



Université de Montréal

**Étude structurale conformationnelle des toxines de l'anthrax par  
cryo-microscopie et dynamique moléculaire**

par  
Lucien Fabre

Département de Biochimie  
Faculté de Médecine

Thèse de maîtrise à la Faculté des études supérieures  
en vue de l'obtention du grade de master en Bioinformatique

Juin, 2011

© Lucien Fabre, 2011.

Université de Montréal  
Faculté des études supérieures

Cette thèse de maîtrise est intitulée :

**Étude structurale conformationnelle des toxines de l'anthrax par  
cryo-microscopie et dynamique moléculaire**

présentée par :

Lucien Fabre

a été évaluée par un jury composé des

Gertraud Burger  
Francois Major  
Isabelle Rouiller  
Nathalie Lamarche

Président  
Directeur de recherche  
Codirecteur  
Examineur externe

Thèse de maîtrise acceptée le : .....

*À mes parents que j'aime énormément*

*À mon frère et toute ma famille*

*À mes professeurs et collègues*

*À tous mes amis*

# Remerciements

*Cette thèse de maîtrise est une expérience riche et passionnante, mais aussi difficile et très prenante. Je ne regretterai jamais de l'avoir tentée, mais je suis aussi, désormais, impatient d'arriver à ce qui constituera un aboutissement.*

*C'est aussi une œuvre collective qui comprend bien sûr le principal intéressé mais aussi tous ceux qui l'entourent. Les quelques mots qui suivent sont les plus attendus et les plus lus, ils représentent l'attachement qu'on porte à tous ceux qui nous entourent. Ces quelques mots sont aussi l'exercice le plus périlleux car il ne faut oublier personne.*

*Mes vifs remerciements s'adressent au Pr. **Gertraud Burger**, Responsable du programme de Bio-informatique et qui a bien voulu accepter de présider le jury de cette thèse de maîtrise. Qu'elle trouve ici l'expression de ma profonde gratitude.*

*Je tiens également à assurer de ma reconnaissance à l'égard de mon directeur de thèse, le Pr. **François Major**, pour m'avoir accompagné et soutenu lors de mes travaux de recherche. Qu'il veuille trouver ici l'expression de mon profond respect et ma haute considération.*

*Je ne saurais exprimer toute ma gratitude et ma profonde reconnaissance à ma codirectrice le Pr. **Isabelle Rouiller** qui m'a chaleureusement accueilli dans son équipe de recherche. Son dynamisme, sa rigueur ainsi que ses qualités humaines et scientifiques suscitent le plus grand respect. Je la remercie pour son aide précieuse, ses conseils et ses encouragements.*

*Je suis très honoré de la participation à ce jury de thèse du Pr. **Nathalie Lamarche**. Je lui adresse un remerciement particulier pour avoir accepté de juger cette thèse.*

*J'ai une pensée particulière à Monsieur **Martin Fleming** pour son aide précieuse concernant la logistique informatique. Qu'il trouve ici l'expression de mon profond respect et ma haute considération. Qu'il trouve ici l'expression de ma profonde gratitude.*

*Mes remerciements vont également à Monsieur **Jean-François Saint-Pierre** pour son aide précieuse concernant la partie Dynamique Moléculaire. Je le remercie vivement pour tout. Qu'il trouve ici l'expression de mon profond respect et ma haute considération.*

*Je ne saurais oublier de remercier tous mes amis et collègues chercheurs. La liste est longue ! Je cite particulièrement **Dr. Driss Mountassif, Dr Mihnea Botsina et Dr. Kaustuv Basu**. Je ne saurais leur exprimer mes remerciements pour leur assistance permanente ainsi que pour l'aide et le soutien qu'ils m'ont apportés tout au long de la réalisation de ce manuscrit. Je les remercie aussi pour leurs encouragements, leur gentillesse et l'excellente ambiance qu'ils ont su créer durant toutes ces années de recherche.*

*J'ai une pensée particulière à **Elaine Meunier et Sandra Botbol**. Je les remercie infiniment pour leur aide et leurs encouragements. Qu'ils trouvent ici l'expression de ma profonde gratitude.*

*Et pour finir, mais j'aurais pu commencer par eux, je suis extrêmement reconnaissant envers **mes parents, mon frère et toute ma famille** pour m'avoir donné tous les éléments qui m'ont permis d'arriver jusqu'ici. Ils m'ont apporté plus que quiconque. Ils sont ma force, ma confiance et ma richesse. Jamais je n'aurai réussi sans eux.*

*Je remercie enfin tous ceux que je n'ai pas nommés et qui ont participé de près ou de loin à l'aboutissement de ce travail.*

*Merci à tous.*

# Résumé

Les **toxines de l'anthrax** font partie de la famille des toxines A-B dans laquelle la moitié B se fixe à la membrane de la cellule permettant par la suite la translocation de la moitié A. Dans le cas de l'anthrax, la moitié B est représentée par le **Protective Antigen** (PA) et la moitié A par les deux protéines Edema Factor (EF) et **Lethal Factor** (LF). Après le recrutement par les récepteurs cellulaires (CMG2 et TEM8), PA s'organise en heptamère. Il peut fixer jusqu'à 3 ligands (EF et LF) avant d'être endocyté. Les modèles actuels de PA suggèrent que la baisse de pH à l'intérieur des endosomes permet un changement de **conformation** de la forme pré-pore vers la forme pore et que les ligands EF et LF passeraient au travers le pore pour entrer dans le cytoplasme. Cependant, le diamètre du pore est environ dix fois inférieur à celui des ligands (10 Å contre 100 Å). Un processus de folding/unfolding a été proposé mais demeure controversé.

Afin d'identifier le processus de passage des facteurs EF et LF dans le cytoplasme, nous avons déterminé par **cryo-microscopie électronique** combinée avec l'analyse d'image les structures tridimensionnelles des complexes formés par PA et LF aux étapes pré-pore et pore. Par la suite, une étude complémentaire par **dynamique moléculaire** nous a permis de modéliser à haute résolution les différentes interactions qui ont lieu au sein du complexe.

La **structure 3D** du complexe pré-pore combiné à 3 LF a été déterminée à une résolution de 14 Å. Nous avons aussi calculé une structure préliminaire du complexe pore également combiné à 3 LF. Celles-ci n'ont jamais été résolues auparavant et leur connaissance permet d'envisager l'étude en profondeur du mécanisme infectieux de l'Anthrax *in vivo*.

---

**Mots clés :** Toxines de l'Anthrax, Protective Antigen, Lethal Factor, Cryo-EM, conformations, structure 3D, dynamique moléculaire.

# *Abstract*

The anthrax toxins are part of the A-B toxin family in which the B moiety binds to the cell membrane allowing subsequent translocation of the A moiety. In the case of anthrax, the B moiety consists of the **Protective Antigen (PA)**, and the A moiety is composed of the two proteins Edema Factor (EF) and the **Lethal Factor (LF)**. After being recruited by the cell receptors (CGM2 or TEM8), PA organizes itself into a heptamer. It can bind up to three ligands (either EF or LF) before being endocytosed. Current models suggest that the decrease of pH inside the endosomes allows a conformational change of PA from a prepore form to a pore form that allows the EF and LF ligands to pass through the pore and enter the cytoplasm. However, the pore diameter is about ten times smaller than the diameter of the ligands (10Å versus 100Å). A process of ligand folding / unfolding has been proposed, but remains controversial.

To identify the mechanism by which EF and LF enter the cytoplasm, we have used **cryo-electron microscopy** and three-dimensional image analysis to determine the **3D structure** of the PA-LF complexes in the pre-pore and pore conformations. Then, we used molecular dynamics to model at high resolution the different interactions that occur within the complex.

The 3D structure of the pre-pore complex bound with three LF ligands has been determined at 14Å resolution. We also calculated a preliminary structure of the LF-bound pore complex. These structures have never been reported before. They provide the necessary information to study in depth the mechanism of anthrax infection *in vivo*.

---

**Keywords :** Anthrax toxins, Protective Antigen, Lethal Factor, Cryo-EM, Molecular Dynamics, 3D structure.



# Table des Matières

## Chapitre 1 : Introduction

I) <i>Bacillus anthracis</i> .....	1
1) Généralités.....	1
2) L'Infection par <i>Bacillus anthracis</i> .....	2
a) Chez l'animal.....	2
b) Chez l'homme.....	2
3) Les toxines de l'anthrax.....	3
a) Caractéristiques.....	3
b) Le mécanisme infectieux de l'anthrax.....	4
c) Présentation des structures résolues des toxines de l'anthrax.....	5
(i) Les récepteurs cellulaires à l'anthrax.....	5
(ii) Le Protective Antigen.....	7
(iii) Le Lethal Factor.....	8
(iv) L'Edema Factor.....	8
(v) Conversion de la forme pré-pore à la forme pore de PA63h.....	9
(vi) Initiation du dépliement de LF au travers du PA.....	11
(vii) Liposomes et Nanodisques.....	12
• Les liposomes.....	13
• Les nanodisques.....	14
II) Cryo-microscopie et reconstruction 3D par particules isolées.....	15
1) Introduction.....	15
2) Transformée de Fourier et son importance dans la reconstruction 3D.....	16
3) Défocus et estimation de la CTF.....	16
4) Alignements.....	17
5) Création d'un volume de référence.....	18
6) Classification 3D.....	19
7) Classification 3D pour le démêlage des mésalignements.....	19
III) Problématique.....	19
IV) Objectifs.....	20

## Chapitre 2 : Matériels et Méthodes

I) Techniques de biologie moléculaire.....	21
A) Amplification par PCR.....	21
B) Analyse de la PCR sur gel d'agarose.....	21
C) Clonage dans le vecteur pET100.....	22
A) Extraction de l'ADN du gel .....	22
B) Clonage dans le vecteur pET100 .....	22
1) Transformation des bactéries compétentes <i>E. coli</i> .....	23
2) Extraction d'ADN plasmidique (Mini-prep).....	23
i) Identification des clones positifs .....	23
ii) Préparation qualitative d'ADN plasmidique (Mini-prep).....	23
C) Séquençage des plasmides recombinants.....	24
II) Techniques de Biochimie.....	24
A) Purification des protéines PA63 et LF.....	24
A) Préparation des bactéries et de l'extrait brut.....	24
B) Chromatographie d'affinité.....	25
C) Production et purification de PA63.....	25
B) Préparation du complexe PAh-pore-LF.....	26
A) Préparation des nanodisques.....	26
B) Assemblage de PA63h avec LF et nanodisques puis conversion de la forme pré-pore en forme pore.....	26
C) Dosage des protéines.....	27
D) Détermination de la pureté et du poids moléculaire des protéines purifiées.....	27
1) Electrophorèse en conditions dénaturantes (PAGE-SDS) .....	27
a) Préparation des gels.....	28
b) Tampon de migration .....	28
c) Solubilisation des échantillons et migration.....	28

d)	Coloration et décoloration des gels.....	29
III)	Microscopie et Analyse d'images.....	29
A)	Microscopie et Instrumentation.....	29
1)	Préparation de la grille pour microscopie.....	29
2)	Microscopie.....	31
3)	Jeux de données.....	32
a)	Complexe PA63h-prépore-3LFs.....	32
b)	Complexe PA63h-pore-3LFs.....	32
B)	Cryo-microscopie et reconstruction 3D par particules isolées.....	32
1)	XMIPP.....	32
2)	Classification 2D.....	33
3)	Classification 3D pour le démêlage des erreurs d'alignement	33
4)	Analyse de la variance 3D.....	34
5)	Analyse individuelle des LF.....	34
6)	Étude de la densité angulaire.....	35
7)	Estimation de la résolution d'un volume 3D obtenu par raffinement.....	35
8)	Docking de structure pdb dans une carte 3D.....	36
VI)	Dynamique moléculaire.....	36
A)	Principe.....	36
B)	Protocole.....	36
1)	Découpage.....	37
2)	Création de la boîte « container ».....	37
3)	Remplissage de la boîte.....	37
4)	Lancement de la simulation.....	38
5)	Déduction de la structure complète.....	38

## Chapitre 3 : Résultats

<b>A) Clonage, Surexpression et Purification de PA63h et LF.....</b>	<b>39</b>
1) Amplification par PCR .....	39
2) Clonage des produits de la PCR dans pET100.....	39
<b>B) Reconstruction de PAh-pore-3LF.....</b>	<b>42</b>
1) Micrographes et particules.....	42
2) Classification 2D.....	43
3) Le volume de référence.....	44
4) Reconstruction 3D avec XMIPP.....	45
5) Etude de la densité angulaire.....	47
6) Variance 3D du volume.....	48
7) Classification 3D.....	49
8) Variance 3D des volumes obtenus par ML3D.....	50
9) Isolation et superposition des trois LF.....	52
10) Amarrage.....	53
11) Analyse du « pied gras » et mise en évidence d'une erreur d'alignement .....	54
12) Raffinement final.....	56
<b>C) Reconstruction de PA63h-pore-3LFs-ND.....</b>	<b>58</b>
1) Micrographes et particules.....	58
2) Classification 2D.....	60
3) Création d'un volume de référence avec EMAN2.....	62
4) Premier essai de raffinement 3D avec XMIPP.....	62
5) Deuxième essai de raffinement 3D.....	64
<b>D) Dynamique Moléculaire de PA63h-pore-3LFs-ND.....</b>	<b>66</b>
1) Jonction PA-LF et les acides aminés impliqués.....	66
2) Amarrage des LFs sur la structure Pa63h-pore de Nguyen.....	68
3) Résultats de la simulation par dynamique moléculaire.....	72
4) Mise en relation des résultats de la MD avec ceux de la cryo.....	73

## **Chapitre 4 : Discussion & Conclusions**

<b>Discussion &amp; Conclusions.....</b>	<b>75</b>
--	-----------

## **Chapitre 5 : Bibliographie**

<b>Bibliographie.....</b>	<b>82</b>
---------------------------	-----------

# *Liste des Figures*

- Figure 1** : Photographie en microscopie optique de bacilles et spores de *Bacillus anthracis*
- Figure 2** : Lésions causées par l'anthrax
- Figure 3** : Famille des toxines A-B
- Figure 4** : Mécanisme infectieux de l'anthrax
- Figure 5** : Monomère de PA complexé avec le domaine I du récepteur CMG2
- Figure 6** : Structure par Cristallographie à rayon X du Protective Antigen
- Figure 7** : Structure par Cristallographie à rayon X du Lethal Factor
- Figure 8** : Structure par Cristallographie à rayon X de l'Edema factor
- Figure 9** : Comparaison entre la structure de PAh et celle de PA-pore obtenue par dynamique moléculaire
- Figure 10** : Reconstruction 3D de PA63h-pore
- Figure 11** : Interaction de LFN avec PA63h et effet du voltage et pH sur cette interaction
- Figure 12** : Domaine de liaison de LF sur PA en conformation octamérique
- Figure 13** : Schéma d'un liposome
- Figure 14** : Structure d'un nanodisque
- Figure 15** : Schéma de protéines piégées dans une couche de glace avec projections
- Figure 16** : Transformée de Fourier d'une image brute et corrigé par CTF
- Figure 17** : Schéma explicatif du raffinement 3D
- Figure 18** : Insertion du produit de la PCR dans le vecteur pET100
- Figure 19** : Grille pour cryo-microscopie (holley grid)
- Figure 20** : Photos de la station de transfert (Gatan 626) (FEMR McGill)
- Figure 21** : Photographie du microscope Technai F20 (FEMR McGill)
- Figure 22** : Schéma caractérisation la rotation horaire de 103° de la position des LFs
- Figure 23** : Détection des densités angulaires
- Figure 24** : Analyse du produit de la PCR de PA83 et LF par électrophorèse sur gel d'agarose
- Figure 25** : Gel d'Agarose montrant la mini-préparation des plasmides
- Figure 26** : Analyse par SDS-PAGE de la purification de PA83 et LF
- Figure 27** : Analyse par SDS-PAGE de la digestion de PA83 par la Trypsine
- Figure 28** : Analyse par SDS-PAGE de la purification de PA63
- Figure 29** : Micrographe et particules sélectionnées dans Boxer
- Figure 30** : XMIPP ML2D

**Figure 31** : Isosurface du volume de référence

**Figure 32** : Position des LFs par rapport aux PAs

**Figure 33** : Isosurface du complexe PA63h-LF filtrée à 16Å

**Figure 34** : Fourier Shell Correlation de la Carte 3D du complexe PA63h-LF

**Figure 35** : Graphique de la densité angulaire

**Figure 36** : Variance 3D de la carte 3D du complexe PA63h-LF

**Figure 37** : Les deux conformations de PA63h-LF obtenues par ML3D

**Figure 38** : Variance 3D des deux volumes obtenus par ML3D

**Figure 39** : Amarrage de la structure pdb de LF dans une carte de densité

**Figure 40** : Superposition des LF

**Figure 41** : Études des conformations de LF

**Figure 42** : Amarrage des volumes avec les structures issues de dynamique moléculaire

**Figure 43** : Mise en évidence du rôle de "la bosse" par superposition avec rotation de 103° du volume initial dupliqué avec lui-même

**Figure 44** : Schéma expliquant la mise en évidence d'une erreur d'alignement

**Figure 45** : Résultats de l'analyse du pied gras

**Figure 46** : Isosurface du volume final de PA63h-prepore-3LF à 14Å

**Figure 47** : Fourier Shell Correlation de la Carte 3D finale du complexe PA63h-LF

**Figure 48** : Transformée de Fourier des micrographes avant et après correction par inversion de phase CTF

**Figure 49** : Images digitales et particules sélectionnées dans Boxer d'EMAN2

**Figure 50** : Aperçu des « mauvaises » particules éliminées par statistique

**Figure 51** : Classification 2D de PA63h-pore-LFs

**Figure 52** : Sélection des projections les plus caractéristiques de la classification 2D de PA63h-pore-LFs

**Figure 53** : Graphique représentant le log de maximum de vraisemblance en fonction du nombre d'itérations

**Figure 54** : Graphique représentant la proportion en pourcentage des projections contenue dans chacune des 32 classes à la dernière itération

**Figure 55** : Isosurface du volume de référence obtenu par « common lines »

**Figure 56** : Reconstruction 3D par raffinement du complexe PA63h-pore-3LFs-ND

**Figure 57** : Fourier Shell Correlation de la Carte 3D du complexe PA63h-pore-3LFs (1er essai)

**Figure 58** : Reconstruction 3D par raffinement du complexe PA63h-pore-3LFs-ND

**Figure 59** : Fourier Shell Correlation de la Carte 3D du complexe PA63h-pore-3LFs (2<sup>ème</sup> essai)

**Figure 60** : Détails des résidus interagissant avec le site de liaison de PA avec LF

**Figure 61** : Détail de la jonction PA-LF

**Figure 62** : Résultats des amarrages par ZDOCK et ROSETTA

**Figure 63** : Superposition des LFs résultant du docking par ROSETTA avec les LFs de la structure PA63h-pore-3LFs

**Figure 64** : Vue rapproché de la zone de collision

**Figure 65** : Comparaison entre l'amarrage fait par ROSETTA et l'amarrage ajusté

**Figure 66** : Gros plan sur la coupure du Lethal Factor

**Figure 67** : Molécule de PA63h-pore-3LFs tronqués pour MD (vue en perspective)

**Figure 68** : Graphique de l'évolution de la RMSD du système au cours du temps

**Figure 69** : Superposition des LFs obtenus par ROSETTA avec ceux obtenus par MD t = 6 ns

**Figure 70** : Amarrage de la structure obtenue par MD avec le volume acquis par le deuxième raffinement de PA63h-pore-3LFs-ND.



## *Liste des abréviations et des symboles*

<b>ADN</b>	:	Acide désoxyribonucléique
<b>AMP</b>	:	Adénosine Monophosphate
<b>Å</b>	:	Angstrom
<b>ApoA1</b>	:	Apo-Lipoprotéine A1
<b>ATP</b>	:	Adénosine Triphosphate
<b>BSA</b>	:	Bovine serum albumin
<b>CMG2</b>	:	Capillary Morphogenesis protein 2
<b>C</b>	:	Celsius
<b>CCD</b>	:	Charge Couple Device
<b>CMG</b>	:	Capillary Morphogenesis protein
<b>CTF</b>	:	Contrast Transfer Fonction
<b>DTT</b>	:	Dithiothréitol
<b>EDTA</b>	:	Acide éthylène diamine tétraacétique
<b>EF</b>	:	Edema Factor
<b>EM</b>	:	Microscopie Electronique
<b>FRC</b>	:	Fourier Ring Correlation
<b>FT</b>	:	Transformée de Fourier
<b>g</b>	:	Gramme
<b>GROMACS</b>	:	GRONingen MACHine for Chemical Simulations
<b>H<sub>2</sub>O<sub>2</sub></b>	:	Eau oxygénée
<b>HCL</b>	:	Acide chlorydrique
<b>HDL</b>	:	High Density Lipid
<b>HEPES</b>	:	2-Hydroxyéthyl-(1-piperazine)-éthane-sulfonique
<b>K<sub>D</sub></b>	:	Constante de dissociation
<b>kDa</b>	:	Kilodalton
<b>L</b>	:	Litre
<b>LF</b>	:	Lethal Factor

<b>MAPKK</b>	:	MAP kinases kinases
<b>MD</b>	:	Dynamique Moléculaire
<b>μl</b>	:	Microlitre
<b>μM</b>	:	Micromolaire
<b>M</b>	:	Molaire
<b>mA</b>	:	Milliampère
<b>min</b>	:	Minute
<b>mm</b>	:	Millimètre
<b>mM</b>	:	Millimolaire
<b>ND</b>	:	Nanodisque
<b>NB</b>	:	Nutrient Broth
<b>nm</b>	:	Nanomètre
<b>nmol</b>	:	Nanomole
<b>POPC</b>	:	Palmitoyl-oleoyl-phosphatidylcholine
<b>PCR</b>	:	Polymerase Chain Reaction
<b>P/V</b>	:	Poids / Volume
<b>PM</b>	:	Poids moléculaire
<b>PA</b>	:	Protective Antigen
<b>PA63h</b>	:	Heptamère PA63
<b>Pdb</b>	:	Protein Data Bank
<b>PSA</b>	:	Persulfate d'ammonium
<b>SDS</b>	:	Dodécylsulfate de sodium
<b>SDS-PAGE</b>	:	Electrophorèse sur gel de polyacrylamide en présence du SDS
<b>Sec</b>	:	Secondes
<b>SNR</b>	:	Signal Noise Ratio
<b>TBS</b>	:	Tris Buffer Saline
<b>TEMED</b>	:	N, N, N', N' tétraméthylène-éthylène-diamine
<b>TEM8</b>	:	Tumor Endothelium Marker 8
<b>Tris</b>	:	Tris (hydroxy-méthyl) amino-méthane
<b>XMIPP</b>	:	X-Window-based Microscopy Image Processing Package

## I) *Bacillus anthracis*

### 1) Généralités

Les espèces du genre *Bacillus* sont des bactéries à Gram positif qui se présentent sous la forme de bâtonnets, aux extrémités carrées ou arrondies (**Figure 1**) [1]. Leur taille est variable entre 0,5 x 1,2 µm et 2,5 x 10 µm. Elles sont généralement mobiles grâce à une ciliature péritriche. Toutefois *Bacillus anthracis* et *Bacillus mycoides* sont immobiles. Ils peuvent être aérobies ou aéro-anaérobies et sont capables d'élaborer une capsule formée d'un polymère d'acide glutamique.

Lorsque les conditions deviennent défavorables les *Bacillus* sporulent et donnent des spores (une seule spore par cellule végétative) souvent très résistantes dans le milieu extérieur. Le phénomène de sporulation, contrairement à ce qui se produit pour les espèces du genre *Clostridium*, n'est pas inhibé par l'oxygène. In vitro, la sporulation dépend des conditions de culture.



**Figure 1 : Photographie en microscopie optique de bacilles et spores de *Bacillus anthracis***  
[http://www.denniskunkel.com/images/StockImages/Images\\_02/21186E.jpg](http://www.denniskunkel.com/images/StockImages/Images_02/21186E.jpg) [1]

Les *Bacillus* sont des germes de l'environnement dont l'habitat principal est le sol où ils joueraient un rôle dans les cycles du carbone et de l'azote. La résistance des spores et la diversité physiologique des formes végétatives en font des bactéries omniprésentes que l'on peut isoler du sol, de l'eau de mer, de l'eau douce ou de denrées alimentaires.

La résistance des spores constitue un problème majeur en médecine et l'épidémiologie de certaines infections repose sur la sporulation.

La menace de maladies émergentes et la réémergence des maladies infectieuses dans la population humaine et l'économie mondiale ne peuvent être sous-estimées. Cette menace, combinée avec l'augmentation de la résistance aux antibiotiques a conduit à plusieurs recherches pour identifier de nouvelles approches pour traiter ces infections bactériennes. De ce fait, il y a un besoin urgent d'une meilleure compréhension des mécanismes utilisés par les agents pathogènes pour pénétrer dans la cellule eucaryote. Cela permettrait d'identifier de nouvelles cibles pour la prévention et le traitement des maladies causées par des micro-organismes pathogènes. En outre, comprendre comment les protéines étrangères entrent dans la cellule pourrait être utilisé pour concevoir et développer de nouvelles applications pharmacologiques telles que la distribution de médicaments, ADN et/ou de protéines dans des cellules spécifiques du corps.

## **2) L'Infection par *Bacillus anthracis***

### **a) Chez l'animal**

En tant que zoonose, le charbon (anthrax) est surtout une maladie d'herbivores : moutons, chèvres, bovidés, chevaux et chameaux. Pratiquement tous les animaux sont susceptibles de l'attraper. Seuls les carnivores et les porcs sont relativement résistants : les premiers (loups, renards, chiens) grâce à leur acidité gastrique plus élevée, les seconds grâce à un facteur antibactérien de leurs leucocytes.

La contamination des animaux semble se faire surtout par l'alimentation ; elle aboutit à une septicémie rapidement mortelle : la rate hypertrophiée, le cœur et les gros vaisseaux contiennent un sang noirâtre fourmillant de bacilles qui s'écoule par les orifices naturels.

### **b) Chez l'homme**

Plusieurs formes de la maladie sont répertoriées chez l'homme (**Figure 2**) [2]:

- **la forme cutanée** (la plus fréquente) donnant des lésions au point de pénétration du bacille sous forme d'une pustule noirâtre. L'infection évolue le plus souvent sans fièvre ni douleur et guérit spontanément.

- **La forme gastro-intestinale** due à l'ingestion de viande contaminée; elle occasionne une diarrhée.
- **La forme pulmonaire** due à l'inhalation de spores donne lieu à une pneumopathie aigue souvent mortelle.
- **La forme septicémique** fréquemment accompagnée d'une atteinte méningée est primitive ou complique l'une des formes précédentes ; elle est souvent mortelle.

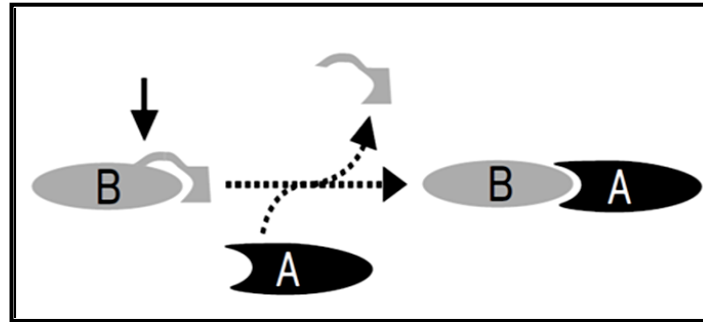


**Figure 2 : Lésions causées par l'anthrax**  
Van den Enden *et al.* [2]

### **3) Les toxines de l'anthrax**

#### **a) Caractéristiques**

Les toxines de l'anthrax font partie de la famille des toxines A-B dans laquelle la moitié B se fixe à la membrane de la cellule permettant par la suite la translocation de la moitié A (**Figure 3**) [3]. Dans le cas de l'anthrax, la moitié B est représentée par la protéine PA (Protective Antigen) et la moitié A par les deux protéines EF (Edema Factor) et LF (Lethal Factor).



**Figure 3 : Famille des toxines A-B**  
Falnes et Sandvig [3]

LF et EF se lient à PA63 avec une forte affinité ( $K_D$  : 1-2 nM) [4] et sont en concurrence pour les mêmes sites sur l'heptamère PA63 (PA63h) [5]. Toutefois, la stœchiométrie et le mode exact de liaison restent controversés. Les études de biochimie initiales (électrophorèse sur gel natif) ont suggéré qu'il y avait une liaison de ligand par monomère PA63 (c'est à dire sept molécules par PA63h; [6]).

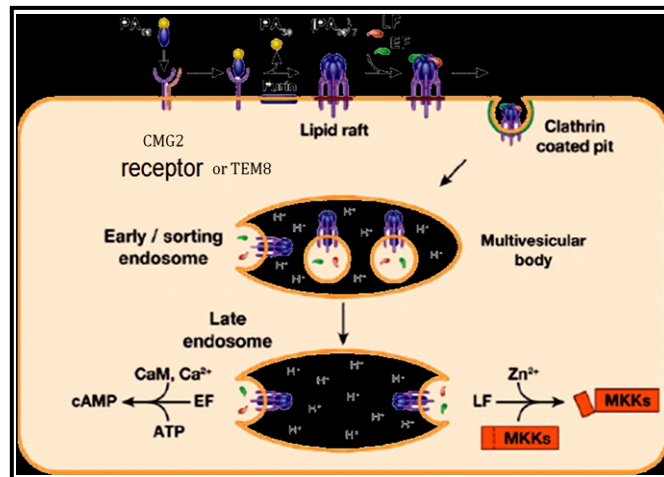
Une étude ultérieure par diffusion multi-angle de la lumière laser a prouvée une occupation maximale de trois molécules LF par heptamère [7,8]. D'importantes interactions de résidus à la fois dans le PA63 [9] et les molécules LF [10, 11] ont été identifiées par des études de mutagenèse. Ensemble, ces études ont conduit à un modèle dans lequel le domaine N-terminal de LF se lie à deux sous-unités adjacentes de l'heptamère PA63 [11, 12].

### **b) Le mécanisme infectieux de l'anthrax (Figure 4)**

Après être recruté par les récepteurs cellulaires "Capillary Morphogenesis protein 2" (CMG2) ou "Tumor Endothelium Marker 8" (TEM8), PA s'organise en heptamère. Il peut alors fixer jusqu'à 3 ligands (EF et LF) avant d'être endocyté. Les modèles actuels suggèrent que la baisse de pH à l'intérieur des endosomes permet un changement de conformation de la forme pré-pore de PA63h vers la forme pore et que les ligands EF et LF passent au travers du pore pour entrer dans le cytoplasme.

Par la suite, EF qui est  $Ca^{2+}$  et calmodulo-dépendent transforme l'ATP en AMP cyclique déstabilisant la cellule. LF qui est une métalloprotéase à zinc capable de couper les MAP kinases (MAPKK1, 2, 3, 4, 6 et 7) qui constituent un des trois modules de la chaîne d'activation

MAPK responsable du transfert d'informations extérieures à la cellule vers le noyau au niveau transcriptionnel [13].



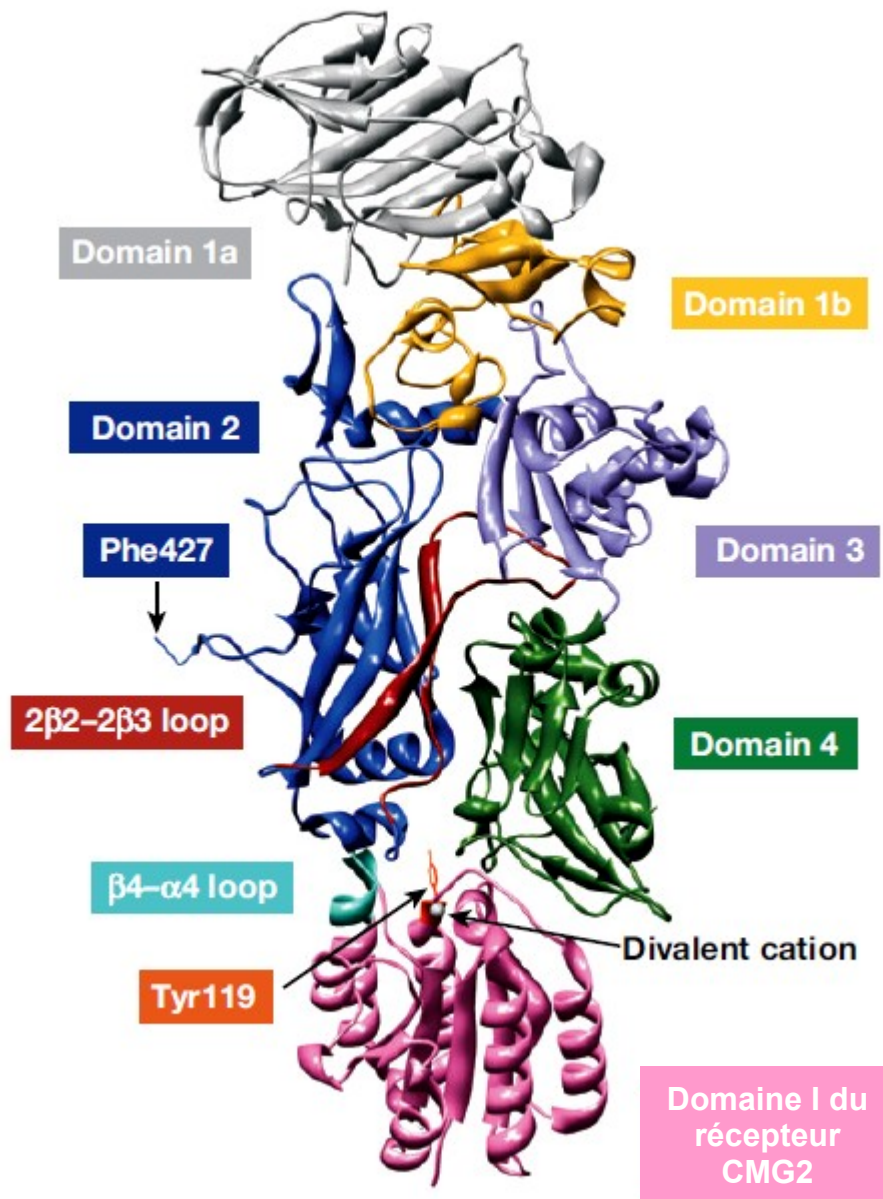
**Figure 4 : Mécanisme infectieux de l'anthrax**  
Young et Collier [13]

### c) Présentation des structures résolues des toxines de l'anthrax

Les différentes protéines responsables de l'infection par l'anthrax ont été résolues par cristallographie à rayons X.

#### (i) Les récepteurs cellulaires à l'anthrax

Comme mentionné plus haut, deux récepteurs cellulaires capable de fixer le Protective Antigen: le "Tumor Endothelial Marker 8" (TEM8) et le "Capillary Morphogenesis protein 2" (CMG2) [14, 15]. PA83 s'attache au domaine I du CMG2 (résidus 35-225) avec un ratio de 1:1 montré par titration calorimétrique isothermale et par résonance de fluorescence par transfert d'énergie (FRET) [16] et par cristallographie à rayon X [13]. Les deux récepteurs à l'anthrax contiennent un site d'adhésion ion-dépendant (MIDAS) compris entre les résidus Asp50, Ser52, Ser54, Thr118, and Asp148 (CGM2)/Asp150 (TEM8) [14, 15]. Le récepteur MIDAS contient un cation divalent qui joue un rôle dans la fixation de PA [16, 17, 18, 19, 20].



**Figure 5 : Monomère de PA complexé avec le domaine I du récepteur CMG2**  
Le cation divalent se fixe au récepteur MIDAS. Adapté de Young et Collier [13].

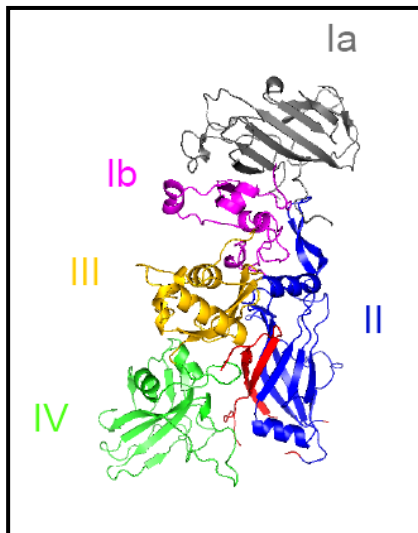
Les récepteurs à l'anthrax sont exprimés dans les cellules endothéliales lors de l'angiogénèse. Contenu du fait que ces récepteurs ont une affinité très importante pour PA, des expériences ont montré que l'utilisation d'un PA muté stimule l'inhibition de l'angiogénèse et inhibe la croissance tumorale *in vivo* [21]. Ainsi, il serait envisageable d'utiliser PA comme inhibiteur pour une thérapie anti-angiogénique.



## (ii) Le Protective Antigen

Protective Antigen (PA) a été cristallisé sous forme monomérique à 2.1 Å et sous forme heptamérique à 4.5 Å. Le monomère a les dimensions  $\sim 100 \times 50 \times 30$  Å et peut être décomposé en quatre domaines (**Figure 6**) [22, 47]:

- **Domaine 1** (résidus 1-258) amino-terminal qui contient deux ions calcium et le site de clivage pour des protéases actives ;
- **Domaine 2** (résidus 259-487) qui permet l'heptamérisation et qui contient une large boucle flexible impliquée dans l'insertion membranaire ;
- **Domaine 3** (résidus 488-595) qui est petit et dont la fonction n'est pas encore connue ;
- **Domaine 4** (résidus 596-735) qui est le récepteur de liaison carboxy-terminal.



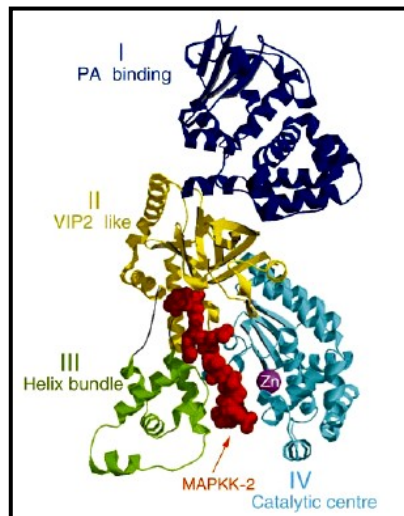
**Figure 6 : Structure par Cristallographie à rayon X du Protective Antigen**  
Petosa *et al.* [22]

PA fait initialement 83 kDa. Le détachement via la furine du fragment de 20 kDa amino-terminal du domaine 1 (au niveau de la séquence RKKR résidus 164-167) permet l'assemblage en heptamère de PA63 et l'exposition des sites de liaison de LF et EF. Il n'y a pas de changements majeurs lors du passage du monomère vers l'heptamère [13, 22].

### (iii) Le Lethal Factor

Le Lethal Factor (LF) est une protéine de 90 kDa hautement pathogène de 100 Å de diamètre et de 70Å de hauteur. Elle a été cristallisée seule à 2.2 Å ainsi qu'avec son substrat (MAPKK-2). Elle se divise également en quatre domaines qui sont les suivants (**Figure 7**) [23, 47]:

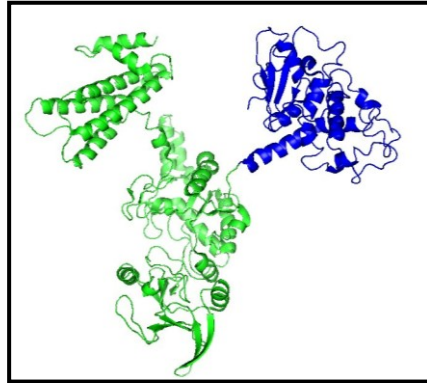
- **Domaine I** se fixe à PA N-terminal ;
- **Domaine II** ressemble à la toxine ADP ribosyme de *Bacillus cereus* mais le site actif a été muté et recruté pour augmenter la reconnaissance du substrat ;
- **Domaine III** est inséré à l'intérieur du domaine II et semble avoir surgi d'une réplication répétée d'un élément structural du domaine II.
- **Domaine IV** est une zinc-métalloprotéase qui contient un site catalytique au centre. Il a été démontré que la mutation E687C inactive la protéine mais n'altère pas la fixation au Zinc [23].



**Figure 7 : Structure par Cristallographie à rayon X du Lethal Factor**  
Pannifer *et al.* [23]

### (iv) L'Edema Factor

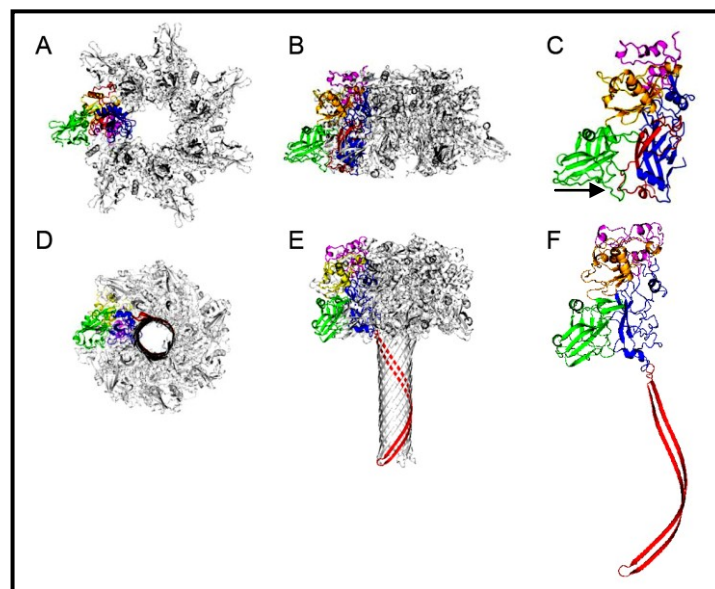
Il ne fait pas partie de l'objet de mon étude. Il a été cristallisé sous forme complexée avec la calmoduline (CaM). EF est une protéine qui fait 89 kDa et qui bloque le domaine N-terminal de CaM indépendamment de la présence du calcium (**Figure 8**) [24].



**Figure 8 : Structure par Cristallographie à rayon X de l'Edema factor**  
Shen *et al.* [24]

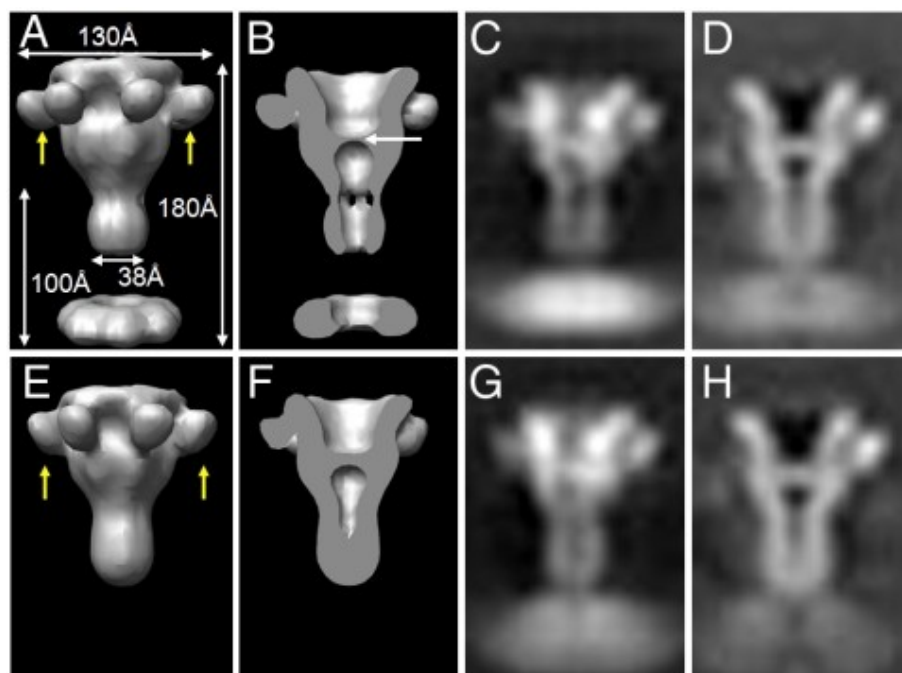
### (v) Conversion de la forme pré-pore à la forme pore de PA63h

Jusqu'à récemment, aucune structure pore de PA n'avait été capturée. Seuls des modèles obtenus par dynamique moléculaire pouvaient servir d'indications. Ainsi le modèle proposé par by **Nguyen [25]** (basé sur la dynamique moléculaire) (**Figure 9**), suggérait que des boucles désordonnées  $2\beta 2-2\beta 3$  provenant du domaine 2 formeraient l'élément pénétrant membranaire de PA63h.



**Figure 9 : Comparaison entre la structure de PAh (A : Vue du dessous ; B : Vue de profile ; C : PA momomérique [14], la flèche indique le déplacement du domaine IV après dépliement de la boucle  $2\beta 2-2\beta 3$  en rouge) et celle de PA-pore obtenue par dynamique moléculaire (D : Vue du dessous ; E : Vue de profile ; F : PA momomérique [25])**

Cependant, une structure obtenue par microscopie électronique en coloration négative est désormais disponible [26] (**Figure 10**). Celle-ci atteint une résolution de 22Å et présente une nouvelle énigme concernant le domaine 2. En effet, une fois PA63h inséré dans une membrane, d'après le modèle de Nguyen et al., 2004, le domaine 2 devrait se déplacer du fait qu'il ne contient plus les boucles 2 $\beta$ 2–2 $\beta$ 3, mais d'après Katayama *et al.* [26], ce n'est pas le cas (**figure 10**). Par contre, cette reconstruction 3D ne remet pas en cause le fait que le diamètre interne du lumen est étroit et varie entre 12 et 17 Å.



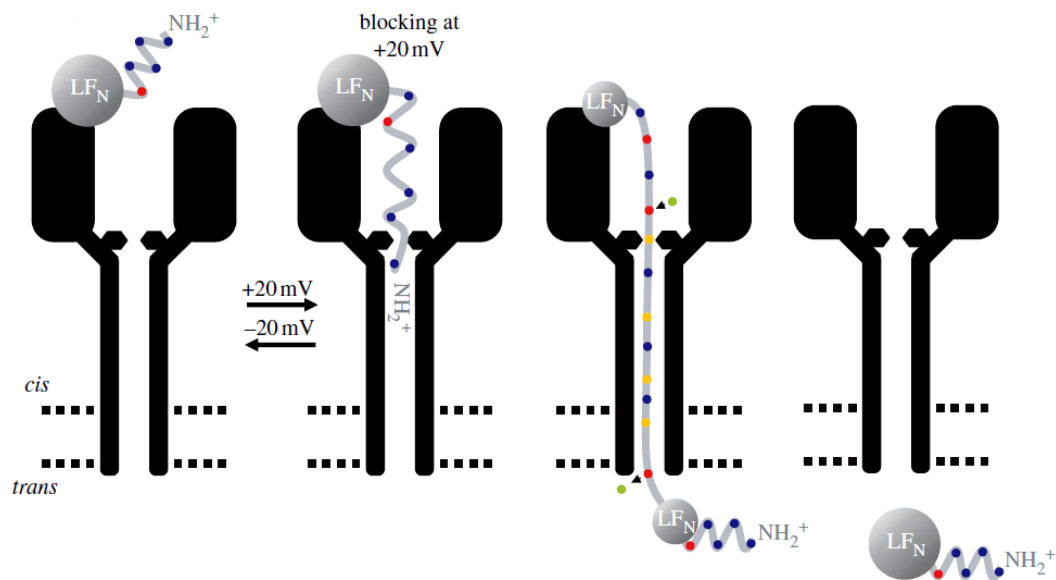
**Figure 10 : Reconstruction 3D de PA63h-pore [26]**

La reconstruction a été filtrée à sa résolution nominale de 22 Å par "low-pass filtered". L'insertion a été faite dans les nanodisques lipidiques (**A à D**) et dans les vésicules (**E à F**).

**A et E** : présentations de surface. **B, D, F et H** : Coupes centrales de la reconstruction pour montrer l'intérieur de la structure. **C, D, G et H** : Projections de la structure de PA-pore. Les flèches jaunes pointent vers le domaine 3. La flèche blanche marque la localisation d'un rebord qui à l'intérieur même du canal.

Aussi, les travaux de Finkelstein [27] ont montré l'existence de PA63h-pore et le passage de LF à travers ce pore. Pour cela, un système à deux compartiments séparés par une bicouche lipidique monolamellaire couverte de PA63h (inséré sous forme de pore) a été utilisé. Dans ces

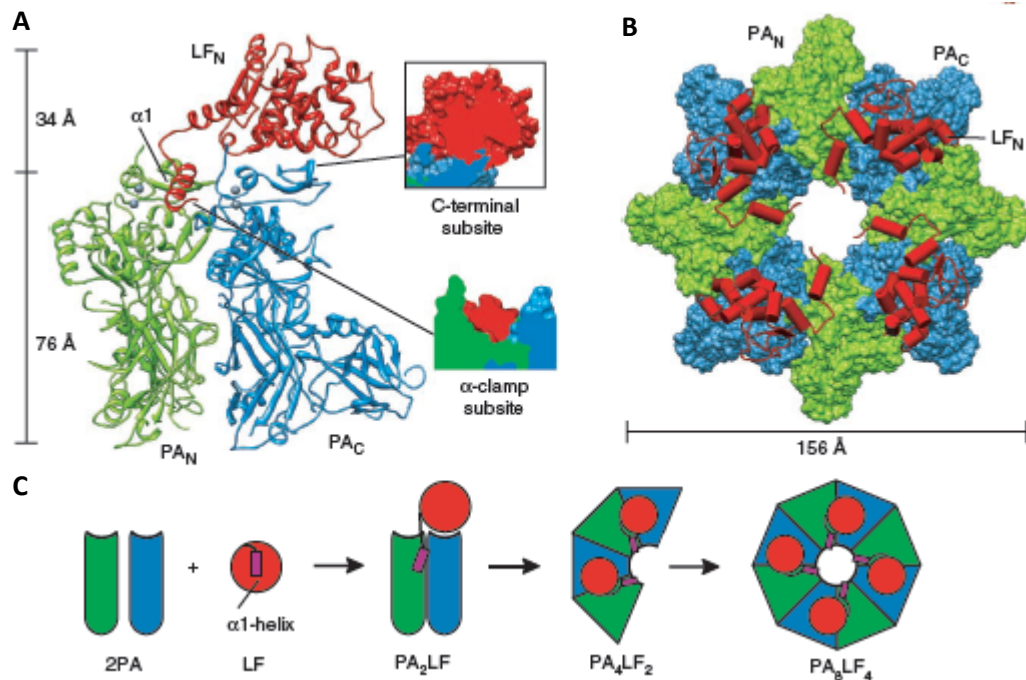
conditions, à 20 mV, la conductance est de 55pS dans 100 mM de KCl à pH 5,5. En ajoutant des quantités nanomolaires de LF<sub>N</sub> (partie N-terminal du LF) dans l'un des deux compartiments, une baisse rapide et forte de la conductance a été observée suivie d'une remontée lente et progressive pour atteindre l'équilibre. D'après l'auteur, la signification de ces changements est directement liée au passage de LF au travers du pore formé par PA63h. Cette traversée s'effectue progressivement (ce qui explique le rétablissement de la conductance). A partir de ces résultats, Finkelstein a proposé un modèle (**Figure 11**) [27].



**Figure 11 : Interaction de LF<sub>N</sub> avec PA63h et effet du voltage et pH sur cette interaction**  
Finkelstein [27]

### (vi) Initiation du dépliage de LF au travers du Protective Antigen

Lorsque LF est associé à PA, il adopte une conformation différente de lorsque qu'il est tout seul (PDB 1J7N, [23]). En effet, pour LF<sub>N</sub> (domaine I dans la **figure 7**), la première hélice alpha ( $\alpha 1$ ) se loge dans la fente formée par la jonction entre deux PA [28]. La surface de liaison de LF au complexe PA est alors composée de deux sous-unités (**figure 12**). Cette association a été démontrée par cristallographie à rayon X sur une structure octamérique PA<sub>8</sub>-LF<sub>4</sub>.



**Figure 12: Domaine de liaison de LF sur PA en conformation octamérique (figure adaptée de [28]).**

A) A gauche, représentation ruban de la PA<sub>2</sub>LF<sub>N</sub>. PA<sub>C</sub> (chaîne A, bleu), PA<sub>N</sub> (chaîne B, vert), LF<sub>N</sub> (chaîne C, rouge) et des ions calcium (sphères grises). A droite, coupe de la surface rendue des deux sous-sites de liaison à LF<sub>N</sub>, avec le sous-site de liaison C-terminal en haut et le sous-site de l'α-clamp en bas.

B) L'octamère PA est représenté comme une surface moléculaire en vue de dessus, et les hélices de LF<sub>N</sub> comme des cylindres. Les couleurs de la partie A) sont respectées.

C) Schéma explicatif du processus de liaison de LF<sub>N</sub> à deux PA.

Les acides aminés de LF impliqués dans la liaison PA-LF ont été déterminés par mutation [10] et sont : Asp-182, Asp-187, Leu-188, Tyr-223, His-229, Leu-235, et Tyr-236 où les mutations de l'alanine produisent un défaut significatif de la liaison avec H229A and Y236A qui inhibe complètement la fixation.

### (vii) Liposomes et nanodisques

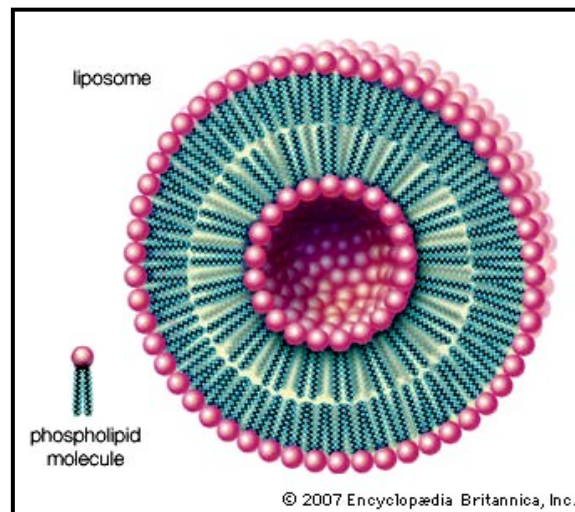
Dans la cellule des mammifères, PA63h s'insère dans la membrane d'un endosome. Ainsi, pour l'étude de la forme pore de PA63h, l'utilisation d'une membrane similaire à celle de l'endosome voire même un endosome lui-même dans le cas d'une étude *in vivo* serait nécessaire. Pour des raisons pratiques, il est plus facile de débiter en utilisant les membranes synthétiques. Ainsi, deux techniques différentes peuvent être intéressantes:

## • Les liposomes

Ce sont des vésicules formées de bicouches concentriques de phospholipides. Les liposomes peuvent être uni-lamellaires ou multi-lamellaires. Ces dernières emprisonnent entre elles des compartiments aqueux. Dans le cas de notre étude, seuls les liposomes uni-lamellaires sont intéressants car ils imitent les endosomes naturels par la présence du bicouche de lipides (**Figure 13**).

Les liposomes peuvent être obtenus de manière simple en dissolvant les lipides dans un tampon aqueux puis en les faisant passer au-travers une membrane via un mini-extrudeur (Avantis Polar Lipids). Le diamètre des pores déterminent la taille des liposomes qui peuvent avoir des diamètres entre 50 nm et 1  $\mu\text{m}$ .

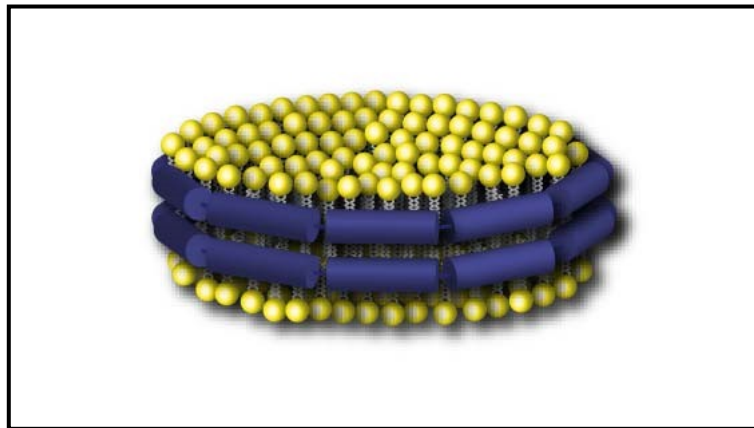
Les liposomes pourraient s'avérer très intéressants pour étudier le passage des ligands au-travers du pore (dans le cas où l'hypothèse de l'entrée de LF/EF à l'intérieur de la cellule se fait via le PA63h-pore serait exacte). En effet, il est possible de recréer les compositions (et donc les conditions) des solutions aqueuses interne et externe au liposome. En effet, il est très simple de modifier le pH interne ou externe des liposomes pour permettre la formation du pore et par la suite vérifier l'entrée des facteurs LF ou EF à travers le pore vers l'intérieur des liposomes [29].



**Figure 13 : Schéma d'un liposome uni-lamellaire**  
Encyclopedia Britania, 2007 [29]

- **Les nanodisques**

Les nanodisques sont composés d'une protéine échafaudage (l'ApoA1) et de lipides. L'ApoA1 est la majeure composante de la protéine du « bon cholestérol » : HDL (High Density Lipid). Ainsi, lorsque l'on met en contact des lipides dissous dans un détergent avec cette protéine, il se crée une association entre les deux produits. En effet, l'ApoA1 « encercle » les fragments de lipides dissous. Il en résulte un disque dont le centre est composé des lipides et la périphérie de la protéine ApoA1 [30] (**Figure 14**).



**Figure 14 : Structure d'un nanodisque**

[http://www.uni-duesseldorf.de/MathNat/ipb/lw\\_resource/datapool/file/item\\_169/nanodisc\\_empty\\_size\\_m.jpg](http://www.uni-duesseldorf.de/MathNat/ipb/lw_resource/datapool/file/item_169/nanodisc_empty_size_m.jpg) [30]

Étant donné que le diamètre du nanodisque est 10 nm, il présente alors un grand intérêt par rapport au liposome. En effet, le nanodisque ne peut permettre que l'insertion d'un seul PA63h (très rarement deux).



## II) Cryo-microscopie et reconstruction 3D par particules isolées

### 1) Introduction

En coloration négative, après le dépôt de l'échantillon, on sèche pour enlever l'excès puis on rajoute le colorant puis on sèche encore. Le séchage cause un aplatissement de la protéine sur la grille et peut également conduire à sa précipitation. Cela modifie largement sa structure tridimensionnelle. Afin d'éviter ce phénomène, on peut utiliser la cryo-microscopie. Avec cette méthode, les protéines sont hydratées et observées directement dans leur tampon physiologique sans aucune interaction avec le colorant.

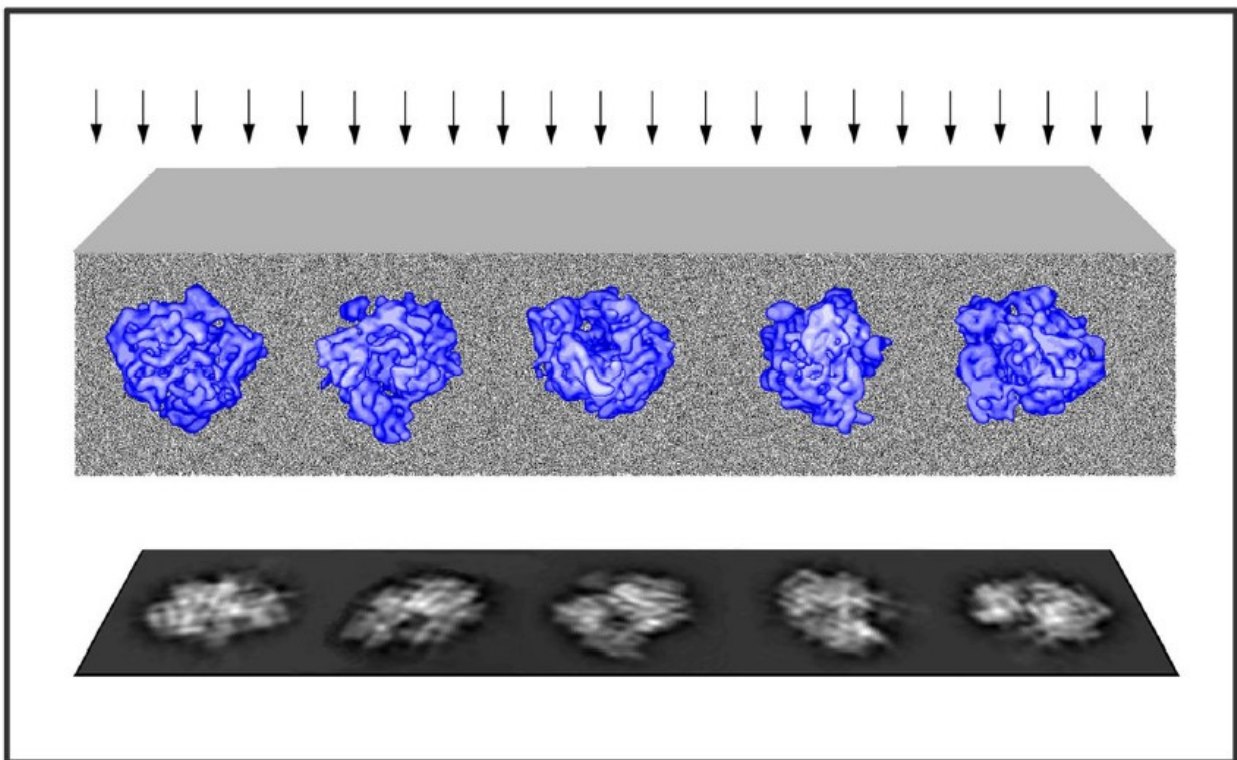


Figure 15 : Schéma de protéines piégées dans une couche de glace avec projections [31]

## 2) Transformée de Fourier et son importance dans la reconstruction 3D

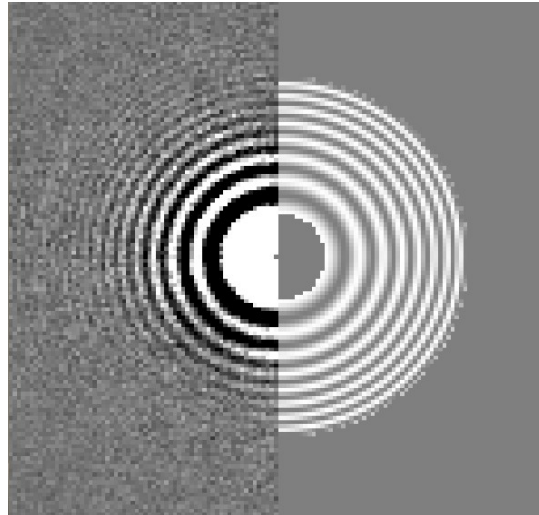
Le mathématicien et physicien français Joseph Fourier a défini au XVIII<sup>e</sup> siècle une fonction permettant de calculer le spectre fréquentiel d'une fonction. Cette fonction peut être à une seule dimension telle que des ondes sonores, à deux dimensions comme pour une image 2D, à trois dimensions pour un volume 3D et même plus. En 1968, DeRosier et Klug ont décrit une méthode générale pour la reconstruction de structures 3D à partir de micrographes en utilisant la transformée de Fourier (FT) [32]. En effet, la transformée de Fourier d'une image permet son analyse en regardant son spectre de fréquence qui est le module de valeurs complexes incluant la phase et l'amplitude.

## 3) Défocus et estimation de la CTF

Pour augmenter le contraste, les photos sont prises à différents défocus (0.5 à 5). En effet, un problème important en cryo-EM est celui du ratio signal bruit (SNR : Signal Noise Ratio). Le faible SNR est principalement dû à l'utilisation du low-dose (faible dose d'électrons reçus pour ne pas détruire l'échantillon). Ceci ne donne pas un fort contraste mais permet cependant de conserver son intégrité dans la glace. Pour augmenter le ratio signal bruit, les images sont prises en défocus. Malheureusement, ce gain du signal s'accompagne par une déformation de la projection du spécimen suivant une fonction non-linéaire qui est la fonction de transfert de contraste ou CTF (pour Contrast Transfer Fonction). La CTF prend compte des caractéristiques propres au microscope (comme la source d'électrons et les aberrations des lentilles) et des conditions de prise d'image (comme le défocus, la charge du carbone...).

Pour permettre la reconstruction 3D avec précision du volume, nous avons travaillé avec des images corrigées par inversion de phase en utilisant Ctfind [33]. Les paramètres suivants ont été utilisés: 2.0 pour l'aberration sphérique ; 200 KV de tension ; 0.07 d'amplitude de contraste ; 50 Kx pour la magnification ; 10  $\mu\text{m}$  pour la taille du pixel scanné (taille réelle :  $4.2\text{\AA}^2$ ).

L'image 2D de la CTF (**figure 16**) est formée de cercles concentriques de différentes intensités: anneaux de Thon. Ces cercles représentent la diffraction de notre échantillon (la concentration en protéine a donc un effet sur l'intensité des anneaux). Plus le défocus est important, plus le nombre de cercle augmente et plus leur largeur diminue. Dans ce travail, la CTF a été estimée par ctfind3.



**Figure 16 : Transformée de Fourier d'une image brute et Représentation de fonction estimée par le logiciel pour l'image**

#### **4) Alignements**

Tout d'abord, des projections références sont calculées à partir d'un volume de référence. Comme ces projections sont issues de calcul, on connaît alors leurs orientations (angles) dans un repère d'Euler. Ces coordonnées sont alors gardées en mémoire. Par la suite, on cherche à aligner les particules (projections provenant des données expérimentales) avec les projections de références (étape nommée : « projection matching »). La correspondance entre ces deux types de données permet de retrouver les orientations (dans le repère d'Euler) des particules expérimentales. Il devient alors possible de recréer la structure de la protéine initiale en reprojétant toutes les particules alignées (projections expérimentales). Le nouveau volume créé servira de référence pour le nouveau calcul de projections de référence. Le cycle itératif peut continuer jusqu'à ce que le réassignement d'angles d'Euler aux images soit minimum et constant. Dans cette situation, la résolution du volume ne s'améliore plus d'une itération à l'autre.

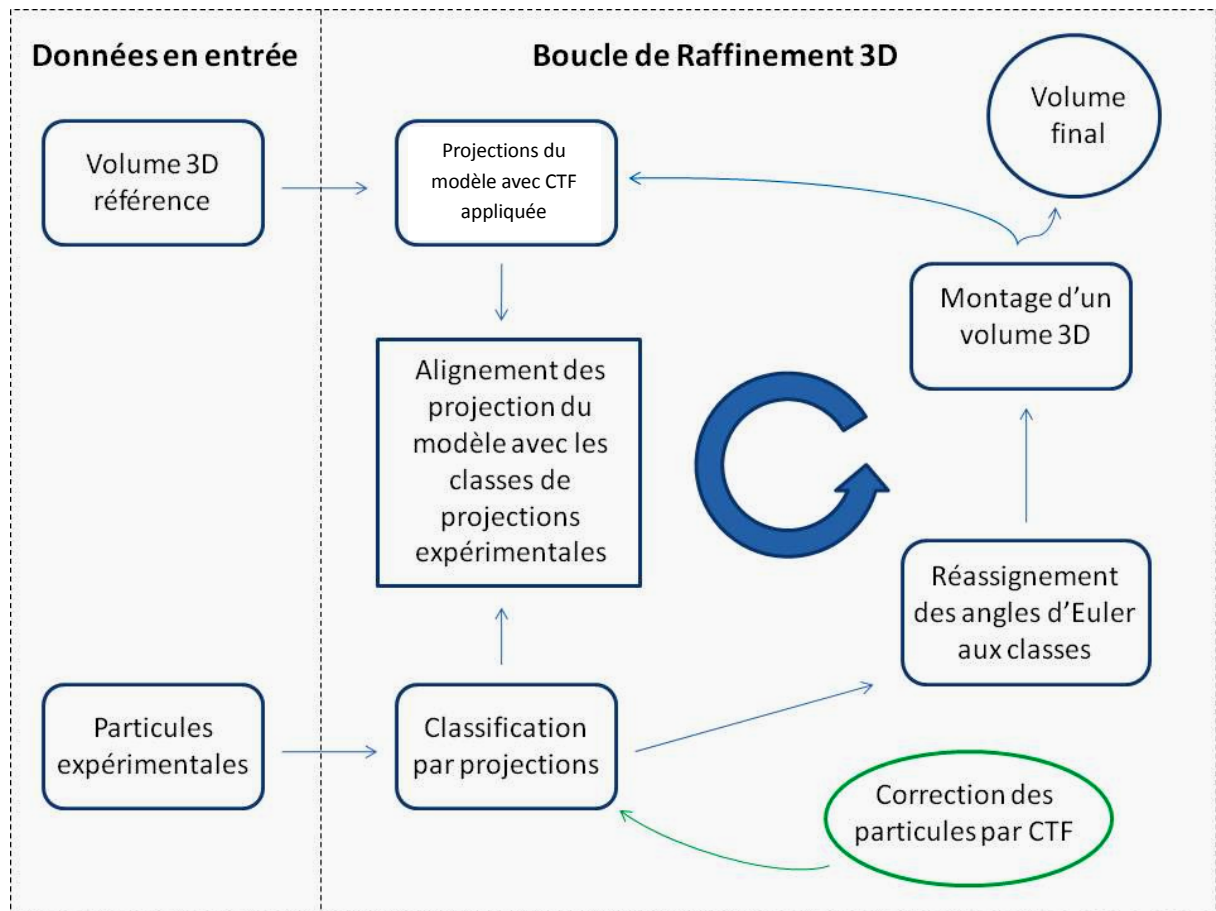


Figure 17 : Schéma explicatif du raffinement 3D (adapté d'EMAN1)

### 5) Création d'un volume de référence

Pour créer le volume de référence, nous avons choisi le « common lines » proposé par le logiciel EMAN2 [34]. Cette méthode est basée le principe suivant : chaque plan constituant un volume ne possède qu'une et seulement une ligne commune (common line) avec un autre plan dans l'espace de Fourier. Ainsi, il devient alors possible de rassembler les plans (projections expérimentales) en un volume en utilisant ces lignes (une classification 2D au préalable permet de limiter le nombre de plan à comparer). Le volume résultant d'une « common line » est assez approximatif. Plusieurs essais avec un même jeu de données donnent des volumes proches. Si ce dernier est de qualité suffisante, les projections qui en découleront pourront se superposer ("matcher") avec les particules et ainsi amorcer le processus raffinement 3D décrit précédemment.

## **6) Classification 3D**

Nous avons évoqué plus haut qu'il est possible de piéger des molécules (protéines) dans plusieurs conformations. Cependant, si tel est le cas, le mélange de ces dernières peut apporter un flou au volume reconstruit. De plus, il peut être intéressant d'étudier chacune de ces conformations séparément. Pour cela, il devient nécessaire de classifier ces particules en fonction de leurs affinités pour telle ou telle conformation.

De la même manière que la classification 2D, l'algorithme compare avec maximum de vraisemblance (ML3D) si la particule se superpose mieux avec la projection de référence correspondante de chacun des volumes de référence. Le système fonctionne de manière itérative et le nombre de reconstruction 3D (« refinement ») est proportionnel au nombre de volumes de référence employés et prend en considération les lobes (« blobs ») des volumes. XMIPP ne propose que la version par maximum de vraisemblance.

## **7) Classification 3D pour le démêlage des mésalignements**

La classification 3D classique permet normalement de retrouver les différentes conformations présentes dans un jeu de données. Dans le cas des mésalignements, l'objectif est de retrouver les particules mal alignées afin de leur permettre de se réaligner de manière correcte.

## ***III) Problématique***

Comme nous l'avons vu précédemment, le diamètre interne du lumen de PA63h demeure serré et varie entre 12 et 17 Å. Ainsi, celui-ci ne permet pas le passage « direct » des ligands LF et EF qui sont environ 10 fois plus gros. Des modèles récents suggéraient que les facteurs LF et EF se déplieraient avant de passer dans le pore [35]. Alternativement, dans le cas où ils ne passeraient pas, le pore servirait au passage des ions à l'intérieur des endosomes, ce qui probablement causerait leur éclatement et la libération des facteurs dans la cellule.

Pour résoudre cette controverse, il est nécessaire de connaître en détails les interactions entre PA et LF/EF.

## ***IV) Objectifs***

Nos objectifs se résument en trois parties:

### **1. L'étude du complexe PA63h-prépore-3LFs :**

La connaissance de ce complexe peut apporter des informations sur la disposition des LFs (attachement) sur la molécule heptamérique.

### **2. L'étude du complexe PA63h-pore-3LFs :**

Elle permet de savoir si la disposition des LFs change depuis la conformation prépore de PA à la conformation pore. Cela donne une idée sur la dynamique du processus.

### **3. L'étude du complexe PA63h-pore-3LFs par dynamique moléculaire :**

Elle permet l'analyse de la stabilité de la structure. De plus une corrélation avec la carte-EM peut conforter les résultats éventuels.

## **I) Techniques de biologie moléculaire**

### **A) Amplification par PCR**

Pour PA et LF, les plasmides ont été acquis via addgene (www.addgene.com). La PCR est réalisée en prenant 2 µl de l'ADN plasmidique contenant le gène souhaité en présence de 5 µl de tampon Taq (10 mM Tris-HCl, pH 9,0 ; 50 mM KCl, 1,5 mM MgCl<sub>2</sub> 25 mM Triton X-100, 1X ; BSA 0,2 mg/ml) (Promega), 1 µl de l'ADN, la High-fidelity Taq Polymerase 2,5 u/µl (Bio-Rad), 1 µl de dNTP 10 mM (dATP, dTTP, dGTP, dCTP) et 1 µl de chacune des amorces (50 pmol/µl), correspondant au fragment à amplifier (pour **PA83**: **For** : 5'-CACCATGAAAAACGAAAAGTGTTAATACC-3' et **Rev** : 5'-TTATCCTATCTCATAGCCTTTTTTAG-3' ; Pour **LF** : **For** : 5'-CACCATGAATATAAAAAAGAATTTATAAAAGTAATT-3' et **Rev** : 5'-TTATGAGTTAATAATGAACTTAATCTG-3'). Le dessin des amorces a été fait par FastPCR. Le volume final est ramené à 25 µl en ajoutant de l'eau milliQ stérile. Ensuite les échantillons sont placés dans le thermocycleur (Bio-Rad) et la PCR est déclenchée en accord avec le protocole suivant : Pré-chauffage des échantillons pendant 1 minute à 95 °C, suivi d'un programme de 35 cycles, chaque cycle se déroule en 3 étapes : une 1ère étape de dénaturation pendant 30 sec. à 95°C, une 2ème étape d'hybridation des amorces de 30 secondes à 51°C pour PA83 et 48°C pour LF et une 3ème étape d'élongation à 72°C pendant 3 minutes. L'amplification est achevée par une incubation de 10 minutes à 72°C.

**NB** : Les amorces ont été désignées en utilisant FastPCR, un logiciel qui permet de calculer l'efficacité des PCR et la synthèse des amorces [36].

### **B) Analyse de la PCR sur gel d'agarose**

L'agarose (Sigma-Aldrich) est dissous, par chauffage aux micro-ondes, dans 100 mL de tampon TAE 0,5X pour une concentration finale de 1% (w/v). Lorsque la solution est à 50°C, 4 µl de Bromure d'ethidium (10 mg/ml, Fisher BioReagent) est ajouté et le gel est ensuite coulé dans la cuve à gel pour l'électrophorèse (Bio-Rad). Le chargement des échantillons se fait en prenant 10 µl de la solution PCR et 2 µl du tampon de charge 6X Dye Solution (Invitrogen). Les mélanges sont chargés avec précaution dans les puits en utilisant 4 µl du marqueur du poids

moléculaire 1 kb Ladder (100 ng/μl, Invitrogen). La migration est faite à 100 V pendant 20 minutes et l'analyse des gels est faite à l'aide du Gel Logic 200.

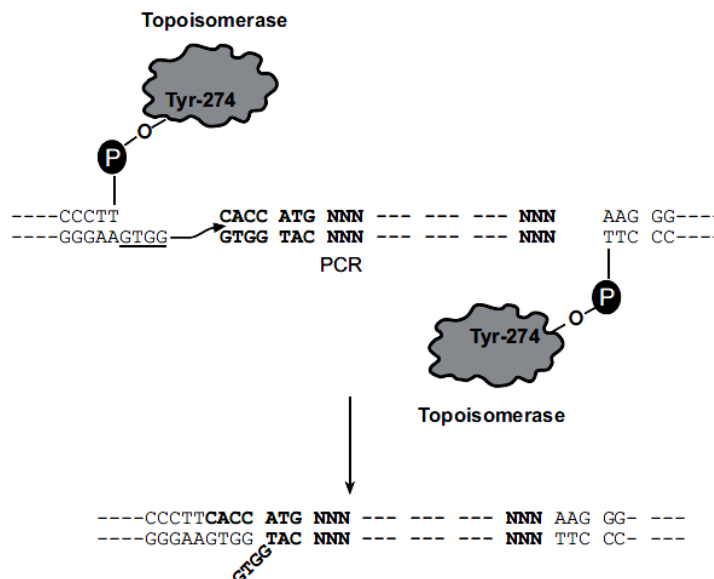
### C) Clonage dans le vecteur pET100

#### 1) Extraction de l'ADN du gel

L'extraction a été faite en utilisant le Kit Nucleospin Extract II (Machery-Nagel) en suivant les étapes suivantes : Pour 100 mg de gel, 200 μl de la solution de fixation sont rajoutés puis incubés 10 minutes à 50°C (pour solubiliser le gel d'agarose et libérer l'ADN). Après totale solubilisation, l'ADN dissous est introduit dans des mini-colonnes avec filtres (qui fixent l'ADN) puis centrifugé à 11000 g pendant 1 min. Après le lavage des mini-colonnes (pour éliminer l'agarose et ne garder que l'ADN fixé sur les filtres de Silice) et le séchage par centrifugation à 11000 g pendant 1 min, l'ADN pure est élué en rajoutant 30 ul de la solution d'élué puis centrifugé à 11000 g pendant 1 min.

#### 2) Clonage dans le vecteur pET100

Nous avons utilisé le kit "pET100 Directional TOPO" pour la surexpression de nos protéines car il permet un clonage directionnel des produits de la PCR sans recours aux enzymes de restriction ou les ligases et ceci grâce à une séquence rajoutée dans l'amorce sens (CACC).



**Figure 18 : Insertion du produit de la PCR dans le vecteur pET100**

<http://products.invitrogen.com/ivgn/product/K10001?ICID=search-product> [37]



Concernant la procédure, elle est comme suit : A 4 µl du produit de la PCR, 1 µl de la solution saline (1.2 M NaCl, 0.06 M MgCl<sub>2</sub>) et 1 µl du vecteur TOPO sont rajoutés. Après un doux mélange, l'ensemble est laissé incubé 5 minutes à température ambiante.

**a) Transformation des bactéries compétentes *E. coli***

100 µl de bactéries compétentes (**DH5α** pour le sous-clonage pour avoir une grande quantité de plasmides ou **BL21** pour la surexpression des protéines) sont mélangées avec 6 µl du produit de la ligation (citée plus haut) puis incubées 30 minutes dans la glace, suivi d'un choc thermique à 42° C pendant 45 secondes puis une incubation de 2 minutes dans la glace. Après addition de 250 µl du milieu SOC (tryptone 2 % ; extrait de levure 0,5 % ; 10 mM NaCl ; 2,5 mM KCl ; 10 mM MgCl<sub>2</sub> ; 10 mM MgSO<sub>4</sub> ; 20 mM glucose), les bactéries sont alors incubées 1 heure à 37°C sous agitation à 220 rpm/min. Ensuite, la culture est étalée sur des boîtes de Pétri LB (Peptone 0,1 % ; extrait de levure 0,05 % ; NaCl 0,05%) (Luria Broth Gibco BRL) contenant l'agar (15 g/l), l'ampicilline à 100µg/ml. Les boîtes sont incubées pendant 16 heures à 37°C pour la croissance des colonies.

**b) Extraction d'ADN plasmidique (Mini-prep)**

**i) Identification des clones positifs**

Plusieurs colonies ont été piquées et mises en culture dans 5 ml de milieu LB liquide en présence de l'ampicilline à 100µg/ml. Les bactéries ont été cultivées à 37°C pendant 16h et sous agitation (220 rpm /min). Les plasmides ont été par la suite extraits puis analysés par PCR et séquençage.

**ii) Préparation qualitative d'ADN plasmidique (Mini-prep)**

L'ADN plasmidique a été extrait par une méthode rapide basée sur la lyse alcaline (Sambrook et Russell, 2001). 1 ml de la culture bactérienne est transféré dans un tube Eppendorf, ensuite centrifugé à 1 500 g pendant 5 minutes. Au culot bactérien 100 µl de solution I (50 mM glucose, 25 mM Tris-HCl pH 8,0, 10 mM EDTA pH 8,0) ont été ajoutés et mélangés par pipetage. Après, 200 µl de solution II sont ajoutés (0,2 N NaOH ; 1% SDS) et mélangés par inversion et finalement 150 µl de solution III sont ajoutés (acétate de potassium 5M pH 6,0), puis

une incubation dans la glace pendant 5 minutes. Les échantillons sont centrifugés à 10000 g pendant 5 minutes et le surnageant transféré dans un autre tube Eppendorf. Ensuite, 900 µl d'éthanol absolu froid déjà incubé à -20°C sont ajoutés. Après avoir bien mélangé l'ensemble on centrifuge à 10000 g pendant 5 minutes. Puis le surnageant est totalement éliminé, et le culot est lavé avec 500 µl d'éthanol 70%, puis centrifugé à 10000 g pendant 5 minutes. Le culot contenant l'ADN est repris dans 30 µl de tampon TE pH 8,0.

### **3) Séquençage des plasmides recombinants**

La détermination de la séquence nucléotidique des plasmides recombinants a été faite par Génome Québec (McGill University) en utilisant les primers : promoteur T7, amorces sens et anti-sens de PA83 et LF (citées plus haut).

## **II) Techniques de Biochimie**

### **A) Purification des protéines PA63 et LF**

#### **1) Préparation des bactéries et de l'extrait brut**

50 µl de culture stock de bactéries (BL21 transformées par les plasmides exprimant PA83 ou LF) sont suspendues dans 5 ml du bouillon nutritif (NB, Nutrient Broth), puis précultivées dans une étuve à 37°C avec agitation (220 tr/min) pendant une nuit. Chaque préculture sert à inoculer une culture de 500 ml du NB. Les cultures ainsiensemencées sont incubées à 37°C avec une agitation de 150 tr/min. Après 3h d'incubation, l'IPTG est ajouté à la concentration finale de 0.5 mM puis les bactéries sont incubées à 30°C pendant 16 h. Ensuite, les cellules bactériennes sont collectées par une centrifugation à 6000 g pendant 30 min, puis lavées 1 fois avec du tampon A (Tris 50 mM, pH 7.5, EDTA 0.5 mM et DTT 1 mM) (tampon A). Les cellules ainsi récupérées sont resuspendues dans le même tampon additionné de lysozyme à 1 mg/ml afin de briser et fragiliser les parois bactériennes.

Pour la préparation de l'extrait brut, la suspension cellulaire est exposée aux ultrasons (3 × 5 min de sonication avec un intervalle de repos de 5 min après chaque sonication pour éviter le réchauffement de l'échantillon) à 90% de puissance. L'appareillage utilisé est Sonic

Dismembrator Model 100 (Fisher Scientifique). La solution est ensuite clarifiée par centrifugation à 6000 g pendant 30 min à 4°C, puis utilisée pour la purification de la protéine désirée.

## **2) Chromatographie d'affinité**

La chromatographie d'affinité est une méthode qui repose sur une interaction moléculaire entre le constituant à purifier et un ligand de structure complémentaire, ce qui se traduit par une fixation spécifique et réversible.

Pour la purification des protéines LF et PA, la Ni<sup>2+</sup>-sépharose a été utilisée. Ce métal est susceptible de fixer spécifiquement les protéines marquées par un poly-histidine (généralement 6). La préparation protéique (Extrait brut) est introduite dans la colonne (2 × 20 cm) préalablement équilibrée avec le tampon A (Tris 50 mM, pH 7.5, EDTA 0.5 mM et DTT 1 mM). La colonne est ensuite abondamment lavée par le même tampon (additionné d'Immidazol 30 mM). L'élution s'effectue dans le même tampon (additionné d'Immidazol 300 mM). Les fractions sont ensuite analysées individuellement par SDS-PAGE puis concentrées et conservées à -80°C jusqu'à utilisation.

La colonne est régénérée avec une solution composée de l'EDTA 100 mM suivie d'un lavage avec le tampon A (voir plus haut) puis une solution de Nickel acétate pour charger la colonne. La conservation de la colonne se fait en présence d'un agent anti-microbien (0,02% azide de sodium).

## **3) Production et purification de PA63**

Pour obtenir PA63, PA83 (20 µM) est digéré par la trypsine (1mg/ml) durant 30 min à 37°C. PA63 est par la suite concentré par des centricons avec des filtres ayant un cut-off de 50 kDa.

## **B) Préparation du complexe PAh-pore-LF**

### **1) Préparation des nanodisques**

11,2  $\mu\text{l}$  de POPC suspendu dans du chloroforme (25  $\mu\text{g}/\mu\text{L}$ ) sont séchés dans un tube à hémolyse (en verre) avec un flux d'argon (bien évaporer le chloroforme). Par la suite, 18,4  $\mu\text{L}$  du tampon B (20 mM de Tris-HCl pH 7.4, 100 mM de NaCl et 40 mM de Cholate de Sodium) sont ajoutés. Le tout est transféré dans un tube eppendorff de 0,5 ml puis chauffé sous un flux d'eau chaude pendant environ 1 minute puis vortexé. Après, le tube est placé dans un bain à ultrasons jusqu'à ce que la solution devienne limpide (environ 10 minutes). 1,6  $\mu\text{l}$  de la protéine ApoA1 diluée 3 fois à partir du tube initial dont la concentration est de 0,5  $\mu\text{g}/\mu\text{L}$  est ajouté aux lipides puis le tout est dialysé une nuit à 4°C dans le tampon C (20 mM de Tris-HCl pH 8, 100 mM de NaCl et 2 mM de  $\text{CaCl}_2$ ) [38].

### **2) Assemblage de PA63h avec LF et nanodisques puis conversion de la forme pré-pore en forme pore**

L'ApoA1, le POPC, le PA63h et le LF ont été mélangés suivant les proportions molaires respectives 1:65:1:6. Cela tient compte du fait qu'il faut 65 molécules de POPC pour 1 protéine d'ApoA1 et que l'on souhaite assembler 1 PA63h par nanodisque. Le ratio PA63h:LF est de 1 pour 6 car évidemment il faut au moins 3 molécules LF pour un PA63h mais aussi, nous ne souhaitons pas atteindre la saturation complète (au moins 10 fois) afin d'obtenir une majorité de PA-3LF mais aussi quelques PA-2LF voire même PA-1LF.

Ainsi, 20  $\mu\text{l}$  des nanodisques dialysés sont récupérés et transférés dans un tube eppendorff de 0,5 ml. A ceux-ci sont rajoutés 2,2  $\mu\text{l}$  de PA63h dilué 3 fois à partir d'un tube à la concentration initiale de 3,0  $\mu\text{g}/\mu\text{l}$  et 2,8  $\mu\text{l}$  de LF à partir d'un tube à la concentration initiale de 1,1  $\mu\text{g}/\mu\text{l}$ . Cela donne un volume total de 25  $\mu\text{l}$  de concentration finale en PA63h de 100  $\mu\text{g}/\text{ml}$ . Le tube est gardé à température ambiante pendant 30 minutes puis son contenu est placé à dialyser pendant 4 heures à température ambiante (ou 1 nuit à 4°C) dans le tampon D (20 mM d'acétate de sodium pH 5.5, 100 mM de NaCl et 2 mM de  $\text{CaCl}_2$ ) pour permettre la conversion de la forme pré-pore à la forme pore de PA63h avec insertion dans un nanodisque.

### **C) Dosage des protéines**

Les protéines ont été quantifiées suivant la technique de Bradford [39]. Elle se base sur l'utilisation du réactif du Bradford (Bio-Rad) qui contient le bleu de coomassie R 250 qui se fixe sur les protéines donnant une coloration allant du marron pour les faibles concentrations jusqu'au bleu pour les fortes concentrations en protéines.

Dans une plaque (96 puits), un volume de 5 µl de l'échantillon à différentes dilutions est rajouté à 200 µl du réactif de Bradford. Après 10 min d'incubation à température ambiante et à l'obscurité, l'absorbance de la coloration développée est mesurée à 595 nm à l'aide du spectrophotomètre (Lecteur de plaques, Beckman Coulter, AD340) et la concentration protéique inconnue est déduite en se référant à la courbe d'étalonnage réalisée à l'aide avec une gamme étalon allant de 0 à 2000 µg de la BSA (Bovine Serum Albumin).

### **D) Détermination de la pureté et du poids moléculaire des protéines purifiées**

Pour vérifier l'homogénéité et la pureté de la protéine purifiée ainsi que pour déterminer son poids moléculaire, nous avons utilisé la technique d'électrophorèse sur gel de polyacrylamide. Cette technique s'applique à des composés biologiques comportant des groupements ionisables. La présence d'un champ électrique entraîne la migration de ces ions macromoléculaires en fonction de leur charge et de leur taille.

#### **1) Electrophorèse en conditions dénaturantes (PAGE-SDS)**

Les protéines sont séparées par électrophorèse verticale sur plaque suivant la technique de Laemmli [40]. Elle est réalisée dans un appareil pour mini-gel (8 x 10 cm), les cales utilisées ont une épaisseur de 1mm.

L'utilisation d'un agent dénaturant comme le dodécylsulfate de sodium (SDS) provoque la dissociation de la structure quaternaire en des formes monomériques, sur lesquelles il se fixe. Dans ces conditions, la séparation des protéines sur le gel s'effectue uniquement en fonction de leur taille, puisque leur densité de charge est rendue globalement négative du fait de la présence du SDS.

**a) Préparation des gels**

• **Gel de séparation à 12 %**

✓ Acrylamide 29,2% - Bisacrylamide 0,8%	4 ml
✓ Tris-HCl 1,5 M pH 8,8	2,5 ml
✓ SDS 10%	100 µl
✓ Eau bidistillée	3,33 ml
✓ Persulfate d'ammonium 10%	50 µl
✓ TEMED	20 µl

• **Gel de concentration à 4 %**

✓ Acrylamide 29,2% - Bisacrylamide 0,8%	0,675 ml
✓ Tris-HCl 0,5M, pH 6,8	1,25 ml
✓ SDS 10%	50 µl
✓ Eau bidistillée	3,05 ml
✓ Persulfate d'ammonium 10%	30 µl
✓ TEMED	20 µl

**b) Tampon de migration**

Il est composé des solutions suivantes :

- ✓ Tris-HCl 25 mM pH 8,3
- ✓ Glycine 320 mM
- ✓ SDS 0,1 %

**c) Solubilisation des échantillons et migration**

Un volume contenant 40 µg de protéines est solubilisé dans 10 µl du tampon de charge dont la composition est la suivante : Tris-HCL 60 mM (pH 6,8), SDS 1% (p/v), Glycérol 10% (v/v), Bleu de Bromophénol 0,01% (p/v) et du β-Mercaptoéthanol 1% (v/v), puis chauffé pendant 3 minutes à 100°C, et chargé dans les puits du gel. La migration est réalisée à température ambiante avec un voltage constant de 100V.

#### **d) Coloration et décoloration des gels**

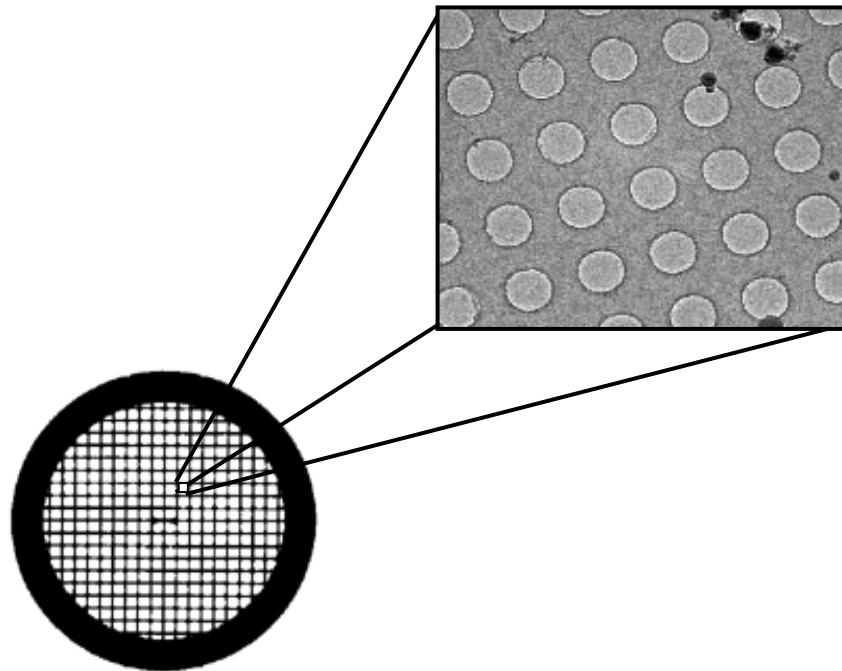
Les gels sont colorés pendant une nuit sous agitation à température ambiante avec une solution de bleu de Coomassie à 0,25%, dans un mélange "Méthanol - Eau distillée - Acide acétique" avec un rapport (4/5/1). Les gels sont ensuite décolorés dans le même mélange sans le bleu de coomassie.

### **III) Microscopie et Analyse d'images**

#### **A) Microscopie et Instrumentation**

##### **1) Préparation de la grille pour microscopie**

Pour la Cryo-microscopie, on utilise des grilles spéciales (Quantifoil Grid R2/2) (**Figure 19**) composées d'un grillage en cuivre. Par dessus ce quadrillage est disposé une mince couche de carbone percée à espace régulier (puits). Avant d'utiliser les grilles, elles sont déchargées grâce à un « glow discharger » pour augmenter l'hydrophilie de la surface. Puis, on dépose 5  $\mu$ l de l'échantillon sur la grille qui est tenue par une pince fine à l'intérieur d'une chambre maintenue à 4°C avec un taux d'humidité de 100%. L'excès de tampon est éliminé afin de garder une couche mince (500 nm) de l'échantillon. 5 secondes de séchage avec une force de 1 (sur l'échelle de l'appareil) semblent optimales. Directement après cette étape, la grille est plongée dans une cuve d'éthane maintenu liquide par de l'azote liquide disposé autour. L'utilisation de l'éthane comme agent cryogène permet d'avoir un film très mince de glace vitreuse contenant l'échantillon étudié. Contrairement à l'azote liquide, l'éthane est un excellent cryogène. Ainsi, dès que plongé dans l'éthane froid, l'échantillon est instantanément congelé. La grille est ensuite transférée dans le cryo porte grille.



**Figure 19 : Grille pour cryo-microscopie (holley grid) [37]**

Le porte grille est ensuite entreposé dans un tube Falcon percé immergé dans de l'azote liquide pour entreposage dans une bonbonne d'azote liquide. En cas de besoin, le tube Falcon est transporté dans un conteneur portable à azote liquide. Le conteneur est déposé juste à côté de la cryo-station de transfert. Celle-ci est formée d'une base et du conteneur à azote liquide qui permet de recevoir le porte grille (Gatan 626). L'extrémité opposée du porte grille peut également recevoir de l'azote liquide. Le froid est ainsi directement acheminé au niveau de la grille via un fil de cuivre.

Il est très important de faire attention à toute forme d'eau. En effet, cette dernière est la principale source de contamination des grilles. Un simple souffle peut amener de la vapeur d'eau à se fixer sur la grille.





**Figure 20 : Photos de la station de transfert (Gatan 626) (FEMR McGill)**

## **2) Microscopie**

Une fois la grille est chargée dans le holder, celui-ci est inséré dans le microscope Tecnai F20 équipé d'une caméra CCD (charge couple device) de 4096 x 4096 pixels. Le microscope est réglé pour une accélération de voltage à 200 keV.



**Figure 21 : Photographie du microscope Tecnai F20 (FEMR McGill)**

### **3) Jeux de données**

#### **a) Complexe PA63h-prépore-3LFs**

Le complexe PA63h-prépore-3LFs a été fourni par Dr. Eugenio Santelli du laboratoire du Dr. Robert Liddington (Burnham Institute). Les photos ont été prises au Burnham Institute par le Dr. I. Rouiller dans le laboratoire de Dr. Dorit Hanein. Pour résoudre des problèmes liés à l'obtention de molécules majoritairement orientées de la même manière (certaines vues ne sont pas bien représentées), une couche de carbone a été rajoutée sur la grille quantifoil.

96 micrographes (sur film photographique) au total ont été pris. La position des complexes sur les micrographes ainsi qu'un volume de référence ont été obtenus lors d'un premier essai de raffinement fait par le Dr I. Rouiller et Avi Biswas.

#### **b) Complexe PA63h-pore-3LFs**

150 images digitales du complexe PA63h-pore-3LFs ont été prises par l'intermédiaire de la caméra CCD du microscope Tecnai F20 au Facility for Electron Microscopy Research (FEMR) de l'Université de McGill.

### **B) Cryo-microscopie et reconstruction 3D par particules isolées**

#### **1) XMIPP**

Pour les alignements et reconstruction 3D, j'ai choisi de travailler avec la suite logicielle XMIPP, "X-Window-based Microscopy Image Processing Package" [41]. Il s'agit d'un programme basé sur SPIDER, un logiciel d'analyse d'images en microscopie électronique. Celui-ci travaille beaucoup avec des listes contenues dans des fichiers texte ce qui permet un accès et un contrôle très étroit du traitement de données. Les résultats de calculs sont contenus également dans des fichiers texte (édition facile) et non directement dans les entêtes d'images « headers » (qui nécessiterait de faire des quantités importantes de copies et donc une perte conséquente

d'espace disque). XMIPP est un logiciel modulaire permettant ainsi son exploitation de manière personnalisée en écrivant des scripts.

## **2) Classification 2D**

Avant de commencer une reconstruction 3D, une analyse 2D est effectuée. Celle-ci permet d'estimer rapidement la qualité des images (et donc des particules) et surtout, si toutes les projections de la molécule (protéine de l'échantillon) sont présentes.

Pour la classification 2D, l'algorithme ML2D (maximum of likelihood) de XMIPP a été utilisé.

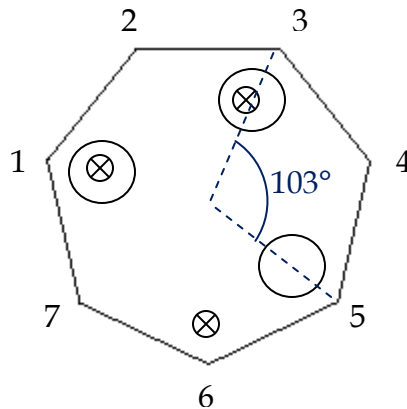
CL2D est un type de classification qui combine alignement et classification en un seul et itératif processus. A chaque itération, toutes les images expérimentales sont alignées et comparées par corrélation croisée à chacune des d'images du groupe de référence (initialement obtenu par alignements au hasard). Par la suite, des moyennes des sous-groupes obtenus composeront le nouveau groupe d'images références. Une autre itération peut alors commencer. La boucle s'arrête lorsqu'aucune image ne peut encore migrer. La corrélation croisée est alors maximale. Le ML2D travaille de la même manière à part qu'au lieu de faire une corrélation croisée, il fait une mesure probabiliste [42].

## **3) Classification 3D pour le démêlage des erreurs d'alignement**

Sur le principe de la classification 3D, j'ai créé un script (**voir annexe 1**) qui fait ce travail en utilisant XMIPP :

D'abord, un volume A de référence a été déterminée à partir de complexe en structure cristalline (pdb). Celui-ci a été aligné avec le volume obtenu par l'alignement d'images "(pied gras)". Par la suite, le volume A a été dupliqué donnant un volume B qui a par la suite été tourné de 103° (**figure 22**). Les projections correspondantes aux deux volumes (A et B) ont été calculées. En gardant la correspondance entre les projections et les particules alignées (issues du volume contenant le pied gras), chaque particule est comparée par cross-corrélation à sa projection correspondante provenant du volume A ou du volume B. Un seuil permet de séparer (dans un troisième groupe) les particules qui correspondent pour les deux volumes. Le processus

est très rapide et s'effectue en seule itération. Ce principe sera plus facile à comprendre durant l'exposition des résultats.



**Figure 22: Schéma caractérisation la rotation horaire de 103° de la position des LFs (vue de dessus).** Les LFs en positions originales sont représentés en croix cerclées. Après rotation de 103° en sens horaire, les LFs se retrouvent aux positions représentées par des cercles simples.

#### 4) Analyse de la variance 3D

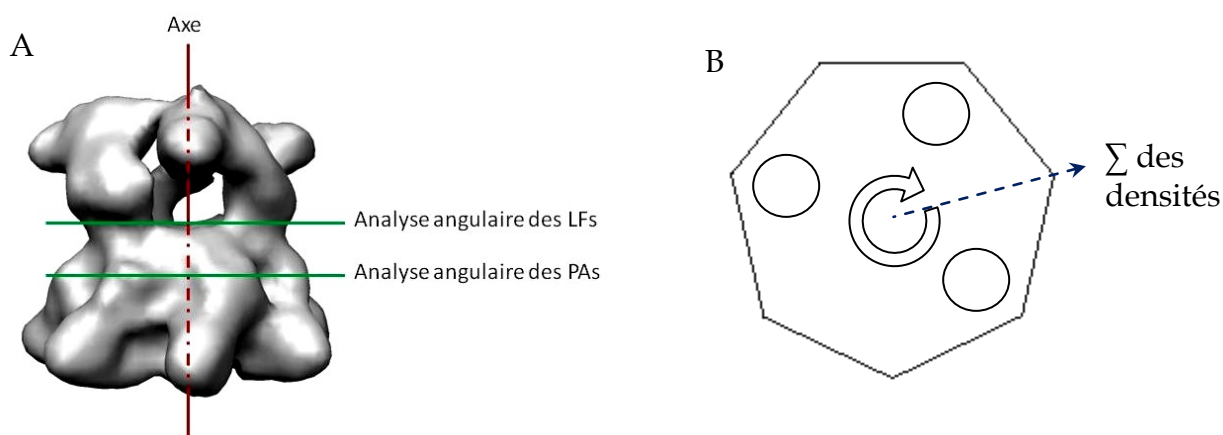
Elle permet la mise en évidence de l'homogénéité de l'alignement des particules. En effet, lors de l'alignement de ces dernières, une corrélation croisée (cc) est calculée. Par l'intermédiaire du logiciel SPIDER, il est possible de mettre en relation les cc avec l'impact positionnel des particules sur le volume reconstruit. Le résultat du calcul peut être interprété avec le logiciel de visionnage des volumes : Chimera de l'Université de Californie à San Francisco (UCSF) [43]. Les zones coloriées en vert correspondent aux endroits où la variance est faible et en rouge celles dont la variance est importante.

#### 5) Analyse individuelle des LF

Les LFs ont d'abord été isolés via mapcarv de la suite COAN (séparation d'éléments par « watershed ») [44]. L'analyse individuelle des différents LFs de la structure PA63h-prepore-3LFs a été faite par Chimera. Les structures sont caractérisées par des angles par l'intermédiaire de ce même logiciel.

## 6) Étude de la densité angulaire

La densité angulaire est définie par la prise de mesures de la somme de la densité d'un rayon autour d'un axe (**figure 23**). Le logiciel SPIDER permet de réaliser cette tâche. Pour cela, l'axe passant au travers du pore a été défini. Le rayon correspond à la distance entre cet axe et un LF.



**Figure 23: Détection des densités angulaires**

A) Hauteurs des plans de coupe

B) Schéma d'un plan de coupe. La densité angulaire est la somme des densités d'un rayon donné (flèche en pointillé bleu).

## 7) Estimation de la résolution d'un volume 3D obtenu par raffinement

La méthode par corrélation de contour de Fourier (FSC : Fourier Shell Correlation) a été utilisée. La résolution est définie comme la fréquence spatiale à laquelle l'échantillonnage annulaire de deux transformées de Fourier est négligeable par corrélation croisée. Le résultat avec XMIPP est un fichier texte contenant plusieurs colonnes dont la FSC en fonction de la résolution en Å. Il est alors possible de tracer la courbe de la FSC en fonction de la résolution. La résolution se lie en prenant la valeur correspondante à une FSC de 0,5.

### **8) Docking de structure pdb dans une carte 3D**

Cela est fait par l'intermédiaire du logiciel Chimera [45]. Les structures cristallines et les carte 3D sont ouvertes puis sont superposées grossièrement à la main. Par la suite, l'algorithme « fit in map » calcule le meilleur docking.

## **IV) Dynamique moléculaire**

### **A) Principe**

La dynamique moléculaire consiste à prédire l'évolution d'un système au cours du temps par le calcul informatique. Cette expérience *in silico* permet la création de modèles qui aident à la compréhension expérimentale. L'aide peut se présenter de deux manières : l'une étant la visualisation sous un autre angle de résultats et l'autre étant d'orienter l'expérience « *out silico* » en créant des points de départ raisonnables (gain de temps et d'argent).

Dans notre étude, le système est une macro-molécule biologique : une protéine. La dynamique moléculaire (MD pour « Molecular Dynamics ») va permettre dans un premier temps de stabiliser un complexe protéique assemblé par amarrage (docking). Dans un deuxième temps, la MD va calculer les interactions entre tous les atomes du système rendant possible la visualisation des mouvements des différentes sous-unités de la protéine. La trajectoire de chaque atome sera enregistrée.

### **B) Protocole**

La MD est réalisée par le logiciel GROMACS, acronyme de « GRONingen MACHine for Chemical Simulations ». La molécule à calculer est au format pdb (Protein Data Bank) a été obtenu par l'assemblage de PA63h-pore (modèle de Nguyen [25]) et de 3 LF par docking via le site de ROSETTA et de ZDOCK. La molécule résultant a été ajustée pour des raisons de légères collisions entre certains atomes par alignement de la zone de contact de PA-LF en conformation

pore avec la zone de contact de PA-LF en conformation pré-pore. Cette manipulation a été faite sur Pymol.

### **1) Découpage**

La molécule PA63h-pore-3LF est d'une taille imposante avec 711 kDa pour 6262 acides aminés et environ 185 000 atomes. Sachant que le temps de calcul de la simulation est directement relié au nombre d'atomes, il va être judicieux de réduire la taille de la protéine en ne gardant que l'essentiel (susceptible de bouger). De plus, ce temps de calcul dépend aussi de l'espace que la molécule prend. A partir de ce principe, le PA63h-pore a été coupé au niveau de la glycine 343 des sept monomères tronquant ainsi de manière significative le tube ( $\beta$ -barrel) de la protéine (partie très stable, non susceptible de bouger). Dans un souci de préserver la structure, la glycine G343 a été virtuellement reliée avec l'acide aspartique D284. De la même manière, les LFs ont été tronqués au niveau de l'arginine R263. Le LF est clairement composé de deux parties reliées au niveau du R263 (extrémité de l'hélice- $\alpha$ ). Le morceau qui interagit directement avec PA est conservé. Ainsi, après réduction, on remarque combien d'espace volumique a été gagné et nous passons de 6262 à 5727 acides aminés pour environ 170 000 atomes.

### **2) Création de la boîte « container »**

Maintenant que le complexe protéique a été réduit, il est placé dans une boîte contenant le solvant dans lequel il doit baigner. Conformément à la taille de la molécule, la boîte cubique a pour dimensions 17x17x17 unités arbitraires. La protéine est centrée aux coordonnées 8,5 8,5 8,5 en utilisant les 14 atomes de calcium comme référence.

### **3) Remplissage de la boîte**

La boîte est remplie de molécules d'eau (solvant) et la molécule est stabilisée dans son nouvel environnement par une courte minimisation d'énergie. On remplace alors des molécules d'eau par des molécules de NaCl pour retrouver les conditions physiologiques en sel et aussi à stabiliser le pH à 5,5.

#### **4) Lancement de la simulation**

La simulation sera faite sur un superordinateur de 16 CPU 64 bits condensés à 2.3 Ghz. Avant de lancer une simulation tenant compte de tous les atomes de la molécule, une première minimisation de 5 fs sur le backbone sera faite. Celle-ci permet de minimiser les collisions possibles entre atomes et apporte une certaine souplesse au système. Si le système ne crashe pas, la minimisation pourra être faite sur la molécule globale. Pour cette étape, j'ai choisi de générer 10 nanosecondes de simulation.

#### **5) Déduction de la structure complète**

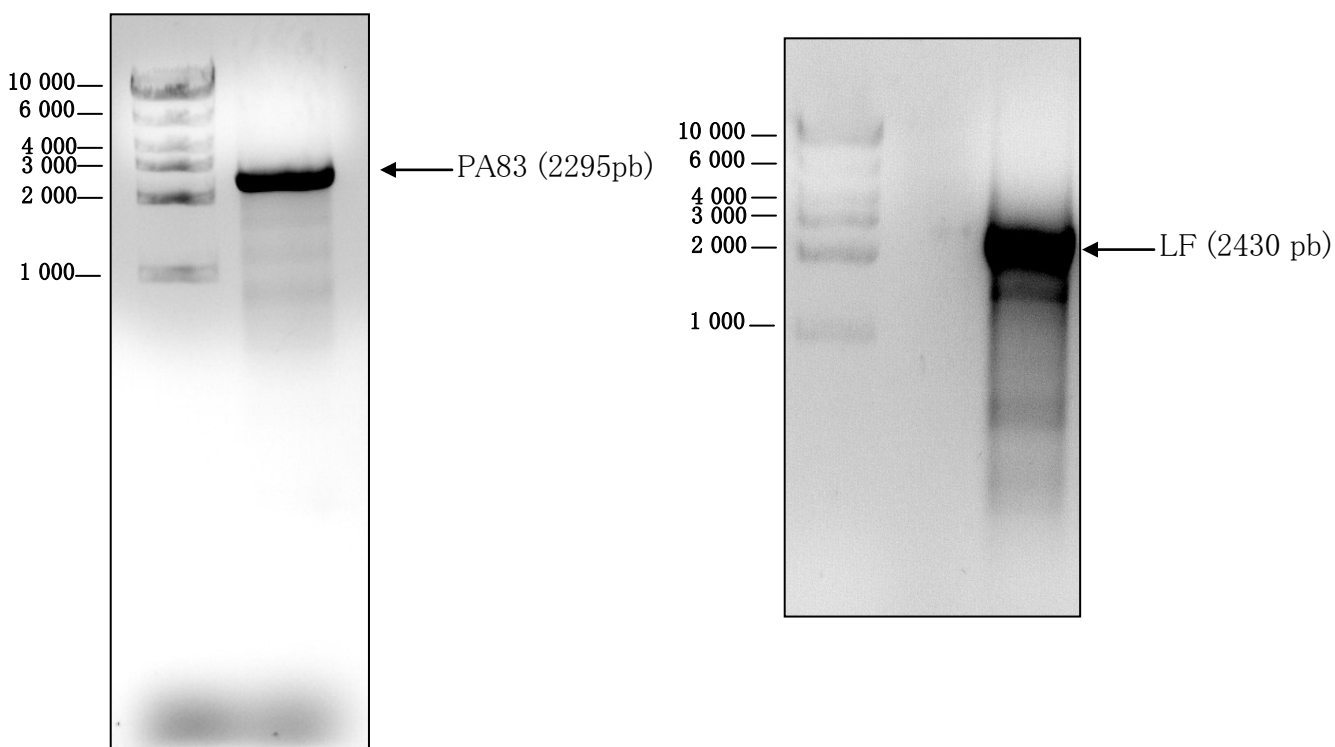
La molécule résultante de la simulation est un système tronqué au niveau des LFs et du  $\beta$ -barrel. Afin d'apprécier le mouvement des LFs complets, les LFs tronqués sont restaurés à leurs états complets par interpolation. Pour cela, des LFs originaux sont alignés sur les LFs tronqués. Cette interpolation ne permet pas de voir les mouvements intra-LFs. Cependant, les études précédentes par L. Alisarae montrent peu de mouvements à ce niveau là et peuvent ainsi être négligés.



## A. Clonage, Surexpression et Purification de PA63h et LF

### 1. Amplification par PCR

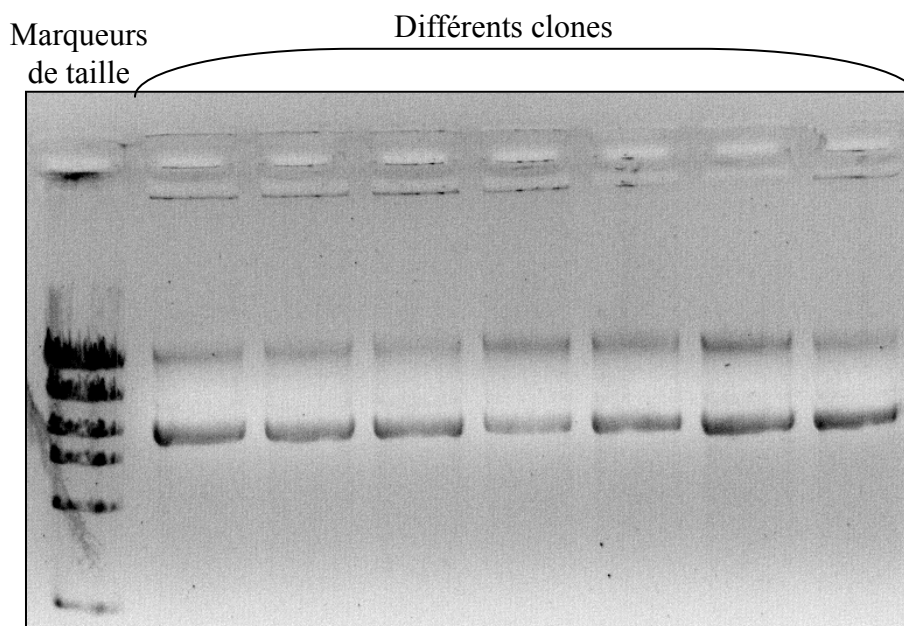
Une PCR a été réalisée pour amplifier les régions codantes de PA83 et LF à l'aide d'amorces spécifiques dessinées par FastPCR [36] afin d'avoir une forte spécificité et un site CACC dans l'amorce sens pour diriger le clonage dans le bon sens. L'analyse par électrophorèse sur gel d'agarose du résultat de la PCR (**Figure 23**) révèle une bande à la taille attendue d'environ 2295pb pour PA83 et 2430pb pour LF.



**Figure 24 : Analyse du produit de la PCR de PA83 et LF par électrophorèse sur gel d'agarose**

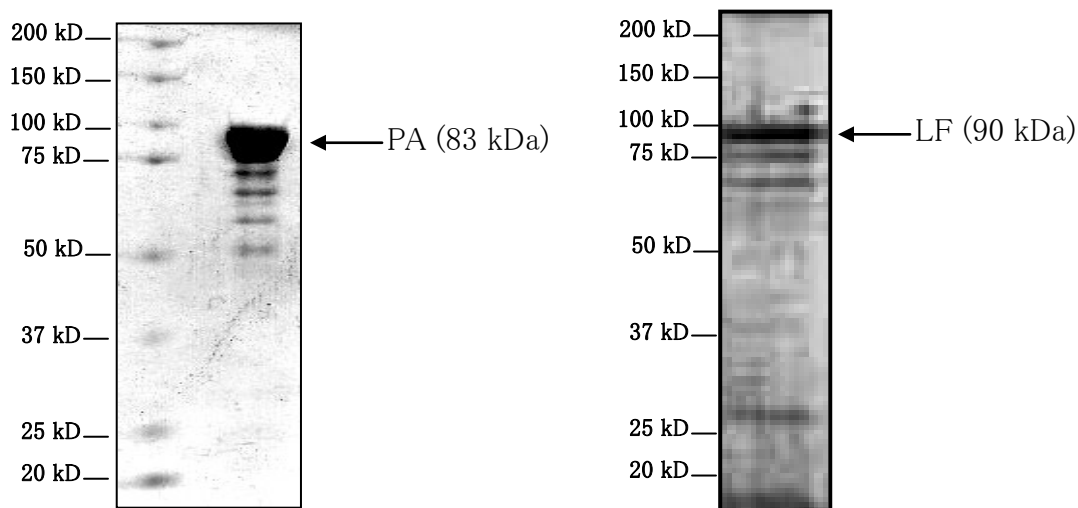
### 2. Clonage des produits de la PCR dans pET100

Afin d'exprimer PA et LF, les gènes, extrait du gel après la PCR, ont été cloné dans le vecteur de surexpression pET100. Une transformation des Bactéries (DH5alpha) suivie d'une culture des clones a été faite. Par la suite, une mini-préparation des plasmides (**Figure 25**) suivie d'une PCR, ont permis de vérifier la présence des régions codantes souhaitées. Avant d'entamer le clonage de ces gènes et la surexpression de PA83 et LF, un séquençage a été effectué afin de s'assurer de l'absence de mutations.



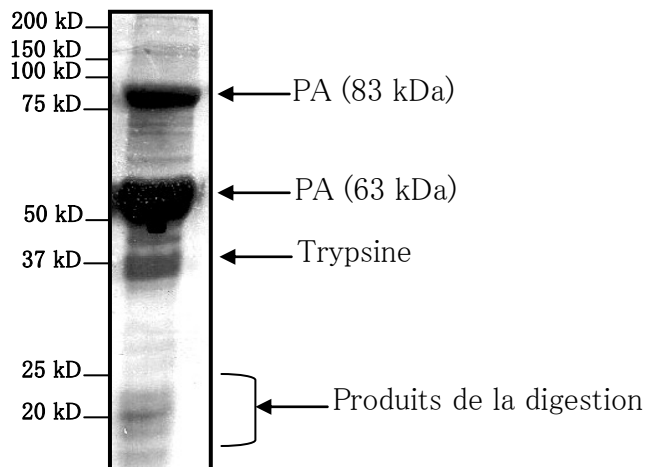
**Figure 25 : Gel d'Agarose montrant la mini-préparation des plasmides**

Une fois le séquençage est fait et les séquences sont vérifiées, les plasmides ont été utilisés pour transformer les bactéries (BL21). Après incubation et sélection sur milieu solide en présence d'ampicilline, de grandes cultures des clones positifs ont été faites avec induction par l'IPTG afin de surexprimer fortement PA83 et LF. Après collecte et lyse des cellules, PA83 et LF avec un his-tag au C-terminal ont chacun été purifiés par chromatographie d'affinité (Ni-sépharose). Les produits de la purification ont été vérifiés par SDS-PAGE (SDS-Polyacrylamide Gel Electrophoresis) (**Figure 26**). On voit dans les gels respectifs une bande dominante à 83 kDa pour PA et une bande de 90 kDa pour LF.

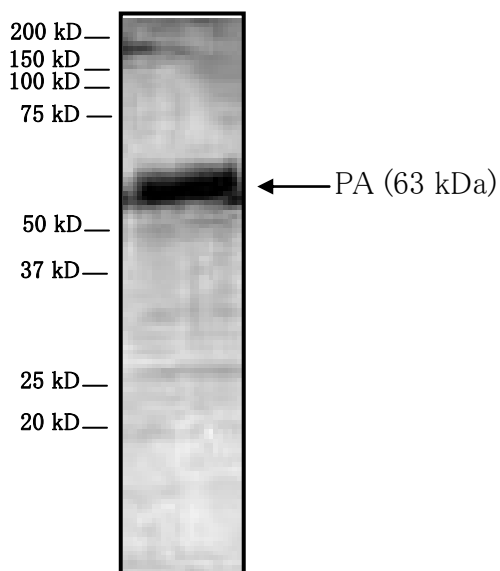


**Figure 25 : Analyse par SDS-PAGE de la purification de PA83 et LF**

Pour obtenir PA sous forme d'un heptamère, une digestion de PA83 était nécessaire. La trypsine digère PA83 en libérant un fragment de 20 kDa qui est lui-même digéré en plusieurs fragments (**Figure 27**). L'élimination de PA20 a été faite en filtrant les produits de la digestion avec des centricons avec un cut-off de 50 kDa (**Figure 28**).



**Figure 27 : Analyse par SDS-PAGE de la digestion de PA83 par la Trypsine**



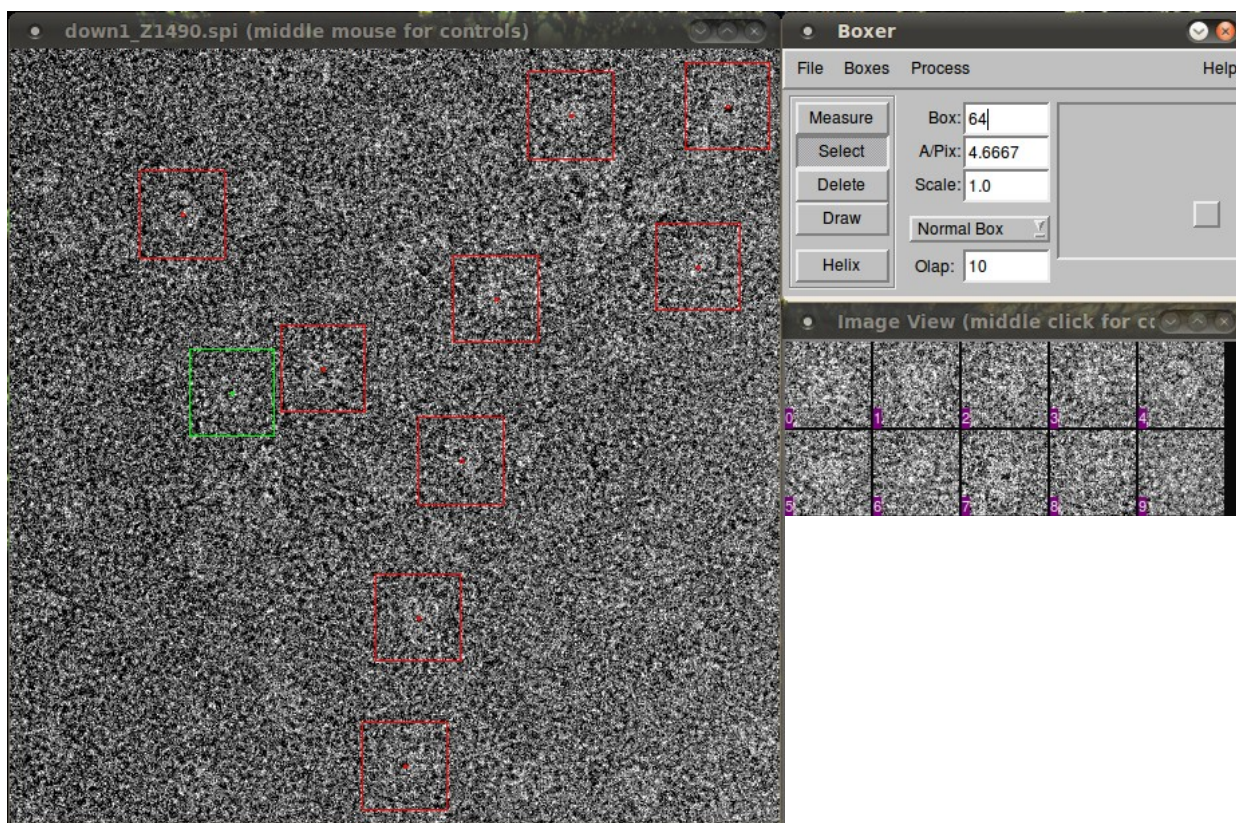
**Figure 28 : Analyse par SDS-PAGE de la purification de PA63**

En conclusion de cette partie, le clonage a permis d'isoler et d'insérer les gènes codant PA83 et LF dans le vecteur d'expression pET100 puis la surexpression et la purification de PA83 et LF ont permis d'obtenir des protéines pures de haute homogénéité.

## B. Reconstruction de PAh-prepore-3LF

### 1. Micrographes et particules

Les particules ont été sélectionnées manuellement par Bsc. Avi Biswas avec le logiciel boxer de la suite d'EMAN1. Chaque micrographe contient en moyenne 550 particules pour un total d'environ 53 000 particules. Celles-ci ont pour dimensions 64x64 (**Figure 29**).

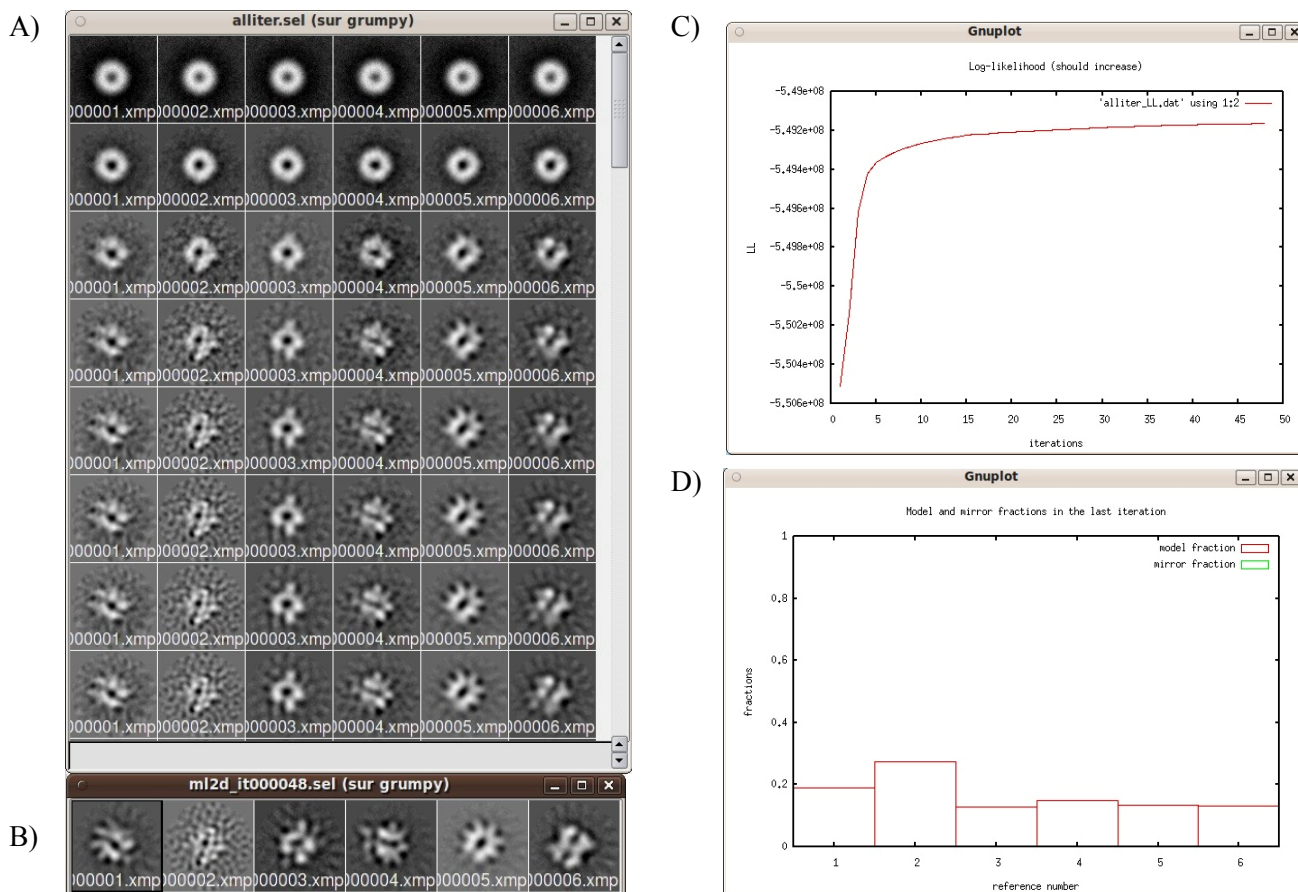


**Figure 29 : Micrographe et particules sélectionnées dans Boxer**

Les images sélectionnées apparaissent dans la fenêtre de droite. La fenêtre en haut à droite permet de se déplacer dans le micrographe et d'avoir accès aux différentes fonctions.

## 2. Classification 2D

Une classification 2D à 6 grains a été effectuée. Les résultats sont présentés dans la **figure 30**.



**Figure 30 : XMIPP ML2D**

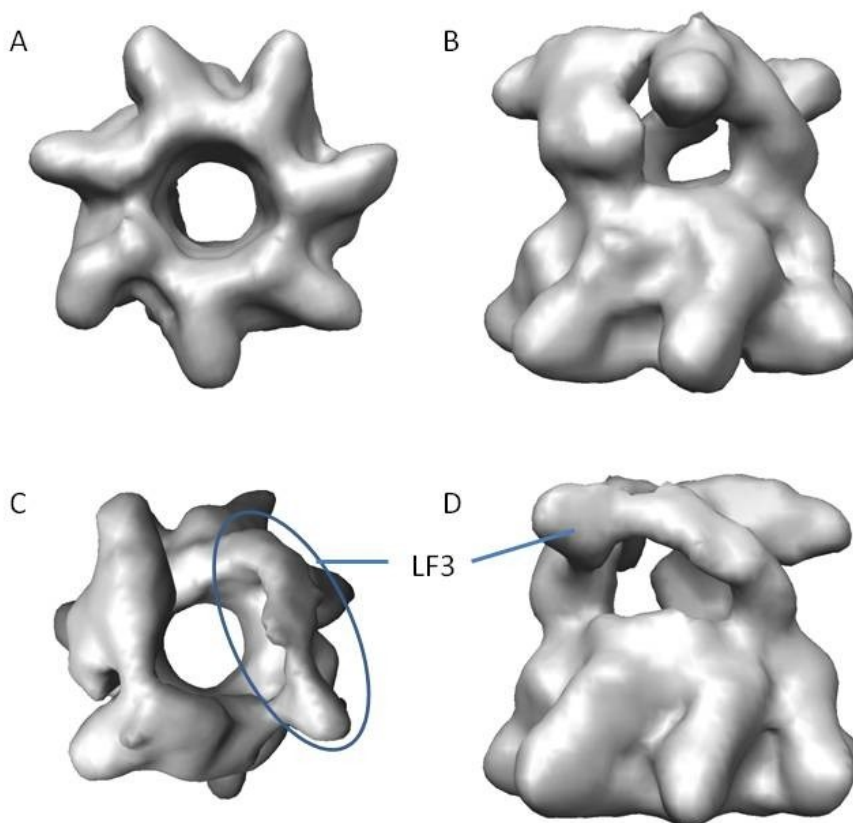
- A) Représentation des huit premières itérations de la classification 2D.
- B) Les classes projections obtenues à la fin de la classification hiérarchique 2D
- C) Graphique représentant le nombre d'itérations en fonction du log de maximum de vraisemblance
- D) Graphique représentant la proportion en pourcentage des images utilisées pour chacune des classes.

Cette classification met en évidence six projections. Celles-ci contiennent respectivement 9967, 14669, 6868, 7819, 7175 et 6949 particules. La cinquième classe est associée à la vue de dessus. La sixième est celle de profil. Les autres classes correspondent à d'autres orientations. La classe numéro 2 est dominée par le bruit ce pourquoi elle contient plus de particules. Le diagramme représentant la répartition des images par vue démontre de la surproportion d'images dans ce groupe. Cela peut s'interpréter par le fait que la classe 2 n'était pas bien définie à cause d'un mélange de plusieurs projections.

D'après le graphique (Figure 27) représentant le nombre d'itérations en fonction du log de maximum de vraisemblance, la classification 2D est complète. La majorité du travail a été faite avant la cinquième itération. Par la suite, on observe un état stationnaire (palier) depuis l'itération 15 jusqu'à l'itération 50.

### 3. Le volume de référence

Le volume de référence est à une résolution de 22Å. La base du volume représente l'heptamère du Protective Antigen (structure en étoile en vue de dessous). Celui-ci est régulier et montre une symétrie de sept. Au-dessus de cette structure sont liés trois ligands Lethal Factor. Les trois LFs ne sont pas symétriques. Alors que deux d'entre eux semblent similaires, le troisième est mal défini (LF3) (**Figure 31**).



**Figure 31 : Isosurface du volume de référence**

A. Vue de dessous

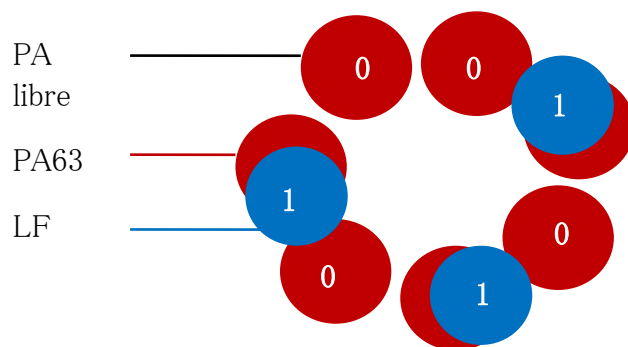
B. Vue de profil

C. Vue de dessus

D. Vue de profil centrée sur le LF3

Le LF mal défini (LF3) a été encerclé sur la vue de dessus (en C) et désigné sur la vue de profil D

Les LFs sont disposés sur les PAs de la manière suivante : chaque LF s'attache à 2 PA et puisqu'il y a 3 LF, il reste un PA libre (**figure 31**).

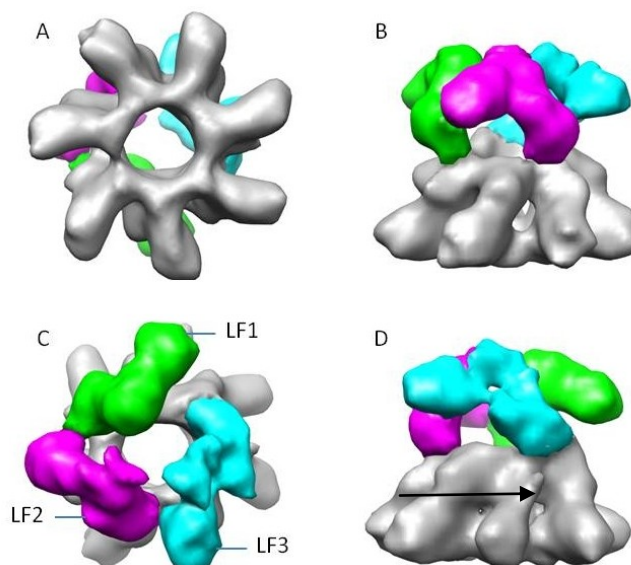


**Figure 32 : Position des LFs par rapport aux PAs**

Si chaque PA libre est représenté par un "0" et chaque PA-LF par un "1", on obtient la suite binaire circulaire "0010101".

#### 4. Reconstruction 3D avec XMIPP

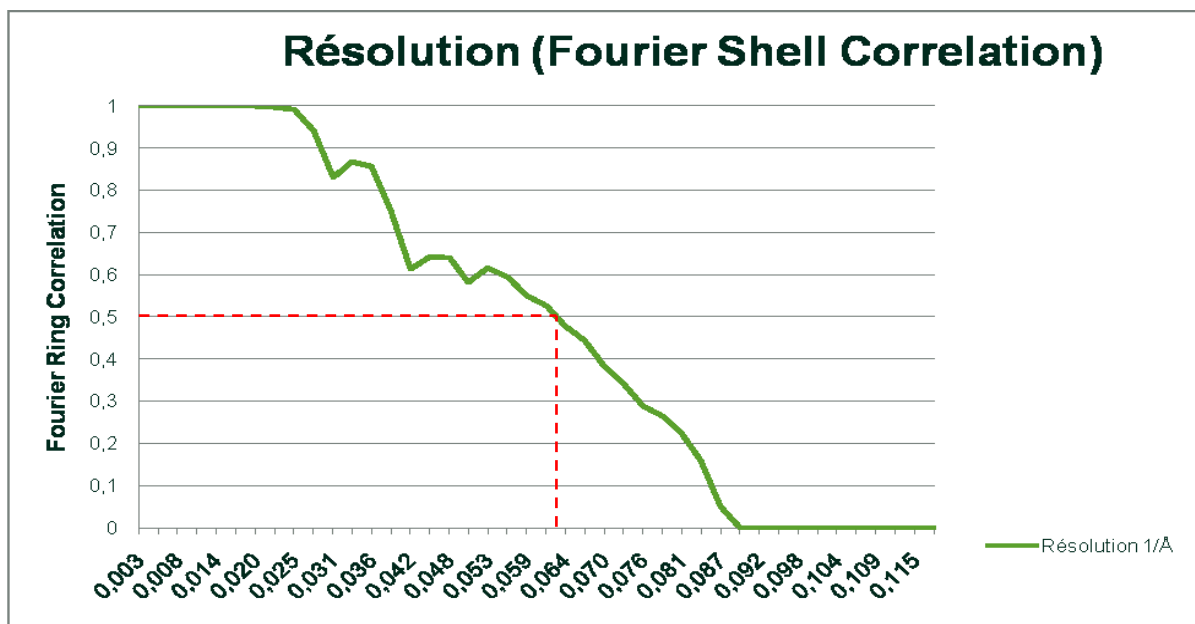
En utilisant le volume du Dr. Rouiller comme référence, une première reconstruction 3D par raffinement a été faite. Un nouveau volume de 16Å de résolution a été obtenu.



**Figure 33 : Isosurface du complexe PA63h-LF filtrée à 16Å segmenté**

- A) Vue de dessous. Les sept bras désignent PA63h-pré-pore
- B) Vue latérale. Au dessus de PA63h se trouvent trois ligands LF colorés de manière différente
- C) Vue de dessus du complexe
- D) Vue latérale sur laquelle on aperçoit un LF mal défini
- E) La flèche montre un trou au centre du LF3

D'après le graphique représentant la "Fourier Shell Correlation", le volume atteint une résolution de 16 Å. Son diamètre approximatif est de 110 Å.

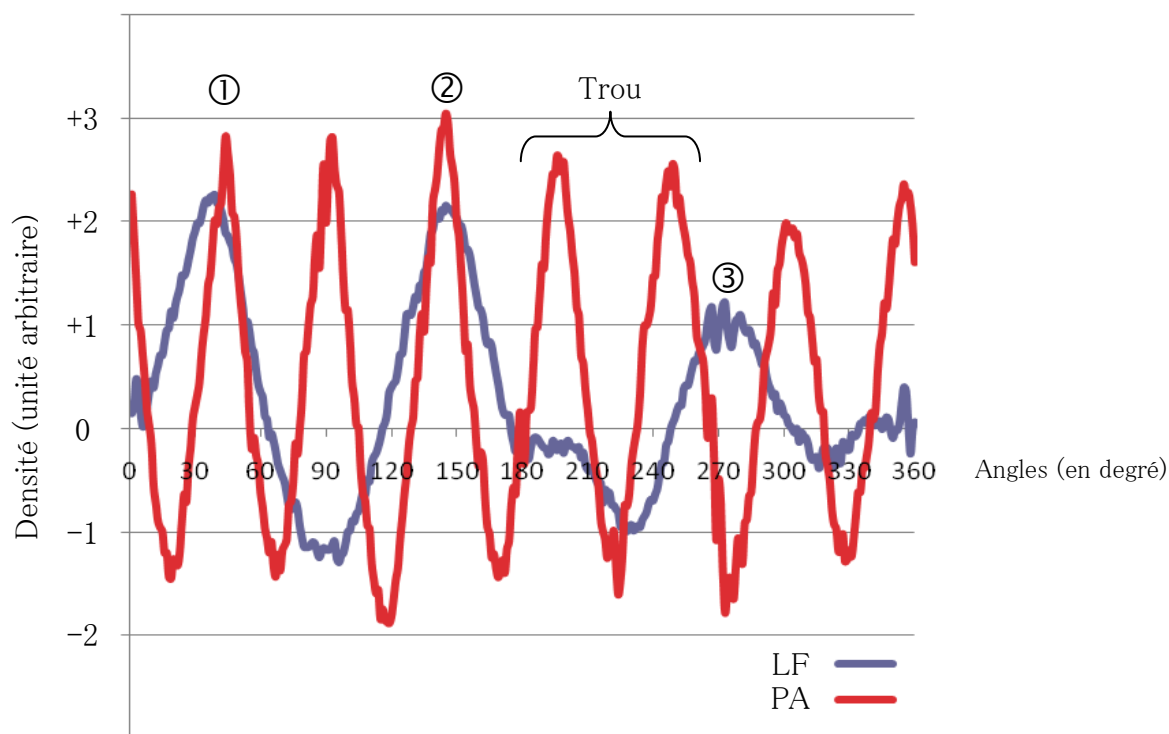


**Figure 34 : Fourier Shell Correlation de la Carte 3D du complexe PA63h-LF**

Dans l'ensemble, les densités du volume sont mieux définies que celles du volume de référence (but du raffinement). D'après la vue de dessus, les 3 LFs ne sont pas identiques. Ainsi, deux d'entre eux sont parfaitement définis et seront nommés LF1 et LF2. Le troisième semble plus grossièrement défini (flou) et prendra pour identité LF3. En effet, les densités en milieu de LF3 sont mal définies et la base de LF3 (jonction avec PA) est plus large que celle des deux autres LFs (**figure 33**).



## 5. Etude de la densité angulaire



**Figure 35 : Graphique de la densité angulaire**

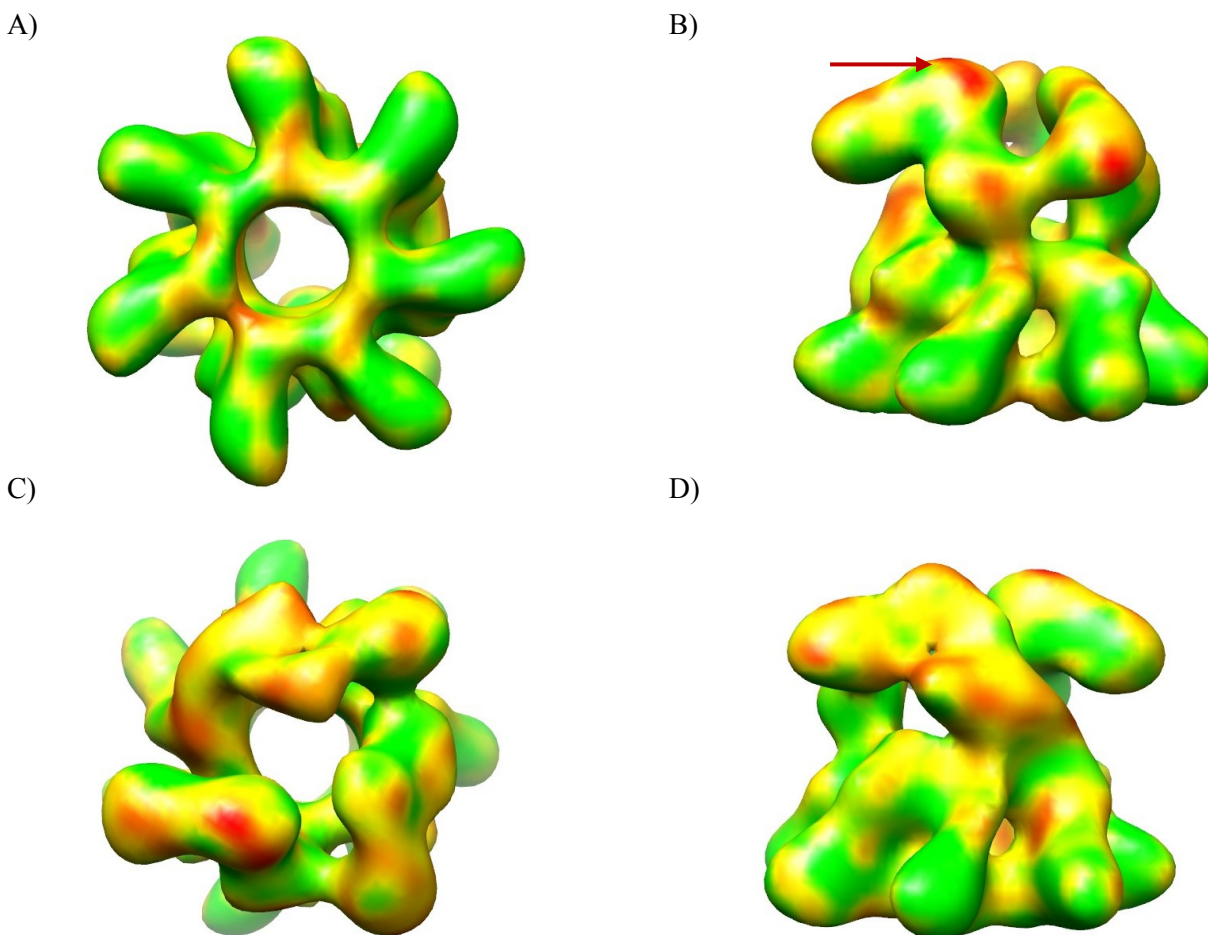
Cette figure représente la densité angulaire calculée à deux hauteurs différentes par rapport à l'axe Z (voir **figure 23**). Les LF1 et LF2 sont en phase avec les PAs. LF3 est à cheval entre deux PAs.

D'après le graphique de la densité angulaire (**figure 35**), on retrouve la correspondance principale d'un LF par PA pour les LF1 et LF2. Cependant, LF3 est situé entre deux PA.

Les différences morphologiques et positionnelles de LF3 par rapport aux deux autres LFs pourraient être dues à la présence de plusieurs conformations du LF3. En effet, le mélange de deux conformations apporte généralement une zone floue. Pour tester cette hypothèse, une classification 3D a été faite, mais avant cela, la variance 3D a été observée.

## 6. Variance 3D du volume

Afin de caractériser avec plus de précision la zone mal définie, une variance 3D a été calculée.



**Figure 36 : Variance 3D de la carte 3D du complexe PA63h-LF**

La variance est exprimée en gradient de couleur allant du vert pour une variance faible à rouge pour une variance élevée.

- A) La vue de dessous ne montre pas de variance significative au niveau de PA63h
- B) La vue latérale montre une variance au niveau des extrémités des LFs ainsi qu'au niveau de leurs pointes.
- C) La vue de dessus montre une variance significative au niveau des LFs
- D) La vue latérale D montre que LF3 est entièrement soumis à une variance non-négligeable  
Flèche rouge : "une pointe".

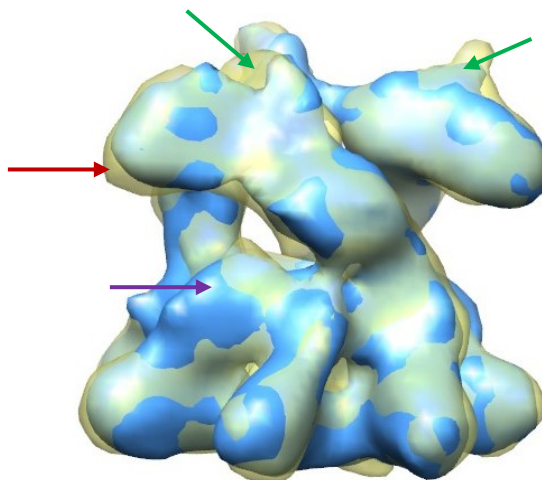
Les zones colorées en rouge représentent les "points chauds" c'est-à-dire, là où d'une projection à l'autre la correspondance diffère (**figure 36**). Alors que la variance est faible dans la base de PA63h, elle est plus élevée dans les densités correspondantes aux LFs. D'abord, tous les LFs présentent des zones de variance au niveau de leurs extrémités ainsi qu'au niveau de leurs

pointes.

La variance est particulièrement élevée pour les densités correspondant au LF qui apparaît flou. En effet, les LFs bien définis possèdent des variances très locales, ce qui suggère de petits mouvements moléculaires. Par contre, le LF3 possède une variance globale laissant imaginer plusieurs conformations. Pour avoir confirmation du phénomène, une classification 3D a été performée.

## 7. Classification 3D

N'ayant aucune idée du nombre de conformations différentes contenues dans nos images, une classification avec 3 graines (seeds) a été calculée pour commencer, bon compromis temps de calcul. Le résultat apporte deux volumes identiques ainsi qu'un volume différent. Afin d'améliorer la différenciation, une deuxième classification a été faite en prenant seulement 2 graines. Les volumes résultants sont montrés ci-dessous :



**Figure 37 : Les deux conformations de PA63h-LF obtenues par ML3D**

En jaune transparent : conformation 1 correspondante à la conformation couchée

En bleu plein : conformation 2 correspondante à la conformation relevée

La flèche rouge marque la majeure différence entre les deux conformations au niveau du LF3.

Les flèches vertes montrent des différences au niveau des « pointes »

La flèche mauve désigne le PA63 "libre" c'est à dire le PA n'étant pas en contact avec un LF.

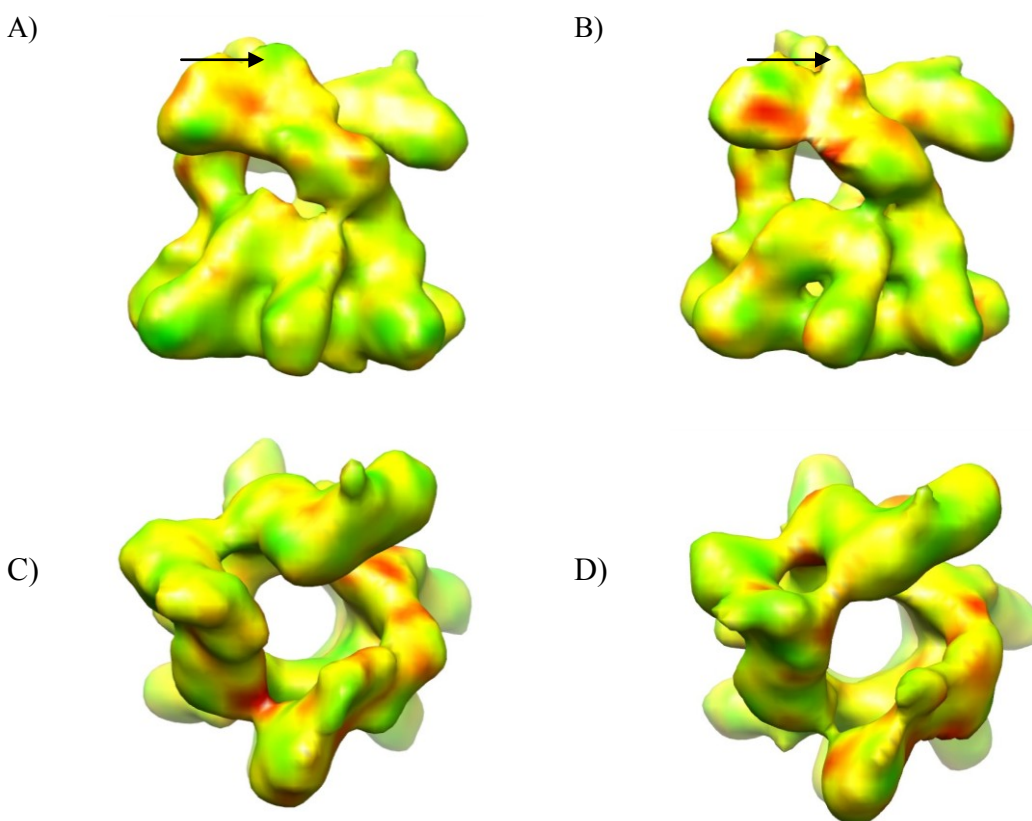
La **figure 37** montre deux volumes dans lesquels le troisième LF a une forme plus similaire avec ses deux homologues. La superposition de ceux-ci laisse voir deux conformations. En effet, l'une semble plus "couchée" que la deuxième que l'on qualifiera de "relevé". Tout le reste de la carte 3D se superposent parfaitement (les densités correspondantes au reste du

complexe sont superposables).

En effet, tous les LF sont fixés à deux PA adjacents. Il parait donc évident, contenu de la structure heptamérique de la base, qu'un PA se retrouve "seul". Cette observation peut servir de base à une hypothèse stipulant que la présence de ce PA "libre" autorise une flexibilité au niveau du LF sous-jacent. En dernière observation, on peut voir que la pointe (voir extrémités des flèches vertes sur la **figure 37** des LFs adaptent également deux conformations. Cela est en accord avec la variance 3D calculée précédemment. La superposition des volumes permet de mettre en évidence une différence unique au niveau du LF3. En effet, le volume 1 montre un LF relevé qui sera appelé « stand up » et le volume 2 un LF couché : « lying ».

### 8. Variance 3D des volumes obtenus par ML3D

Le fait que la classification 3D en prenant 3 graines donne deux structures identiques et une différente suggère qu'il y a principalement 2 conformations.



**Figure 38 : Variance 3D des deux volumes obtenus par ML3D**

A) Vue latérale de la conformation 1

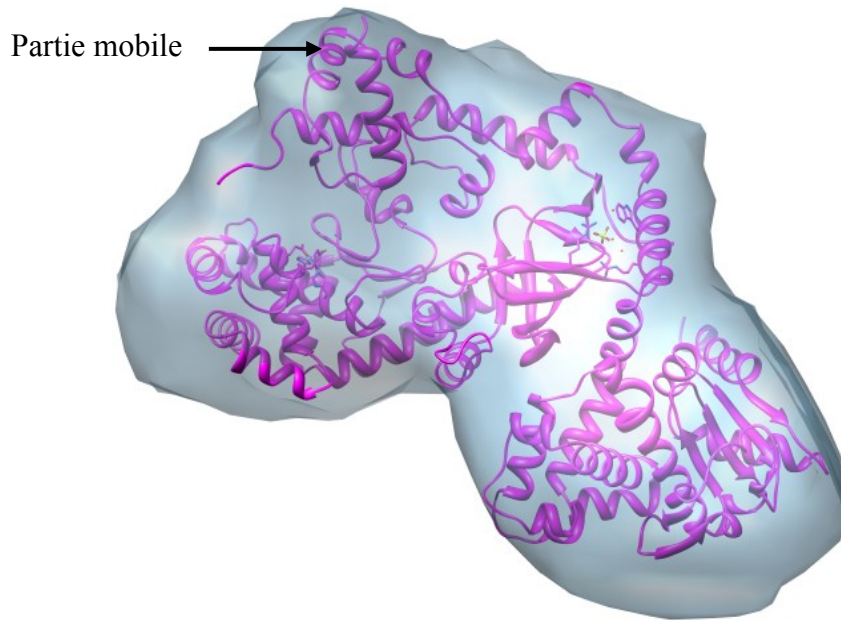
B) Vue latérale de la conformation 2

C) Vue de dessus de la conformation 1

D) Vue de dessus de la conformation 2

La variance est caractérisée par un gradient allant du vert pour un niveau faible au rouge un niveau élevée. Les flèches désignent les pointes.

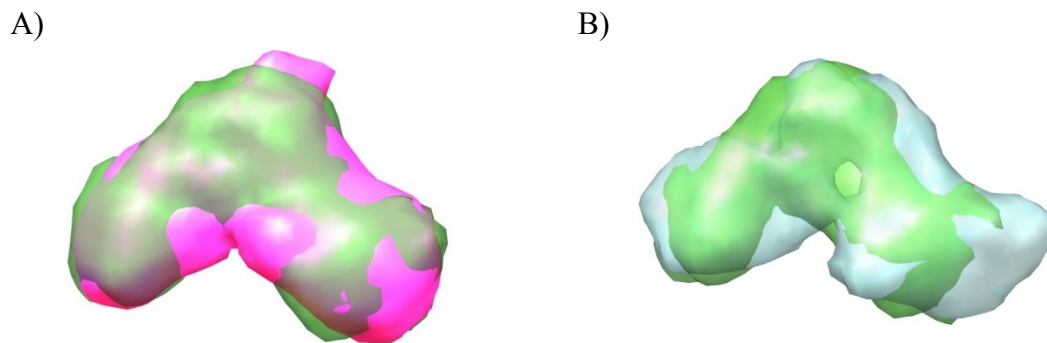
Comparativement à la variance initiale (**figure 36**), une amélioration est perçue. Cependant, cette amélioration des cartes obtenues après la ML3D n'est pas suffisante pour affirmer l'existence de deux conformations uniques. La conformation couchée possède une variance moins importante que celle relevée. De plus, la conformation couchée contient plus de projections que celle relevée ce qui peut biaiser l'interprétation de la variance. Par contre, les pointes répondent à notre hypothèse de départ (existence de deux conformations) en montrant deux positions différentes. La vue du dessus montre tout de même une nette amélioration de la variance en comparaison avec celle obtenue avec la carte 3D initiale. Par contre, on s'aperçoit que le haut du LF est rouge. Néanmoins, en regardant la structure de cristallographie par rayon X (pdb), on s'aperçoit que cette zone correspond à une hélice alpha mobile (**Figure 39**).



**Figure 39 : Amarrage de la structure pdb de LF dans une carte de densité**  
La structure de LF résolue par cristallographie à rayon X est représenté en magenta. La carte 3D correspondante est représenté en gris transparent. La flèche désigne la "partie mobile" de LF.

## 9. Isolation et superposition des trois LF

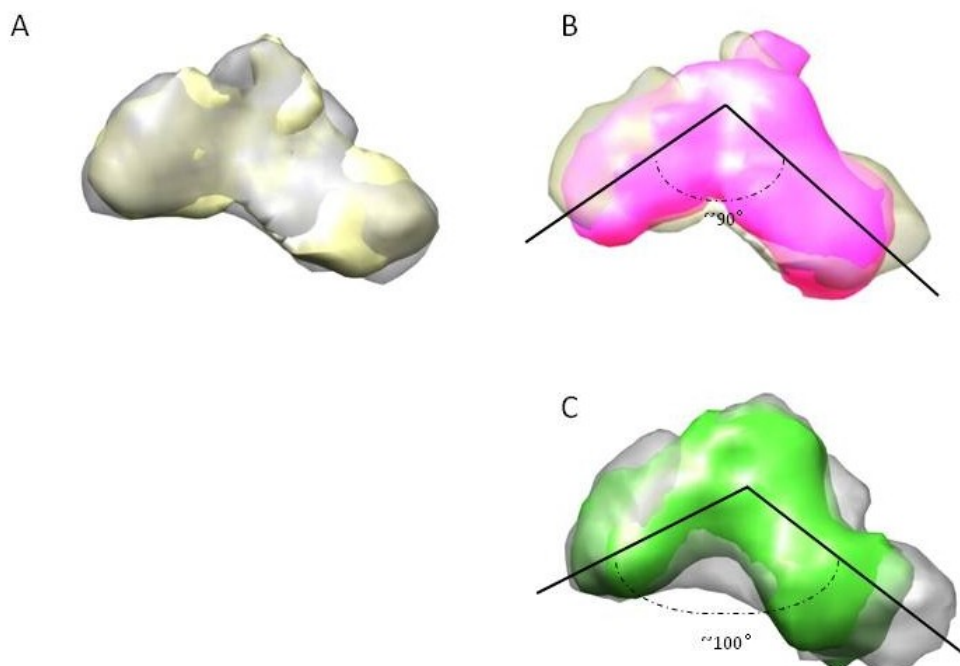
Les LF ont été isolés pour faciliter leur étude. Ici, on peut remarquer que les LF1 et LF2 se superposent parfaitement (**figure 40A**) alors que LF1 est différent de LF3 (**figure 40B**).



**Figure 40 : Superposition des LFs**

A) LF1 (vert) + LF2 (rose)

B) LF1 (vert) + LF3 (bleu)



**Figure 41 : Études des conformations de LF**

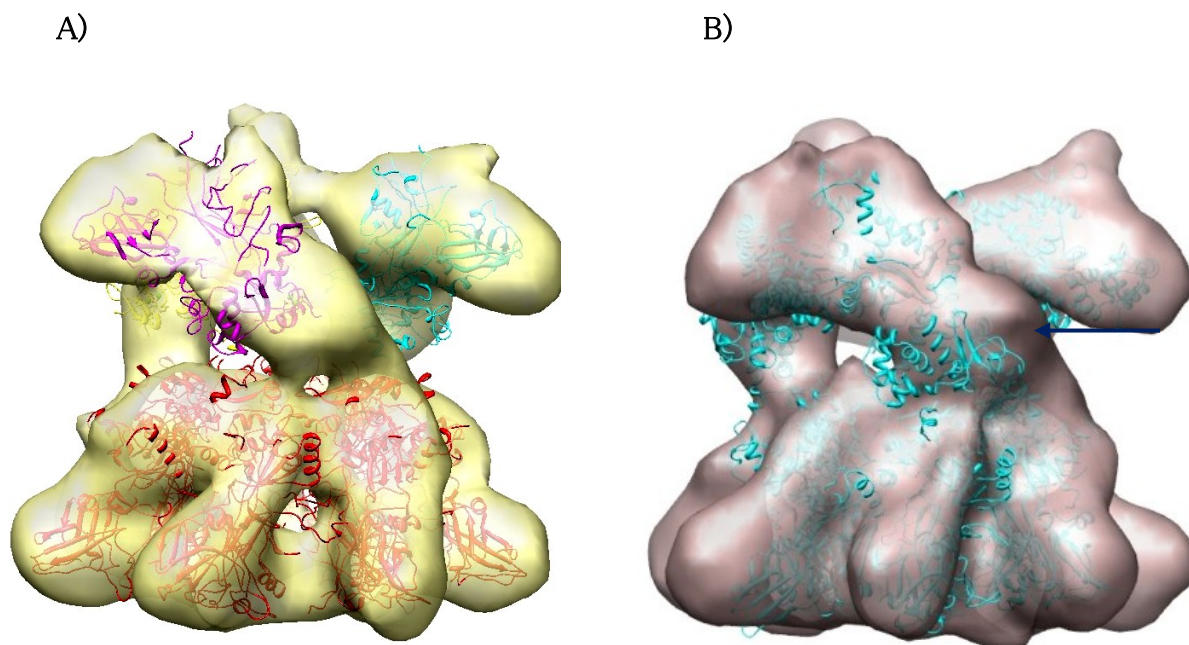
A) Superposition des deux conformations de LF3 (conformation couchée en gris et conformation relevée en jaune) ; B) Superposition de la conformation relevée de LF3 avec LF2 ; C) Superposition de la conformation couchée de LF3 avec LF2. LF1 possède un angle d'environ  $90^\circ$  alors que le LF3 de la conformation couchée semble avoir un angle d'environ  $100^\circ$ .

La **figure 41A** présente la superposition des deux conformations de LF3 qui sont visiblement différentes. En fait, la conformation relevée de LF3 se superpose avec celles de LF1 et LF2 (**figure 41B**). On observe un angle d'environ  $90^\circ$  caractérisant la version relevée. Dans le cas de la conformation couchée, on voit qu'elle présente plutôt un angle d'environ  $100^\circ$  et ne se superpose pas avec LF1 (**figure 41C**). Tout cela pourrait étayer l'hypothèse selon laquelle la présence de ce PA "libre" autoriserait une flexibilité au niveau du LF sous-jacent dans le sens où, l'espace créé par le PA libre déstabiliserait le LF sous-jacent.

## 10. Amarrage (Docking)

Une simulation par dynamique moléculaire du complexe PA63h-prepore-3LF (Dr Laleh Alisarae, postdoctorante au laboratoire) a mis en évidence des mouvements de LFs. Ainsi, deux structures ont été révélées.

La superposition des structures de Dr. Alisarae et des volumes issus de la classification 3D donne une bonne corrélation. En effet, la conformation relevée s'adapte plutôt bien avec le pdb de la structure obtenue à 0ns qui elle aussi contient un LF relevé alors que la conformation couchée s'adapte mieux avec le pdb de la structure obtenue à 10ns qui elle aussi contient un LF couché.



**Figure 42 : Amarrage des volumes avec les structures issues de dynamique moléculaire**

A) Conformation relevée amarrée le pdb de la structure obtenue à 0ns.

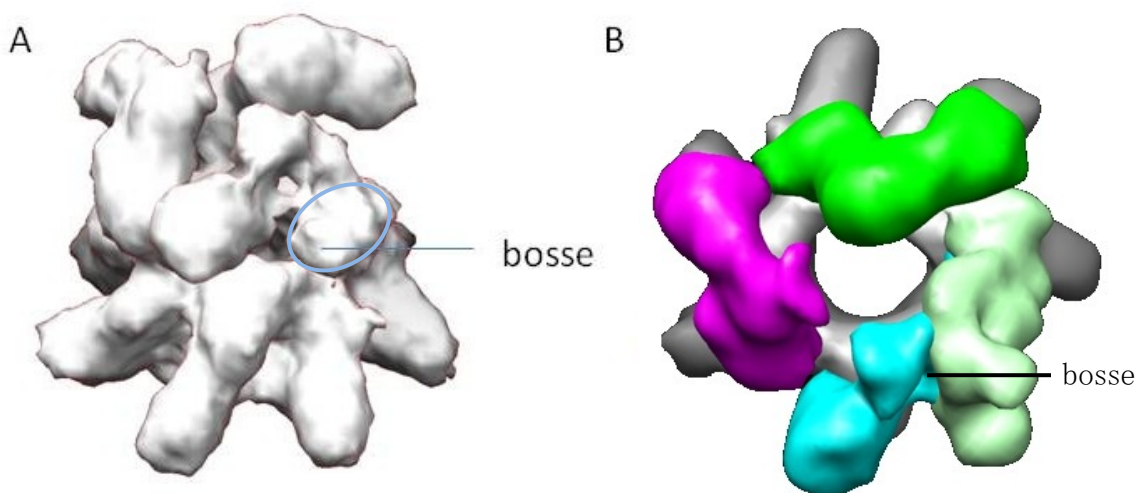
B) Conformation couchée amarrée avec le pdb de la structure obtenue à 10ns.

La flèche désigne le pied gras.

Les volumes A et B de la **figure 42** présentent un LF3 sans trou central (**voir figure 33**) mais contiennent tous les deux un pied gras. De plus, le volume initial possédait également un pied gras. Il est alors possible que ce soit un artéfact amplifié par alignement des particules.

### 11. Analyse du « pied gras » et mise en évidence d'une erreur d'alignement

Les résultats précédents focalisent sur la résolution de la mauvaise définition de LF3, mais n'explique pas la présence du pied gras. La superposition des volumes A et B de la **figure 42** peut être expliquée par une erreur de rotation de  $103^\circ$  (**Figure 43**). Ainsi, on peut voir qu'un LF « sort au niveau de la bosse » du LF correspondant.



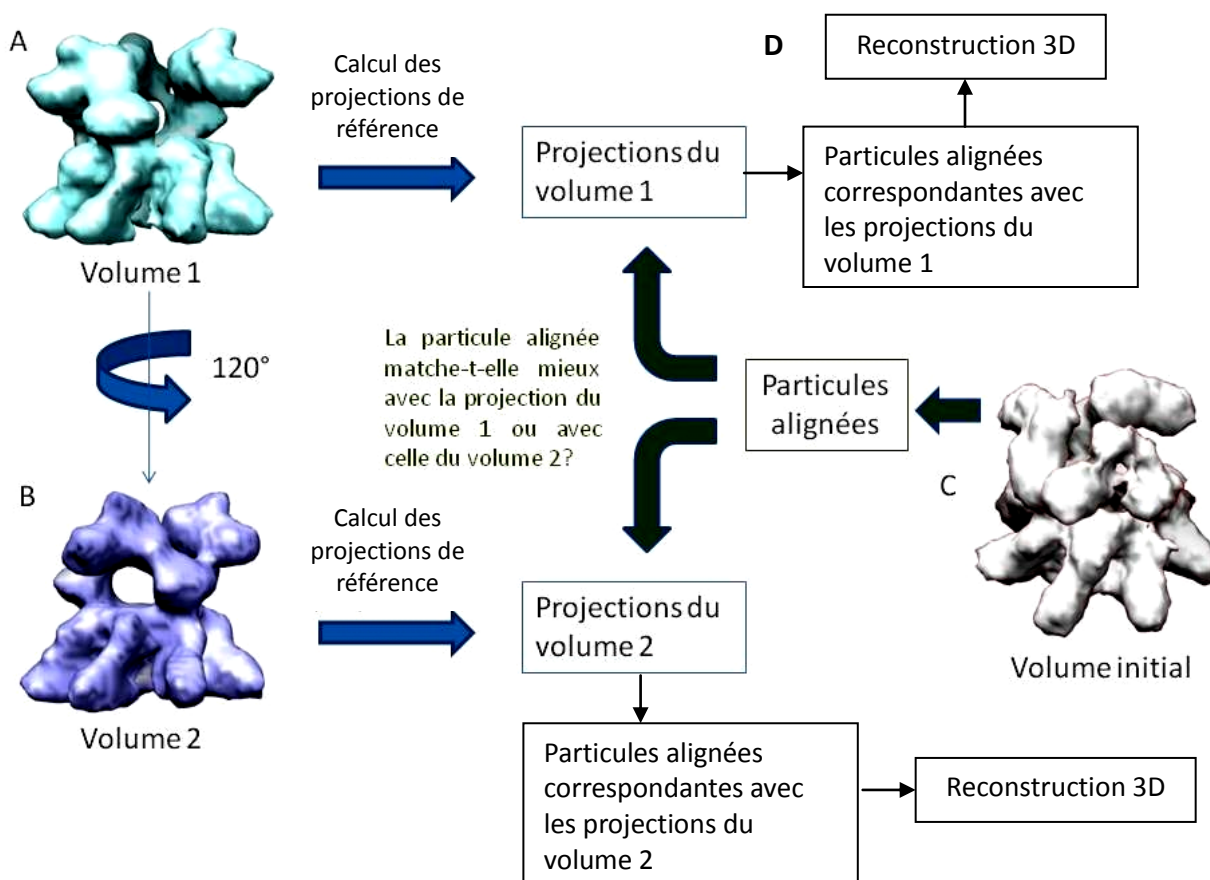
**Figure 43 : Mise en évidence du rôle de "la bosse" par superposition avec rotation de  $103^\circ$  du volume initial dupliqué avec lui-même**

- A) Volume obtenu par raffinement 3D
- B) Après duplication puis pivotement de  $103^\circ$  de la structure A, la superposition du nouveau volume avec le volume A montre un LF supplémentaire (en vert clair) au niveau de la bosse. (vue de dessus)

Le pivotement du volume de  $103^\circ$  donne un nouveau volume fils qui se superpose avec son volume père excepté au niveau du LF3 qui se chevauche avec le LF1 (vert clair). Pour examiner cette ambiguïté, une classification 3D particulière a été effectuée.

Lors des amarrages précédents, on a pu se rendre compte que la structure cristalline se superpose assez bien avec notre volume de raffinement. Ainsi, à partir du pdb, un volume d'une résolution de  $30\text{\AA}$  a été créé. Celui-ci contient donc trois LFs identiques et a été aligné avec le volume de raffinement (volume 1 de la **figure 44**). Par la suite, il a été dupliqué puis aligné avec l'erreur de  $103^\circ$  (volume 2 de la **figure 42**).





**Figure 44 : Schéma expliquant la mise en évidence d'une erreur d'alignement**

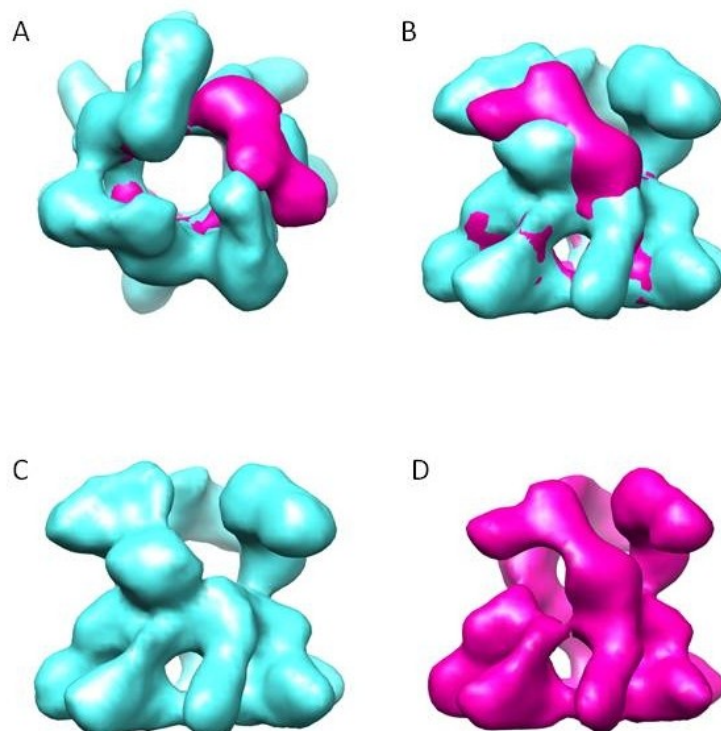
- A) Volume de référence en position 1
- B) Volume de référence en position 2 obtenu en tournant de 103° anti-horaire le volume 1
- C) Volume obtenu par raffinement 3D
- D) Reconstruction 3D à partir des particules alignées séparées

D'abord, un volume de référence (volume 1) a été déterminée à partir de complexe en structure cristalline (pdb). Celui-ci a été aligné avec le volume initial C "(pied gras)". Par la suite, le volume 1 a été dupliqué et tourné de 103° donnant le volume 2 . Les projections correspondantes aux deux volumes (1 et 2) ont été calculées. En gardant la correspondance entre les projections et les particules alignées (issues du volume initial), chaque particule alignée est comparée par cross-corrélation à sa projection correspondante provenant du volume 1 ou du volume 2.

Les projections des deux volumes ont été calculées. Ainsi, les particules expérimentales déjà alignées lors du raffinement ont pu être comparé avec les projections de chacun des deux volumes.

Les résultats montrent deux groupes de particules : l'un correspond pour au volume 1 et l'autre au volume 2. La reconstruction séparé des deux groupes obtenus donne deux volumes distincts ayant chacun un « pied fin » (**figure 45**). Seule la position du trou (PA qui n'est pas en contact avec un LF) change prouvant l'erreur d'alignement de 103°. Les particules étant à cheval

entre les deux positions du LF ont été éliminées (13654). Les deux volumes distincts contiennent 27655 et 12106 particules.

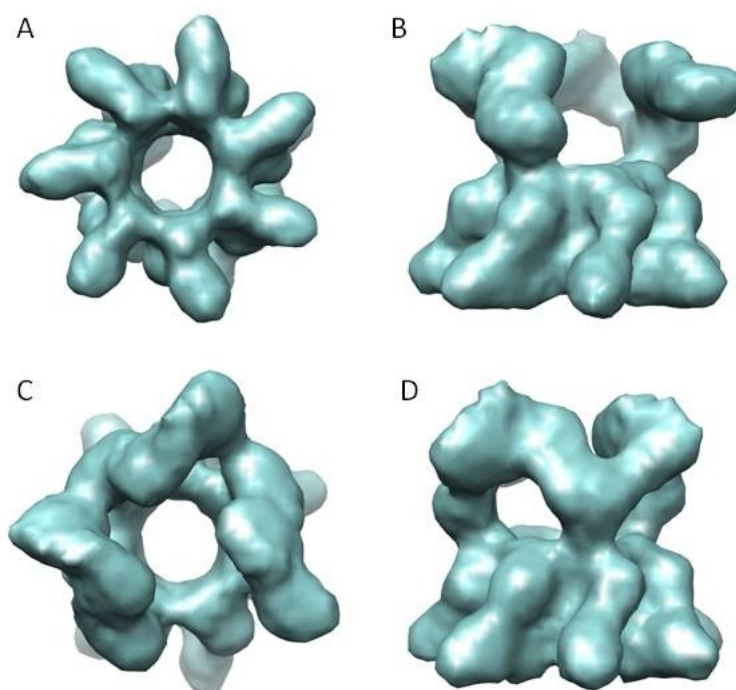


**Figure 45 : Résultats de l'analyse du « pied gras »**

- A) Superposition des reconstructions 3D avec les particules alignées du groupe 1 (en bleu) et du groupe 2 (en rose), vue de dessus
- B) Superposition des reconstructions 3D avec les particules alignées du groupe 1 (en bleu) et du groupe 2 (en rose), vue latérale
- C) Reconstruction 3D reconstructions 3D avec les particules alignées du groupe 1
- D) Reconstruction 3D reconstructions 3D avec les particules alignées du groupe 2

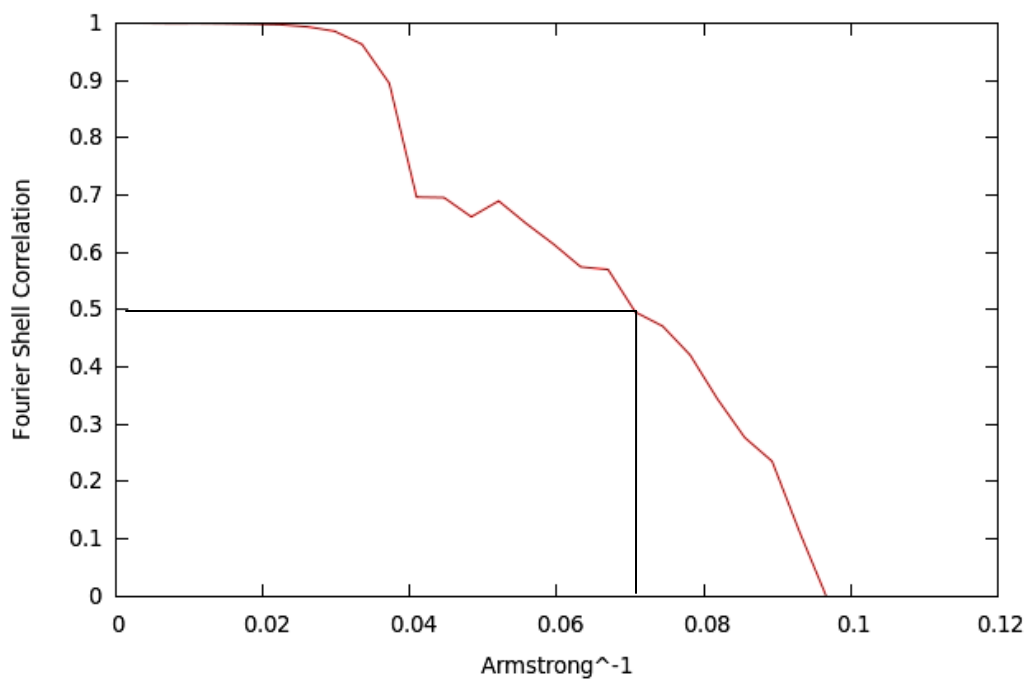
## 12. Raffinement final

Afin de prévenir l'apparition d'une nouvelle erreur d'alignement, il a été nécessaire de travailler sur le volume de référence. En effet, le volume de référence employé jusqu'à présent contenait déjà un "pied gras" provoquant une confusion au niveau de l'alignement des projections. Ainsi, un nouveau volume de référence a été créé en symétrisant les LFs. En utilisant cette nouvelle référence (filtré à 30 Å), le raffinement donne un nouveau volume (**Figure 46**) sans "pied gras". La résolution a été estimée par "Fourier Shell Correlation" à 0.5 et nous donne une résolution de 14Å (**Figure 47**).



**Figure 46 : Isosurface du volume final de PA63h-prepore-3LF à 14Å**

- A) Vue de dessous
- B) Vue latérale 1
- C) Vue de dessus
- D) Vue latérale 2

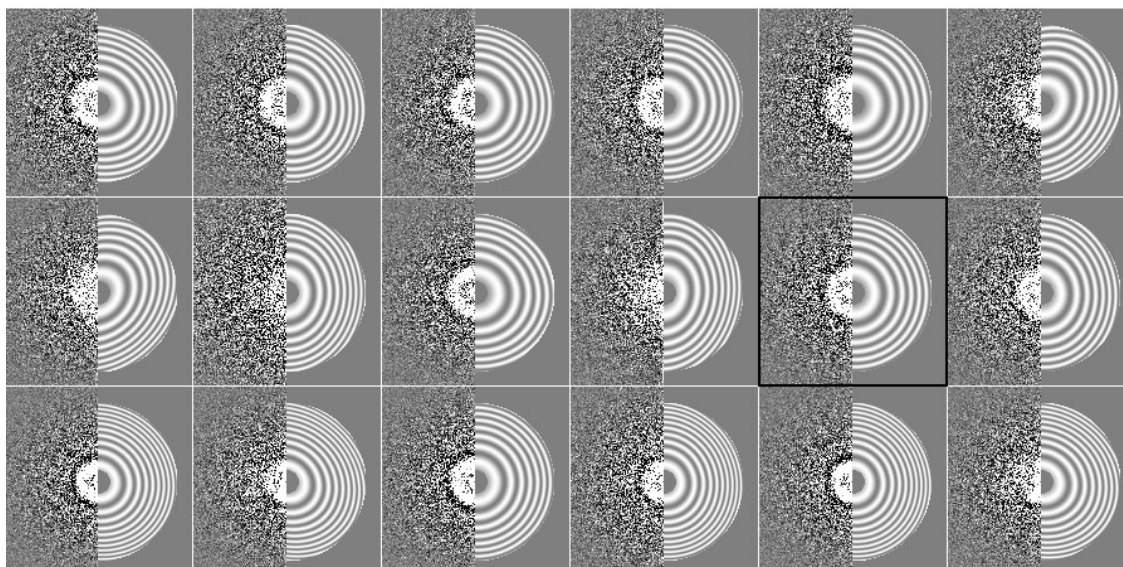


**Figure 47 : Fourier Shell Correlation de la Carte 3D finale du complexe PA63h-LF**  
 En accord avec le critère "0.5", la résolution estimée de la carte finale du complexe PA63h-LF est de  $1/0.07$  soit environ 14Å.

## C. Reconstruction de PA63h-pore-3LFs-ND

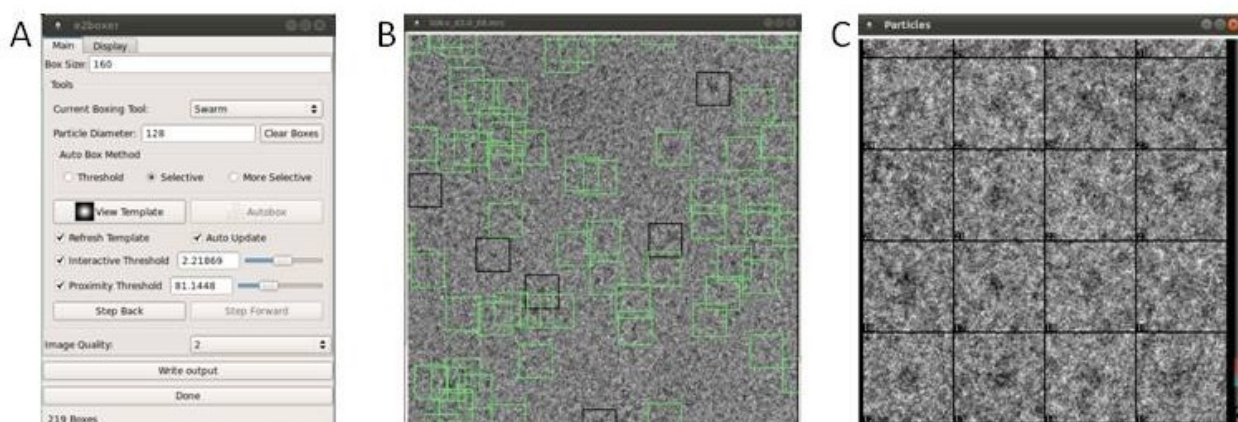
### 1. Images digitales et particules

150 micrographes de l'échantillon de PA63h-pore-3LFs-ND ont été capturés sur le microscope Tecnai F20 avec une magnification de 50kx. Les images ont été corrigées par inversion de phase CTF (CTF phase flipped) avant la sélection des particules. Un aperçu des transformées de Fourier des images avant et après correction montrent les anneaux de Thon (Thon rings).



**Figure 48 : Transformée de Fourier des micrographes avant et après correction par inversion de phase CTF**

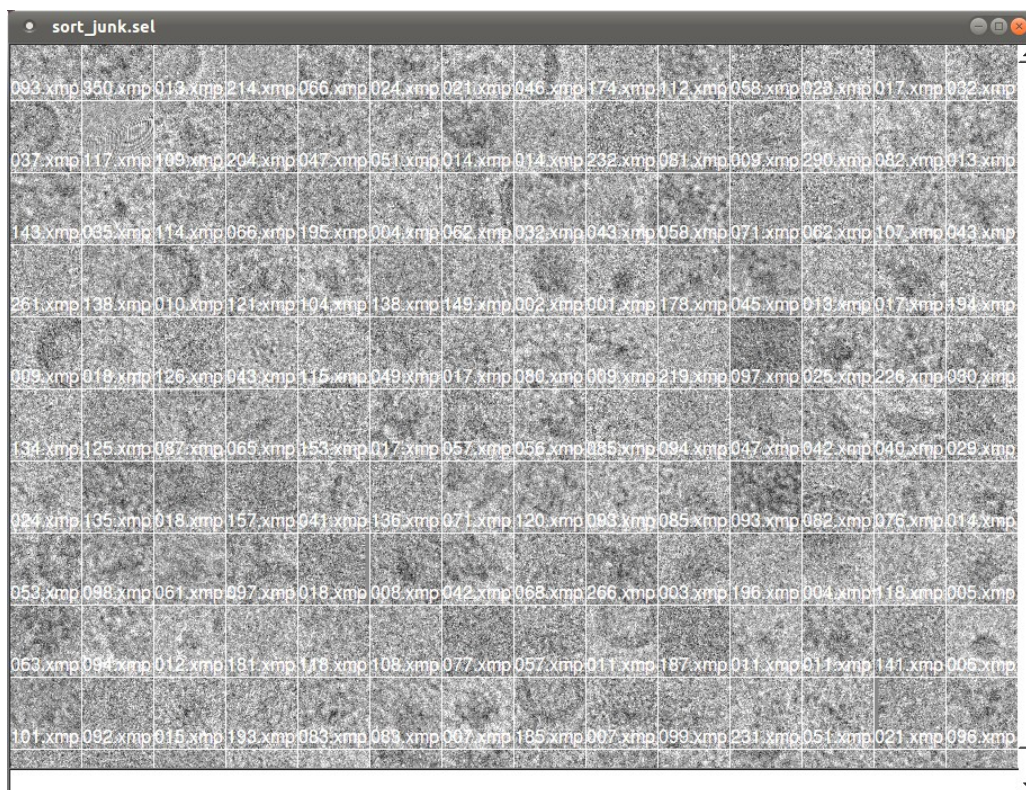
Les particules sont assez difficiles à voir à l'œil nu. De ce fait, il a été choisi de faire une sélection semi-automatique avec le logiciel Boxer de la suite Eman 2. En effet, en sélectionnant quelques particules à la main, le logiciel se crée une base (template) et va par la suite être capable de trouver de manière efficace les autres particules. Chaque micrographe contient en moyenne 220 particules pour un total d'environ 33 000 particules.



**Figure 49 : Images digitales et particules sélectionnées dans Boxer d'EMAN2**

- A) Interphase de Boxer d'EMAN2
- B) Micrographie avec particules sélectionnées
- C) Particules sélectionnées

Du fait de la sélection semi-automatique, certaines particules ne sont en fait que des débris de protéines. Ceux-ci ont été éliminés par une classification statistique pour un reste total d'environ 29 000 particules.

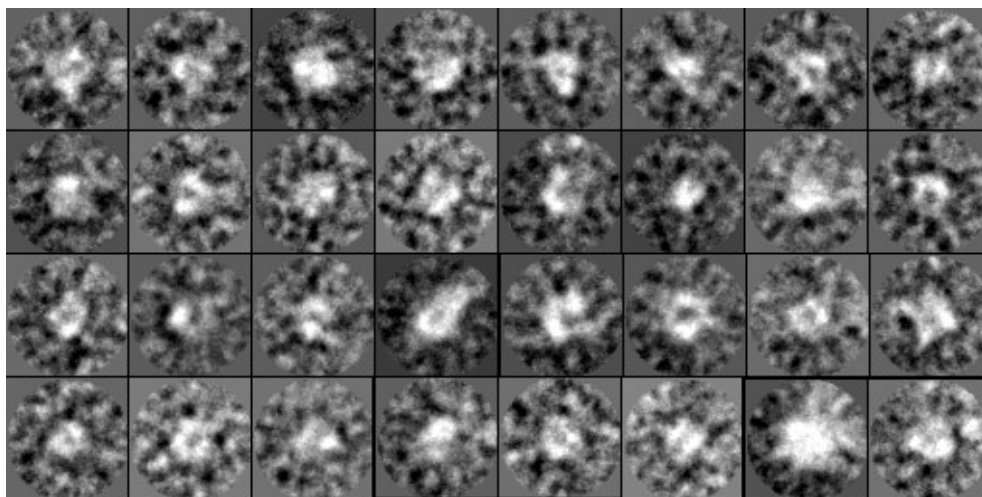


**Figure 50 : Aperçu des « mauvaises » particules éliminées par statistique**

Le complexe PA63h-pore-3LFs-ND étant volumineux, les particules ont pour dimensions 128x128. Pour les travaux préliminaires, des particules sous-échantillonnées en 64x64 ont été utilisées pour limiter le temps de calcul.

## 2. Classification 2D

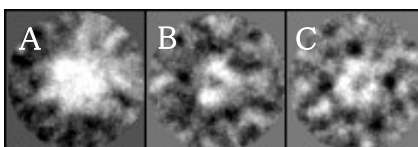
De la même manière que pour le complexe PA63h-prepore-3LFs, plusieurs classifications 2D ont été effectuées. Seule celle de 32 graines est présentée ici car elle contient des projections caractéristiques différentes (**figure 51 et 52**).



**Figure 51: Classification 2D de PA63h-pore-LFs**

Table représentant les projections finales obtenues par classification 2D. Chacune des 32 projections est une moyenne des projections (particules) la composant.

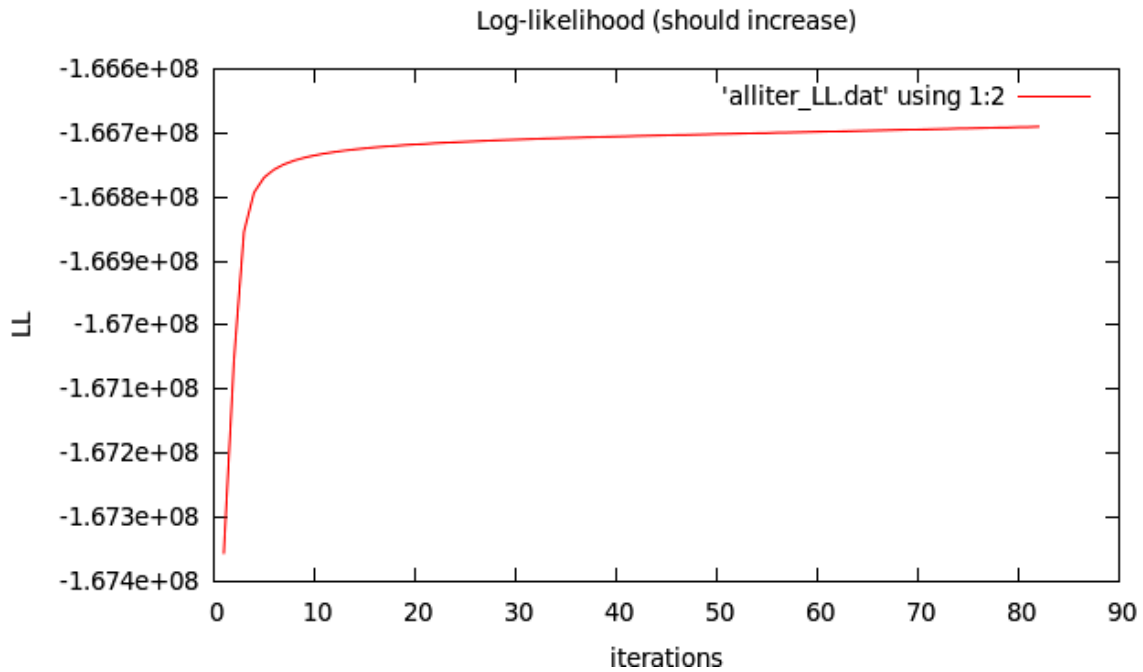
Une sélection des classes les plus caractéristiques a été faite ci-dessous.



**Figure 52: Sélection des projections les plus caractéristiques de la classification 2D de PA63h-pore-LFs**

- A) Classe représentant la projection de dessus
- B) Classe représentant projection latérale 1 (en ‘Y’, la base étant le  $\beta$ -barrel)
- C) Classe représentant projection latérale 2 (en ‘Y’)

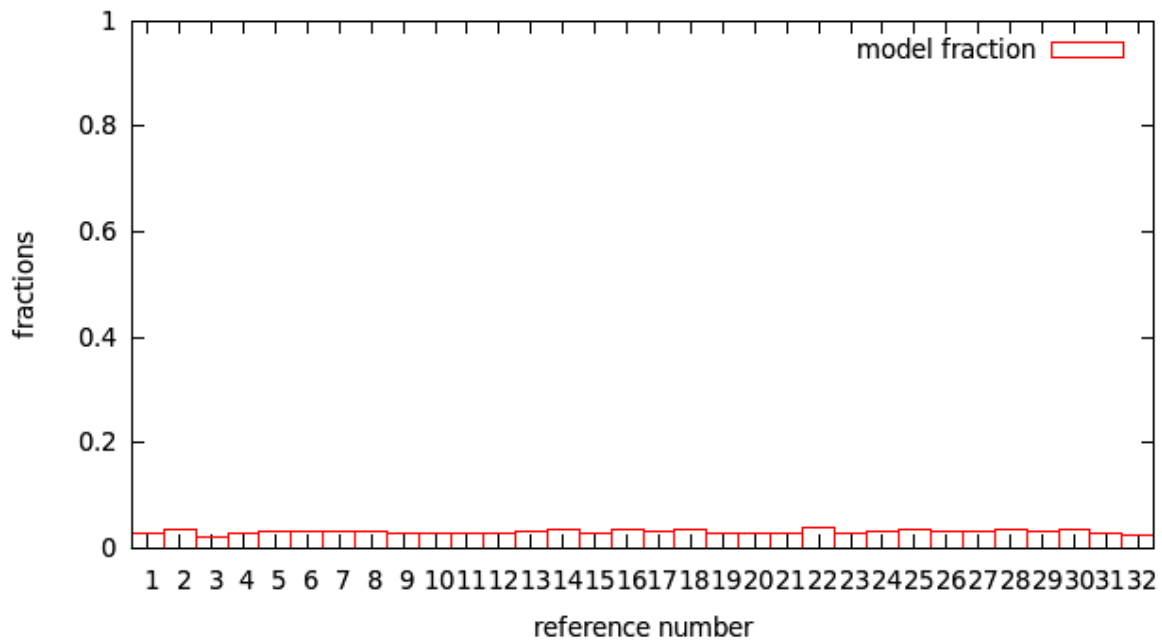
Ici, on peut voir les projections caractéristiques de dessus (en A) ainsi que des vues latérales (en B et C) en « Y ». D'après le graphique représentant le log de maximum de vraisemblance en fonction du nombre d'itérations, on peut affirmer que la classification 2D est complète. En effet, on atteint un plateau stationnaire (le palier).



**Figure 53 : Graphique représentant le log de maximum de vraisemblance en fonction du nombre d'itérations**

La phase ascendante est comprise entre les itérations 0 à 10. Par la suite, de 10 à 83, les maximum de vraisemblance atteint un plateau.

Les vues semblent équitablement réparties comme le montre le graphique suivant :

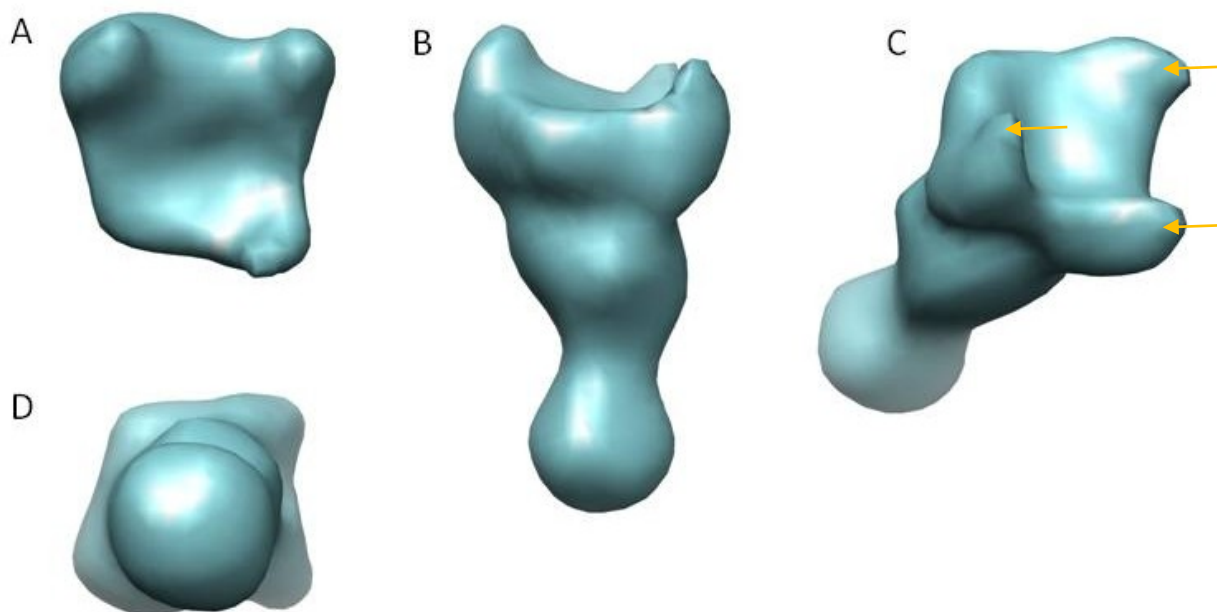


**Figure 54 : Graphique représentant la proportion en pourcentage des projections contenue dans chacune des 32 classes à la dernière itération**

Les différentes classes contiennent un nombre homogène de projections.

### 3. Création d'un volume de référence avec EMAN2

Un volume de référence a été créé en utilisant le théorème des lignes communes « common lines ». Après une vingtaine d'essais et la création de centaines de volumes de toutes les formes, certains d'entre eux se sont distingués par leur similitude avec le volume de PA63h-pore résolu par Katayama *et al.*, [26]. Le meilleur étant le volume suivant :



**Figure 55 : Isosurface du volume de référence obtenu par « common lines »**

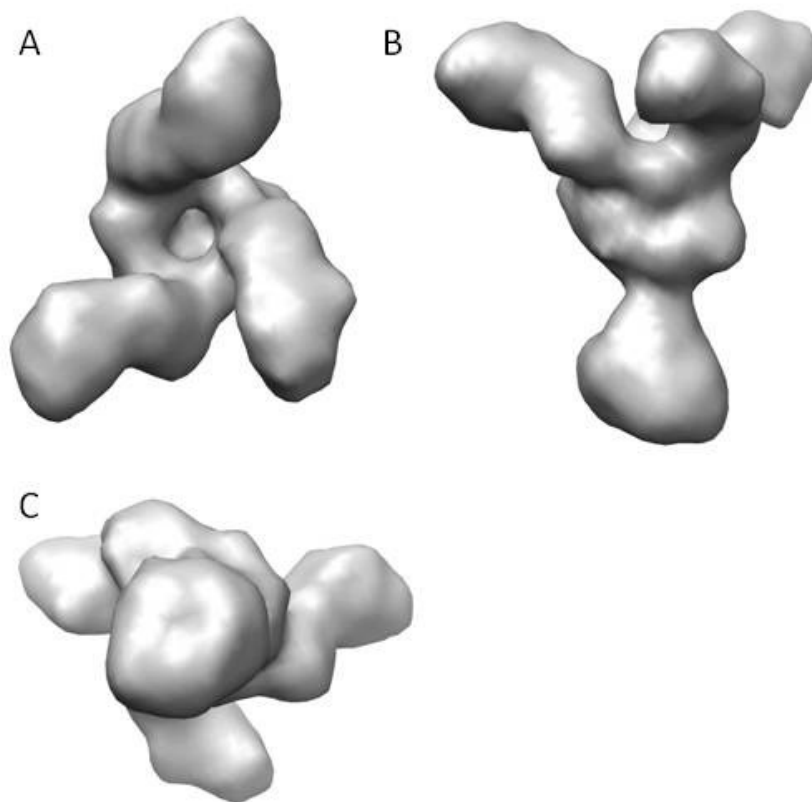
A) Vue de dessus ; B) Vue latérale 1 ; C) Vue latérale 2 ; D) Vue de dessous  
Les flèches jaunes pointent vers les bourgeons des 3 LFs

Le raffinement 3D va permettre de vérifier si le volume de référence est suffisamment proche du volume réel.

### 4. Premier essai de raffinement 3D avec XMIPP

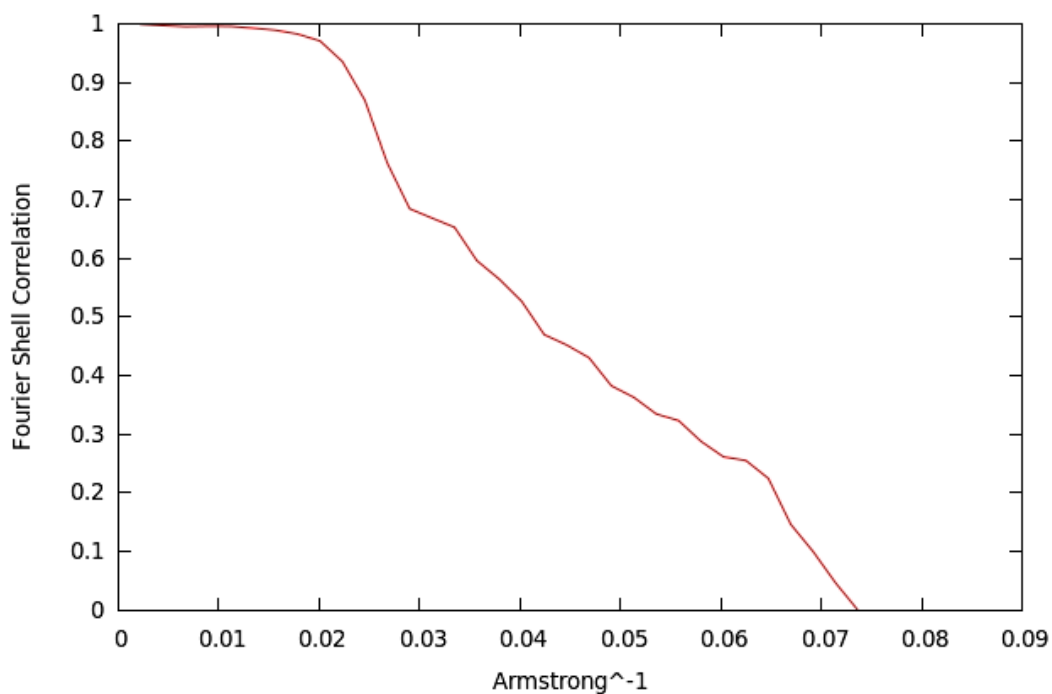
Le volume précédent de référence a été utilisé comme volume de référence pour la reconstruction 3D. D'ailleurs, au cours des itérations les appendices (bourgeons) se sont allongés pour former les trois LFs. Le résultat final à partir des particules 128x128 nous donne un volume (**figure 56**) avec une résolution de 24Å d'après la courbe de FSC (**figure 57**).





**Figure 56 : Reconstruction 3D par raffinement du complexe PA63h-pore-3LFs-ND**

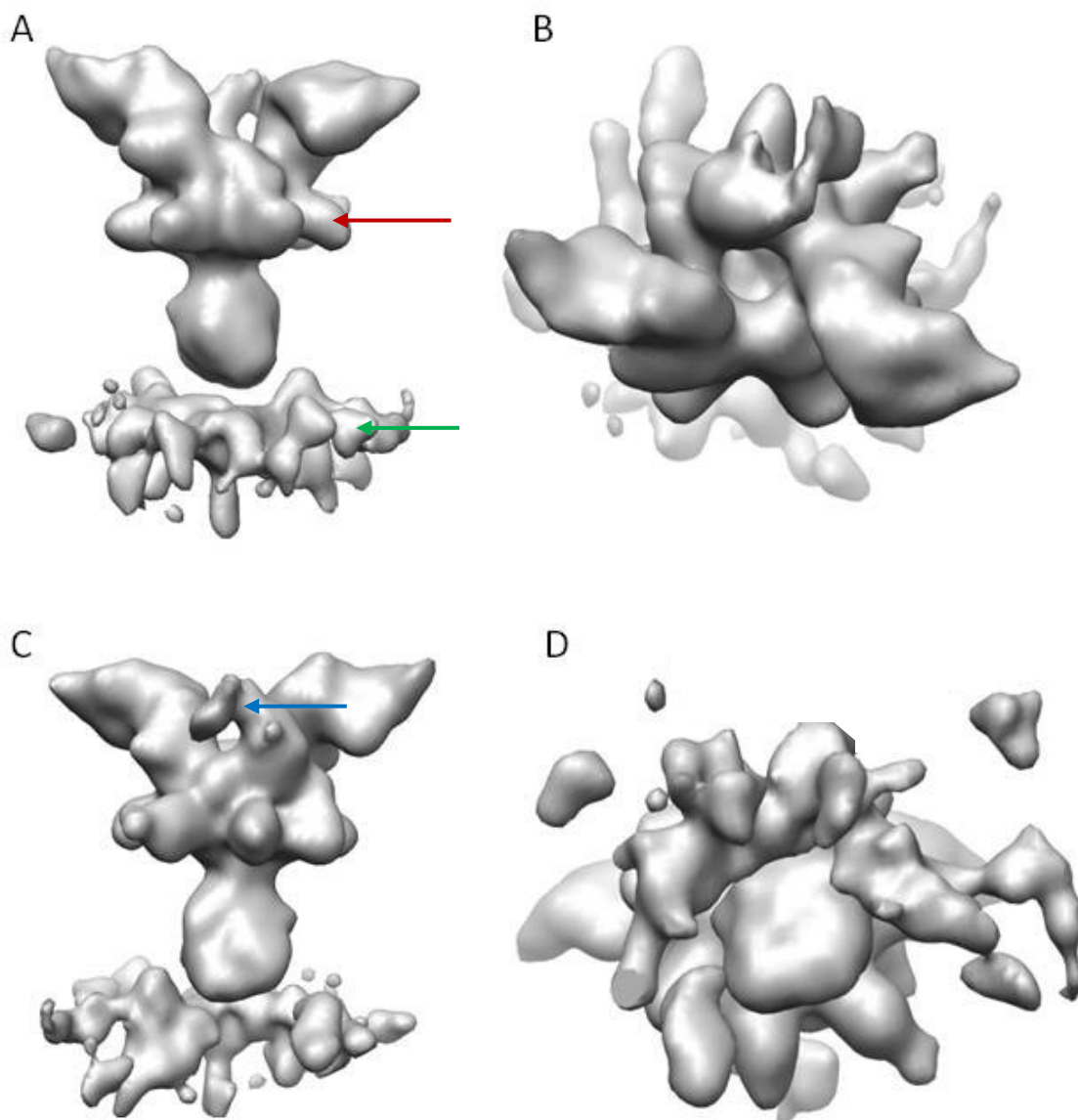
A) Vue de dessus      B) Vue latérale 1      C) Vue de dessous



**Figure 57 : Fourier Shell Correlation de la Carte 3D du complexe PA63h-pore-3LFs (premier essai)**

## 5. Deuxième essai de raffinement 3D

Le volume obtenu précédemment lors d'une première tentative de raffinement 3D a été utilisé comme référence pour un deuxième essai donnant le résultat suivant :

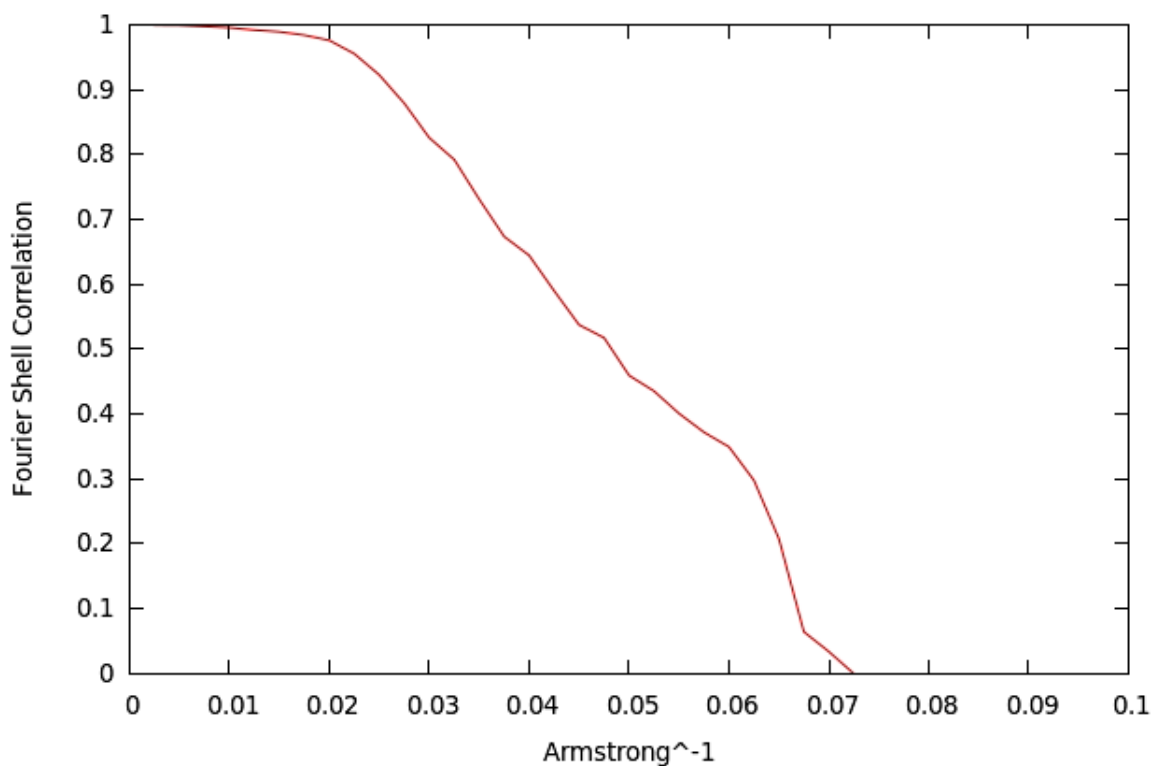


**Figure 58 : Reconstruction 3D par raffinement du complexe PA63h-pore-3LFs-ND**

- A) Vue latérale 1
- B) Vue de dessus
- C) Vue latérale 2
- D) Vue de dessous

La flèche rouge pointe vers couronne en étoile de 7 bras formé par l'heptamère de PAs. La flèche verte désigne une densité correspondante à la partie protéique d'un nanodisc. Enfin, la flèche bleu montre un LF moins bien défini que ces deux frères.

Dans cette reconstruction, les PAs apparaissent mieux définis puisqu'il est possible de distinguer la couronne en étoile des 7 bras. De plus, une densité correspondante à la position du nanodisque est maintenant visible. Par contre, on observe un LF moins bien défini par rapport à ses frères. D'après la courbe de RSC, la résolution atteint 21Å.



**Figure 59 : Fourier Shell Correlation de la Carte 3D du complexe PA63h-pore-3LFs, deuxième essai**

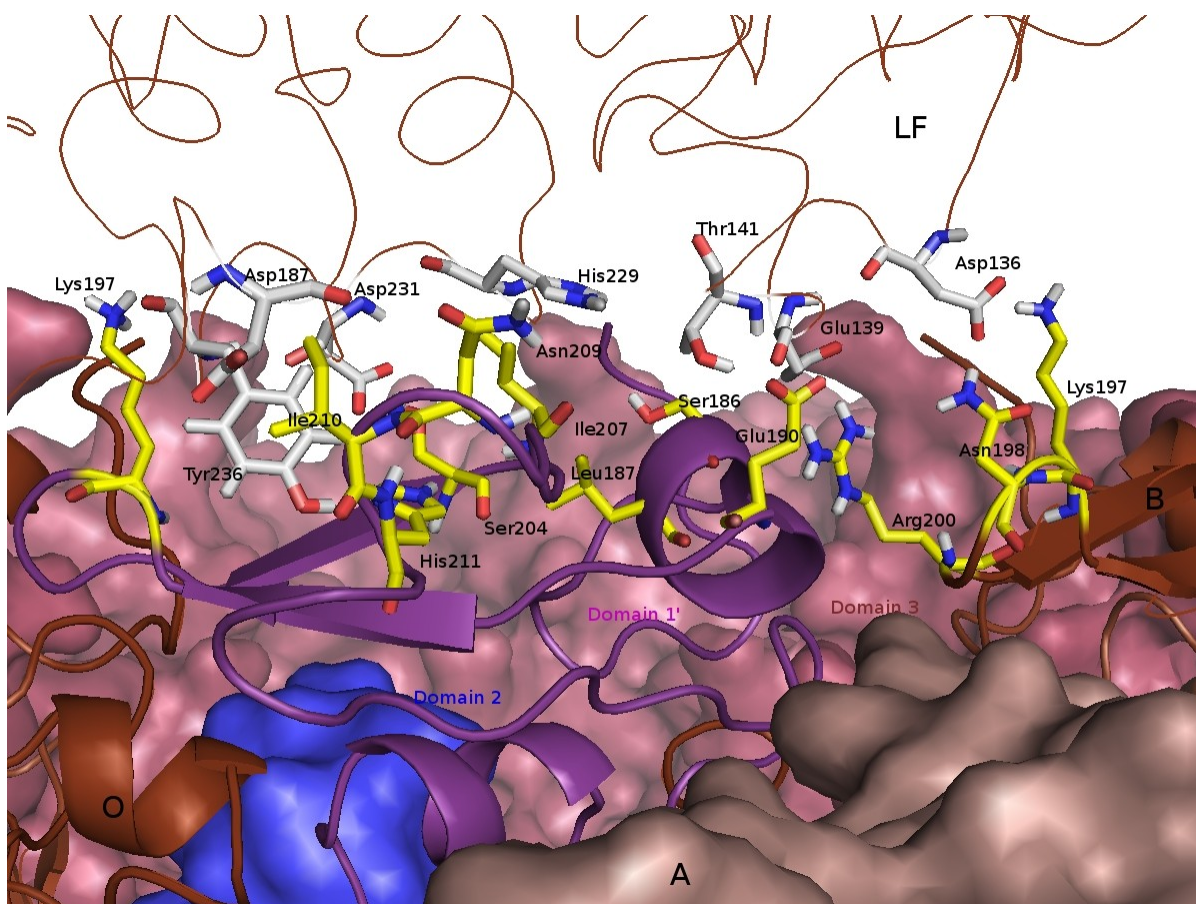
Les résultats précédents ont permis de déterminer une structure préliminaire du complexe PA63h-pore-3LFs-ND. On peut déjà voir une forte similitude entre la partie heptamérique de notre complexe et son homologue dans la carte 3D de Katayama et al., [26]. En effet, le domaine III est très visible avec sa forme en étoile juste au dessous de la liaison PA-LF. D'un autre côté, le domaine IV n'est pas très visible. Néanmoins, la structure actuelle n'est pas la structure finale et nécessite également l'apport de particules supplémentaires pour gagner en précision. De plus, la structure de Katayama a été obtenue par coloration négative. Elle contient peut-être des déformations dues au fait que les protéines ne sont pas hydratées, ce qui risque de beaucoup modifier la structure du complexe.

## D. Dynamique Moléculaire de PA63h-pore-3LFs-ND

### 1. Jonction PA-LF et les acides aminés impliqués

Les acides aminés impliqués dans le contact entre PA et LF ont été identifiés sur le complexe PA63h-prepore-3LFs par une ancienne post-doctorante de notre laboratoire : Dr. L. Alisarae.

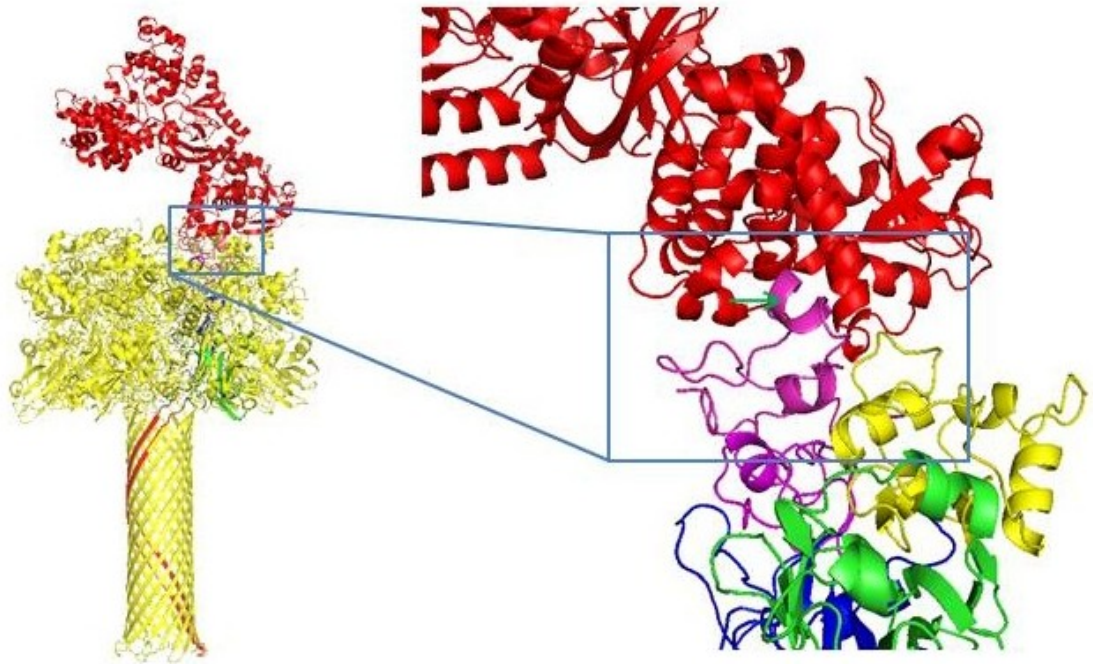
Le pdb 1TZO représentant la structure pré-pore du complexe PA63h a nommé les monomères respectivement A, B, C, D, E, F et O. On peut ainsi remarquer que LF est fixé sur deux PA, mais majoritairement sur une seule sous-unité (ici, la sous-unité A). Les acides aminés de la sous-unité A sont Lys197, Ile210, ASN209, Ile207, Leu187, Ser186 et Glu190 ; et ceux de la sous-unité B sont Arg200, Asn198 et Lys197.



**Figure 60 : Détails des résidus interagissant avec le site de liaison de PA avec LF par L. Alisarae**  
Les résidus de LF sont représentés en « stick » blancs et ceux de PA en « stick » jaune

Le pdb 1J7N représente la structure du LF dont les acides aminés en contact avec PA sont Asp187, Tyr236, Asp231, His229, Thr141, Glu139 et Asp136. les études *in vitro* faites par Cunningham *et al.*, [9] et Lacy *et al.*, [11] ont prouvé l'implication de ces aminoacides dans ces interactions.

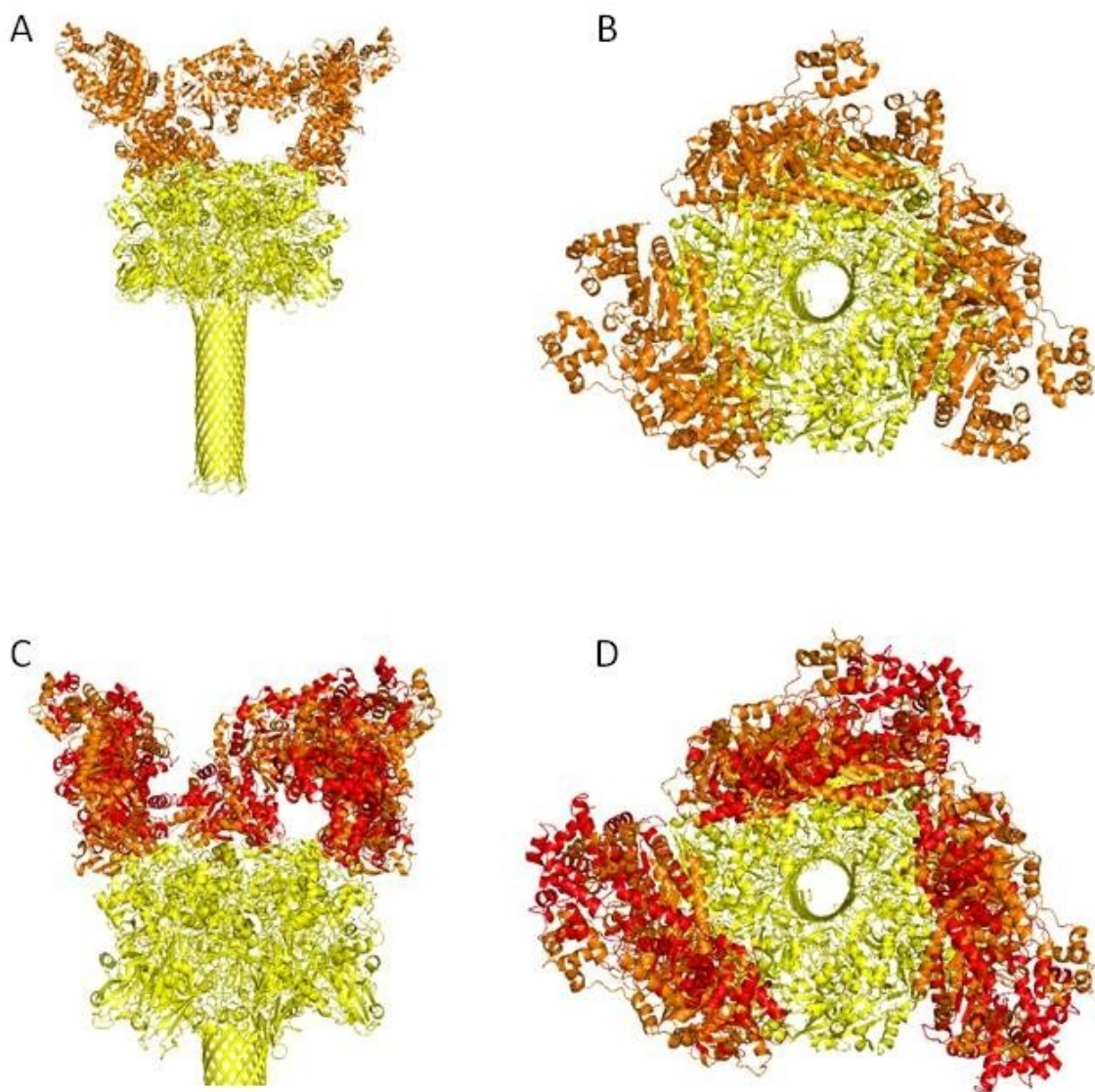
La jonction PA-LF est caractérisée par l'inclusion d'un motif boucle-hélice-boucle dans un sillon de LF (**figure 61**).



**Figure 61 : Détail de la jonction PA-LF.**

La protéine LF est représenté en rouge et le domaine 1b du protective antigen en magenta. La flèche verte pointe vers le motif boucle-hélice-boucle.

## 2. Amarrage des LFs sur la structure Pa63h-pore de Nguyen



**Figure 62 : Résultats des amarrages par ZDOCK et ROSETTA**

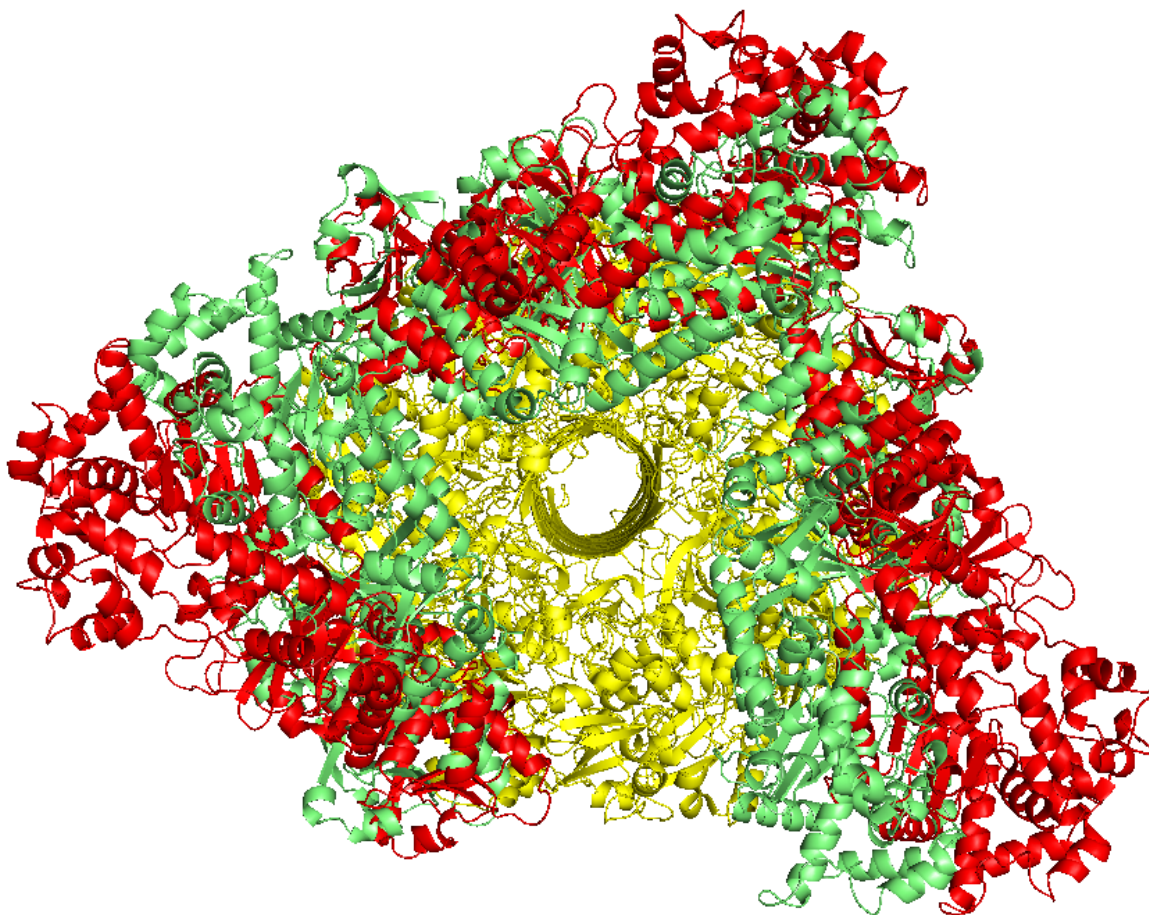
A) Amarrage par ZDOCK, vue latérale

B) Amarrage par ZDOCK, vue de dessus

C) Superposition des amarrage par ZDOCK en orange et par ROSETTA en rouge, vue latérale

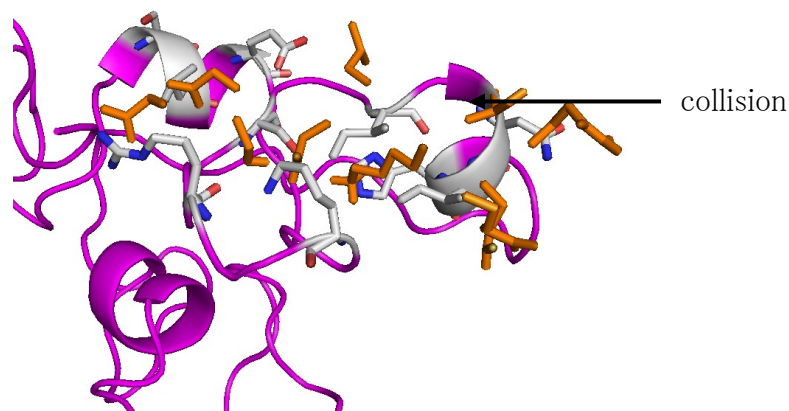
D) Superposition des amarrage par ZDOCK en orange et par ROSETTA en rouge, vue de dessus

Les résultats des amarrage via ZDOCK et ROSETTA donnent des résultats assez similaires avec des orientations « normales », inverses, et couchés. On appelle (par hypothèse) orientations « normales » les orientations de LF dans le même sens que ceux obtenues par l'étude du complexe pré-pore.



**Figure 63 : Superposition des LFs résultant de l'amarrage par ROSETTA (en rouge) avec les LFs de la structure PA63h-prepore-3LFs à  $t = 0$  ns calculé par L. Alisarae (en vert). Seul le complexe PA63h-pore est représenté jaune. Le PA63h-prepore a été aligné au PA63h-pore avec une RMSD de 8.9Å**

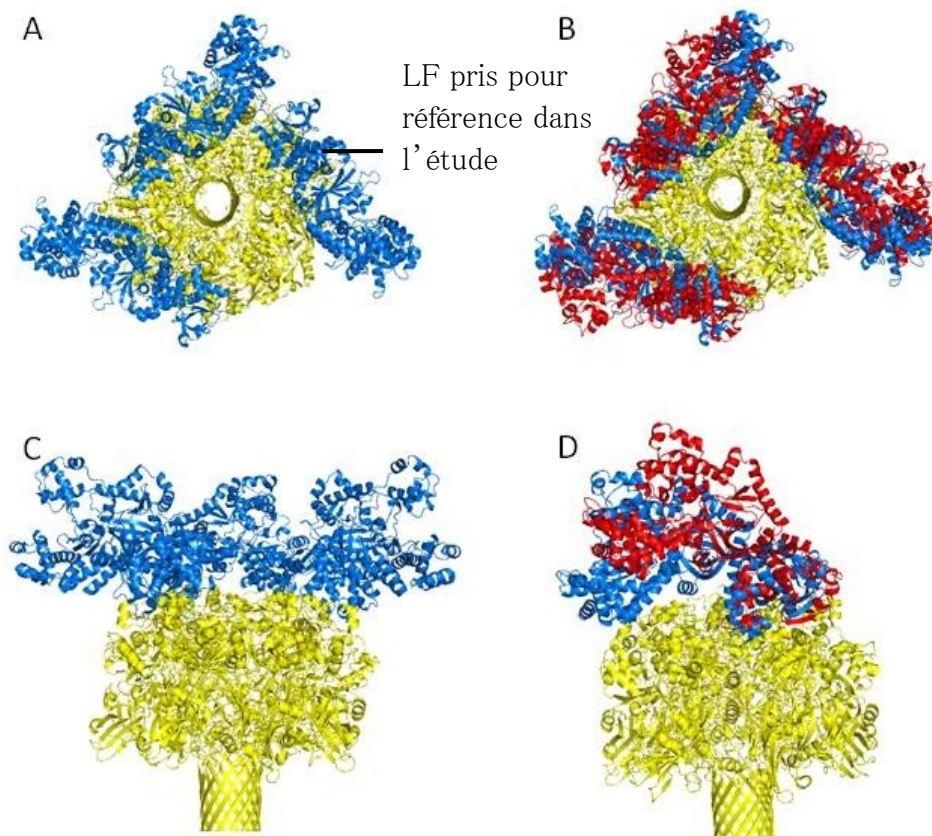
On remarque que les solutions proposées par ZDOCK et ROSETTA donnent des LFs qui ne se superposent pas avec celles de la forme prepore (en vert). Ils se dirigent plus vers l'extérieur du pore (si on prend l'axe du beta-barrel comme repère). On observe également quelques collisions d'atomes au niveau de la zone de jonction.



**Figure 64 : Vue rapproché de la zone de collision**

PA est représenté en « cartoon » magenta et les acides aminés de LF en orange.

Les positions relatives de LF par rapport à PA en conformation ont été réutilisées pour docké LF sur PA-pore. Pour cela, la partie de PA-prépore en contact avec LF a été aligné avec la partie de PA-pore en contact avec LF. La RMSD de l'alignement est de 2.9Å. Cet amarrage sera appelé « amarrage ajusté ».



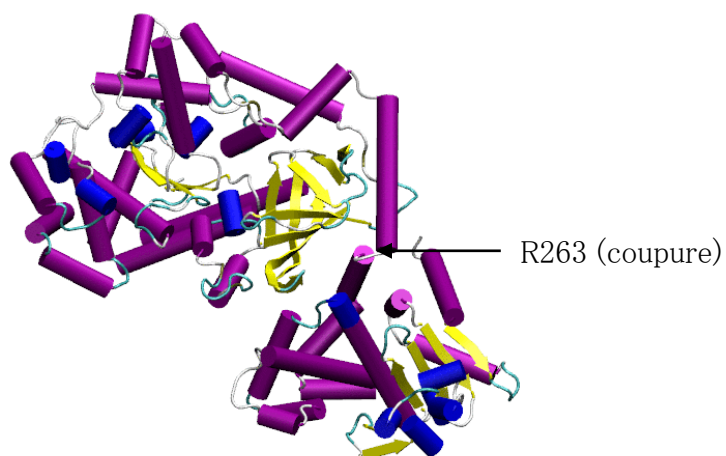
**Figure 65 : Comparaison entre l'amarrage fait par ROSETTA et l'amarrage ajusté**

- A) LFs ajustés (en bleu) en vue de dessus ; Superposition des LFs issus de l'amarrage avec ROSETTA (en rouge) et des LFs issus de l'amarrage ajusté (en bleu), vue de dessus ; Vue latérale des LFs ajustés ; Superposition d'un LF issus de l'amarrage avec ROSETTA (en rouge) et d'un LFs issus de l'amarrage ajusté (en bleu), vue de dessus

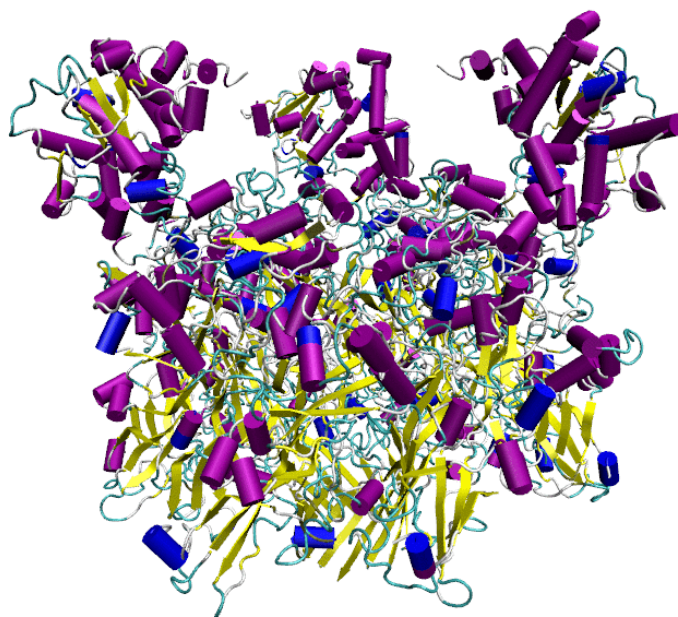


Cet amarrage corrigé montre un LF plus couché qu'avec ROSETTA mais aucune collision atomique n'existe.

La simulation a été lancée en utilisant des LFs coupés (**figure 65 et 66**) pour réduire le temps de calcul.



**Figure 66 : Gros plan sur la coupure du Lethal Factor**



**Figure 67 : Molécule de PA63h-pore-3LFs tronqués pour MD (vue en perspective)**

### 3. Résultats de la simulation par dynamique moléculaire

La simulation de 10ns a été faite avec la structure issue de l'amarrage ajusté. Le graphique ci-dessus représente l'évolution de la RMSD au cours du temps :

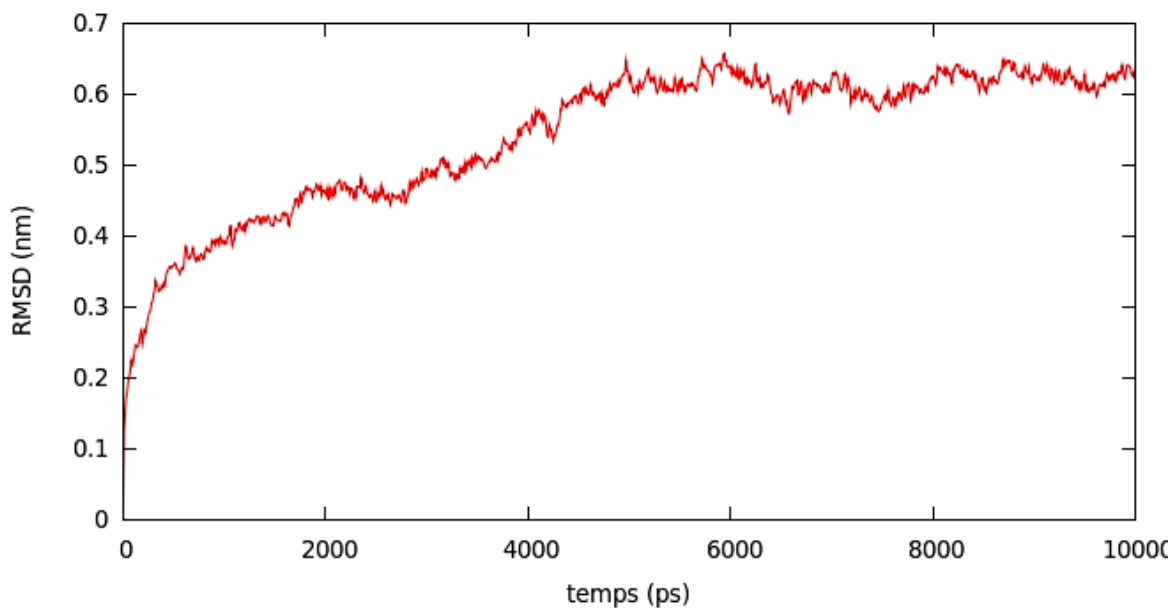
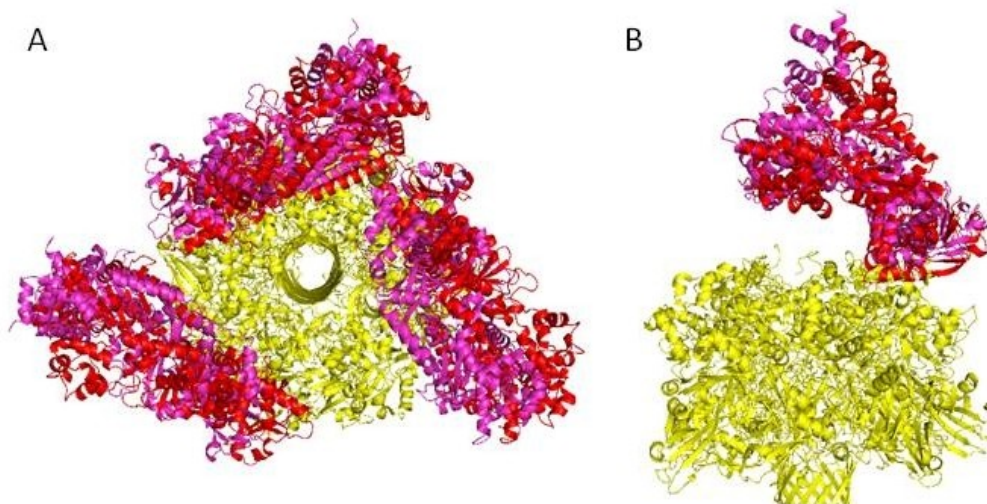


Figure 68 : Graphique de l'évolution de la RMSD du système au cours du temps

On remarque que le système prend environ 5 ns avant d'atteindre un état stable. Cela coïncide avec le temps nécessaire au LF pris pour référence pour se stabiliser. En effet, après 5 ns, celui-ci garde sa position haute en bougeant très peu. On peut cependant remarquer que cette position est très proche de celle prédite par ROSETTA (sans les collisions).



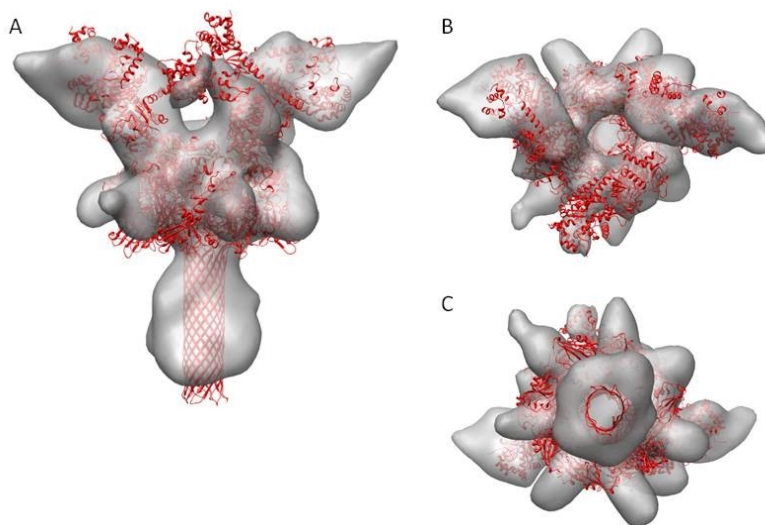
**Figure 69 : Superposition des LFs obtenus par ROSETTA (en rouge) avec ceux obtenus par MD à l'étape  $t = 6$  ns (en magenta)**

- A) Vue globale de dessus
- B) Vue latérale avec le LF de référence de l'étude

Le résultat de la simulation apparaît donc ici comme un amarrage de PA63h-pore-3LF stabilisé par dynamique moléculaire.

#### 4. Mise en relation des résultats de la MD avec ceux de la cryo-EM

La structure obtenue par MD a été amarrée dans le volume acquis par le deuxième raffinement de PA63h-pore-3LFs-ND.



**Figure 70 : Amarrage de la structure obtenue par MD avec le volume acquis par le deuxième raffinement de PA63h-pore-3LFs-ND. Le nanodisc a été retiré pour plus de clarté. A) Vue latérale ; B) Vue de dessus ; C) Vue de dessous**

On observe une bonne correspondance positionnelle au niveau des LFs. La couronne formée par le domaine III de l'heptamère de PAs est plus constrictif que sur la carte 3D. Ce problème est également visible sur la carte 3D de Katayama *et al.*, [26]. Le domaine IV n'est pas bien représenté. Le  $\beta$ -barrel continue au-delà de la carte de densité. Cela est dû au fait que l'extrémité débordante correspond à la partie du tube insérée dans le nanodisque.

L'objectif de ce mémoire est de mieux connaître les interactions entre PA et LF. Pour cela, les travaux suivants ont été effectués :

### **1<sup>ère</sup> partie : Étude du complexe PA63h-prépore-3LFs**

Dans un premier temps, le complexe PA63h-3LFs a été purifié de manière homogène par gel filtration.

Mes travaux ont commencé avec l'étude de la conformation prépore du complexe PA-3LFs. La carte de densité calculée a révélé une symétrie de 7 au niveau du Protective Antigen. Par contre, les LFs ne le sont pas. En effet, chacun contient des acides aminés en contact avec deux PAs, il s'en suit alors que six PAs sont nécessaires à la fixation des trois LFs. En outre, le contact de LF avec les deux PAs n'est pas équilibré car il est réparti de la manière suivante : 5 acides aminés en contact avec le premier PA et 2 avec le second.

Si l'on représente la succession des PAs en une suite binaire circulaire (dernière chiffre connecté au premier chiffre) avec '0' pour PA seul (ou connecté avec seulement deux acides aminés avec un LF) et '1' pour un PA connecté à un LF (5 acides aminés en contact avec un LF), la carte 3D montrerait la répartition suivante : 0101001 (voir **figure 32**). Cependant, dans notre première reconstruction 3D, la disposition apparaissait plutôt comme cela : 0101011. C'est-à-dire que le troisième LF est à cheval entre deux PAs. De plus, ce LF apparaît flou. Ces observations ont orienté l'étude vers l'hypothèse stipulant qu'il existerait plusieurs conformations de cette sous-unité. Afin de vérifier si la zone floue correspondait à une variabilité au niveau de la correspondance « matching » des particules (point chaud), une variance 3D a été calculée. Celle-ci a effectivement révélée des points chauds au niveau du LF3. Par la suite, la classification 3D a confirmé l'existence de plusieurs conformations et les

résultats trouvés ont montré une amélioration au niveau de la forme du LF3 (disparition du trou central) mais pas au niveau de la jonction (pied gras).

Le pied gras a demeuré suspect à deux niveaux :

- d'où vient la densité supplémentaire ?
- pourquoi est-ce que les études par MD de L. Alisarae ne montrent pas un LF positionné entre deux PAs ?

C'est alors qu'ont été entreprises « les études du pied gras » avec une nouvelle hypothèse selon laquelle il y aurait peut-être une erreur d'alignement. La preuve du mésalignement a permis de retrouver un volume final dont le LF3 est entièrement défini. Le volume final est meilleur que ceux issus de la classification 3D car il présente une résolution de 14Å contre 16Å.

Il demeure tout de même intrigant que les différences de densité des LFs soient seulement dues à une erreur d'alignement puisque les reconstructions devraient donner le même résultat peu importe le modèle initial [39]. Cependant, ce n'est pas le cas pour nous. Cela pourrait peut-être provenir d'une utilisation de paramètres pas assez restrictifs en rapport avec l'angle de recherche lors de l'alignement.

Néanmoins, notre volume final est en accord avec les structures prédites par les études faites par dynamique moléculaire par L. Alisarae. Il est également en accord avec les études biochimiques de Mogridge *et al*, [7] stipulant une stœchiométrie de 3 LFs pour 7 PAs.

Afin de terminer l'étude, il pourrait être intéressant de refaire un calcul de variance 3D pour montrer que cette fois-ci, le volume final est vraiment celui attendu.

La résolution de la carte 3D du complexe PA63h-préopore-3LFs apporte de nouvelles informations dans une littérature très conversée. En effet, les travaux d'électrophorèses sur

gel natif de Singh *et al.* [6] ont montré que 7 LFs se fixaient sur l'heptamère. Or, nos travaux montrent que seul 3 LFs s'attachent à PA63h. Il est de même avec les résultats de Ren *et al.* [46] qui présentent une structure résolue par cryo-microscopie avec PA63h-1LF. Ces derniers ont trouvé qu'avec saturation en LF, il n'y a que des complexes PA-1LF.

## 2<sup>ème</sup> partie: Étude du complexe PA63h-pore-3LFs

Les résultats concernant le complexe PA63h-pore-3LFs-ND sont encore qu'à un stade très préliminaire. Cependant, on peut déjà observer que l'heptamère PA en conformation pore supporte toujours trois ligands LFs. Cela confirme la théorie d'une liaison forte PA-LF [4]. De plus, la forme pore de PA63h présente une forte similitude avec la carte 3D de Katayama *et al.* [26]. En effet, on perçoit le domaine III avec sa forme en caractéristique heptamérique en étoile juste au dessous de la liaison PA-LF (voir flèches jaunes à la **figure 10**). Ce motif caractéristique est suivi par un tube, le  $\beta$ -barrel.

D'un autre côté, le domaine IV n'est pas visible. Néanmoins, la structure actuelle n'est pas la structure finale et nécessite également l'apport de particules supplémentaires pour gagner en précision. De plus, il serait nécessaire de vérifier la validité de l'ensemble de nos projections expérimentales. En effet, contenu de leur faible visibilité sur les images digitales, il se pourrait que certaines aient été sélectionnées par erreur rajoutant ainsi du bruit dans notre carte 3D. En outre, la structure de Katayama [26] a été obtenue par coloration négative. Elle peut de ce fait contenir des déformations dues au fait que les protéines ne sont pas dans leur état natif. En effet, elles sont déshydratées. Les études en cryo devraient apporter de meilleurs résultats.

En vue de créer une synergie avec nos résultats de cryo-EM, une étude en parallèle par dynamique moléculaire semblait pertinente. En effet, dans le cas où les résultats convergeraient dans le même sens, la dynamique moléculaire permettrait l'étude des interactions entre PA et LF au niveau moléculaire.

### **3<sup>ème</sup> partie : Dynamique Moléculaire du complexe PA63h-pore-3LFs**

Les résultats du docking avec ZDOCK et ROSETTA ont donné une structure qui semblait crédible du point de vue que les LFs étaient dans le même sens (horaire) que ceux amarrés et stabilisés par dynamique moléculaire lors des travaux de L. Alisarae. Cependant, une collision entre atomes a remis en cause cette issue. Sans doute que le lancement de la minimisation sur le backbone seul aurait réajusté le complexe. En tant que novice, une autre issue a été choisie. Celle-ci se base sur un réajustement par alignement de structures connues. Pour cela, le monomère PA-prépore a été aligné par sa partie liant le LF avec la partie homologue du PA-pore. De ce fait, le LF a été disposé sur le PA-pore de la même manière que sur le PA-prépore en gardant les contacts entre acides aminés. Il en résulte un complexe PA63h-pore-3LFs dont les LFs sont en position plus couchés qu'avec les amarrages par ROSETTA et ZDOCK. Cependant, la simulation par dynamique moléculaire a stabilisé le complexe et les LFs ont retrouvé leur position (relevés) prédites avec ROSETTA et ZDOCK. Il s'agit du résultat actuel de la simulation qui démontre la stabilité de la liaison PA-LF. Cela a été étayé par l'étude de la courbe de RMSD qui présente un palier à partir de 5ns. La simulation totale étant de 10ns, le complexe a eu le temps d'acquiescer une conformation dynamiquement stable. Il serait maintenant intéressant d'étudier en détail les mouvements



des trois LFs séparément. De plus, le complexe n'ayant pas explosé pendant les calculs lors de cette première simulation, il est alors possible de remplacer les LFs tronqués par des LFs complets même si le temps de calcul sera beaucoup plus long (taille de la boîte).

Les résultats préliminaires actuels du complexe PA63h-pore-LF résolu par cryo-EM montrent déjà des similitudes avec la simulation par MD. En effet, dans les deux cas, les LFs se superposent.

Les travaux précédents ont tous démontré un résultat commun qui est celui de la stabilité de la liaison PA-LF. Cependant, l'hypothèse actuellement la plus admise sur le mécanisme infectieux de l'anthrax est celle qui stipule que les ligands EF et LF passeraient au travers du pore. Or, la stabilité de la liaison PA-LF pose un problème. Si LF ne se détache pas de PA, comment celui-ci peut-il passer au-travers du pore? De plus le site actif de LF n'est pas en contact avec PA lui permettant de maintenir une activité catalytique malgré leur liaison. D'un autre côté, les résultats présentés dans ce mémoire montrent que les LFs contenus dans le complexe PA63h-pore-3LFs ont leurs extrémités qui s'éloignent du lumen.

Pour finir, tous nos résultats montrent que l'heptamère de PA supporte 3 LFs. Cela a été comparé avec d'autres travaux pour l'étude du complexe pré-pore. Cependant, nos résultats sur la conformation pore avec trois LFs sont inédits pour le moment.

Au stade actuel, les résultats de la dynamique moléculaire montrent une structure qui se superpose bien au niveau des LFs. Contenu de la précocité des résultats de la cryo-EM du complexe pore, il est encore trop tôt pour faire des comparaisons et des conclusions précises. Cependant, il est intéressant de constater que la différence d'orientation des LFs entre la conformation pré-pore du complexe et celle pore observée par EM est respectée par le MD.

## **Perspectives :**

La reconstruction du complexe PA63h-pore-3LFs doit-être améliorée. Pour cela, des particules supplémentaires doivent-être ajoutées. Il est donc nécessaire de prendre de nouvelles photos. L'idéal serait de doubler le nombre de particules actuelles. Celles-ci étant contenues dans 150 micrographes, on peut déduire qu'il en resterait la moitié à acquérir.

Pour la dynamique moléculaire, il serait utile de poursuivre l'étude notamment en s'intéressant aux mouvements des trois LF's séparément. De plus, le système actuel contenant des LF's tronqués n'ayant pas explosé, il serait envisageable de lancer une simulation avec des LF's complets (le temps de calcul serait beaucoup plus long).

Les travaux actuels ont permis une insertion de PA63h dans un nanodisque. Maintenant, il serait très pertinent de reproduire ces manipulations en remplaçant les nanodisques par des liposomes. Néanmoins, ces expériences ont déjà été essayé au laboratoire mais sans succès. En effet, un problème d'insertion des PA dans les membranes ont été observé. Le problème était que l'échantillon ne contenait jamais des complexes PA63h-liposome homogène. C'est à dire que certains liposomes ne contenaient aucun PA et d'autres étaient sursaturés. Le screening des concentrations n'a pas apporté une solution au problème. Nos expériences avec les liposomes portaient sur des vésicules de diamètre environ 200nm. Il pourrait être fructueux de travailler à l'avenir avec des liposomes plus petits (50nm) qui permettrait, de par sa petite taille, de fixer seulement un ou deux PA. Cela rendrait possible l'étude par particule isolée pratiquée jusqu'à lors.

En parallèle, une optimisation de la purification du complexe PA63h pourrait

permettre plusieurs choses. Tout d'abord, l'ajout d'un His-tag en C-terminal va nous permettre de mieux réguler et répartir l'insertion de PA63h dans les liposomes. En effet, il est possible de rajouter des lipides contenant des ions nickel dans la composition de la membrane des liposomes. Ainsi, l'ajustement de la concentration de ces lipides par rapport aux autres lipides initialement utilisée va permettre, par saturation en PA, de contrôler le nombre de PA à fixer par liposome. D'un autre côté, la production de PA et de LF serait nécessaire pour faire une filtration par gel des complexes PA-LF. On pourrait alors étudier par cryo-microscopie les complexes PA63h-1LF et PA63h-2LFs.

Le remplacement des nanodisques par des liposomes permettrait l'étude d'un éventuel passage au travers du pore grâce à la reproduction d'un potentiel ionique (via la compartimentation). On pourrait également créer des mutants pour bloquer le passage (dans le cas où il aurait lieu).

L'étape ultime serait d'étudier le phénomène *in vivo* en utilisant des cellules de mammifères. Il serait alors possible d'identifier les facteurs cellulaires interagissant dans le processus infectieux.

- [1] [http://www.denniskunkel.com/images/StockImages/Images\\_02/21186E.jpg](http://www.denniskunkel.com/images/StockImages/Images_02/21186E.jpg)
- [2] Van den Enden E, Van Gompel A, Van Esbroeck M. (2006). Cutaneous anthrax, Belgian traveler. *Emerg Infect Dis.* 12(3): 523-5.
- [3] Falnes PO, Sandvig K. (2000). Penetration of protein toxins into cells. *Curr Opin Cell Biol.* 12(4):407-13.
- [4] Elliott, J.L., Mogridge, J. & Collier, R.J. (2000). A quantitative study of the interactions of *Bacillus anthracis* edema factor and lethal factor with activated protective antigen. *Biochemistry*, 39(22): 6706-6713.
- [5] Pimental, R.-A.L., Christensen, K.A., Krantz, B.A. & Collier, R.J. (2004). Anthrax toxin complexes: heptameric protective antigen can bind lethal factor and edema factor simultaneously. *Biochem Biophys Res Commun*, 322(1): 258-262.
- [6] Singh, Y., Klimpel, K.R., Goel, S., Swain, P.K. & Leppla, S.H. (1999). Oligomerization of anthrax toxin protective antigen and binding of lethal factor during endocytic uptake into mammalian cells. *Infect Immun*, 67(4) : 1853-1859.
- [7] Mogridge, J., Cunningham, K. & Collier, R.J. (2002). Stoichiometry of anthrax toxin complexes. *Biochemistry*, 41(3): 1079-1082.
- [8] Mogridge, J., Cunningham, K., Lacy, D.B., Mourez, M. & Collier, R.J. (2002). The lethal and edema factors of anthrax toxin bind only to oligomeric forms of the protective antigen. *Proc Natl Acad Sci U S A*, 99(10): 7045-7048.
- [9] Cunningham, K., Lacy, D.B., Mogridge, J. & Collier, R.J. (2002). Mapping the lethal factor and edema factor binding sites on oligomeric anthrax protective antigen. *Proc Natl Acad Sci U S A*, 99(10): 7049-7053.
- [10] Lacy, D.B., Mourez, M., Fouassier, A. & Collier, R.J. (2002). Mapping the anthrax protective antigen binding site on the lethal and edema factors. *J Biol Chem*, 277(4): 3006-3010.
- [11] Lacy, D.B., Lin, H.C., Melnyk, R.A., Schueler-Furman, O., Reither, L., Cunningham, K., Baker, D. & Collier, R.J. (2005). A model of anthrax toxin lethal factor bound to protective antigen. *Proc Natl Acad Sci U S A*, 102(45): 16409-16414.
- [12] Melnyk, R.A., Hewitt, K.M., Lacy, D.B., Lin, H.C., Gessner, C.R., Li, S., Woods, V.L. & Collier, R.J. (2006). Structural determinants for the binding of anthrax lethal factor to oligomeric protective antigen. *J Biol Chem*, 281(3): 1630-1635.
- [13] Young, J.A.T. & Collier, R.J. (2007). Anthrax toxin: receptor binding, internalization, pore formation, and translocation. *Annu Rev Biochem*, 76: 243-265.

- [14] Bradley, K.A., Mogridge, J., Mourez, M., Collier, R.J., Young, J.A. (2001). *Nature* 414:225–29.
- [15] Scobie, H.M., Rainey, G.J., Bradley, K.A., Young, J.A. (2003). *Proc. Natl. Acad. Sci. USA* 100:5170–74.
- [16] Wigelsworth, D.J., Krantz, B.A., Christensen, K.A., Lacy, D.B., Juris, S.J., Collier, R.J. (2004). *J. Biol. Chem.* 279:23349–56.
- [17] Lacy, D.B., Wigelsworth, D.J., Melnyk, R.A., Harrison, S.C., Collier, R.J. (2004). *Proc. Natl. Acad. Sci. USA* 101:13147–51.
- [18] Santelli, E., Bankston, L.A., Leppla, S.H., Liddington, R.C. (2004). *Nature* 430:905–8.
- [19] Bradley, K.A., Mogridge, J., Rainey, G.J.A., Batty, S., Young, J.A.T. (2003). *J. Biol. Chem.* 278:49342–47.
- [20] Scobie, H.M., Wigelsworth, D.J., Marlett, J.M., Thomas, D., Rainey, G.J.A., (2006). *PLoS Pathog.* 2:949–55.
- [21] Rogers, M.S., Christensen, K.A., Birsner, A.E., Short, S.M., Wigelsworth, D.J., Collier, R.J., D'Amato, R.J. (2007). Mutant Anthrax Toxin B Moiety (Protective Antigen) Inhibits Angiogenesis and Tumor Growth. *Cancer Res*; 67:(20).
- [22] Petosa, C., Collier, R.J., Klimpel, K.R., Leppla, S.H. & Liddington, R.C. (1997). Crystal structure of the anthrax toxin protective antigen. *Nature, Biochemistry*, 385(6619): 833-838.
- [23] Pannifer, A.D., Wong, T.Y., Schwarzenbacher, R., Renatus, M., Petosa, C., Bienkowska, J., Lacy, D.B., Collier, R.J., Park, S., Leppla, S.H., Hanna, P. & Liddington, R.C. (2001). Crystal structure of the anthrax lethal factor. *Nature, Biochemistry*, 414(6860): 229-233.
- [24] Shen, Y., Zhukovskaya, N.L., Guo, Q., Florián, J. & Tang, W.-J. (2005). Calcium-independent calmodulin binding and two-metal-ion catalytic mechanism of anthrax edema factor. *EMBO J*, 24(5): 929-941.
- [25] Nguyen, T.L. (2004). Three-dimensional model of the pore form of anthrax protective antigen. Structure and biological implications. *J Biomol Struct Dyn*, 22(3): 253-265.
- [26] Katayama, H., Wang, J., Tama, F., Chollet, L., Gogol, E.P., Collier, R.J. & Fisher, M.T. (2010). Three-dimensional structure of the anthrax toxin pore inserted into lipid nanodiscs and lipid vesicles. *Proc Natl Acad Sci U S A*, 107(8): 3453-3457.
- [27] Finkelstein, A. (2009). Proton-coupled protein transport through the anthrax toxin channel. *Phil. Trans. R. Soc. B*; 364, 209–215.

- [28] Feld, G.K., Thoren, K.L., Kintzer, A.F., Sterling, H.J. , Tang, I.I., Greenberg, S.G., Williams, E.R., Krantz, B.A. (2010). Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nature Structural & Molecular Biology*; 17, 1383–1390.
- [29] Encyclopedia Britania, 2007.
- [30] [http://www.uni-duesseldorf.de/MathNat/ipb/lw\\_resource/datapool/file/item\\_169/nanodisc\\_empty\\_size\\_m.jpg](http://www.uni-duesseldorf.de/MathNat/ipb/lw_resource/datapool/file/item_169/nanodisc_empty_size_m.jpg).
- [31] <http://www.scripps.edu> (Frank course, 2005).
- [32] DeRosier, D.J., and Klug, A. (1968), *Nature* 217, 130-134.
- [33] Mindell, JA, Grigorieff N. 2003. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* 142:334-47.
- [34] Tang G., Peng L., Baldwin P.R., Mann D.S., Jiang W., Rees I., Ludtke S.J. (2007). EMAN2: an extensible image processing suite for electron microscopy. *J Struct Biol.*, 157(1):38-46.
- [35] Collier, R.J. (2009). Membrane translocation by anthrax toxin. *Mol Aspects Med*, 30(6): 413-422.
- [36] Kalendar R, Lee D, Schulman AH 2009. FastPCR Software for PCR Primer and Probe Design and Repeat Search. *Genes, Genomes and Genomics*, 3(1): 1-14.
- [37] <http://products.invitrogen.com/ivgn/product/K10001?ICID=search-product>
- [38] <http://sligarlab.life.uiuc.edu/nanodisc/protocols.html>.
- [39] Bradford M. (1976). A rapid and sensitive method for the quantization of microgram quantities of protein utilizing the principle of protein dye binding, *Analytical Biochem.*, 72: 248-254.
- [40] Laemmli U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4, *Nature*, 227: 4668-4673.
- [41] Sorzano CO, Marabini R, Velázquez-Muriel J, Bilbao-Castro JR, Scheres SH, Carazo JM, Pascual-Montano A. (2004). XMIPP: a new generation of an open-source image processing package for electron microscopy. *J Struct Biol.*, 148(2):194-204.
- [42] Scheres SH, Valle M, Carazo JM. (2005). Fast maximum-likelihood refinement of electron microscopy images. *Bioinformatics.*, 21(2): 243-4.

- [43] <http://www.cgl.ucsf.edu/chimera/>
- [44] <http://coan.burnham.org/>
- [45] <http://www.cgl.ucsf.edu/chimera/>
- [46] Ren G, Quispe J, Leppla SH, Mitra AK. (2004). Large-scale structural changes accompany binding of lethal factor to anthrax protective antigen: a cryo-electron microscopic study. *Structure*, 12(11): 2059-66.
- [47] Lacy, D.B., Wigelsworth, D.J., Melnyk, R.A., Harrison, S.C., Collier, R.J. (2004). Structure of heptameric protective antigen bound to an anthrax toxin receptor: a role for receptor in pH-dependent pore formation. *Proc.Natl.Acad.Sci.USA* 101: 13147-13151
- [48] Turk, B.E., Wong, T.Y., Schwarzenbacher, R., Jarrell, E.T., Leppla, S.H., Collier, R.J., Liddington, R.C., Cantley, L.C. (2004). The structural basis for substrate and inhibitor selectivity of the anthrax lethal factor. *Nat.Struct.Mol.Biol.* 11: 60-66

# Annexe 1

## Script pour le démêlage des mésalignements

```
#!/usr/bin/perl

#© Lucien Fabre - The "cross correlation filter". July 2010.

#questions and keyboard input here:
print "Enter docfile from Projection Matching : \n";
$docfile = <>;
print "Enter repertory of the ReferenceLibrary 1 : \n";
$selffile_1 = <>;
print "Enter repertory of the ReferenceLibrary 2 : \n";
$selffile_2 = <>;
print "Enter threshold : \n";
$threshold = <>;

#opening docfile
open doc, $docfile or die "no file";
while(<doc>){
    push(@doc_lignes,$_);
}
close(doc);

#write results
open global,">global.match" or die "E/S : $!\n";
print global " ; Headerinfo columns: rot (1), tilt (2), psi (3), Xoff
(4), Yoff (5), Ref (6), Flip (7), maxCC1 (8), maxCC2 (9), diff (10)\n";
open match1,">lib_1.match" or die "E/S : $!\n";
print match1 " ; Headerinfo columns: rot (1), tilt (2), psi (3), Xoff
(4), Yoff (5), Ref (6), Flip (7), maxCC (8)\n";
open match2,">lib_2.match" or die "E/S : $!\n";
print match2 " ; Headerinfo columns: rot (1), tilt (2), psi (3), Xoff
(4), Yoff (5), Ref (6), Flip (7), maxCC (8)\n";

#docfile process
for (@doc_lignes) {
    $ligne = $_;
    if ($ligne =~ /(;/){
        push(@ligne);
        @i = split(/ ; /,$ligne);
        $image = $i[1];
        chomp $image;
    }

    else {

        @colone = split (/\\b\\s/, $ligne);

        $lib1 = &extract_ref($ligne, $selffile_1);
        $lib2 = &extract_ref($ligne, $selffile_2);

        $sortie1 = `xmipp_correlation -ref $image -i $lib1 -cc`;
        $sortie2 = `xmipp_correlation -ref $image -i $lib2 -cc`;
```



```

    $cc1 = &extract_cc($sortiel);
    $cc2 = &extract_cc($sortie2);

    print global " ; " . $image . "\n";
    print global &formatationDouble($ligne, $cc1, $cc2);

    if (($cc1 > $cc2) && (abs($cc1-$cc2) > $threshold))
    {
        print match1 " ; " . $image . "\n";
        print match1 &formatation($ligne, $cc1);
    }
    elsif (($cc1 < $cc2) && (abs($cc1-$cc2) > $threshold))
    {
        print match2 " ; " . $image . "\n";
        print match2 &formatation($ligne, $cc2);
    }
    else
    {
        print match1 " ; " . $image . "\n";
        print match1 &formatation($ligne, $cc1);
        print match2 " ; " . $image . "\n";
        print match2 &formatation($ligne, $cc2);
    }
}

sub formatation{
    my($ligne, $cc) = @_ ;
    {
        $temp = substr($ligne, 0, 88);
        $ligne = $temp . " " . $cc . "\n";
        return $ligne;
    }
}

sub formatationDouble{
    my($ligne, $cc1, $cc2) = @_ ;
    {
        $diff = abs($cc1-$cc2);
        $diff = sprintf("%1.3f", $diff);
        $temp = substr($ligne, 0, 88);
        $ligne = $temp . " " . $cc1 . " " . $cc2 . " " . $diff
    }
}

sub extract_ref{
    my ($ligne, $selfile) = @_ ;
    @colone = split (/\\b\\s/, $ligne);
    $lib = sprintf("%d", $colone[7]);
    $lib = sprintf("%6g", $lib);
    $lib =~s/ /0/g;
    $lib = "$selfile" . "ref" . "$lib" . ".xmp";
    return $lib;
}

```

```
sub extract_cc {
  my ($var1) = @_ ;
  chomp $var1;
  @var2 = split (/cc= /,$var1);
  $var2[1] = sprintf("%1.5f",$var2[1]);
  if ($var2[1] =~ "-0."){
    $var2[1] = "0.00000";
  }
  return $var2[1];
}
close(match1);
close(match2);
close(global);
```