein that will be recovered. On the other hand, a shorter incubation time might restrict the size of the proteins interacting with the resin. As large particles such as soluble oligomers diffuse slower in solution, their binding on affinity resin should thus be limited when incubation time is short.

To assess this hypothesis, we selected two resins showing the highest monomer ratio for both the reference MBP-16E6mut and the challenging MBP-8E6 proteins, namely N1 and N2. Since the “decant and take up” elution method already improved the fraction of monomeric MBP-16E6mut by up to 70–90%, this third step in the optimization protocol was focused on MBP-8E6, the protein most prone to self-association in our study. Incubation time was set at either 30 min or 5 min and systematically compared to the 2-h reference time used in the previous trials. Elution was done by pipetting the supernatant after low-speed centrifugation of the resin resuspended in elution buffer. The results presented in Fig. 6 show that the shorter the contact time between the cleared lysate and the resin, the higher the proportion of monomer in the final purified fraction, confirming our assumption. With the N1 resin, about 75% of monomers were purified with either 30 min or 5 min contact time whereas this ratio dropped down to 56% for the 2-h incubation delay. The best monomer ratio for monomeric MBP-8E6 (92%) was obtained by using the N2 resin for purification, with a 5 min contact time and the “decant and take up” elution method. Considering that only 30% of monomers were obtained using this same resin and applying the initial protocol (i.e. 2-h contact time and elution by filtration), this result shows that our optimization brought a significant improvement to purify a biochemically problematic HPV E6 protein. One should note that this quality refinement of the purified protein is at the expense of a decrease of the total amount of the MBP-8E6 protein. It is particularly visible from the purifications with the N2 resin, for which a fivefold decrease of the protein quantity was observed between 2-h and 5-min contact time.

For MBP-16E6mut, the absolute amount of monomeric protein was further increased by applying the above

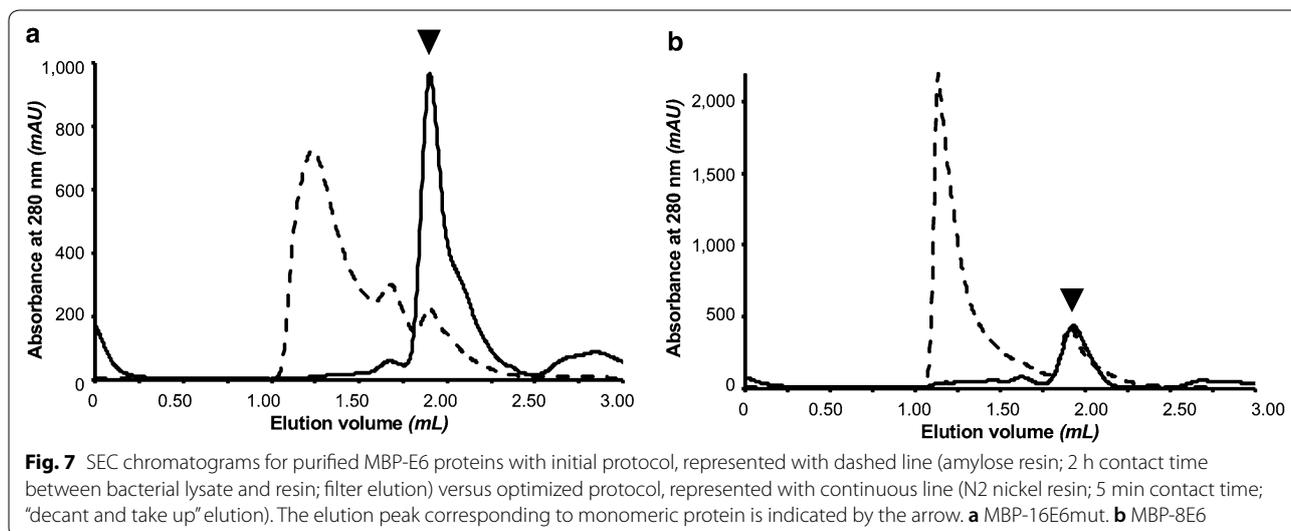


refined protocol (Fig. 7). In contrast, the amount of monomeric MBP-8E6 remained approximately the same with both purification strategies. The increase of the ratio of monomeric protein with the optimized protocol can be explained by the drastic elimination of oligomeric species (in particular those from the void volume at 1.18 mL).

**Probing the activity of the purified protein by SPR**

In order to assess the activity of protein samples purified according to the optimized protocol, we performed a protein–peptide interaction assay by SPR. This was achieved using 16E6mut in order to compare with SPR data previously obtained on a related 16E6 mutant using the original large-scale approach [30].

We purified MBP-16E6mut according to our batch protocol: 5 min contact time between the bacterial

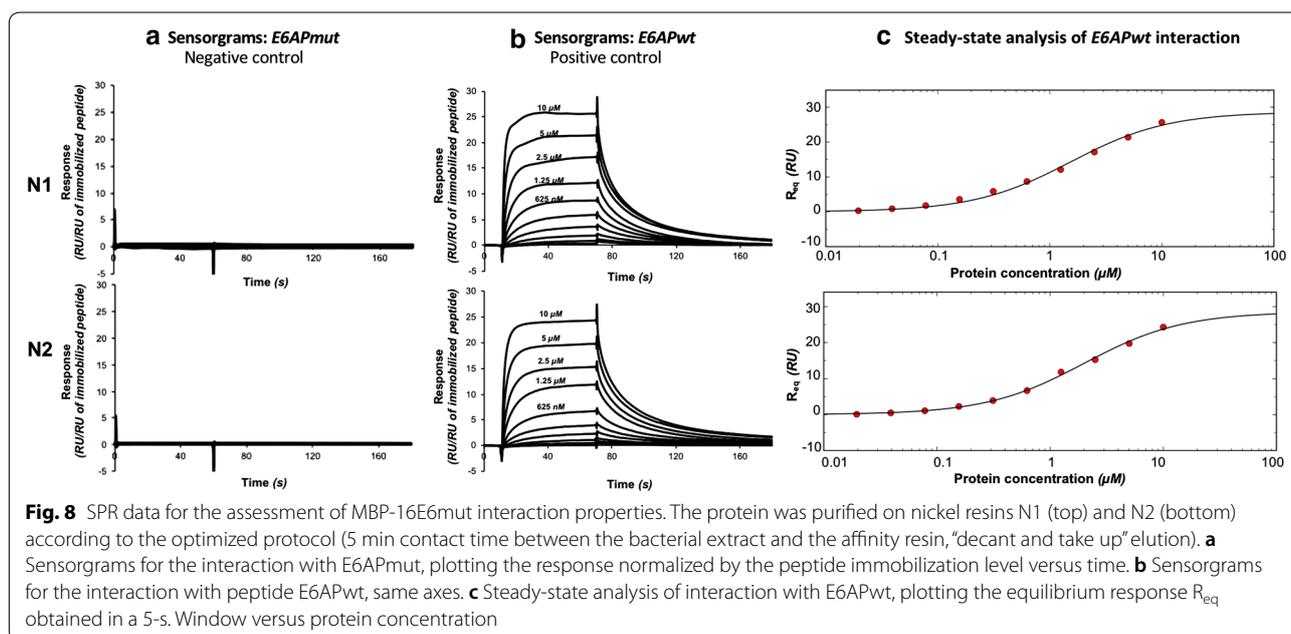


lysate and the affinity resin, followed by “decant and take up” elution by pipetting the supernatant after short centrifugation in Eppendorf tube. We performed this protocol with N1 and N2, the two resins that gave the highest monomer ratio after purification with the optimized elution protocol. In order to remove imidazole from protein solution, the final fractions were desalted in the SPR running buffer by bench desalting column.

Following this protocol, we purified the equivalent of 8.5 μmol of MBP-16E6mut per liter of bacterial culture, which is largely sufficient to perform parallel SPR assays

(in our setup, the minimum protein amount required for one SPR assay including 10 cycles is 2 nmole).

MBP-E6 was tested against its prototypical target peptide, the 16-mer LXXLL motif of E6AP (PESELTLQELL-GEER, called thereafter *E6APwt*) [21]. The CAPture kit system (GE Healthcare) was used to reversibly immobilize biotinylated peptides on the chip. Each cycle consists of an immobilization of the peptide, followed by the injection of the protein at a given concentration, and a regeneration step to wash out both the remaining protein and the immobilized peptide of the surface. Figure 8b shows the superimposition of time-aligned sensorgrams



obtained for different concentrations after reference subtraction and normalization according to the peptide immobilization level. As a negative control, we tested the interaction against an E6AP mutant (PESELTAQELLGEER, called thereafter *E6APmut*) abolishing the binding (Fig. 8a). No significant signal was observed for this control. In particular, the association phase does not display any concentration-dependent signal.

Steady-state analysis was performed to estimate dissociation constants ( $K_D$ ) of the MBP-16E6mut interacting with *E6APwt* (Fig. 8c). For the protein samples purified on resins N1 and N2, the estimated  $K_D$  were  $1.6 \pm 0.3 \mu\text{M}$  and  $2.0 \pm 0.3 \mu\text{M}$ , respectively. These data are in close agreement with a previously published SPR analysis of the binding of purified HPV16 E6 to the LXXLL motif of E6AP [30]. In that study, the obtained  $K_D$  was  $4.15 \pm 0.52 \mu\text{M}$ . The moderate difference in the absolute affinity constant measured in that previous work might be due to several changes in the experimental setup. In the former publication, the recombinant peptide was fused to GST and further captured on CM5 chip via an anti-GST antibody, whereas biotinylated synthetic peptides used in the present study were reversibly captured on CAP chip. In addition, the previous publication used a distinct mutant of HPV16 E6, called E6 6C/6S. This former mutant included two additional cys/ser mutations that were later shown to slightly destabilize the N-terminal domain of HPV16 E6 [31]. Despite these experimental differences, the good agreement between affinity measurements indicates that the optimized single-step purification protocol is efficient to obtain protein sample in an active form.

## Discussion

Quality of recombinant protein sample can be improved by optimizing either the expression or the purification process. The present work mainly focused on the purification process to maximize the yield of soluble overexpressed protein. Our purpose was to obtain an MBP-E6 sample as homogeneous as possible from the available material based on the expression conditions in *E. coli* that we previously published [9].

E6 oncoproteins from papillomaviruses contain two zinc-binding domains and have been found to interact with some host proteins through LXXLL motif recognition, such as E6AP ubiquitin ligase [32]. After overexpression in *E. coli* and purification, it was reported that MBP-E6 forms mixtures of soluble active monomers and soluble oligomers comprising misfolded E6 moieties (thus unable to specifically bind LXXLL motifs) [10]. These oligomers can be eliminated by overnight ultracentrifugation followed by size-exclusion chromatography [22]. Although this approach is appropriate for X-ray

crystallography, a faster and easier protocol was designed here for the purpose of performing protein-peptide interaction assays in parallel. As compared to structural studies, binding assays generally need less material, yet high homogeneity and solubility are required. Thus, our ultimate quality criterion was the percentage of active protein rather than the overall purification yield. We designed a single-step batch purification process maximizing the proportion of monomeric and active MBP-E6 protein. This approach is compatible with parallel purifications and generates sufficient amounts of protein for our biophysical assays. In the present study, we investigated the purification of two HPV E6 oncoproteins fused to MBP: the solubility-enhanced mutant HPV16 E6 F47R 4C/4S (16E6mut) whose structure was solved by X-ray crystallography, and the HPV8 E6 (8E6) which, in our experience, was particularly recalcitrant to purify as a monomer.

This study aimed at the preferential selection of monomeric protein species within a single purification step. Oligomers can be distinguished from monomers based on their size, more precisely their volume in solution. The radius of a particle in solution is correlated with the diffusion coefficient  $D$  according to the Stokes–Einstein formula [33]:

$$D = kT/6\pi\eta r$$

with  $k$  the Boltzmann constant,  $T$  the temperature,  $\eta$  the viscosity coefficient of the medium and  $r$  the radius. Thus, larger protein species diffuse slower than monomers in solution. Based on these size and diffusion properties, we optimized three parameters in the purification protocol: (i) reticulation of affinity resin, (ii) contact time between the bacterial lysate and the resin and (iii) elution by “decant and take up” method. Adapting the resin reticulation is a way to select particles based on their size in solution that is faster and easier to parallelize than SEC. The two other optimization lines rely on the distinct diffusion properties of oligomers and monomers. First, a reduced contact time limits the binding of oligomeric species on the resin. Second, a fast elution method by pipetting only the supernatant fraction enables the preferential recovery of monomeric species.

To discriminate particles based on their volume in solution, reticulated resins can be selected to set a specific exclusion limit by adapting the pore size and the agarose concentration. Protein species exceeding a threshold hydrodynamic volume in solution will not be able to enter the affinity resin pores. This property allows a double selection based on the protein size and the affinity tag, within a single step.

We compared the size of particles retained on different types of resin, amylose resin and nickel resins with

distinct reticulation levels. In particular, resin N2 was a 10% cross-linked agarose resin excluding proteins above 500 kDa (10% BCL Agarose Bead Standard, ABT) that was functionalized to bind Nickel, whereas all other nickel-chelating resins have an exclusion limit of about 4000 kDa. Our results (notably with MBP-16E6mut) showed that amylose resin leads to the purification of a larger fraction of oligomers than nickel resins. N2 resin was the most efficient resin to purify the monomeric form of both MBP-16E6mut and MBP-8E6, which confirms the efficiency of adapting affinity resin pore size to limit the purification of aggregates. N1 resin was the second most efficient resin to maximize the ratio of monomeric protein. In addition, the highest total protein amount (including both monomer and oligomers) was purified with N1 resin.

The proportion of purified unwanted oligomers can be minimized by reducing incubation times at key steps of the process. First, the contact time between the bacterial crude extract and the affinity resin can be reduced in order to minimize the binding of oligomers, as demonstrated with the recalcitrant MBP-8E6 construct for which the highest monomer fraction was obtained for the minimum contact time. Second, the elution method can be designed to recover a predominantly monomeric species capable to rapidly diffuse to the supernatant. This approach increased by two-fold the ratio of purified monomeric MBP-8E6, despite its natural tendency to form a majority of oligomeric species.

Finally, we were able to confirm by SPR the binding affinity between *E6APwt* LXXLL peptide and MBP-16E6mut purified according to our optimized protocol and using the two nickel reticulated resins that allowed to purify highest proportion of monomeric protein (N1 and N2). The protein did not interact with the mutated motif *E6APmut* (AXLL), which rules out the possibility of unspecific interactions that might happen with misfolded proteins and/or aggregated proteins.

## Conclusions

In the present work, we designed a single-step purification strategy optimized for maximal recovery of monomeric MBP-E6 that we accurately quantified by means of analytical SEC. By using a customized reticulated resin, adapting the contact time and performing fast elution, we were able to reach high monomer ratios in the final purified fraction. This method was particularly efficient for the solubility-enhanced mutant MBP-16E6mut and the highly aggregation-prone MBP-8E6. As assessed by SPR, the purified MBP-16E6mut protein interacts with its known partner within the expected range of affinity. This method is promising for forthcoming biophysical studies with recalcitrant proteins, while allowing to perform

parallel purifications. The approach should also be easily adaptable to robotization. It would be interesting to automatize this protocol for purifying at high throughput families of proteins, which, like E6 oncoproteins, are difficult to fold except when fused to MBP or other soluble carrier proteins. The overexpression of such constructs leads to a limited proportion of properly folded and active monomers that require to be separated from non-active oligomers, which is now possible using the protocol presented here.

## Methods

### Customized Nickel resin N2

#### Preparation

High Density Nickel 10BCL has been manufactured as a customized product by Agarose Bead Technologies. The resin consists of crosslinked 10% agarose beads to which a Nickel chelating group has been immobilized. This chelating group has been obtained by a modification of the procedures described elsewhere [34, 35] using 10% BCL Agarose Bead Standard as raw material. The epoxide generated has been treated with iminodiacetic acid and charged with  $\text{Ni}^{2+}$  ions according to a modification of the method previously described.

#### Characterization

The characteristics of 10% BCL Agarose Bead Standard for protein separation based on their molecular weight are the following: the fractionation range for globular proteins is  $1 \times 10^4$ – $5 \times 10^5$  Da and the exclusion limit is  $>5 \times 10^5$  Da. Nickel was quantified by spectrophotometric assay, resulting in  $34 \mu\text{mol Ni}^{2+}/\text{mL}$  gel.

#### Expression and fast batch purification procedure

The sequences encoding for 16E6mut (already described in [22]) and for 8E6 were cloned into a pETM-41 vector by using *NcoI* and *Acc65I* sites, allowing the production of MBP-6His-TEV site-E6 fusions (Fig. 1). The sequence encoding for 8E6 was kindly provided by Yves Jacob from Pasteur Institute (Paris, France) and corresponds to a variant with three substitutions as compared to the reference protein sequence (Uniprot ID: P06428), namely Y9N, A26E and S36L.

Both constructs were expressed in *Escherichia coli* BL21 (DE3) cells, grown in LB medium supplemented with kanamycine  $50 \mu\text{g}/\text{mL}$  at  $37^\circ\text{C}$  until  $\text{OD}_{600} \approx 0.7$ . Expression was induced by adding  $0.5 \text{ mM}$  IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) and  $100 \mu\text{M}$   $\text{ZnSO}_4$ . Cells were grown overnight at  $16^\circ\text{C}$  and harvested by centrifugation. Pellets were stored at  $-20^\circ\text{C}$ .

Purification buffers were thoroughly degassed and equilibrated with argon before adding  $2 \text{ mM}$  DTT. Culture pellets were then resuspended in lysis buffer, which

was Buffer A (Tris 50 mM pH 8; NaCl 400 mM; DTT 2 mM) supplemented with 5% (w/v) glycerol, 0.25 µg/mL RNase I, 0.25 µg/mL DNase I, lysozyme at approximately 1 µg/mL and anti-protease cocktail EDTA-free (Roche) according to the manufacturer instruction. This buffer was supplemented with 10 mM imidazole for purification on Nickel resins. Pellet equivalent to 250 mL culture volume was resuspended in 10 mL lysis buffer. Cells were lysed by sonication on ice and then centrifuged at 100,000×g at 4 °C for 45 min. The supernatant was then incubated at 4 °C with either nickel (100 µL resin for 250 mL culture pellet) or amylose resins (300 µL resin for 250 mL culture pellet), equilibrated in Buffer A supplemented with 20% of the recommended concentration of anti-protease cocktail. The incubation time of the cleared lysate with the resin was 2 h in the initial purification tests, before to be evaluated as a variable parameter (5 min, 30 min or 2 h) and reduced to 5 min in the final optimized protocol. The resins were pelleted by 5 min-centrifugation at 500×g at 4 °C; the supernatant was discarded and the resins were first washed in Buffer A (30 min incubation at 4 °C), then in Buffer B (Tris 50 mM pH 8, NaCl 1 M, DTT 2 mM, 20% of the recommended concentration of anti-protease cocktail and supplemented with 10 mM imidazole for Nickel resins) with same incubation time and temperature. Resins were re-equilibrated in Buffer A during 10 min at 4 °C before elution. Resins were then transferred in Eppendorf tubes, pelleted by centrifugation for 5 min at 500×g, 4 °C and supernatants were discarded. Finally, resins were resuspended in elution buffer (Buffer A supplemented with 20% anti-protease cocktail and, for Nickel resins: imidazole 600 mM; for Amylose resin: maltose 15 mM). 100 µL Nickel resin were eluted with 140 µL imidazole elution buffer while 300 µL amylose resin were eluted with 200 µL maltose elution buffer.

For standard elution, resins were transferred in 96-well filter plate and centrifuged 5 min at 500×g and 4 °C for maximal recovery of liquid phase by filtration. For the optimized elution protocol, the tube content was then rapidly mixed by gentle vortex and immediately centrifuged during 5 min at 500×g and 4 °C. The supernatant was recovered and used as final purified protein solution. Protein concentration was determined by Nanodrop measurement, based on absorbance at 280 nm. Final protein samples were systematically analyzed either by SDS-PAGE or by microfluidics capillary gel electrophoresis in order to check the presence of protein contaminants in denaturing conditions. Prior to SPR assays, protein samples were desalted on Illustra NAP-5 columns (GE Healthcare) equilibrated in Buffer A allowing to eliminate imidazole or maltose from the solution.

#### Microfluidics capillary gel electrophoresis

Purification samples (from the washing steps and elution) were transferred in 96-well plates, mixed with the Caliper sample buffer and boiled according to the manufacturer instructions. The plates were loaded and measured on LabChip GX II device (Caliper, Perkin Elmer), with HT Protein Express 100 High Sensitivity protocol (10–100 kDa). The kit allows to label the proteins with a fluorescent dye. This dye is excited by a laser during protein capillary electrophoresis in denaturing conditions. The emitted fluorescence signal is plotted versus migration time, leading to the protein separation according to their molecular weight. Data processing is performed with LabChip GX II software. Migration time is first converted into molecular weight using standard protein markers before a quantitative analysis is performed thanks to empirical linear correlation between fluorescence intensity and protein amount.

#### Analytical size-exclusion chromatography

Samples used for analytical SEC were centrifuged 5 min at 13,700×g, 4 °C and only the supernatant was loaded on the column Superdex 200 Increase 5/150 GL (GE Healthcare) previously equilibrated in Buffer A. SEC run was performed on ÄKTA purifier (GE Healthcare), by injecting 50 µL per sample, with flow set at 0.150 mL/min. Spikes of absorbance 280 nm signal due to air bubbles were excluded from the chromatogram before performing data analysis. The original chromatograms are provided in Additional file 2.

We calibrated the column with Gel Filtration Calibration Kit High Molecular Weight (GE Healthcare), using a solution of dextran Blue (0.1 mg/mL) and two mixes. Mix 1 was composed of ovoalbumine (4 mg/mL), aldolase (4 mg/mL) and thyroglobuline (5 mg/mL). Mix 2 contained ferritine (0.3 mg/mL) and conalbumine (3 mg/mL). Each of the three calibration solution was loaded in a 50 µL loop and injected separately at a flow rate of 0.150 mL/min on the column previously equilibrated in Buffer A. The resulting elution volumes were used to establish a calibration plot to determine the molecular weight  $MW$ , based on the following equation [33]:

$$K_{av} = \frac{V_e - V_o}{V_t - V_o}$$

where  $K_{av}$  stands for the available distribution coefficient,  $V_e$  corresponds to the elution volume of each injected protein,  $V_o$  is the void volume of the column (determined by the elution volume of the dextran blue) and  $V_t$  is the total bed volume (3 mL with this specific column). Thus, the coefficients  $a$  and  $b$  were determined by linear regression using the following equation [33]:

$$K_{av} = a \times \log(MW) + b \quad (1)$$

Moreover, the hydrodynamic radius  $R_h$  was determined according to:

$$\sqrt{-\log_{10}(K_{av})} = c \times R_h + d \quad (2)$$

where  $c$  and  $d$  coefficients are determined by linear regression [33].

The area of elution peaks (based on absorbance at 280 nm) was calculated using the evaluation module of Unicorn 5.31 software.

According to the column calibration, we inferred the elution volume of monomeric (approximately 1.92 mL) and oligomeric MBP-E6 ( $V_e < 1.90$  mL). For each assay, the area of the elution peak corresponding to the monomer was divided by the area of peaks corresponding to all eluted protein species (monomeric and oligomeric), allowing us to measure the ratio of monomer and oligomer in the purified protein sample.

Regarding the ratios of monomeric protein, we observed that the duplication of the data leads to standard deviations never exceeding 15% unit. This value has consequently been considered as the uncertainty of monomer/oligomer ratios in this work. This uncertainty, although given as a percentage, corresponds to the absolute variation of the ratio of monomeric protein and not as a relative variation coefficient.

### Peptide synthesis

To test the ability of the purified protein to interact with LXXLL peptides, we used two synthetic peptides: *E6APwt* (PESELTLQELLGEER) and *E6APmut* (PESELTAQELLGEER). Both peptides were N-terminally biotinylated with a TTDS spacer (*N*-(13-amino-4,7,10-trioxa-tridecyl)-succinamic acid) (Iris Biotech GmbH) and synthesized either by JPT Innovative Peptide Solutions or by the peptide synthesis service at IGBMC with 70–80% purity. The lyophilized peptides were resuspended in water at a final stock concentration of 5 mM and stored at  $-20$  °C.

### Surface plasmon resonance

Peptide–protein interaction assays were performed by surface plasmon resonance (SPR) on a Biacore T200 instrument (GE Healthcare—Biacore) at 25 °C. The running buffer was Tris 50 mM pH 8, NaCl 400 mM, DTT 2 mM and 0.005% (v/v) surfactant polysorbate 20 (GE Healthcare). We used Biotin CAPture kit (GE Healthcare—Biacore) for the reversible capture of biotinylated peptides on a chip. Briefly, the chip is coated with a deoxyribooligonucleotide that hybridizes with the

complementary strand bound to streptavidin. The biotinylated peptide binds to streptavidin and can be washed by dehybridizing the two oligonucleotides. An empty control surface was systematically included on every cycle to serve as a reference for non-specific binding of the analyte to the matrix and for monitoring changes in solution refractive index. This reference surface was treated as the peptide surfaces except that peptide injection was omitted. Kinetic runs were performed by injecting series of two-fold cascade dilutions of the analyte MBP-E6 samples, starting from 10  $\mu$ M.

Thus, at the beginning of each binding cycle, CAPture reagent (diluted 5 times in running buffer) was injected on all channels during 300 s at 2  $\mu$ L/min. For peptide surfaces, peptide solution (50 nM) was injected during 15–30 s at 10  $\mu$ L/min in order to reach 4–15 RU immobilization level. Protein sample was then injected during 60 s at 30  $\mu$ L/min, followed by 120 s buffer flow. The surfaces were then regenerated by injecting regeneration solution as indicated by the manufacturer (guanidine hydrochloride 6 M; sodium hydroxide 250 mM) for 60 s at 5  $\mu$ L/min.

The SPR signals from the regions corresponding to the protein injection and post-injection phases were plotted as RU versus time. Data were first processed using the Biacore T200 Evaluation 1.0 software (GE Healthcare, Biacore Life Science, Uppsala, Sweden). Sensorgrams obtained for the different protein concentrations were corrected for buffer effects and bulk refractive index changes by subtracting the empty cell signal, and subsequently normalized according to the peptide levels which can slightly differ from cycle to cycle due to the immobilization process. The steady-state binding signal was derived by averaging the signals at equilibrium within a five second-window ( $R_{eq}$ ). Steady-state analysis was performed using an in-house Python script by fitting the average and normalized signal  $R_{eq}$  as a function of total analyte concentration, and assuming a simple 1:1 interaction binding isotherm model. Uncertainties of the derived parameters were estimated using Monte-Carlo approach, considering an experimental uncertainty of 5 RU. This value was estimated by duplicating the full experimental cycle corresponding to a single protein concentration.

### Additional files

**Additional file 1.** Corrected size-exclusion chromatograms. Prior to protein peak integration, spikes due to air bubbles were excluded. **A:** Chromatograms for analytical SEC of MBP-16E6mut. In order to control to oligomeric state of the purified protein, we performed systematic analytical SEC on the final protein sample. The calibration of the column allowed

us to estimate the size of the particles at different elution volumes, thus the elution peak of monomeric protein is indicated by an arrow. Note the increase of the monomer fraction compared to the total amount of purified protein when shifting from the filter elution to the “decant and take up” elution method. **B:** Chromatograms for analytical SEC of MBP-8E6. On this protein challenging to purify, the increase of the monomer fraction (indicated by an arrow) between the two elution methods is dramatic. **C:** Chromatograms for analytical SEC of MBP-8E6: optimization of bacterial extract contact time. The elution peak of monomeric protein is indicated by an arrow. The decrease of the oligomeric fraction (eluted between the void volume  $V_0$  and the monomer fraction) is particularly visible for nickel resin N1.

**Additional file 2.** Raw size-exclusion chromatograms. We provide in this file the raw data of size-exclusion chromatograms used in the study (absorbance at 280 nm versus elution volume). The protein (MBP-16E6mut, MBP-8E6), resin (N1...N6, A), elution method (filter, decant) and contact time for optimization assays (2 h, 30 min, 5 min) are indicated in the worksheet names. The first worksheet entitled Readme contains explanations on the nomenclature used for worksheet titles.

### Abbreviations

DTT: dithiothreitol; HPV: human papillomavirus; IMAC: immobilized-metal affinity chromatography; IPTG: isopropyl  $\beta$ -D-1-thiogalactopyranoside; MBP: maltose-binding protein; SEC: size-exclusion chromatography; SPR: surface plasmon resonance.

### Authors' contributions

GT and AB designed the experiments; AD, AB and CK performed the experiments, SPR; YN and AB analyzed data and generated figures (for SPR and analytical SEC respectively); PA, LSJ and EDC designed and produced custom affinity resins; AB; YN and GT wrote the manuscript. All authors read and approved the final manuscript.

### Author details

<sup>1</sup> Équipe Labellisée Ligue 2015, Department of Integrated Structural Biology, Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), INSERM U1258/CNRS UMR 7104/Université de Strasbourg, 1 rue Laurent Fries, BP 10142, 67404 Illkirch, France. <sup>2</sup> ABT-Agarose Bead Technologies, C/La Forja, 9, Torrejón de Ardoz, 28850 Madrid, Spain.

### Acknowledgements

The authors thank D. Altschuh for the optimization of Biacore assays, P. Eberling for peptide synthesis, A. Vanden Broeck for his valuable help on size-exclusion chromatography columns, P. Klein for careful reading and colleagues of the Travé team for helpful discussion and support.

### Competing interests

L.S.J., E.D.C. and P.A. declare competing financial interests since ABT (Agarose Bead Technologies), the company they belong to, provided affinity beads. Other authors declare that they have no competing interests.

### Availability of data and materials

The datasets analyzed during the current study are included within the article and additional files.

### Consent for publication

Not applicable.

### Ethics approval and consent to participate

Not applicable.

### Funding

This work received institutional support from le Centre National de la Recherche Scientifique (CNRS), Université de Strasbourg, Institut National de la Santé et de la Recherche Médicale (INSERM) and Région Alsace. The work was supported in part by grants from Ligue contre le Cancer (équipe labellisée 2015 and fellowship to A.B.), National Institutes of Health (Grant R01CA134737), Instruct (ESFRI), and the French Infrastructure for Integrated Structural Biology (FRISBI). The authors declare that the content is solely

their responsibility and does not represent the official views of the National Institutes of Health.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 10 October 2018 Accepted: 23 November 2018

Published online: 01 December 2018

### References

1. Waugh DS. The remarkable solubility-enhancing power of *Escherichia coli* maltose-binding protein. *Postepy Biochem*. 2016;62:377–82.
2. Berrow NS, Büsow K, Coutard B, Diprose J, Ekberg M, Folkers GE, et al. Recombinant protein expression and solubility screening in *Escherichia coli*: a comparative study. *Acta Crystallogr D Biol Crystallogr*. 2006;62(10):1218–26.
3. Riggs P. Expression and purification of recombinant proteins by fusion to maltose-binding protein. *Mol Biotechnol*. 2000;15(1):51–63.
4. Riggs P. Expression and purification of maltose-binding protein fusions. In: Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, et al., editors. *Current protocols in molecular biology*. Hoboken: John Wiley & Sons, Inc.; 2001 (**Chapter 16:Unit 16.6**).
5. Sachdev D, Chirgwin JM. [20] Fusions to maltose-binding protein: control of folding and solubility in protein purification. In: Thorner J, Emr SD, Abelson JN, editors. *Applications of chimeric genes and hybrid proteins part A: gene expression and protein purification*. Cambridge: Academic Press; 2000. p. 312–21 (**Methods Enzymol**; vol. 326).
6. Sachdev D, Chirgwin JM. Properties of soluble fusions between mammalian aspartic proteinases and bacterial maltose-binding protein. *J Protein Chem*. 1999;18(1):127–36.
7. Sachdev D, Chirgwin JM. Order of fusions between bacterial and mammalian proteins can determine solubility in *Escherichia coli*. *Biochem Biophys Res Commun*. 1998;244(3):933–7.
8. Sachdev D, Chirgwin JM. Solubility of proteins isolated from inclusion bodies is enhanced by fusion to maltose-binding protein or thioredoxin. *Protein Expr Purif*. 1998;12(1):122–32.
9. Nominé Y, Ristriani T, Laurent C, Lefèvre J-F, Weiss É, Travé G. A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein. *Protein Eng*. 2001;14(4):297–305.
10. Nominé Y, Ristriani T, Laurent C, Lefèvre J-F, Weiss É, Travé G. Formation of soluble inclusion bodies by HPV E6 oncoprotein fused to maltose-binding protein. *Protein Expr Purif*. 2001;23(1):22–32.
11. Zanier K, Nominé Y, Charbonnier S, Ruhlmann C, Schultz P, Schweizer J, et al. Formation of well-defined soluble aggregates upon fusion to MBP is a generic property of E6 proteins from various human papillomavirus species. *Protein Expr Purif*. 2007;51(1):59–70.
12. Zanier K, Ruhlmann C, Melin F, Masson M, Ould M'hamed Ould Sidi AO, Bernard X, et al. E6 proteins from diverse papillomaviruses self-associate both in vitro and in vivo. *J Mol Biol*. 2010;396(1):90–104.
13. Raran-Kurussi S, Waugh DS. The ability to enhance the solubility of its fusion partners is an intrinsic property of maltose-binding protein but their folding is either spontaneous or chaperone-mediated. *PLoS ONE*. 2012;7(11):e49589.
14. de Marco A. Recombinant polypeptide production in *E. coli*: towards a rational approach to improve the yields of functional proteins. *Microb Cell Fact*. 2013;12(1):101.
15. de Marco A. Optimization of purification protocols based on the step-by-step monitoring of the protein aggregates in soluble fractions. In: Lorence A, editor. *Recombinant gene expression*. Totowa: Humana Press; 2012. p. 145–54 (**Methods Mol Biol**; vol 824).
16. Suarez I, Trave G. Structural insights in multifunctional papillomavirus oncoproteins. *Viruses*. 2018;10(1):37.
17. Wallace NA, Galloway DA. Novel functions of the human papillomavirus E6 oncoproteins. *Annu Rev Virol*. 2015;2(1):403–23.

18. Vande Pol SB, Klingelutz AJ. Papillomavirus E6 oncoproteins. *Virology*. 2013;445(1–2):115–37.
19. Ould M'hamed Ould Sidi A, Ould Babah K, Brimer N, Nominé Y, Romier C, Kieffer B, et al. Strategies for bacterial expression of protein–peptide complexes: application to solubilization of papillomavirus E6. *Protein Expr Purif*. 2011;80(1):8–16.
20. White EA, Howley PM. Proteomic approaches to the study of papillomavirus–host interactions. *Virology*. 2013;435(1):57–69.
21. Scheffner M, Huibregtse JM, Vierstra RD, Howley PM. The HPV-16 E6 and E6-AP complex functions as a ubiquitin–protein ligase in the ubiquitination of p53. *Cell*. 1993;75(3):495–505.
22. Zanier K, Charbonnier S, Sidi AOMO, McEwen AG, Ferrario MG, Poussin-Courmontagne P, et al. Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science*. 2013;339(6120):694–8.
23. Martinez-Zapien D, Ruiz FX, Poirson J, Mitschler A, Ramirez J, Forster A, et al. Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. *Nature*. 2016;529(7587):541–5.
24. Tommasino M. The human papillomavirus family and its role in carcinogenesis. *Semin Cancer Biol*. 2014;26:13–21.
25. Bzhalava D, Eklund C, Dillner J. International standardization and classification of human papillomavirus types. *Virology*. 2015;476:341–4.
26. Van Doorslaer K, Tan Q, Xirasagar S, Bandaru S, Gopalan V, Mohamoud Y, et al. The Papillomavirus Episteme: a central resource for papillomavirus sequence data and analysis. *Nucleic Acids Res*. 2012;41(D1):D571–8.
27. de Sanjose S, Quint WG, Alemany L, Geraets DT, Klaustermeier JE, Lloveras B, et al. Human papillomavirus genotype attribution in invasive cervical cancer: a retrospective cross-sectional worldwide study. *Lancet Oncol*. 2010;11(11):1048–56.
28. Gillison ML, Alemany L, Snijders PJF, Chaturvedi A, Steinberg BM, Schwartz S, et al. Human papillomavirus and diseases of the upper airway: head and neck cancer and respiratory papillomatosis. *Vaccine*. 2012;30:F34–54.
29. Tommasino M. The biology of beta human papillomaviruses. *Virus Res*. 2017;231:128–38.
30. Zanier K, Charbonnier S, Baltzinger M, Nominé Y, Altschuh D, Travé G. Kinetic analysis of the interactions of human papillomavirus E6 oncoproteins with the ubiquitin ligase E6AP using surface plasmon resonance. *J Mol Biol*. 2005;349(2):401–12.
31. Zanier K, Ould M'hamed Ould Sidi A, Boulade-Ladame C, Rybin V, Chappelle A, Atkinson A, et al. Solution structure analysis of the HPV16 E6 oncoprotein reveals a self-association mechanism required for E6-mediated degradation of p53. *Structure*. 2012;20(4):604–17.
32. Brimer N, Drews CM, Vande Pol SB. Association of papillomavirus E6 proteins with either MAML1 or E6AP clusters E6 proteins by structure, function, and evolutionary relatedness. *PLoS Pathog*. 2017;13(12):e1006781.
33. Tayyab S, Qamar S, Islam M. Size exclusion chromatography and size exclusion HPLC of proteins. *Biochem Educ*. 1991;19(3):149–52.
34. Hubert P, Porath J. Metal chelate affinity chromatography: I. Influence of various parameters on the retention of nucleotides and related compounds. *J Chromatogr A*. 1980;198(3):247–55.
35. Sundberg L, Porath J. Preparation of adsorbents for biospecific affinity chromatography: I. Attachment of group-containing ligands to insoluble polymers by means of bifunctional oxiranes. *J Chromatogr A*. 1974;90(1):87–98.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

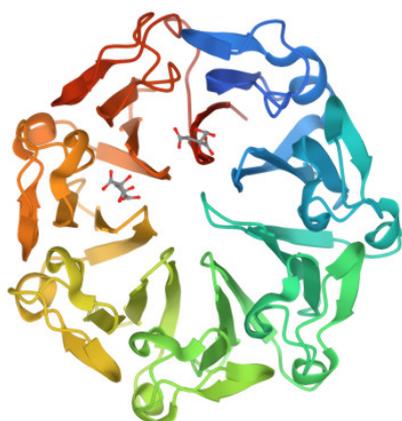


## **2 | Annexe 2**

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2

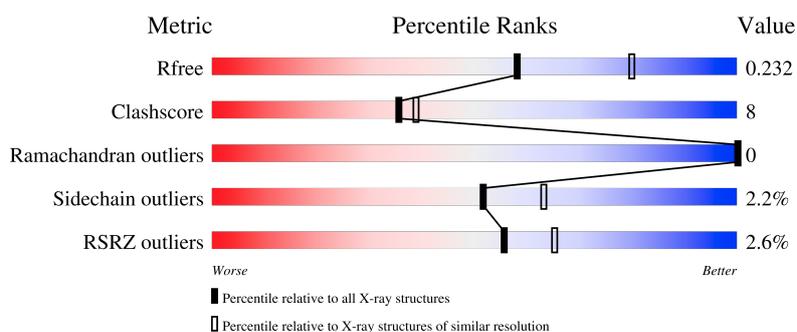


7Q40



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 2.35 Å  
 R-Value Free: 0.230  
 R-Value Work: 0.179  
 R-Value Observed: 0.181

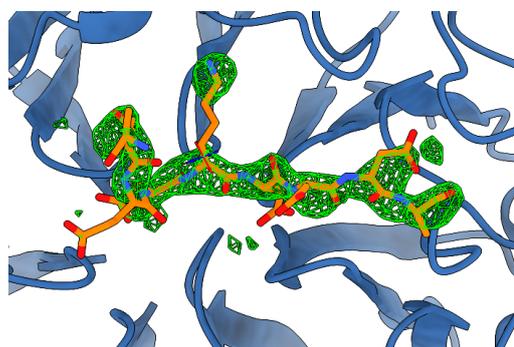
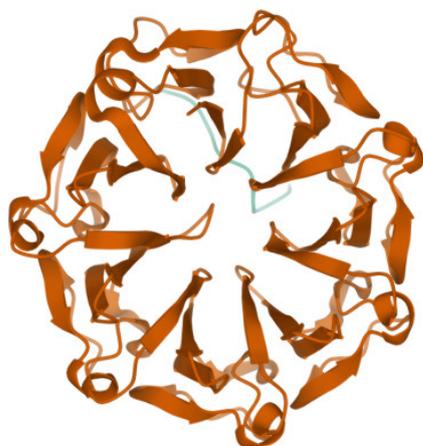


ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	2.35	54.27	98.8	0.998	13.21	26.6657389414		60722			48.3943368623

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2 in complex with DXDKDXD motif of ubiquitin-protein ligase E3A (E6AP)

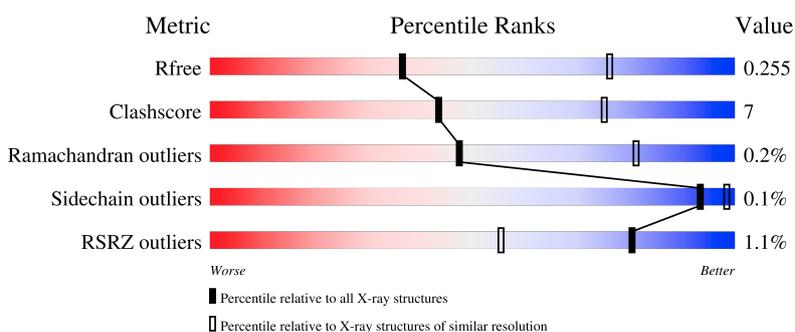


7Q41



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 3.01 Å  
 R-Value Free: 0.255  
 R-Value Work: 0.198  
 R-Value Observed: 0.201

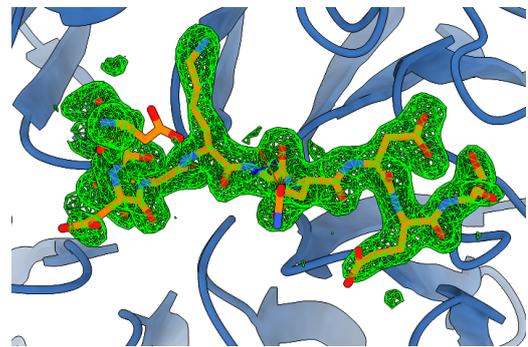


ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	3.01	48.725	99.2	0.968	5.81	5.91558595531		54825			58.2859737188

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2 in complex with DXDKDXD motif of chromatin reader BAZ2B

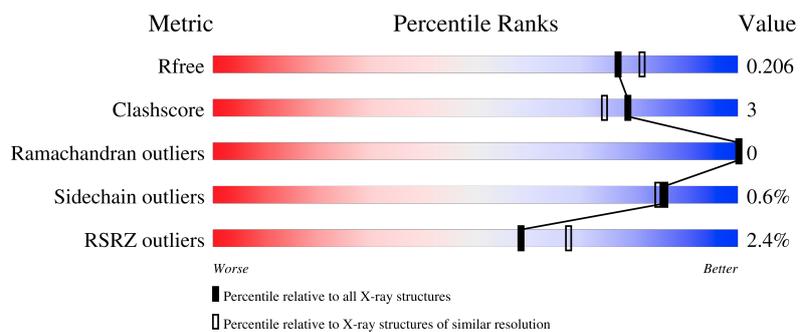


7Q42



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 1.95 Å  
 R-Value Free: 0.206  
 R-Value Work: 0.168  
 R-Value Observed: 0.170

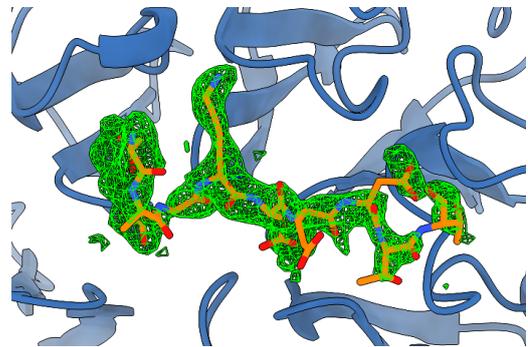
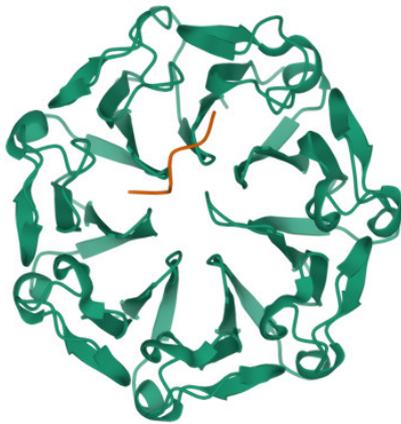


ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	1.95	49.29	99.92	0.998	10.93	27.0621601735		104649			29.4271306625

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2 in complex with DXDKDXD motif of dedicator of cytokinesis protein 10

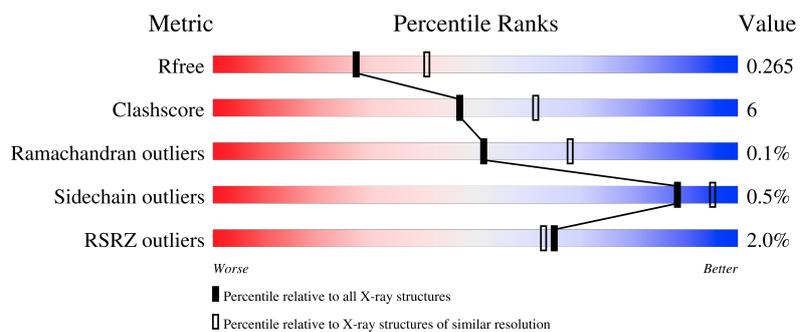


7Q43



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 2.40 Å  
 R-Value Free: 0.265  
 R-Value Work: 0.202  
 R-Value Observed: 0.206

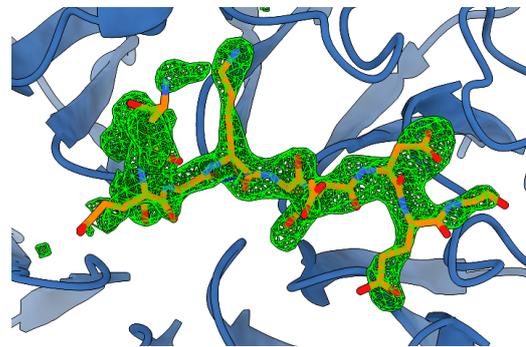


ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	2.4	49.665	99.7	0.983	5.77	37.6713780309		57945			33.6856059616

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2 in complex with DXDKDXD motif of deubiquitinase USP35

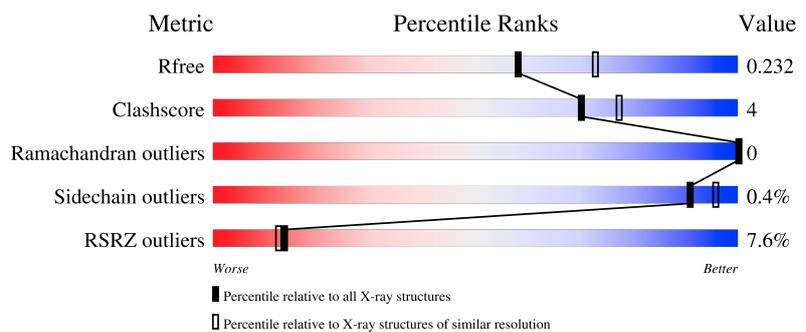


7Q44



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 2.20 Å  
 R-Value Free: 0.231  
 R-Value Work: 0.182  
 R-Value Observed: 0.185

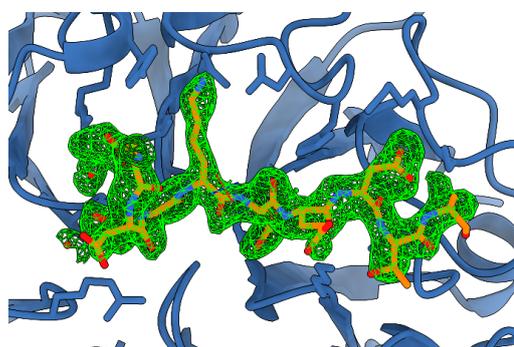
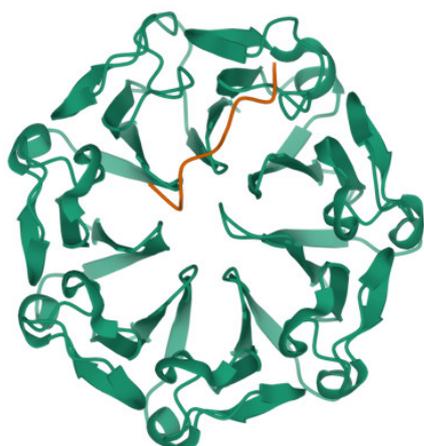


ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	2.2	49.516	99.9	0.994	8.65	26.7427710681		74354			32.0006866889

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2 in complex with DXDKDED motif of Myelin transcription factor 1

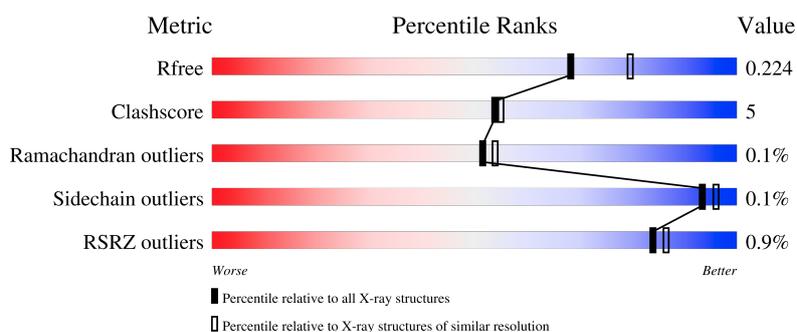


7Q45



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 2.10 Å  
 R-Value Free: 0.224  
 R-Value Work: 0.184  
 R-Value Observed: 0.186

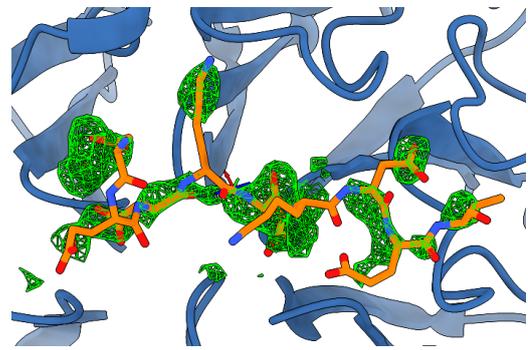
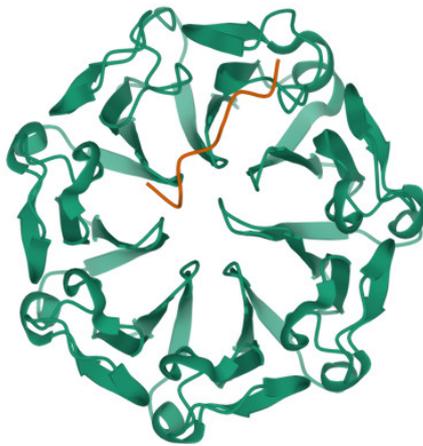


ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	2.1	48.473	99.2	0.995	9.67	26.794735657		84379			30.3915849439

# Crystal structure of RCC1-Like domain 2 of ubiquitin ligase HERC2 in complex with DXDKDED motif of pericentriolar material 1 protein

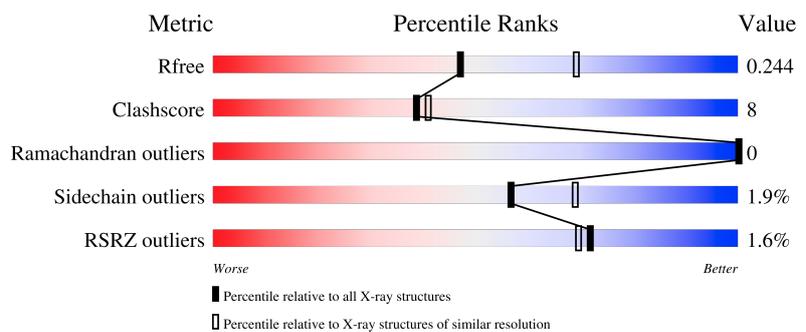


7Q46



## Experimental Data Snapshot

Method: X-RAY DIFFRACTION  
 Resolution: 2.46 Å  
 R-Value Free: 0.245  
 R-Value Work: 0.179  
 R-Value Observed: 0.182



ID #	Resolution (High)	Resolution (Low)	Percent Possible (Observed)	CC (Half)	Net I Over Average Sigma (I)	Redundancy	Number Reflections (All)	Number Reflections (Observed)	Observed Criterion Sigma (F)	Observed Criterion Sigma (I)	B (Isotropic) From Wilson Plot
1	2.46	49.156	100	0.984	6.33	26.5269079675		52438			32.7291669667

## **3 | Annexe 3**

# Interaction between UBE3A and HERC2 : Two E3 ubiquitin ligases involved in Neurodevelopmental Syndromes

Demenge A., Howard E., Kostmann C., Cousido-Siah A. Mitschler A. Bonhoure A., and Travé G.



## Angelman Syndrome (AS) :

Mental retardation, speechlessness, ataxia, hyperactivity.

Genetic abnormalities involving the region 15q11-13 (Region containing UBE3A and HERC2) regulated by an imprinting center (PWS/AS).

Lack of UBE3A → Angelman syndrome

Lack of HERC2 → Angelman like syndrome

## HERC2 and UBE3A:

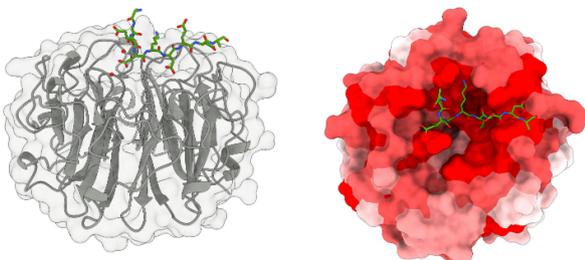


HERC2 protein binds to UBE3A

The RLD2 domain of HERC2 is necessary and sufficient for the interaction

HERC2 increases ubiquitin ligase activity of UBE3A

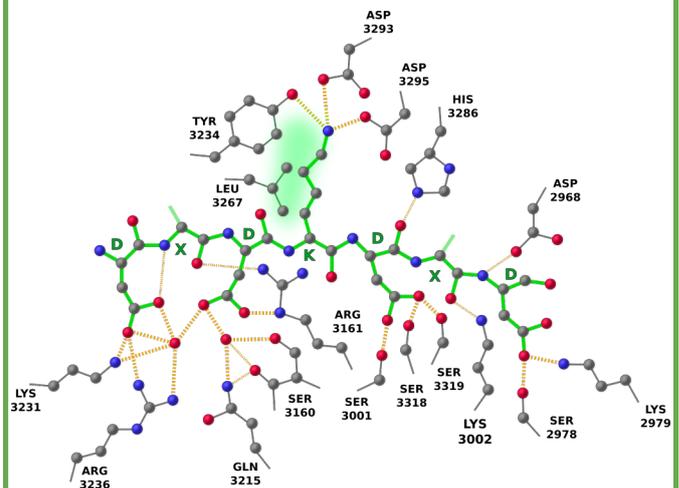
## Structure of the interaction:



Crystal structure of RLD2 domain of HERC2 in complex with UBE3A peptide.

7 bladed beta propeller, **high conserved spot** on the UBE3A binding site in different species.

## Identification of the motif:



Identification of motif **D.DKD.D** by analyzing the structure

Aspartic acids take place in a positively charged groove.

The lysine of the motif is embedded in a negatively charged pocket.

Identification of human protein containing this motif using SLiMSearch software.

## Identification of interaction network:

Protein	Motif	Kd in $\mu\text{M}$
UBE3A	LQAK <b>DEDKDE</b> DEKEK	4.59
DOCK10	IDHE <b>DADKDE</b> TTSH	5.29
PCM1	GEID <b>DEDKDK</b> DETET	8.22
USP35	GRGF <b>DEDKDE</b> EGSP	9.73
BAZ2B	EEED <b>DDDKDQ</b> DES	12.18
ARID4A	KKQE <b>DSDKD</b> SDEEEE	21.07
ARIP4	SSND <b>DEDKDD</b> DVIEV	26.94
MYT1	EVR <b>SDDDKDE</b> DTHSR	31.88
RERE	MTAD <b>KDKDK</b> DKEKD	32.21

Measurement of affinity of different peptides for RLD2 by fluorescence polarization.

DOCK10, MYT1 are involved in neurodevelopmental mechanism.

ARID4A, BAZ2B, ARIP4, and MYT1 are involved in chromatin remodeling.

ARID4A Plays a role in the regulation of epigenetic modifications at the PWS/AS imprinting center.

## Conclusion:

We obtained the structure of RLD2 domain apoform and in complex with different peptides.

Identification of a network of protein interacting with RLD2 domain, including proteins involved in neurodevelopmental mechanisms.

## Perspective:

Understanding the impact of these different interactions and their alteration



Sixième partie

Bibliographie



- 
1. STEINBERG, D. & VAUGHAN, M. Observations on intracellular protein catabolism studied in vitro. en. *Archives of Biochemistry and Biophysics* **65**, 93-105. ISSN : 00039861. <https://linkinghub.elsevier.com/retrieve/pii/0003986156901801> (2022) (nov. 1956).
  2. SIMPSON, M. V. THE RELEASE OF LABELED AMINO ACIDS FROM THE PROTEINS OF RAT LIVER SLICES. en. *Journal of Biological Chemistry* **201**, 143-154. ISSN : 00219258. <https://linkinghub.elsevier.com/retrieve/pii/S0021925818713562> (2022) (mars 1953).
  3. CIECHANOVER, A., HOD, Y. & RERSHKOL, A. A HEAT-STABLE POLYPEPTIDE COMPONENT OF AN ATP-DEPENDENT PROTEOLYTIC SYSTEM FROM RETICULOCYTES. en. *BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS* **81** (1978).
  4. CIECHANOVER, A., HELLER, H., ELIAS, S., HAAS, A. L. & HERSHKO, A. ATP-dependent conjugation of reticulocyte proteins with the polypeptide required for protein degradation. en. *Proceedings of the National Academy of Sciences* **77**, 1365-1368. ISSN : 0027-8424, 1091-6490. <https://pnas.org/doi/full/10.1073/pnas.77.3.1365> (2022) (mars 1980).
  5. HERSHKO, A., CIECHANOVER, A., HELLER, H., HAAS, A. L. & ROSE, I. A. Proposed role of ATP in protein breakdown : conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis. en. *Proceedings of the National Academy of Sciences* **77**, 1783-1786. ISSN : 0027-8424, 1091-6490. <https://pnas.org/doi/full/10.1073/pnas.77.4.1783> (2022) (avr. 1980).
  6. WILKINSON, K. D., URBAN, M. K. & HAAS, A. L. Ubiquitin is the ATP-dependent proteolysis factor I of rabbit reticulocytes. *Journal of Biological Chemistry* **255**. Publisher : ASBMB, 7529-7532 (1980).
  7. SCHLESINGER, D. H., GOLDSTEIN, G., SCHEID, M. P. & BOYSE, E. A. Chemical synthesis of a peptide fragment of thymopoietin II that induces selective T cell differentiation. *Cell* **5**, 367-370. ISSN : 0092-8674. <https://www.sciencedirect.com/science/article/pii/0092867475900550> (1975).
  8. CAPPADOCIA, L. & LIMA, C. D. Ubiquitin-like Protein Conjugation : Structures, Chemistry, and Mechanism. en. *Chemical Reviews* **118**, 889-918. ISSN : 0009-2665, 1520-6890. <https://pubs.acs.org/doi/10.1021/acs.chemrev.6b00737> (2022) (fév. 2018).
  9. ÖZKAYNAK, E., FINLEY, D. & VARSHAVSKY, A. The yeast ubiquitin gene : head-to-tail repeats encoding a polyubiquitin precursor protein. en. *Nature* **312**. Number : 5995 Publisher : Nature Publishing Group, 663-666. ISSN : 1476-4687. <https://www.nature.com/articles/312663a0> (2022) (déc. 1984).

- 
10. VIJAY-KUMAR, S., BUGG, C. E. & COOK, W. J. Structure of ubiquitin refined at 1.8 Å resolution. eng. *Journal of Molecular Biology* **194**, 531-544. ISSN : 0022-2836 (avr. 1987).
  11. LARSEN, C. N., KRANTZ, B. A. & WILKINSON, K. D. Substrate specificity of deubiquitinating enzymes : ubiquitin C-terminal hydrolases. eng. *Biochemistry* **37**, 3358-3368. ISSN : 0006-2960 (mars 1998).
  12. KOMANDER, D. & RAPE, M. The Ubiquitin Code. en. *Annual Review of Biochemistry* **81**, 203-229. ISSN : 0066-4154, 1545-4509. <https://www.annualreviews.org/doi/10.1146/annurev-biochem-060310-170328> (2022) (juill. 2012).
  13. DIKIC, I., WAKATSUKI, S. & WALTERS, K. J. Ubiquitin-binding domains - from structures to functions. eng. *Nature Reviews. Molecular Cell Biology* **10**, 659-671. ISSN : 1471-0080 (oct. 2009).
  14. SLOPER-MOULD, K. E., JEMC, J. C., PICKART, C. M. & HICKE, L. Distinct Functional Surface Regions on Ubiquitin\*. en. *Journal of Biological Chemistry* **276**, 30483-30489. ISSN : 0021-9258. <https://www.sciencedirect.com/science/article/pii/S0021925820897634> (2022) (août 2001).
  15. KAMADURAI, H. B. *et al.* Insights into ubiquitin transfer cascades from a structure of a UbcH5B approximately ubiquitin-HECT(NEDD4L) complex. eng. *Molecular Cell* **36**, 1095-1102. ISSN : 1097-4164 (déc. 2009).
  16. REYES-TURCU, F. E. *et al.* The ubiquitin binding domain ZnF UBP recognizes the C-terminal diglycine motif of unanchored ubiquitin. eng. *Cell* **124**, 1197-1208. ISSN : 0092-8674 (mars 2006).
  17. HU, M. *et al.* Crystal structure of a UBP-family deubiquitinating enzyme in isolation and in complex with ubiquitin aldehyde. eng. *Cell* **111**, 1041-1054. ISSN : 0092-8674 (déc. 2002).
  18. RAHIGHI, S. *et al.* Specific recognition of linear ubiquitin chains by NEMO is important for NF-kappaB activation. eng. *Cell* **136**, 1098-1109. ISSN : 1097-4172 (mars 2009).
  19. JIN, L., WILLIAMSON, A., BANERJEE, S., PHILIPP, I. & RAPE, M. Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex. eng. *Cell* **133**, 653-665. ISSN : 1097-4172 (mai 2008).
  20. WEBB, G. C., BAKER, R. T., FAGAN, K. & BOARD, P. G. Localization of the human UbB polyubiquitin gene to chromosome band 17p11.1-17p12. *American Journal of Human Genetics* **46**, 308-315. ISSN : 0002-9297. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1684968/> (2022) (fév. 1990).
  21. BOARD, P. G., COGGAN, M., BAKER, R. T., VUUST, J. & WEBB, G. C. Localization of the human UBC polyubiquitin gene to chromosome band 12q24.3. en. *Genomics*

- 
- 12**, 639-642. ISSN : 0888-7543. <https://www.sciencedirect.com/science/article/pii/S0888754392902873> (2022) (avr. 1992).
22. WEBB, G. C., BAKER, R. T., COGGAN, M. & BOARD, P. G. Localization of the Human UBA52 Ubiquitin Fusion Gene to Chromosome Band 19p13.1-p12. en. *Genomics* **19**, 567-569. ISSN : 0888-7543. <https://www.sciencedirect.com/science/article/pii/S0888754384711086> (2022) (fév. 1994).
23. KIRSCHNER, L. S. & STRATAKIS, C. A. Structure of the Human Ubiquitin Fusion Gene Uba80 (RPS27a) and One of Its Pseudogenes. en. *Biochemical and Biophysical Research Communications* **270**, 1106-1110. ISSN : 0006-291X. <https://www.sciencedirect.com/science/article/pii/S0006291X00925683> (2022) (avr. 2000).
24. KIMURA, Y. & TANAKA, K. Regulatory mechanisms involved in the control of ubiquitin homeostasis. *The Journal of Biochemistry* **147**, 793-798. ISSN : 0021-924X. <https://doi.org/10.1093/jb/mvq044> (2022) (juin 2010).
25. WELCHMAN, R. L., GORDON, C. & MAYER, R. J. Ubiquitin and ubiquitin-like proteins as multifunctional signals. en. *Nature Reviews Molecular Cell Biology* **6**. Number : 8 Publisher : Nature Publishing Group, 599-609. ISSN : 1471-0080. <https://www.nature.com/articles/nrm1700> (2022) (août 2005).
26. PADALA, P. *et al.* Novel insights into the interaction of UBA5 with UFM1 via a UFM1-interacting sequence. eng. *Scientific Reports* **7**, 508. ISSN : 2045-2322 (mars 2017).
27. WHITBY, F. G., XIA, G., PICKART, C. M. & HILL, C. P. Crystal structure of the human ubiquitin-like protein NEDD8 and interactions with ubiquitin pathway enzymes. eng. *The Journal of Biological Chemistry* **273**, 34983-34991. ISSN : 0021-9258 (déc. 1998).
28. AICHEM, A. *et al.* The structure of the ubiquitin-like modifier FAT10 reveals an alternative targeting mechanism for proteasomal degradation. *Nature Communications* **9**, 3321. ISSN : 2041-1723. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6102260/> (2023) (août 2018).
29. NARASIMHAN, J. *et al.* Crystal structure of the interferon-induced ubiquitin-like protein ISG15. eng. *The Journal of Biological Chemistry* **280**, 27356-27365. ISSN : 0021-9258 (juill. 2005).
30. MCNALLY, T. *et al.* Structural analysis of UBL5, a novel ubiquitin-like modifier. *Protein Science : A Publication of the Protein Society* **12**, 1562-1566. ISSN : 0961-8368. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2323916/> (2023) (juill. 2003).
31. CAPPADOCIA, L. *et al.* Structural and functional characterization of the phosphorylation-dependent interaction between PML and SUMO1. eng. *Structure (London, England : 1993)* **23**, 126-138. ISSN : 1878-4186 (jan. 2015).

- 
32. HUANG, W.-C., KO, T.-P., LI, S. S.-L. & WANG, A. H.-J. Crystal structures of the human SUMO-2 protein at 1.6 Å and 1.2 Å resolution : implication on the functional differences of SUMO proteins. eng. *European Journal of Biochemistry* **271**, 4114-4122. ISSN : 0014-2956 (oct. 2004).
  33. DING, H. *et al.* Solution structure of human SUMO-3 C47S and its binding surface for Ubc9. eng. *Biochemistry* **44**, 2790-2799. ISSN : 0006-2960 (mars 2005).
  34. VAUGHAN, R. M., KUPAI, A. & ROTHBART, S. B. Chromatin Regulation through Ubiquitin and Ubiquitin-like Histone Modifications. en. *Trends in Biochemical Sciences* **46**, 258-269. ISSN : 09680004. <https://linkinghub.elsevier.com/retrieve/pii/S0968000420302760> (2023) (avr. 2021).
  35. YE, Y. *et al.* Polyubiquitin binding and cross-reactivity in the USP domain deubiquitinase USP21. eng. *EMBO reports* **12**, 350-357. ISSN : 1469-3178 (avr. 2011).
  36. SCHULMAN, B. A. & WADE HARPER, J. Ubiquitin-like protein activation by E1 enzymes : the apex for downstream signalling pathways. en. *Nature Reviews Molecular Cell Biology* **10**, 319-331. ISSN : 1471-0072, 1471-0080. <https://www.nature.com/articles/nrm2673> (2022) (mai 2009).
  37. BARGHOUT, S. H. & SCHIMMER, A. D. E1 Enzymes as Therapeutic Targets in Cancer. en. *Pharmacological Reviews* **73** (éd. BARKER, E.) 1-56. ISSN : 0031-6997, 1521-0081. <http://pharmrev.aspetjournals.org/lookup/doi/10.1124/pharmrev.120.000053> (2022) (jan. 2021).
  38. HUANG, D. T., WALDEN, H., DUDA, D. & SCHULMAN, B. A. Ubiquitin-like protein activation. en. *Oncogene* **23**, 1958-1971. ISSN : 0950-9232, 1476-5594. <https://www.nature.com/articles/1207393> (2022) (mars 2004).
  39. LV, Z. *et al.* Domain alternation and active site remodeling are conserved structural features of ubiquitin E1. en. *Journal of Biological Chemistry* **292**, 12089-12099. ISSN : 00219258. <https://linkinghub.elsevier.com/retrieve/pii/S0021925820429813> (2022) (juill. 2017).
  40. HAAS, A. L. & ROSE, I. A. The mechanism of ubiquitin activating enzyme. A kinetic and equilibrium analysis. en. *Journal of Biological Chemistry* **257**, 10329-10337. ISSN : 00219258. <https://linkinghub.elsevier.com/retrieve/pii/S0021925818340249> (2022) (sept. 1982).
  41. SCHÄFER, A., KUHN, M. & SCHINDELIN, H. Structure of the ubiquitin-activating enzyme loaded with two ubiquitin molecules. eng. *Acta Crystallographica. Section D, Biological Crystallography* **70**, 1311-1320. ISSN : 1399-0047 (mai 2014).
  42. CLAGUE, M. J., HERIDE, C. & URBÉ, S. The demographics of the ubiquitin system. en. *Trends in Cell Biology* **25**, 417-426. ISSN : 09628924. <https://linkinghub.elsevier.com/retrieve/pii/S0962892415000549> (2022) (juill. 2015).

- 
43. PICKART, C. M. & ROSE, I. A. Functional heterogeneity of ubiquitin carrier proteins. en. *Journal of Biological Chemistry* **260**, 1573-1581. ISSN : 00219258. <https://linkinghub.elsevier.com/retrieve/pii/S0021925818896326> (2022) (fév. 1985).
44. SUMMERS, M. K., PAN, B., MUKHYALA, K. & JACKSON, P. K. The Unique N Terminus of the UbcH10 E2 Enzyme Controls the Threshold for APC Activation and Enhances Checkpoint Regulation of the APC. en. *Molecular Cell* **31**, 544-556. ISSN : 10972765. <https://linkinghub.elsevier.com/retrieve/pii/S1097276508005030> (2022) (août 2008).
45. SILVER, E., GWOZD, T., PTAK, C., GOEBL, M. & ELLISON, M. A chimeric ubiquitin conjugating enzyme that combines the cell cycle properties of CDC34 (UBC3) and the DNA repair properties of RAD6 (UBC2) : implications for the structure, function and evolution of the E2s. en. *The EMBO Journal* **11**, 3091-3098. ISSN : 02614189. <https://onlinelibrary.wiley.com/doi/10.1002/j.1460-2075.1992.tb05381.x> (2022) (août 1992).
46. HALDEMAN, M. T., XIA, G., KASPEREK, E. M. & PICKART, C. M. Structure and Function of Ubiquitin Conjugating Enzyme E2-25K : The Tail Is a Core-Dependent Activity Element. en. *Biochemistry* **36**, 10526-10537. ISSN : 0006-2960, 1520-4995. <https://pubs.acs.org/doi/10.1021/bi970750u> (2022) (août 1997).
47. BICH, G., MONSELLIER, E., TRAVÉ, G. & NOMINÉ, Y. ProFeatMap : a customizable tool for 2D feature representation of protein sets. *bioRxiv*. Publisher : Cold Spring Harbor Laboratory \_eprint : <https://www.biorxiv.org/content/early/2022/04/14/2022.04.14.488325.full.pdf> <https://www.biorxiv.org/content/early/2022/04/14/2022.04.14.488325> (2022).
48. SHENG, Y. *et al.* A human ubiquitin conjugating enzyme (E2)-HECT E3 ligase structure-function screen. eng. *Molecular & cellular proteomics : MCP* **11**, 329-341. ISSN : 1535-9484 (août 2012).
49. BURROUGHS, A. M., JAFFEE, M., IYER, L. M. & ARAVIND, L. Anatomy of the E2 ligase fold : Implications for enzymology and evolution of ubiquitin/Ub-like protein conjugation. en. *Journal of Structural Biology* **162**, 205-218. ISSN : 10478477. <https://linkinghub.elsevier.com/retrieve/pii/S1047847707003152> (2022) (mai 2008).
50. OLSEN, S. K. & LIMA, C. D. Structure of a Ubiquitin E1-E2 Complex : Insights to E1-E2 Thioester Transfer. en. *Molecular Cell* **49**, 884-896. ISSN : 10972765. <https://linkinghub.elsevier.com/retrieve/pii/S1097276513000464> (2023) (mars 2013).
51. LEE, I. & SCHINDELIN, H. Structural Insights into E1-Catalyzed Ubiquitin Activation and Transfer to Conjugating Enzymes. *Cell* **134**, 268-278. ISSN : 0092-8674. <https://www.sciencedirect.com/science/article/pii/S0092867408007095> (2008).

- 
52. YE, Y. & RAPE, M. Building ubiquitin chains : E2 enzymes at work. en. *Nature Reviews Molecular Cell Biology* **10**, 755-764. ISSN : 1471-0072, 1471-0080. <http://www.nature.com/articles/nrm2780> (2022) (nov. 2009).
53. STEWART, M. D., RITTERHOFF, T., KLEVIT, R. E. & BRZOVIC, P. S. E2 enzymes : more than just middle men. en. *Cell Research* **26**, 423-440. ISSN : 1001-0602, 1748-7838. <http://www.nature.com/articles/cr201635> (2022) (avr. 2016).
54. MIURA, T., KLAUS, W., GSELL, B., MIYAMOTO, C. & SENN, H. Characterization of the binding interface between ubiquitin and class I human ubiquitin-conjugating enzyme 2b by multidimensional heteronuclear NMR spectroscopy in solution<sup>11</sup>Edited by P. E. Wright. en. *Journal of Molecular Biology* **290**, 213-228. ISSN : 0022-2836. <https://www.sciencedirect.com/science/article/pii/S0022283699928590> (2023) (juill. 1999).
55. SAKATA, E. *et al.* Crystal Structure of UbcH5bUbiquitin Intermediate : Insight into the Formation of the Self-Assembled E2Ub Conjugates. *Structure* **18**, 138-147. ISSN : 0969-2126. <https://www.sciencedirect.com/science/article/pii/S096921260900464X> (2010).
56. FREEMONT, P. S., HANSON, I. M. & TROWSDALE, J. A novel cysteine-rich sequence motif. *Cell* **64**, 483-484. ISSN : 0092-8674. <https://www.sciencedirect.com/science/article/pii/009286749190229R> (1991).
57. BARLOW, P. N., LUISI, B., MILNER, A., ELLIOTT, M. & EVERETT, R. Structure of the C3HC4 Domain by 1H-nuclear Magnetic Resonance Spectroscopy. : A New Structural Class of Zinc-finger. *Journal of Molecular Biology* **237**, 201-211. ISSN : 0022-2836. <https://www.sciencedirect.com/science/article/pii/S0022283684712228> (1994).
58. DESHAIES, R. J. & JOAZEIRO, C. A. RING Domain E3 Ubiquitin Ligases. en. *Annual Review of Biochemistry* **78**, 399-434. ISSN : 0066-4154, 1545-4509. <https://www.annualreviews.org/doi/10.1146/annurev.biochem.78.101807.093809> (2022) (juin 2009).
59. SAURIN, A. J., BORDEN, K. L. B., BODDY, M. N. & FREEMONT, P. S. Does this have a familiar RING? *Trends in Biochemical Sciences* **21**, 208-214. ISSN : 0968-0004. <https://www.sciencedirect.com/science/article/pii/S096800049680017X> (1996).
60. COSCOY, L. & GANEM, D. PHD domains and E3 ubiquitin ligases : viruses make the connection. *Trends in Cell Biology* **13**, 7-12. ISSN : 0962-8924. <https://www.sciencedirect.com/science/article/pii/S0962892402000053> (2003).
61. ARAVIND, L., IYER, L. & KOONIN, E. V. Scores of RINGS but No PHDs in Ubiquitin Signaling. en. *Cell Cycle* **2**, 123-126. ISSN : 1538-4101, 1551-4005. <http://www.tandfonline.com/doi/abs/10.4161/cc.2.2.335> (2022) (mars 2003).

- 
62. ZHENG, N., WANG, P., JEFFREY, P. D. & PAVLETICH, N. P. Structure of a c-Cbl-UbcH7 Complex : RING Domain Function in Ubiquitin-Protein Ligases. *Cell* **102**, 533-539. ISSN : 0092-8674. <https://www.sciencedirect.com/science/article/pii/S009286740000057X> (2000).
63. KOZLOV, G. *et al.* Structural Basis for UBA-mediated Dimerization of c-Cbl Ubiquitin Ligase\*. *Journal of Biological Chemistry* **282**, 27547-27555. ISSN : 0021-9258. <https://www.sciencedirect.com/science/article/pii/S0021925820588261> (2007).
64. MACE, P. D. *et al.* Structures of the cIAP2 RING Domain Reveal Conformational Changes Associated with Ubiquitin-conjugating Enzyme (E2) Recruitment\*. *Journal of Biological Chemistry* **283**, 31633-31640. ISSN : 0021-9258. <https://www.sciencedirect.com/science/article/pii/S0021925820569068> (2008).
65. LIEW, C. W., SUN, H., HUNTER, T. & DAY, C. L. RING domain dimerization is essential for RNF4 function. en. *Biochemical Journal* **431**, 23-29. ISSN : 0264-6021, 1470-8728. <https://portlandpress.com/biochemj/article/431/1/23/45275/RING-domain-dimerization-is-essential-for-RNF4> (2022) (oct. 2010).
66. DOU, H., BUETOW, L., SIBBET, G. J., CAMERON, K. & HUANG, D. T. BIRC7-E2 ubiquitin conjugate structure reveals the mechanism of ubiquitin transfer by a RING dimer. en. *Nature Structural & Molecular Biology* **19**, 876-883. ISSN : 1545-9993, 1545-9985. <http://www.nature.com/articles/nsmb.2379> (2022) (sept. 2012).
67. HASHIZUME, R. *et al.* The RING Heterodimer BRCA1-BARD1 Is a Ubiquitin Ligase Inactivated by a Breast Cancer-derived Mutation\*. *Journal of Biological Chemistry* **276**, 14537-14540. ISSN : 0021-9258. <https://www.sciencedirect.com/science/article/pii/S002192581956728X> (2001).
68. WANG, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. en. *Nature* **431**, 873-878. ISSN : 0028-0836, 1476-4687. <http://www.nature.com/articles/nature02985> (2022) (oct. 2004).
69. LINARES, L. K., HENGSTERMANN, A., CIECHANOVER, A., MÜLLER, S. & SCHEFFNER, M. HdmX stimulates Hdm2-mediated ubiquitination and degradation of p53. en. *Proceedings of the National Academy of Sciences* **100**, 12009-12014. ISSN : 0027-8424, 1091-6490. <https://pnas.org/doi/full/10.1073/pnas.2030930100> (2022) (oct. 2003).
70. PETROSKI, M. D. & DESHAIES, R. J. Mechanism of Lysine 48-Linked Ubiquitin-Chain Synthesis by the Cullin-RING Ubiquitin-Ligase Complex SCF-Cdc34. *Cell* **123**, 1107-1120. ISSN : 0092-8674. <https://www.sciencedirect.com/science/article/pii/S0092867405010391> (2005).
71. MORREALE, F. E. & WALDEN, H. Types of Ubiquitin Ligases. en. *Cell* **165**, 248-248.e1. ISSN : 00928674. <https://linkinghub.elsevier.com/retrieve/pii/S0092867416302598> (2022) (mars 2016).

- 
72. ELETR, Z. M., HUANG, D. T., DUDA, D. M., SCHULMAN, B. A. & KUHLMAN, B. E2 conjugating enzymes must disengage from their E1 enzymes before E3-dependent ubiquitin and ubiquitin-like transfer. en. *Nature Structural & Molecular Biology* **12**, 933-934. ISSN : 1545-9993, 1545-9985. <http://www.nature.com/articles/nsmb984> (2022) (oct. 2005).
73. PRUNEDA, J. N., STOLL, K. E., BOLTON, L. J., BRZOVIC, P. S. & KLEVIT, R. E. Ubiquitin in Motion : Structural Studies of the Ubiquitin-Conjugating Enzyme-Ubiquitin Conjugate. en. *Biochemistry* **50**, 1624-1633. ISSN : 0006-2960, 1520-4995. <https://pubs.acs.org/doi/10.1021/bi101913m> (2022) (mars 2011).
74. LAMPERT, F. *et al.* The multi-subunit GID/CTLH E3 ubiquitin ligase promotes cell proliferation and targets the transcription factor Hbp1 for degradation. en. *eLife* **7**, e35528. ISSN : 2050-084X. <https://elifesciences.org/articles/35528> (2023) (juin 2018).
75. ITOH, M. *et al.* Mind Bomb Is a Ubiquitin Ligase that Is Essential for Efficient Activation of Notch Signaling by Delta. en. *Developmental Cell* **4**, 67-82. ISSN : 15345807. <https://linkinghub.elsevier.com/retrieve/pii/S1534580702004094> (2023) (jan. 2003).
76. ARAVIND, L. & KOONIN, E. V. The U box is a modified RING finger — a common domain in ubiquitination. en. *Current Biology* **10**, R132-R134. ISSN : 09609822. <https://linkinghub.elsevier.com/retrieve/pii/S0960982200003985> (2022) (fév. 2000).
77. OHI, M. D., VANDER KOOI, C. W., ROSENBERG, J. A., CHAZIN, W. J. & GOULD, K. L. Structural insights into the U-box, a domain associated with multi-ubiquitination. en. *Nature Structural Biology* **10**, 250-255. ISSN : 1072-8368 (avr. 2003).
78. ZHANG, M. *et al.* Chaperoned Ubiquitylation—Crystal Structures of the CHIP U Box E3 Ubiquitin Ligase and a CHIP-Ubc13-Uev1a Complex. en. *Molecular Cell* **20**, 525-538. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S1097276505016448> (2023) (nov. 2005).
79. OHI, M. D. *et al.* Structural and Functional Analysis of Essential pre-mRNA Splicing Factor Prp19p. en. *Molecular and Cellular Biology* **25**, 451-460. ISSN : 0270-7306, 1098-5549. <https://journals.asm.org/doi/10.1128/MCB.25.1.451-460.2005> (2022) (jan. 2005).
80. HUIBREGTSE, J. M., SCHEFFNER, M., BEAUDENON, S. & HOWLEY, P. M. A family of proteins structurally and functionally related to the E6-AP ubiquitin-protein ligase. *Proceedings of the National Academy of Sciences of the United States of America* **92**, 2563-2567. ISSN : 0027-8424. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC42258/> (2022) (mars 1995).

- 
81. HUANG, L. *et al.* Structure of an E6AP-UbcH7 Complex : Insights into Ubiquitination by the E2-E3 Enzyme Cascade. en. *Science* **286**, 1321-1326. ISSN : 0036-8075, 1095-9203. <https://www.science.org/doi/10.1126/science.286.5443.1321> (2022) (nov. 1999).
  82. VERDECIA, M. A. *et al.* Conformational Flexibility Underlies Ubiquitin Ligation Mediated by the WWP1 HECT Domain E3 Ligase. *Molecular Cell* **11**, 249-259. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S1097276502007748> (2003).
  83. OGUNJIMI, A. A. *et al.* Regulation of Smurf2 Ubiquitin Ligase Activity by Anchoring the E2 to the HECT Domain. *Molecular Cell* **19**, 297-308. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S1097276505014309> (2005).
  84. NUBER, U. & SCHEFFNER, M. Identification of Determinants in E2 Ubiquitin-conjugating Enzymes Required for hect E3 Ubiquitin-Protein Ligase Interaction \*. English. *Journal of Biological Chemistry* **274**. Publisher : Elsevier, 7576-7582. ISSN : 0021-9258, 1083-351X. [https://www.jbc.org/article/S0021-9258\(18\)37037-6/abstract](https://www.jbc.org/article/S0021-9258(18)37037-6/abstract) (2023) (mars 1999).
  85. MASPERO, E. *et al.* Structure of a ubiquitin-loaded HECT ligase reveals the molecular basis for catalytic priming. eng. *Nature Structural & Molecular Biology* **20**, 696-701. ISSN : 1545-9985 (juin 2013).
  86. KAMADURAI, H. B. *et al.* Mechanism of ubiquitin ligation and lysine prioritization by a HECT E3. *eLife* **2** (éd. KURIYAN, J.) Publisher : eLife Sciences Publications, Ltd, e00828. ISSN : 2050-084X. <https://doi.org/10.7554/eLife.00828> (2022) (août 2013).
  87. SALVAT, C., WANG, G., DASTUR, A., LYON, N. & HUIBREGTSE, J. M. The -4 Phenylalanine Is Required for Substrate Ubiquitination Catalyzed by HECT Ubiquitin Ligases \*. English. *Journal of Biological Chemistry* **279**. Publisher : Elsevier, 18935-18943. ISSN : 0021-9258, 1083-351X. [https://www.jbc.org/article/S0021-9258\(19\)75454-4/abstract](https://www.jbc.org/article/S0021-9258(19)75454-4/abstract) (2022) (avr. 2004).
  88. RONCHI, V. P., KLEIN, J. M. & HAAS, A. L. E6AP/UBE3A ubiquitin ligase harbors two E2~ubiquitin binding sites. eng. *The Journal of Biological Chemistry* **288**, 10349-10360. ISSN : 1083-351X (avr. 2013).
  89. MARÍN, I., LUCAS, J. I., GRADILLA, A.-C. & FERRÚS, A. Parkin and relatives : the RBR family of ubiquitin ligases. eng. *Physiological Genomics* **17**, 253-263. ISSN : 1531-2267 (mai 2004).
  90. LECHTENBERG, B. C. *et al.* Structure of a HOIP/E2~ubiquitin complex reveals RBR E3 ligase mechanism and regulation. eng. *Nature* **529**, 546-550. ISSN : 1476-4687 (jan. 2016).

- 
91. COOK, W. J., JEFFREY, L. C., CARSON, M., CHEN, Z. & PICKART, C. M. Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). eng. *The Journal of Biological Chemistry* **267**, 16467-16471. ISSN : 0021-9258 (août 1992).
  92. RYABOV, Y. & FUSHMAN, D. Structural assembly of multidomain proteins and protein complexes guided by the overall rotational diffusion tensor. eng. *Journal of the American Chemical Society* **129**, 7894-7902. ISSN : 0002-7863 (juin 2007).
  93. BREMM, A., FREUND, S. M. V. & KOMANDER, D. Lys11-linked ubiquitin chains adopt compact conformations and are preferentially hydrolyzed by the deubiquitinase Cezanne. eng. *Nature Structural & Molecular Biology* **17**, 939-947. ISSN : 1545-9985 (août 2010).
  94. MATSUMOTO, M. L. *et al.* K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. eng. *Molecular Cell* **39**, 477-484. ISSN : 1097-4164 (août 2010).
  95. VIRDEE, S., YE, Y., NGUYEN, D. P., KOMANDER, D. & CHIN, J. W. Engineered diubiquitin synthesis reveals Lys29-isopeptide specificity of an OTU deubiquitinase. eng. *Nature Chemical Biology* **6**, 750-757. ISSN : 1552-4469 (oct. 2010).
  96. KOMANDER, D. *et al.* Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. eng. *EMBO reports* **10**, 466-473. ISSN : 1469-3178 (mai 2009).
  97. SWANEY, D. L., RODRÍGUEZ-MIAS, R. A. & VILLÉN, J. Phosphorylation of ubiquitin at Ser65 affects its polymerization, targets, and proteome-wide turnover. *EMBO reports* **16**. Publisher : John Wiley & Sons, Ltd, 1131-1144. ISSN : 1469-221X. <https://www.embopress.org/doi/full/10.15252/embr.201540298> (2023) (sept. 2015).
  98. OHTAKE, F. *et al.* Ubiquitin acetylation inhibits polyubiquitin chain elongation. *EMBO reports* **16**. Publisher : John Wiley & Sons, Ltd, 192-201. ISSN : 1469-221X. <https://www.embopress.org/doi/full/10.15252/embr.201439152> (2023) (fév. 2015).
  99. HERHAUS, L. & DIKIC, I. Expanding the ubiquitin code through post-translational modification. *EMBO reports* **16**. Publisher : John Wiley & Sons, Ltd, 1071-1083. ISSN : 1469-221X. <https://www.embopress.org/doi/full/10.15252/embr.201540891> (2023) (sept. 2015).
  100. ELSASSER, S., CHANDLER-MILITELLO, D., MÜLLER, B., HANNA, J. & FINLEY, D. Rad23 and Rpn10 Serve as Alternative Ubiquitin Receptors for the Proteasome \*. English. *Journal of Biological Chemistry* **279**. Publisher : Elsevier, 26817-26822. ISSN : 0021-9258, 1083-351X. [https://www.jbc.org/article/S0021-9258\(20\)85335-6/abstract](https://www.jbc.org/article/S0021-9258(20)85335-6/abstract) (2023) (juin 2004).

- 
101. HUSNJAK, K. *et al.* Proteasome subunit Rpn13 is a novel ubiquitin receptor. en. *Nature* **453**. Number : 7194 Publisher : Nature Publishing Group, 481-488. ISSN : 1476-4687. <https://www.nature.com/articles/nature06926> (2023) (mai 2008).
  102. DANIELI, A. & MARTENS, S. p62-mediated phase separation at the intersection of the ubiquitin-proteasome system and autophagy. *Journal of Cell Science* **131**, jcs214304. ISSN : 0021-9533. <https://doi.org/10.1242/jcs.214304> (2023) (oct. 2018).
  103. ORDUREAU, A. *et al.* Quantitative Proteomics Reveal a Feedforward Mechanism for Mitochondrial PARKIN Translocation and Ubiquitin Chain Synthesis. en. *Molecular Cell* **56**, 360-375. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S109727651400714X> (2023) (nov. 2014).
  104. CUNNINGHAM, C. N. *et al.* USP30 and parkin homeostatically regulate atypical ubiquitin chains on mitochondria. en. *Nature Cell Biology* **17**. Number : 2 Publisher : Nature Publishing Group, 160-169. ISSN : 1476-4679. <https://www.nature.com/articles/ncb3097> (2023) (fév. 2015).
  105. POHL, C. & DIKIC, I. Cellular quality control by the ubiquitin-proteasome system and autophagy. *Science* **366**. Publisher : American Association for the Advancement of Science, 818-822. <https://www.science.org/doi/10.1126/science.aax3769> (2023) (nov. 2019).
  106. WAGNER, S. A. *et al.* A Proteome-wide, Quantitative Survey of In Vivo Ubiquitylation Sites Reveals Widespread Regulatory Roles\*. en. *Molecular & Cellular Proteomics* **10**, M111.013284. ISSN : 1535-9476. <https://www.sciencedirect.com/science/article/pii/S1535947620301298> (2023) (oct. 2011).
  107. KIM, W. *et al.* Systematic and Quantitative Assessment of the Ubiquitin-Modified Proteome. en. *Molecular Cell* **44**, 325-340. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S1097276511006757> (2023) (oct. 2011).
  108. ELIA, A. E. H. *et al.* Quantitative Proteomic Atlas of Ubiquitination and Acetylation in the DNA Damage Response. eng. *Molecular Cell* **59**, 867-881. ISSN : 1097-4164 (sept. 2015).
  109. WU-BAER, F. & BAER, R. Effect of DNA damage on a BRCA1 complex. en. *Nature* **414**. Number : 6859 Publisher : Nature Publishing Group, 36-36. ISSN : 1476-4687. <https://www.nature.com/articles/35102118> (2023) (nov. 2001).
  110. MORRIS, J. R. & SOLOMON, E. BRCA1 :BARD1 induces the formation of conjugated ubiquitin structures, dependent on K6 of ubiquitin, in cells during DNA replication and repair. *Human Molecular Genetics* **13**, 807-817. ISSN : 0964-6906. <https://doi.org/10.1093/hmg/ddh095> (2023) (avr. 2004).
  111. WU, T. *et al.* UBE2S drives elongation of K11-linked ubiquitin chains by the anaphase-promoting complex. eng. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1355-1360. ISSN : 1091-6490 (jan. 2010).

- 
112. GATTI, M. *et al.* RNF168 Promotes Noncanonical K27 Ubiquitination to Signal DNA Damage. en. *Cell Reports* **10**, 226-238. ISSN : 2211-1247. <https://www.sciencedirect.com/science/article/pii/S221112471401050X> (2023) (jan. 2015).
113. ZHOU, Q. & ZHANG, J. K27-linked noncanonic ubiquitination in immune regulation. en. *Journal of Leukocyte Biology* **111**. \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/JLB.4397RR.223-235>. ISSN : 1938-3673. <https://onlinelibrary.wiley.com/doi/abs/10.1002/JLB.4RU0620-397RR> (2023) (2022).
114. YU, Y. *et al.* K29-linked ubiquitin signaling regulates proteotoxic stress response and cell cycle. en. *Nature Chemical Biology* **17**, 896-905. ISSN : 1552-4450, 1552-4469. <http://www.nature.com/articles/s41589-021-00823-5> (2023) (août 2021).
115. HUANG, H. *et al.* K33-Linked Polyubiquitination of T Cell Receptor- $\zeta$  Regulates Proteolysis-Independent T Cell Signaling. en. *Immunity* **33**, 60-70. ISSN : 1074-7613. <https://www.sciencedirect.com/science/article/pii/S107476131000244X> (2023) (juill. 2010).
116. LIN, M. *et al.* USP38 Inhibits Type I Interferon Signaling by Editing TBK1 Ubiquitination through NLRP4 Signalosome. en. *Molecular Cell* **64**, 267-281. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S1097276516304762> (2023) (oct. 2016).
117. RIESER, E., CORDIER, S. M. & WALCZAK, H. Linear ubiquitination : a newly discovered regulator of cell signalling. *Trends in Biochemical Sciences* **38**, 94-102. ISSN : 0968-0004. <https://www.sciencedirect.com/science/article/pii/S0968000412001909> (2013).
118. SPIT, M., RIESER, E. & WALCZAK, H. Linear ubiquitination at a glance. en. *Journal of Cell Science* **132**, jcs208512. ISSN : 1477-9137, 0021-9533. <https://journals.biologists.com/jcs/article/132/2/jcs208512/57247/Linear-ubiquitination-at-a-glance> (2023) (jan. 2019).
119. TRACZ, M. & BIALEK, W. Beyond K48 and K63 : non-canonical protein ubiquitination. *Cellular & Molecular Biology Letters* **26**, 1. ISSN : 1689-1392. <https://doi.org/10.1186/s11658-020-00245-6> (2023) (jan. 2021).
120. KOMANDER, D., CLAGUE, M. J. & URBÉ, S. Breaking the chains : structure and function of the deubiquitinases. en. *Nature Reviews Molecular Cell Biology* **10**. Number : 8 Publisher : Nature Publishing Group, 550-563. ISSN : 1471-0080. <https://www.nature.com/articles/nrm2731> (2023) (août 2009).
121. RIVETT, A. J., PALMER, A. & KNECHT, E. Electron microscopic localization of the multicatalytic proteinase complex in rat liver and in cultured cells. en. *Journal of Histochemistry & Cytochemistry* **40**, 1165-1172. ISSN : 0022-1554, 1551-5044. <http://journals.sagepub.com/doi/10.1177/40.8.1619280> (2023) (août 1992).

- 
122. WÓJCIK, C. & DEMARTINO, G. N. Intracellular localization of proteasomes. en. *The International Journal of Biochemistry & Cell Biology* **35**, 579-589. ISSN : 13572725. <https://linkinghub.elsevier.com/retrieve/pii/S1357272502003801> (2023) (mai 2003).
123. KOPP, F., STEINER, R., DAHLMANN, B., KUEHN, L. & REINAUER, H. Size and shape of the multicatalytic proteinase from rat skeletal muscle. en. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **872**, 253-260. ISSN : 0167-4838. <https://www.sciencedirect.com/science/article/pii/0167483886902785> (2023) (août 1986).
124. DING, Z. *et al.* High-resolution cryo-EM structure of the proteasome in complex with ADP-AIFx. eng. *Cell Research* **27**, 373-385. ISSN : 1748-7838 (mars 2017).
125. UNNO, M. *et al.* The structure of the mammalian 20S proteasome at 2.75 Å resolution. eng. *Structure (London, England : 1993)* **10**, 609-618. ISSN : 0969-2126 (mai 2002).
126. GROLL, M. *et al.* Structure of 20S proteasome from yeast at 2.4Å resolution. en. *Nature* **386**. Number : 6624 Publisher : Nature Publishing Group, 463-471. ISSN : 1476-4687. <https://www.nature.com/articles/386463a0> (2023) (avr. 1997).
127. TANAKA, K. The proteasome : Overview of structure and functions. en. *Proceedings of the Japan Academy, Series B* **85**, 12-36. ISSN : 0386-2208, 1349-2896. [http://www.jstage.jst.go.jp/article/pjab/85/1/85\\_1\\_12/\\_article](http://www.jstage.jst.go.jp/article/pjab/85/1/85_1_12/_article) (2023) (2009).
128. DEMARTINO, G. N. & GILLETTE, T. G. Proteasomes : Machines for All Reasons. English. *Cell* **129**. Publisher : Elsevier, 659-662. ISSN : 0092-8674, 1097-4172. [https://www.cell.com/cell/abstract/S0092-8674\(07\)00598-3](https://www.cell.com/cell/abstract/S0092-8674(07)00598-3) (2023) (mai 2007).
129. MAUPIN-FURLOW, J. A. *et al.* Proteasomes from structure to function : perspectives from Archaea. eng. *Current Topics in Developmental Biology* **75**, 125-169. ISSN : 0070-2153 (2006).
130. ABI HABIB, J., LESENFANTS, J., VIGNERON, N. & VAN DEN EYNDE, B. J. Functional Differences between Proteasome Subtypes. en. *Cells* **11**, 421. ISSN : 2073-4409. <https://www.mdpi.com/2073-4409/11/3/421> (2023) (jan. 2022).
131. LANDER, G. C. *et al.* Complete subunit architecture of the proteasome regulatory particle. en. *Nature* **482**. Number : 7384 Publisher : Nature Publishing Group, 186-191. ISSN : 1476-4687. <https://www.nature.com/articles/nature10774> (2023) (fév. 2012).
132. TOMKO, R. J., FUNAKOSHI, M., SCHNEIDER, K., WANG, J. & HOCHSTRASSER, M. Heterohexameric Ring Arrangement of the Eukaryotic Proteasomal ATPases : Implications for Proteasome Structure and Assembly. en. *Molecular Cell* **38**, 393-403. ISSN : 10972765. <https://linkinghub.elsevier.com/retrieve/pii/S109727651000314X> (2023) (mai 2010).

- 
133. TANAKA, K. Role of proteasomes modified by interferon-  $\gamma$  in antigen processing. en. *Journal of Leukocyte Biology* **56**, 571-575. ISSN : 07415400. <https://onlinelibrary.wiley.com/doi/10.1002/jlb.56.5.571> (2023) (nov. 1994).
134. MURATA, S. *et al.* Regulation of CD8+ T Cell Development by Thymus-Specific Proteasomes. *Science* **316**. Publisher : American Association for the Advancement of Science, 1349-1353. <https://www.science.org/doi/full/10.1126/science.1141915> (2023) (juin 2007).
135. KNOWLTON, J. R. *et al.* Structure of the proteasome activator REG $\alpha$  (PA28 $\alpha$ ). en. *Nature* **390**. Number : 6660 Publisher : Nature Publishing Group, 639-643. ISSN : 1476-4687. <https://www.nature.com/articles/37670> (2023) (déc. 1997).
136. ZHOU, P. REG $\gamma$  : A Shortcut to Destruction. English. *Cell* **124**. Publisher : Elsevier, 256-257. ISSN : 0092-8674, 1097-4172. [https://www.cell.com/cell/abstract/S0092-8674\(06\)00051-1](https://www.cell.com/cell/abstract/S0092-8674(06)00051-1) (2023) (jan. 2006).
137. KHOR, B. *et al.* Proteasome activator PA200 is required for normal spermatogenesis. eng. *Molecular and Cellular Biology* **26**, 2999-3007. ISSN : 0270-7306 (avr. 2006).
138. USTRELL, V., HOFFMAN, L., PRATT, G. & RECHSTEINER, M. PA200, a nuclear proteasome activator involved in DNA repair. eng. *The EMBO journal* **21**, 3516-3525. ISSN : 0261-4189 (juill. 2002).
139. FINLEY, D., CHEN, X. & WALTERS, K. J. Gates, Channels, and Switches : Elements of the Proteasome Machine. eng. *Trends in Biochemical Sciences* **41**, 77-93. ISSN : 0968-0004 (jan. 2016).
140. SHI, Y. *et al.* Rpn1 provides adjacent receptor sites for substrate binding and deubiquitination by the proteasome. eng. *Science (New York, N. Y.)* **351**, 10.1126/science.aad9421 aad9421. ISSN : 1095-9203 (fév. 2016).
141. BARD, J. A. M. *et al.* Structure and Function of the 26S Proteasome. eng. *Annual Review of Biochemistry* **87**, 697-724. ISSN : 1545-4509 (juin 2018).
142. LIVNEH, I., COHEN-KAPLAN, V., COHEN-ROSENZWEIG, C., AVNI, N. & CIECHANOVER, A. The life cycle of the 26S proteasome : from birth, through regulation and function, and onto its death. eng. *Cell Research* **26**, 869-885. ISSN : 1748-7838 (août 2016).
143. VERMA, R. *et al.* Role of Rpn11 metalloprotease in deubiquitination and degradation by the 26S proteasome. eng. *Science (New York, N. Y.)* **298**, 611-615. ISSN : 1095-9203 (oct. 2002).
144. BHATTACHARYYA, S., YU, H., MIM, C. & MATOUSCHEK, A. Regulated protein turnover : snapshots of the proteasome in action. en. *Nature Reviews Molecular Cell Biology* **15**. Number : 2 Publisher : Nature Publishing Group, 122-133. ISSN : 1471-0080. <https://www.nature.com/articles/nrm3741> (2023) (fév. 2014).
145. SMITH, D. M. *et al.* ATP Binding to PAN or the 26S ATPases Causes Association with the 20S Proteasome, Gate Opening, and Translocation of Unfolded Proteins. en.

- 
- Molecular Cell* **20**, 687-698. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S1097276505017144> (2023) (déc. 2005).
146. VOGES, D., ZWICKL, P. & BAUMEISTER, W. The 26S proteasome : a molecular machine designed for controlled proteolysis. eng. *Annual Review of Biochemistry* **68**, 1015-1068. ISSN : 0066-4154 (1999).
147. CARDOZO, C. Catalytic components of the bovine pituitary multicatalytic proteinase complex (proteasome). eng. *Enzyme & Protein* **47**, 296-305. ISSN : 1019-6773 (1993).
148. GROLL, M. & HUBER, R. Inhibitors of the eukaryotic 20S proteasome core particle : a structural approach. en. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* **1695**, 33-44. ISSN : 01674889. <https://linkinghub.elsevier.com/retrieve/pii/S0167488904002435> (2023) (nov. 2004).
149. SARIC, T., GRAEF, C. I. & GOLDBERG, A. L. Pathway for Degradation of Peptides Generated by Proteasomes. en. *Journal of Biological Chemistry* **279**, 46723-46732. ISSN : 00219258. <https://linkinghub.elsevier.com/retrieve/pii/S0021925820698866> (2023) (nov. 2004).
150. YAMAMOTO, Y., HUIBREGTSE, J. M. & HOWLEY, P. M. The Human E6-AP Gene (UBE3A) Encodes Three Potential Protein Isoforms Generated by Differential Splicing. en. *Genomics* **41**, 263-266. ISSN : 0888-7543. <https://www.sciencedirect.com/science/article/pii/S0888754397946176> (2023) (avr. 1997).
151. SIROIS, C. L. *et al.* Abundance and localization of human UBE3A protein isoforms. *Human Molecular Genetics* **29**, 3021-3031. ISSN : 0964-6906. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7645711/> (2023) (août 2020).
152. LASALLE, J. M., REITER, L. T. & CHAMBERLAIN, S. J. Epigenetic regulation of UBE3A and roles in human neurodevelopmental disorders. *Epigenomics* **7**, 1213-1228. ISSN : 1750-1911. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4709177/> (2023) (oct. 2015).
153. KIM, H. C. & HUIBREGTSE, J. M. Polyubiquitination by HECT E3s and the determinants of chain type specificity. eng. *Molecular and Cellular Biology* **29**, 3307-3318. ISSN : 1098-5549 (juin 2009).
154. KUMAR, S., TALIS, A. L. & HOWLEY, P. M. Identification of HHR23A as a Substrate for E6-associated Protein-mediated Ubiquitination. en. *Journal of Biological Chemistry* **274**, 18785-18792. ISSN : 00219258. <https://linkinghub.elsevier.com/retrieve/pii/S0021925819872216> (2023) (juin 1999).
155. MULHERKAR, S. A., SHARMA, J. & JANA, N. R. The ubiquitin ligase E6-AP promotes degradation of  $\alpha$ -synuclein. en. *Journal of Neurochemistry* **110**. \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1471-4159.2009.06293.x>, 1955-1964. ISSN : 1471-4159. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1471-4159.2009.06293.x> (2023) (2009).

- 
156. MARGOLIS, S. S. *et al.* EphB-Mediated Degradation of the RhoA GEF Ephexin5 Relieves a Developmental Brake on Excitatory Synapse Formation. en. *Cell* **143**, 442-455. ISSN : 00928674. <https://linkinghub.elsevier.com/retrieve/pii/S0092867410011244> (2023) (oct. 2010).
157. ZAAROR-REGEV, D. *et al.* Regulation of the polycomb protein Ring1B by self-ubiquitination or by E6-AP may have implications to the pathogenesis of Angelman syndrome. eng. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 6788-6793. ISSN : 1091-6490 (avr. 2010).
158. SUN, A. X. *et al.* Potassium channel dysfunction in human neuronal models of Angelman syndrome. eng. *Science (New York, N.Y.)* **366**, 1486-1492. ISSN : 1095-9203 (déc. 2019).
159. LEMAK, A., YEE, A., BEZSONOVA, I., DHE-PAGANON, S. & ARROWSMITH, C. H. Zn-binding AZUL domain of human ubiquitin protein ligase Ube3A. eng. *Journal of biomolecular NMR* **51**, 185-190. ISSN : 1573-5001 (sept. 2011).
160. BUEL, G. R. *et al.* Structure of E3 ligase E6AP with a proteasome-binding site provided by substrate receptor hRpn10. eng. *Nature Communications* **11**, 1291. ISSN : 2041-1723 (mars 2020).
161. LEE, S. Y. *et al.* Ube3a, the E3 ubiquitin ligase causing Angelman syndrome and linked to autism, regulates protein homeostasis through the proteasomal shuttle Rpn10. eng. *Cellular and molecular life sciences : CMLS* **71**, 2747-2758. ISSN : 1420-9071 (juill. 2014).
162. *Zn-binding AZUL domain of human ubiquitin protein ligase Ube3A - PubMed* <https://pubmed.ncbi.nlm.nih.gov/21947926/> (2023).
163. KLEIJNEN, M. F. *et al.* The hPLIC Proteins May Provide a Link between the Ubiquitination Machinery and the Proteasome. en. *Molecular Cell* **6**, 409-419. ISSN : 1097-2765. <https://www.sciencedirect.com/science/article/pii/S109727650000040X> (2023) (août 2000).
164. BUEL, G. R. *et al.* *Discovery of E6AP AZUL binding to UBQLN1/2 in cells, phase-separated droplets, and an AlphaFold-NMR integrated structure* en. preprint (Biophysics, sept. 2022). <http://biorxiv.org/lookup/doi/10.1101/2022.09.29.510132> (2023).
165. MARTÍNEZ-NOËL, G. *et al.* Network Analysis of UBE3A/E6AP-Associated Proteins Provides Connections to Several Distinct Cellular Processes. *Journal of Molecular Biology* **430**, 1024-1050. ISSN : 0022-2836. <https://www.sciencedirect.com/science/article/pii/S0022283618300470> (2018).
166. KO, H. S., UEHARA, T., TSURUMA, K. & NOMURA, Y. Ubiquilin interacts with ubiquitylated proteins and proteasome through its ubiquitin-associated and ubiquitin-like domains. eng. *FEBS letters* **566**, 110-114. ISSN : 0014-5793 (mai 2004).

- 
167. LIANG, R.-Y. *et al.* Rad23 Interaction with the Proteasome Is Regulated by Phosphorylation of Its Ubiquitin-Like (UbL) Domain. en. *Journal of Molecular Biology* **426**, 4049-4060. ISSN : 0022-2836. <https://www.sciencedirect.com/science/article/pii/S0022283614005269> (2023) (déc. 2014).
168. GOSSAN, N. C. *et al.* The E3 ubiquitin ligase UBE3A is an integral component of the molecular circadian clock through regulating the BMAL1 transcription factor. eng. *Nucleic Acids Research* **42**, 5765-5775. ISSN : 1362-4962 (mai 2014).
169. MISHRA, A., GODAVARTHI, S. K., MAHESHWARI, M., GOSWAMI, A. & JANA, N. R. The ubiquitin ligase E6-AP is induced and recruited to aggresomes in response to proteasome inhibition and may be involved in the ubiquitination of Hsp70-bound misfolded proteins. eng. *The Journal of Biological Chemistry* **284**, 10537-10545. ISSN : 0021-9258 (avr. 2009).
170. DHANANJAYAN, S. C. *et al.* WW domain binding protein-2, an E6-associated protein interacting protein, acts as a coactivator of estrogen and progesterone receptors. eng. *Molecular Endocrinology (Baltimore, Md.)* **20**, 2343-2354. ISSN : 0888-8809 (oct. 2006).
171. VATSA, N. & JANA, N. R. UBE3A and Its Link With Autism. *Frontiers in Molecular Neuroscience* **11**, 448. ISSN : 1662-5099. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6290346/> (2023) (déc. 2018).
172. RONCHI, V. P., KLEIN, J. M., EDWARDS, D. J. & HAAS, A. L. The active form of E6-associated protein (E6AP)/UBE3A ubiquitin ligase is an oligomer. eng. *The Journal of Biological Chemistry* **289**, 1033-1048. ISSN : 1083-351X (jan. 2014).
173. IKENBERG, H., GISSMANN, L., GROSS, G., GRUSSENDORF-CONEN, E.-I. & HAUSEN, H. Z. Human papillomavirus type-16-related DNA in genital Bowen's disease and in bowenoid papulosis. de. *International Journal of Cancer* **32**. \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ijc.2910320507>, 563-565. ISSN : 1097-0215. <https://onlinelibrary.wiley.com/doi/abs/10.1002/ijc.2910320507> (2023) (1983).
174. De MARTEL, C., PLUMMER, M., VIGNAT, J. & FRANCESCHI, S. Worldwide burden of cancer attributable to HPV by site, country and HPV type. eng. *International Journal of Cancer* **141**, 664-670. ISSN : 1097-0215 (août 2017).
175. SUAREZ, I. & TRAVE, G. Structural Insights in Multifunctional Papillomavirus Oncoproteins. *Viruses* **10**, 37. ISSN : 1999-4915. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5795450/> (2023) (jan. 2018).
176. BONHOURE, A. *Interférence de l'oncoprotéine E6 de HPV avec l'interactome humain : étude quantitative et structurale* These de doctorat (Strasbourg, oct. 2020). <https://www.theses.fr/2020STRAJ027> (2023).
177. SCHEFFNER, M., HUIBREGTSE, J. M., VIERSTRA, R. D. & HOWLEY, P. M. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination

- 
- of p53. en. *Cell* **75**, 495-505. ISSN : 0092-8674. <https://www.sciencedirect.com/science/article/pii/S0092867493903843> (2023) (nov. 1993).
178. CUBILLOS-ROJAS, M. *et al.* The E3 Ubiquitin Protein Ligase HERC2 Modulates the Activity of Tumor Protein p53 by Regulating Its Oligomerization \*. English. *Journal of Biological Chemistry* **289**. Publisher : Elsevier, 14782-14795. ISSN : 0021-9258, 1083-351X. [https://www.jbc.org/article/S0021-9258\(20\)38681-6/abstract](https://www.jbc.org/article/S0021-9258(20)38681-6/abstract) (2023) (mai 2014).
179. ZANIER, K. *et al.* Structural basis for hijacking of cellular LxxLL motifs by papillomavirus E6 oncoproteins. *Science (New York, N.Y.)* **339**, 694-698. ISSN : 0036-8075. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3899395/> (2023) (fév. 2013).
180. MARTINEZ-ZAPIEN, D. *et al.* Structure of the E6/E6AP/p53 complex required for HPV-mediated degradation of p53. en. *Nature* **529**. Number : 7587 Publisher : Nature Publishing Group, 541-545. ISSN : 1476-4687. <https://www.nature.com/articles/nature16481> (2023) (jan. 2016).
181. DREWS, C. M., BRIMER, N. & VANDE POL, S. B. Multiple regions of E6AP (UBE3A) contribute to interaction with papillomavirus E6 proteins and the activation of ubiquitin ligase activity. en. *PLOS Pathogens* **16** (éd. LAMBERT, P. F.) e1008295. ISSN : 1553-7374. <https://dx.plos.org/10.1371/journal.ppat.1008295> (2023) (jan. 2020).
182. MORTENSEN, F. *et al.* Role of ubiquitin and the HPV E6 oncoprotein in E6AP-mediated ubiquitination. eng. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 9872-9877. ISSN : 1091-6490 (août 2015).
183. KAO, W. H., BEAUDENON, S. L., TALIS, A. L., HUIBREGTSE, J. M. & HOWLEY, P. M. Human papillomavirus type 16 E6 induces self-ubiquitination of the E6AP ubiquitin-protein ligase. eng. *Journal of Virology* **74**, 6408-6417. ISSN : 0022-538X (juill. 2000).
184. GOGL, G. *et al.* Quantitative fragmentomics allow affinity mapping of interactomes. en. *Nature Communications* **13**, 5472. ISSN : 2041-1723. <https://www.nature.com/articles/s41467-022-33018-0> (2023) (sept. 2022).
185. LEHMAN, A. L. *et al.* A very large protein with diverse functional motifs is deficient in rjs (runty, jerky, sterile) mice. *Proceedings of the National Academy of Sciences* **95**. Publisher : Proceedings of the National Academy of Sciences, 9436-9441. <https://www.pnas.org/doi/full/10.1073/pnas.95.16.9436> (2023) (août 1998).
186. KAYSER, M. *et al.* Three Genome-wide Association Studies and a Linkage Analysis Identify HERC2 as a Human Iris Color Gene. en. *The American Journal of Human Genetics* **82**, 411-423. ISSN : 0002-9297. <https://www.sciencedirect.com/science/article/pii/S0002929708000748> (2023) (fév. 2008).

- 
187. MENGEL-FROM, J., BØRSTING, C., SANCHEZ, J. J., EIBERG, H. & MORLING, N. Human eye colour and HERC2, OCA2 and MATP. en. *Forensic Science International : Genetics* **4**, 323-328. ISSN : 1872-4973. <https://www.sciencedirect.com/science/article/pii/S1872497309002117> (2023) (oct. 2010).
188. DONNELLY, M. P. *et al.* A global view of the OCA2-HERC2 region and pigmentation. en. *Human Genetics* **131**, 683-696. ISSN : 1432-1203. <https://doi.org/10.1007/s00439-011-1110-x> (2023) (mai 2012).
189. BEKKER-JENSEN, S. *et al.* HERC2 coordinates ubiquitin-dependent assembly of DNA repair factors on damaged chromosomes. en. *Nature Cell Biology* **12**. Number : 1 Publisher : Nature Publishing Group, 80-86. ISSN : 1476-4679. <https://www.nature.com/articles/ncb2008> (2023) (jan. 2010).
190. MATHIEU, N. A., LEVIN, R. H. & SPRATT, D. E. Exploring the Roles of HERC2 and the NEDD4L HECT E3 Ubiquitin Ligase Subfamily in p53 Signaling and the DNA Damage Response. *Frontiers in Oncology* **11**. ISSN : 2234-943X. <https://www.frontiersin.org/articles/10.3389/fonc.2021.659049> (2023) (2021).
191. ZHANG, Z., YANG, H. & WANG, H. The Histone H2A Deubiquitinase USP16 Interacts with HERC2 and Fine-tunes Cellular Response to DNA Damage \*. English. *Journal of Biological Chemistry* **289**. Publisher : Elsevier, 32883-32894. ISSN : 0021-9258, 1083-351X. [https://www.jbc.org/article/S0021-9258\(20\)47487-3/abstract](https://www.jbc.org/article/S0021-9258(20)47487-3/abstract) (2023) (nov. 2014).
192. AL-HAKIM, A. K., BASHKUROV, M., GINGRAS, A.-C., DUROCHER, D. & PELLETIER, L. Interaction Proteomics Identify NEURL4 and the HECT E3 Ligase HERC2 as Novel Modulators of Centrosome Architecture\*. en. *Molecular & Cellular Proteomics* **11**, M111.014233. ISSN : 1535-9476. <https://www.sciencedirect.com/science/article/pii/S1535947620304199> (2023) (juin 2012).
193. IMAI, Y. *et al.* The Parkinson's Disease-Associated Protein Kinase LRRK2 Modulates Notch Signaling through the Endosomal Pathway. en. *PLOS Genetics* **11** (éd. LIM, K.-L.) e1005503. ISSN : 1553-7404. <https://dx.plos.org/10.1371/journal.pgen.1005503> (2023) (sept. 2015).
194. WEERSMA, R. K. *et al.* Confirmation of Multiple Crohn's Disease Susceptibility Loci in a Large Dutch-Belgian Cohort. en-US. *Official journal of the American College of Gastroenterology / ACG* **104**, 630. ISSN : 0002-9270. [https://journals.lww.com/ajg/Abstract/2009/03000/Confirmation\\_of\\_Multiple\\_Crohn\\_s\\_Disease.22.aspx](https://journals.lww.com/ajg/Abstract/2009/03000/Confirmation_of_Multiple_Crohn_s_Disease.22.aspx) (2023) (mars 2009).
195. WANG, K. *et al.* Comparative genetic analysis of inflammatory bowel disease and type 1 diabetes implicates multiple loci with opposite effects. *Human Molecular Genetics* **19**, 2059-2067. ISSN : 0964-6906. <https://doi.org/10.1093/hmg/ddq078> (2023) (mai 2010).

- 
196. VALNEGRI, P. *et al.* RNF8/UBC13 ubiquitin signaling suppresses synapse formation in the mammalian brain. en. *Nature Communications* **8**. Number : 1 Publisher : Nature Publishing Group, 1271. ISSN : 2041-1723. <https://www.nature.com/articles/s41467-017-01333-6> (2023) (nov. 2017).
197. CHEN, C. K.-M., CHAN, N.-L. & WANG, A. H.-J. The many blades of the  $\beta$ -propeller proteins : conserved but versatile. en. *Trends in Biochemical Sciences* **36**, 553-561. ISSN : 09680004. <https://linkinghub.elsevier.com/retrieve/pii/S0968000411001150> (2023) (oct. 2011).
198. MARGUERON, R. *et al.* Role of the polycomb protein EED in the propagation of repressive histone marks. eng. *Nature* **461**, 762-767. ISSN : 1476-4687 (oct. 2009).
199. SUGASE, K., DYSON, H. J. & WRIGHT, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* **447**, 1021-1025. ISSN : 0028-0836, 1476-4687. <http://www.nature.com/doifinder/10.1038/nature05858> (2018) (juill. 2007).
200. JENNINGS, B. H. *et al.* Molecular recognition of transcriptional repressor motifs by the WD domain of the Groucho/TLE corepressor. eng. *Molecular Cell* **22**, 645-655. ISSN : 1097-2765 (juin 2006).
201. COUTURE, J.-F., COLLAZO, E. & TRIEVEL, R. C. Molecular recognition of histone H3 by the WD40 protein WDR5. eng. *Nature Structural & Molecular Biology* **13**, 698-703. ISSN : 1545-9993 (août 2006).
202. OLIVER, A. W., SWIFT, S., LORD, C. J., ASHWORTH, A. & PEARL, L. H. Structural basis for recruitment of BRCA2 by PALB2. eng. *EMBO reports* **10**, 990-996. ISSN : 1469-3178 (sept. 2009).
203. SONG, J.-J., GARLICK, J. D. & KINGSTON, R. E. Structural basis of histone H4 recognition by p53. eng. *Genes & Development* **22**, 1313-1318. ISSN : 0890-9369 (mai 2008).
204. DAVIS, T. L., BONACCI, T. M., SPRANG, S. R. & SMRCKA, A. V. Structural and molecular characterization of a preferred protein interaction surface on G protein beta gamma subunits. eng. *Biochemistry* **44**, 10593-10604. ISSN : 0006-2960 (août 2005).
205. HUANG, Z. *et al.* Two FGF Receptor Kinase Molecules Act in Concert to Recruit and Transphosphorylate Phospholipase C $\gamma$ . en. *Molecular Cell* **61**, 98-110. ISSN : 10972765. <http://linkinghub.elsevier.com/retrieve/pii/S1097276515009028> (2017) (jan. 2016).
206. ORLICKY, S., TANG, X., WILLEMS, A., TYERS, M. & SICHERI, F. Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase. eng. *Cell* **112**, 243-256. ISSN : 0092-8674 (jan. 2003).

- 
207. WU, G. *et al.* Structure of a beta-TrCP1-Skp1-beta-catenin complex : destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase. eng. *Molecular Cell* **11**, 1445-1456. ISSN : 1097-2765 (juin 2003).
208. STIRNIMANN, C. U., PETSALAKI, E., RUSSELL, R. B. & MÜLLER, C. W. WD40 proteins propel cellular networks. *Trends in Biochemical Sciences* **35**, 565-574. ISSN : 0968-0004. <https://www.sciencedirect.com/science/article/pii/S0968000410000721> (2010).
209. GALLIGAN, J. T. *et al.* Proteomic Analysis and Identification of Cellular Interactors of the Giant Ubiquitin Ligase HERC2. en. *Journal of Proteome Research* **14**, 953-966. ISSN : 1535-3893, 1535-3907. <https://pubs.acs.org/doi/10.1021/pr501005v> (2023) (fév. 2015).
210. KÜHNLE, S. *et al.* Physical and Functional Interaction of the HECT Ubiquitin-protein Ligases E6AP and HERC2\*. en. *Journal of Biological Chemistry* **286**, 19410-19416. ISSN : 0021-9258. <https://www.sciencedirect.com/science/article/pii/S0021925820510181> (2023) (juin 2011).
211. WANG, T.-S., TSAI, W.-H., TSAI, L.-P. & WONG, S.-B. Clinical characteristics and epilepsy in genomic imprinting disorders : Angelman syndrome and Prader-Willi syndrome. en. *Tzu Chi Medical Journal* **32**, 137. ISSN : 1016-3190. <http://www.tcmjmed.com/text.asp?2020/32/2/137/270166> (2023) (2020).
212. REIS, A. *et al.* Imprinting mutations suggested by abnormal DNA methylation patterns in familial Angelman and Prader-Willi syndromes. eng. *American Journal of Human Genetics* **54**, 741-747. ISSN : 0002-9297 (mai 1994).
213. SAITOH, S. *et al.* Minimal definition of the imprinting center and fixation of chromosome 15q11-q13 epigenotype by imprinting mutations. eng. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 7811-7815. ISSN : 0027-8424 (juill. 1996).
214. BUTING, K. Prader-Willi syndrome and Angelman syndrome. en. *American Journal of Medical Genetics Part C : Seminars in Medical Genetics* **154C**. \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/ajmg.c.30273>, 365-376. ISSN : 1552-4876. <https://onlinelibrary.wiley.com/doi/abs/10.1002/ajmg.c.30273> (2023) (2010).
215. BUTING, K. *et al.* Epimutations in Prader-Willi and Angelman Syndromes : A Molecular Study of 136 Patients with an Imprinting Defect. *The American Journal of Human Genetics* **72**, 571-577. ISSN : 0002-9297. <https://www.sciencedirect.com/science/article/pii/S0002929707605734> (2003).
216. HORSTHEMKE, B. & WAGSTAFF, J. Mechanisms of imprinting of the Prader-Willi/Angelman region. en. *American Journal of Medical Genetics Part A* **146A**. \_eprint : <https://on->

- 
- linelibrary.wiley.com/doi/pdf/10.1002/ajmg.a.32364, 2041-2052. ISSN : 1552-4833. <https://onlinelibrary.wiley.com/doi/abs/10.1002/ajmg.a.32364> (2023) (2008).
217. JOSHI, R. S. *et al.* DNA Methylation Profiling of Uniparental Disomy Subjects Provides a Map of Parental Epigenetic Bias in the Human Genome. *The American Journal of Human Genetics* **99**, 555-566. ISSN : 0002-9297. <https://www.sciencedirect.com/science/article/pii/S0002929716302701> (2016).
218. MATSUBARA, K. *et al.* Exploring the unique function of imprinting control centers in the PWS/AS-responsible region : finding from array-based methylation analysis in cases with variously sized microdeletions. eng. *Clinical epigenetics* **11**, 36. ISSN : 1868-7083. <https://europepmc.org/articles/PMC6396496> (2023) (fév. 2019).
219. ROUGEULLE, C., CARDOSO, C., FONTÉS, M., COLLEAUX, L. & LALANDE, M. An imprinted antisense RNA overlaps UBE3A and a second maternally expressed transcript. en. *Nature Genetics* **19**. Number : 1 Publisher : Nature Publishing Group, 15-16. ISSN : 1546-1718. <https://www.nature.com/articles/ng0598-15> (2023) (mai 1998).
220. ANGELMAN, H. 'Puppet' Children A Report on Three Cases. en. *Developmental Medicine & Child Neurology* **7**. \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1469-8749.1965.tb07844.x>, 681-688. ISSN : 1469-8749. <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1469-8749.1965.tb07844.x> (2023) (1965).
221. BOWER, B. D. & JEAUVONS, P. M. The "happy puppet" syndrome. en. *Archives of Disease in Childhood* **42**, 298-302. ISSN : 0003-9888, 1468-2044. <https://adc.bmj.com/lookup/doi/10.1136/adc.42.223.298> (2023) (juin 1967).
222. MAGENIS, R. E. *et al.* Is angelman syndrome an alternate result of del(15)(qllq13)? en. *American Journal of Medical Genetics* **28**, 829-838. ISSN : 0148-7299, 1096-8628. <https://onlinelibrary.wiley.com/doi/10.1002/ajmg.1320280407> (2023) (déc. 1987).
223. KISHINO, T., LALANDE, M. & WAGSTAFF, J. UBE3A/E6-AP mutations cause Angelman syndrome. en. **15** (1997).
224. *Orphanet : Syndrome d'Angelman* [https://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?lng=FR&Expert=72](https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=FR&Expert=72) (2023).
225. ELGERSMA, Y. & SONZOGNI, M. UBE3A reinstatement as a disease-modifying therapy for Angelman syndrome. *Developmental Medicine and Child Neurology* **63**, 802-807. ISSN : 0012-1622. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8248324/> (2023) (juill. 2021).
226. PETERS, S. U., HOROWITZ, L., BARBIERI-WELGE, R., TAYLOR, J. L. & HUNDLEY, R. J. Longitudinal follow-up of autism spectrum features and sensory behaviors in Angelman syndrome by deletion class. eng. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* **53**, 152-159. ISSN : 1469-7610 (fév. 2012).

- 
227. BIRD, L. M. Angelman syndrome : review of clinical and molecular aspects. *The Application of Clinical Genetics* **7**, 93-104. ISSN : 1178-704X. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4036146/> (2023) (mai 2014).
228. GENTILE, J. K. *et al.* A neurodevelopmental survey of Angelman syndrome with genotype-phenotype correlations. eng. *Journal of developmental and behavioral pediatrics : JDBP* **31**, 592-601. ISSN : 1536-7312 (sept. 2010).
229. TAN, W.-H. *et al.* Angelman syndrome : Mutations influence features in early childhood. eng. *American Journal of Medical Genetics. Part A* **155A**, 81-90. ISSN : 1552-4833 (jan. 2011).
230. VALENTE, K. D. *et al.* Angelman syndrome caused by deletion : a genotype-phenotype correlation determined by breakpoint. eng. *Epilepsy Research* **105**, 234-239. ISSN : 1872-6844 (juill. 2013).
231. SILVA-SANTOS, S. *et al.* Ube3a reinstatement identifies distinct developmental windows in a murine Angelman syndrome model. *The Journal of Clinical Investigation* **125**, 2069-2076. ISSN : 0021-9738. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4463212/> (2023) (mai 2015).
232. DAILY, J. L. *et al.* Adeno-associated virus-mediated rescue of the cognitive defects in a mouse model for Angelman syndrome. eng. *PloS One* **6**, e27221. ISSN : 1932-6203 (2011).
233. ADHIKARI, A. *et al.* Functional rescue in an Angelman syndrome model following treatment with lentivector transduced hematopoietic stem cells. eng. *Human Molecular Genetics* **30**, 1067-1083. ISSN : 1460-2083 (juin 2021).
234. DODGE, A. *et al.* Identification of UBE3A Protein in CSF and Extracellular Space of the Hippocampus Suggest a Potential Novel Function in Synaptic Plasticity. eng. *Autism Research : Official Journal of the International Society for Autism Research* **14**, 645-655. ISSN : 1939-3806 (avr. 2021).
235. MARKATI, T., DUIS, J. & SERVAIS, L. Therapies in preclinical and clinical development for Angelman syndrome. *Expert Opinion on Investigational Drugs* **30**. Publisher : Taylor & Francis, 709-720. ISSN : 1354-3784. <https://www.tandfonline.com/doi/ref/10.1080/13543784.2021.1939674> (2023) (juill. 2021).
236. MENG, L. *et al.* Towards a therapy for Angelman syndrome by targeting a long non-coding RNA. eng. *Nature* **518**, 409-412. ISSN : 1476-4687 (fév. 2015).
237. WOLTER, J. M. *et al.* Cas9 gene therapy for Angelman syndrome traps Ube3a-ATS long non-coding RNA. eng. *Nature* **587**, 281-284. ISSN : 1476-4687 (nov. 2020).
238. BAILUS, B. J. *et al.* Protein Delivery of an Artificial Transcription Factor Restores Widespread Ube3a Expression in an Angelman Syndrome Mouse Brain. *Molecular Therapy* **24**, 548-555. ISSN : 1525-0016. <https://www.sciencedirect.com/science/article/pii/S1525001616309728> (2016).

- 
239. HUANG, H.-S. *et al.* Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. eng. *Nature* **481**, 185-189. ISSN : 1476-4687 (déc. 2011).
240. EGAWA, K. *et al.* Decreased tonic inhibition in cerebellar granule cells causes motor dysfunction in a mouse model of Angelman syndrome. eng. *Science Translational Medicine* **4**, 163ra157. ISSN : 1946-6242 (déc. 2012).
241. CRUZ, E. *et al.* CIM6P/IGF-2 Receptor Ligands Reverse Deficits in Angelman Syndrome Model Mice. eng. *Autism Research : Official Journal of the International Society for Autism Research* **14**, 29-45. ISSN : 1939-3806 (jan. 2021).
242. WANG, J. *et al.* UBE3A-mediated PTPA ubiquitination and degradation regulate PP2A activity and dendritic spine morphology. eng. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 12500-12505. ISSN : 1091-6490 (juin 2019).
243. SUN, J. *et al.* PKA and Ube3a regulate SK2 channel trafficking to promote synaptic plasticity in hippocampus : Implications for Angelman Syndrome. en. *Scientific Reports* **10**. Number : 1 Publisher : Nature Publishing Group, 9824. ISSN : 2045-2322. <https://www.nature.com/articles/s41598-020-66790-4> (2023) (juin 2020).
244. HERBER, D. L., WEEBER, E. J., D'AGOSTINO, D. P. & DUIS, J. Evaluation of the safety and tolerability of a nutritional Formulation in patients with ANgelman Syndrome (FANS) : study protocol for a randomized controlled trial. *Trials* **21**, 60. ISSN : 1745-6215. <https://doi.org/10.1186/s13063-019-3996-x> (2023) (jan. 2020).
245. PUFFENBERGER, E. G. *et al.* A homozygous missense mutation in HERC2 associated with global developmental delay and autism spectrum disorder. en. *Human Mutation* **33**. \_eprint : <https://onlinelibrary.wiley.com/doi/pdf/10.1002/humu.22237>, 1639-1646. ISSN : 1098-1004. <https://onlinelibrary.wiley.com/doi/abs/10.1002/humu.22237> (2023) (2012).
246. HARLALKA, G. V. *et al.* Mutation of HERC2 causes developmental delay with Angelman-like features. eng. *Journal of Medical Genetics* **50**, 65-73. ISSN : 1468-6244 (fév. 2013).
247. MORICE-PICARD, F. *et al.* Complete loss of function of the ubiquitin ligase HERC2 causes a severe neurodevelopmental phenotype. en. *European Journal of Human Genetics* **25**. Number : 1 Publisher : Nature Publishing Group, 52-58. ISSN : 1476-5438. <https://www.nature.com/articles/ejhg2016139> (2023) (jan. 2017).
248. ELPIDOROU, M. *et al.* Novel loss-of-function mutation in HERC2 is associated with severe developmental delay and paediatric lethality. eng. *Journal of Medical Genetics* **58**, 334-341. ISSN : 1468-6244 (mai 2021).
249. NGUYEN, L. S. *et al.* A nonsense variant in HERC1 is associated with intellectual disability, megalencephaly, thick corpus callosum and cerebellar atrophy. en. *European Journal of Human Genetics* **24**, 455-458. ISSN : 1018-4813, 1476-5438. <http://www.nature.com/articles/ejhg2015140> (2023) (mars 2016).

- 
250. TAN, W.-H., BIRD, L. M., THIBERT, R. L. & WILLIAMS, C. A. If not Angelman, what is it? a review of Angelman-like syndromes. en. *American Journal of Medical Genetics Part A* **164**, 975-992. ISSN : 15524825. <https://onlinelibrary.wiley.com/doi/10.1002/ajmg.a.36416> (2023) (avr. 2014).
251. PRADER, A. Ein Syndrom von Adipositas, Kleinwuchs, Kryptorchismus und Oligophrenie nach myatonieartigem Zustand im Neugeborenenalter. *Schweizerische Medizinische Wochenschrift*. <https://www.semanticscholar.org/paper/Ein-Syndrom-von-Adipositas%2C-Kleinwuchs%2C-und-nach-im-Prader/5ab1fb5919e45186c3b4d8316fd58bbfd8> (2023) (1956).
252. *Orphanet : Syndrome de Prader Willi* [https://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?lng=fr&Expert=739](https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=fr&Expert=739) (2023).
253. CASSIDY, S. B., SCHWARTZ, S., MILLER, J. L. & DRISCOLL, D. J. Prader-Willi syndrome. en. *Genetics in Medicine* **14**, 10-26. ISSN : 10983600. <https://linkinghub.elsevier.com/retrieve/pii/S1098360021032846> (2023) (jan. 2012).
254. COOK, E. H. *et al.* Autism or atypical autism in maternally but not paternally derived proximal 15q duplication. *American Journal of Human Genetics* **60**, 928-934. ISSN : 0002-9297. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1712464/> (2023) (avr. 1997).
255. BATTAGLIA, A. The inv dup (15) or idic (15) syndrome (Tetrasomy 15q). en. *Orphanet Journal of Rare Diseases* **3**, 30. ISSN : 1750-1172. <https://ojrd.biomedcentral.com/articles/10.1186/1750-1172-3-30> (2023) (déc. 2008).
256. *Orphanet : Syndrome de duplication inversée du chromosome 15* [https://www.orpha.net/consor/cgi-bin/OC\\_Exp.php?lng=FR&Expert=3306](https://www.orpha.net/consor/cgi-bin/OC_Exp.php?lng=FR&Expert=3306) (2023).
257. HOGART, A., WU, D., LASALLE, J. M. & SCHANEN, N. C. The Comorbidity of Autism with the Genomic Disorders of Chromosome 15q11.2-q13. *Neurobiology of disease* **38**, 181-191. ISSN : 0969-9961. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2884398/> (2023) (mai 2010).
258. MAO, R. *et al.* Characteristics of two cases with dup(15) (q 11.2-q 12) : one of maternal and one of paternal origin. en. *Genetics in Medicine* **2**. Number : 2 Publisher : Nature Publishing Group, 131-135. ISSN : 1530-0366. <https://www.nature.com/articles/gim2000234> (2023) (mars 2000).
259. BUITING, K., WILLIAMS, C. & HORSTHEMKE, B. Angelman syndrome - insights into a rare neurogenetic disorder. eng. *Nature Reviews. Neurology* **12**, 584-593. ISSN : 1759-4766 (oct. 2016).
260. SATO, M. Early Origin and Evolution of the Angelman Syndrome Ubiquitin Ligase Gene Ube3a. *Frontiers in Cellular Neuroscience* **11**. ISSN : 1662-5102. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5339648/> (2023) (mars 2017).

- 
261. VOS, R. M., ALTREUTER, J., WHITE, E. A. & HOWLEY, P. M. The ubiquitin-specific peptidase USP15 regulates human papillomavirus type 16 E6 protein stability. eng. *Journal of Virology* **83**, 8885-8892. ISSN : 1098-5514 (sept. 2009).
262. URRACA, N. *et al.* Significant transcriptional changes in 15q duplication but not Angelman syndrome deletion stem cell-derived neurons. eng. *Molecular Autism* **9**, 6. ISSN : 2040-2392 (2018).
263. KANG, T.-H., REARDON, J. T. & SANCAR, A. Regulation of nucleotide excision repair activity by transcriptional and post-transcriptional control of the XPA protein. *Nucleic Acids Research* **39**, 3176-3187. ISSN : 0305-1048. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3082913/> (2023) (avr. 2011).
264. RENAULT, L. *et al.* The 1.7 Å crystal structure of the regulator of chromosome condensation (RCC1) reveals a seven-bladed propeller. eng. *Nature* **392**, 97-101. ISSN : 0028-0836 (mars 1998).
265. PEI, J. & GRISHIN, N. V. AL2CO : calculation of positional conservation in a protein sequence alignment. eng. *Bioinformatics (Oxford, England)* **17**, 700-712. ISSN : 1367-4803 (août 2001).
266. JAUDON, F. *et al.* The RhoGEF DOCK10 is essential for dendritic spine morphogenesis. *Molecular Biology of the Cell* **26**, 2112-2127. ISSN : 1059-1524. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4472020/> (2023) (juin 2015).
267. HORI, A. & TODA, T. Regulation of centriolar satellite integrity and its physiology. eng. *Cellular and molecular life sciences : CMLS* **74**, 213-229. ISSN : 1420-9071 (jan. 2017).
268. WU, M.-Y., TSAI, T.-F. & BEAUDET, A. L. Deficiency of Rbbp1/Arid4a and Rbbp1l1/Arid4b alters epigenetic modifications and suppresses an imprinting defect in the PWS/AS domain. *Genes & Development* **20**, 2859-2870. ISSN : 0890-9369. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1619944/> (2023) (oct. 2006).
269. GOGL, G. *et al.* Hierarchized phosphotarget binding by the seven human 14-3-3 isoforms. en. *Nature Communications* **12**. Number : 1 Publisher : Nature Publishing Group, 1677. ISSN : 2041-1723. <https://www.nature.com/articles/s41467-021-21908-8> (2023) (mars 2021).
270. GUPTA, G. D. *et al.* A Dynamic Protein Interaction Landscape of the Human Centrosome-Cilium Interface. eng. *Cell* **163**, 1484-1499. ISSN : 1097-4172 (déc. 2015).
271. GE, X., FRANK, C. L., CALDERON DE ANDA, F. & TSAI, L.-H. Hook3 interacts with PCM1 to regulate pericentriolar material assembly and the timing of neurogenesis. eng. *Neuron* **65**, 191-203. ISSN : 1097-4199 (jan. 2010).
272. MONROE, T. O. *et al.* PCM1 is necessary for focal ciliary integrity and is a candidate for severe schizophrenia. en. *Nature Communications* **11**. Number : 1 Publisher : Na-

- 
- ture Publishing Group, 5903. ISSN : 2041-1723. <https://www.nature.com/articles/s41467-020-19637-5> (2023) (nov. 2020).
273. ZOUBOVSKY, S. *et al.* Neuroanatomical and behavioral deficits in mice haploinsufficient for Pericentriolar material 1 (Pcm1). en. *Neuroscience Research* **98**, 45-49. ISSN : 0168-0102. <https://www.sciencedirect.com/science/article/pii/S016801021500036X> (2023) (sept. 2015).
274. GONG, W. *et al.* Structural insight into recognition of methylated histone tails by retinoblastoma-binding protein 1. eng. *The Journal of Biological Chemistry* **287**, 8531-8540. ISSN : 1083-351X (mars 2012).
275. WU, R.-C., JIANG, M., BEAUDET, A. L. & WU, M.-Y. ARID4A and ARID4B regulate male fertility, a functional link to the AR and RB pathways. *Proceedings of the National Academy of Sciences* **110**. Publisher : Proceedings of the National Academy of Sciences, 4616-4621. <https://www.pnas.org/doi/full/10.1073/pnas.1218318110> (2023) (mars 2013).
276. IOSSIFOV, I. *et al.* The contribution of de novo coding mutations to autism spectrum disorder. eng. *Nature* **515**, 216-221. ISSN : 1476-4687 (nov. 2014).
277. KRUPP, D. R. *et al.* Exonic Mosaic Mutations Contribute Risk for Autism Spectrum Disorder. eng. *American Journal of Human Genetics* **101**, 369-390. ISSN : 1537-6605 (sept. 2017).
278. SCOTT, T. M. *et al.* BAZ2B haploinsufficiency as a cause of developmental delay, intellectual disability, and autism spectrum disorder. eng. *Human Mutation* **41**, 921-925. ISSN : 1098-1004 (mai 2020).
279. DOMANSKYI, A. *et al.* Expression and localization of androgen receptor-interacting protein-4 in the testis. eng. *American Journal of Physiology. Endocrinology and Metabolism* **292**, E513-522. ISSN : 0193-1849 (fév. 2007).
280. SITZ, J. H., TIGGES, M., BAUMGÄRTEL, K., KHASPEKOV, L. G. & LUTZ, B. Dyrk1A Potentiates Steroid Hormone-Induced Transcription via the Chromatin Remodeling Factor Arip4. *Molecular and Cellular Biology* **24**, 5821-5834. ISSN : 0270-7306. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC480880/> (2023) (juill. 2004).
281. MENON, V. R. *et al.* DYRK1A regulates the recruitment of 53BP1 to the sites of DNA damage in part through interaction with RNF169. eng. *Cell Cycle (Georgetown, Tex.)* **18**, 531-551. ISSN : 1551-4005 (mars 2019).
282. HUTTLIN, E. L. *et al.* Dual proteome-scale networks reveal cell-specific remodeling of the human interactome. eng. *Cell* **184**, 3022-3040.e28. ISSN : 1097-4172 (mai 2021).
283. HIRAOKA, D., HOSODA, E., CHIBA, K. & KISHIMOTO, T. SGK phosphorylates Cdc25 and Myt1 to trigger cyclin B-Cdk1 activation at the meiotic G2/M transition. *The Journal of Cell Biology* **218**, 3597-3611. ISSN : 0021-9525. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6829662/> (2023) (nov. 2019).

- 
284. VASCONCELOS, F. F. *et al.* MyT1 Counteracts the Neural Progenitor Program to Promote Vertebrate Neurogenesis. eng. *Cell Reports* **17**, 469-483. ISSN : 2211-1247 (oct. 2016).
285. YOKOYAMA, A. *et al.* Identification of myelin transcription factor 1 (MyT1) as a subunit of the neural cell type-specific lysine-specific demethylase 1 (LSD1) complex. eng. *The Journal of biological chemistry* **289**, 18152-18162. ISSN : 1083-351X. <https://europepmc.org/articles/PMC4140267> (2023) (juin 2014).
286. CHEN, K. *et al.* MYT1 attenuates neuroblastoma cell differentiation by interacting with the LSD1/CoREST complex. eng. *Oncogene* **39**, 4212-4226. ISSN : 1476-5594 (mai 2020).
287. LIU, J. *et al.* Structural Insight into Binding of the ZZ Domain of HERC2 to Histone H3 and SUMO1. en. *Structure* **28**, 1225-1230.e3. ISSN : 0969-2126. <https://www.sciencedirect.com/science/article/pii/S0969212620302367> (2023) (nov. 2020).
288. JI, Y. *et al.* The ancestral gene for transcribed, low-copy repeats in the Prader-Willi/Angelman region encodes a large protein implicated in protein trafficking, which is deficient in mice with neuromuscular and spermiogenic abnormalities. eng. *Human Molecular Genetics* **8**, 533-542. ISSN : 0964-6906 (mars 1999).
289. CAO, J. *et al.* USP35, regulated by estrogen and AKT, promotes breast tumorigenesis by stabilizing and enhancing transcriptional activity of estrogen receptor  $\alpha$ . *Cell Death & Disease* **12**, 619. ISSN : 2041-4889. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8206120/> (2023) (juin 2021).
290. TANG, Z. *et al.* Deubiquitinase USP35 modulates ferroptosis in lung cancer via targeting ferroportin. *Clinical and Translational Medicine* **11**, e390. ISSN : 2001-1326. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8087931/> (2023) (mai 2021).
291. ZHANG, J. *et al.* Deubiquitinase USP35 restrains STING-mediated interferon signaling in ovarian cancer. *Cell Death and Differentiation* **28**, 139-155. ISSN : 1350-9047. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7853139/> (2023) (jan. 2021).
292. ESPOSITO, M. *et al.* USP13 controls the stability of Aurora B impacting progression through the cell cycle. *ONCOGENE* **39**, 6009-6023. ISSN : 0950-9232 (sept. 2020).
293. PARK, J., KWON, M.-S., KIM, E. E., LEE, H. & SONG, E. J. USP35 regulates mitotic progression by modulating the stability of Aurora B. eng. *Nature Communications* **9**, 688. ISSN : 2041-1723 (fév. 2018).
294. TSAI, I.-C. *et al.* Genome-wide suppressor screen identifies USP35/USP38 as therapeutic candidates for ciliopathies. eng. *JCI insight* **4**, e130516, 130516. ISSN : 2379-3708 (nov. 2019).
295. PRILUSKY, J. *et al.* FoldIndex© : a simple tool to predict whether a given protein sequence is intrinsically unfolded. *Bioinformatics* **21**, 3435-3438. ISSN : 1367-4803. <https://doi.org/10.1093/bioinformatics/bti537> (2023) (août 2005).

- 
296. JUMPER, J. *et al.* Highly accurate protein structure prediction with AlphaFold. en. *Nature* **596**. Number : 7873 Publisher : Nature Publishing Group, 583-589. ISSN : 1476-4687. <https://www.nature.com/articles/s41586-021-03819-2> (2023) (août 2021).
297. ZEMLA, A. LGA : a method for finding 3D similarities in protein structures. *Nucleic Acids Research* **31**, 3370-3374. ISSN : 0305-1048. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC168977/> (2023) (juill. 2003).
298. *AlphaFold : a solution to a 50-year-old grand challenge in biology* en. <https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology> (2023).
299. RODRIGUEZ, M. S., DESTERRO, J. M., LAIN, S., LANE, D. P. & HAY, R. T. Multiple C-terminal lysine residues target p53 for ubiquitin-proteasome-mediated degradation. eng. *Molecular and Cellular Biology* **20**, 8458-8467. ISSN : 0270-7306 (nov. 2000).
300. LOHRUM, M. A. E., WOODS, D. B., LUDWIG, R. L., BÁLINT, É. & VOUSDEN, K. H. C-Terminal Ubiquitination of p53 Contributes to Nuclear Export. *Molecular and Cellular Biology* **21**, 8521-8532. ISSN : 0270-7306. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC100015/> (2023) (déc. 2001).
301. SAILER, C. *et al.* Structural dynamics of the E6AP/UBE3A-E6-p53 enzyme-substrate complex. en. *Nature Communications* **9**, 4441. ISSN : 2041-1723. <https://www.nature.com/articles/s41467-018-06953-0> (2023) (oct. 2018).
302. KANE, E. I. *et al.* Redefining the catalytic HECT domain boundaries for the HECT E3 ubiquitin ligase family. *Bioscience Reports* **42**, BSR20221036. ISSN : 0144-8463. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9547173/> (2023) (oct. 2022).
303. MASCARENHAS, N. M. & GOSAVI, S. Understanding protein domain-swapping using structure-based models of protein folding. *Progress in Biophysics and Molecular Biology* **128**, 113-120. ISSN : 0079-6107. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7127520/> (2023) (sept. 2017).
304. PANDYA, R. K., PARTRIDGE, J. R., LOVE, K. R., SCHWARTZ, T. U. & PLOEGH, H. L. A Structural Element within the HUWE1 HECT Domain Modulates Self-ubiquitination and Substrate Ubiquitination Activities. *The Journal of Biological Chemistry* **285**, 5664-5673. ISSN : 0021-9258. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2820794/> (2023) (fév. 2010).
305. O, S. *ColabFold - v1.5.1* original-date : 2021-07-19T23 :24 :37Z. Fév. 2023. <https://github.com/sokrypton/ColabFold> (2023).
306. *Clustal Omega < Multiple Sequence Alignment < EMBL-EBI* <https://www.ebi.ac.uk/Tools/msa/clustalo/> (2023).

- 
307. GEER, L. Y., DOMRACHEV, M., LIPMAN, D. J. & BRYANT, S. H. CDART : protein homology by domain architecture. eng. *Genome Research* **12**, 1619-1623. ISSN : 1088-9051 (oct. 2002).
308. *T-COFFEE Multiple Sequence Alignment Server* <https://tcoffee.crg.eu/> (2023).
309. *Calculate accessible surface area and accessibility of protein* <http://cib.cf.ocha.ac.jp/bitool/ASA/> (2023).
310. CHARBONNIER, S., ZANIER, K., MASSON, M. & TRAVÉ, G. Capturing protein–protein complexes at equilibrium : The holdup comparative chromatographic retention assay. en. *Protein Expression and Purification* **50**, 89-101. ISSN : 1046-5928. <https://www.sciencedirect.com/science/article/pii/S1046592806001859> (2023) (nov. 2006).
311. VINCENTELLI, R. *et al.* Quantifying domain–ligand affinities and specificities by high-throughput holdup assay. eng. *Nature Methods* **12**, 787-793. ISSN : 1548-7105 (août 2015).



## Auguste Demenge

# Analyse structurale et interactomique des ubiquitine-ligases E3, Ube3A et HERC2, impliquées dans le syndrome d'Angelman.

---

### Résumé :

Le syndrome d'Angelman ainsi que le syndrome du spectre autistique Dup15q sont causés par des dérégulations d'origine génétique de l'expression ou de l'activité de la protéine "E3 ubiquitine-ligase" Ube3a (E6AP). Ube3a lie et est modulée par une autre E3 ubiquitine-ligase, HERC2, elle-même mutée dans une maladie rare, proche du syndrome d'Angelman. Nous avons identifié un nouveau réseau d'interactions qui pourrait contribuer aux altérations du développement neurologique dans ces syndromes. L'étude a révélé que le domaine RLD2 de HERC2 interagit avec un motif DxDKDxD présent dans plusieurs protéines. Nous avons résolu les structures cristallographiques du domaine RLD2 de HERC2 en complexe avec les peptides du motif DxDKDxD de différentes protéines. Nous avons aussi obtenu des informations structurales de la protéine Ube3a grâce à la purification de différentes constructions de cette protéine et à l'utilisation de prédictions de structure. Ces résultats participent à une meilleure compréhension des mécanismes mis en cause dans ces syndromes.

**Mots-clés :** Angelman, Angelman-like, dup15q, Ube3a, E6AP, HERC2, Ubiquitine-ligase, Neurodéveloppement, Système ubiquitine-protéasome.

---

### Summary :

Angelman syndrome as well as the autism spectrum syndrome Dup15q are caused by genetically based deregulations of the expression or activity of the E3 ubiquitin ligase protein Ube3a (E6AP). Ube3a binds and is modulated by another E3 ubiquitin ligase, HERC2, which is itself mutated in a rare disease similar to Angelman syndrome. We have identified a novel network of interactions that may contribute to the neurodevelopmental alterations in these syndromes. The study revealed that the RLD2 domain of HERC2 interacts with a DxDKDxD motif present in several proteins. We solved the crystallographic structures of the RLD2 domain of HERC2 in complex with the DxDKDxD motif peptides of different proteins. We also obtained structural information of the Ube3a protein by purifying different constructs of this protein and using structure predictions. These results contribute to a better understanding of the mechanisms involved in these syndromes.

**Keywords :** Angelman, Angelman-like, dup15q, Ube3a, E6AP, HERC2, Ubiquitin ligase, Neurodevelopment, Ubiquitin-proteasome system