

**ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES (ED222)**

**Institut de Chimie de Strasbourg UMR7177**

Laboratoire de RMN et biophysique des membranes

**THÈSE** présentée par :

**Omar RIFI**

soutenue le : **31 janvier 2014**

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

Discipline/ Spécialité : Biochimie/Biophysique

**Production des polypeptides issus  
des glycoprotéines d'enveloppe du  
VIH-1 pour des études biophysique et  
structurale par RMN et DC.**

**THÈSE dirigée par :**  
**M. BECHINGER Burkhard**

Professeur, université de Strasbourg

**RAPPORTEURS :**  
**M. KUHN Andreas**  
**M. DUFOURC Erick**

Professeur, université de Hohenheim  
Directeur de recherche, CNRS

---

**AUTRES MEMBRES DU JURY :**  
**Mme LEIZE-WAGNER Emmanuelle**

Directeur de recherche, CNRS

# Remerciements

Mes premiers remerciements vont à mon directeur de thèse, Pr. Burkhard BECHINGER, de m'avoir accueilli au sein de son laboratoire et pour la grande liberté dont j'ai bénéficié durant ces quatre années incluant le stage de Master 2.

Je remercie l'institut du martyr Rafic EL-HARIRI d'avoir financé ma thèse.

Emmanuelle LEIZE-WAGNER, Andreas KUHN et Erick DUFOURC ont accepté d'évaluer ce travail. Qu'ils y trouvent toute ma gratitude.

Je remercie de tout mon cœur Elise pour son encadrement, sa disponibilité,... et pour son soutien moral surtout dans les moments les plus difficiles de ma vie que j'ai eu durant cette thèse. Je serai toujours reconnaissant à toi.

Je remercie également Evgeniy et Philippe pour m'avoir initié à la spectroscopie RMN et Arnaud aux expériences de spectroscopies optiques.

Un grand merci aussi à Christopher, pour sa grande disponibilité, sa gentillesse...

Merci aussi au reste de l'équipe : Delphine, Patricia, Natalia, Jésus... et à toi Caroline un grand Merci, non pas seulement pour avoir toujours été là pour me faciliter les tâches administratives mais aussi pour les paroles de vie. Et mille mercis à Hiba pour ton support, ta gentillesse, et bien sûr pour ton amitié.

Mes remerciements vont aussi au Dr Martine HEINRICH pour avoir été toujours à l'écoute, pour sa disponibilité, pour ses conseils...

J'ai quelques remerciements plus personnels à formuler... Tout d'abord à ma mère qui va rester éternellement dans mon cœur en me supportant, m'encourageant, me dirigeant... A mon père qui grâce à lui que je suis arrivé ici.... A mes sœurs et à mes frères, que j'aime de tout mon cœur. A toute ma famille pour m'avoir soutenu durant toute ma vie et notamment mon oncle Achraf. Un merci spécial à ma cousine, Mayssa, pour sa sollicitude, sa curiosité...

Je tiens également à remercier mes amis. Tout d'abord mon meilleur ami Ahmad, qui m'écoute, me comprend, me fait rire depuis vingt ans maintenant, partage mes joies et mes peines... Merci à mes

chers, Mme et M. DHAYBI, pour être toujours à l'écoute, pour leur support inconditionnel... Merci à Marianne pour sa grande disponibilité, sa gentillesse, sa bonté, sa générosité... Je remercie aussi mes « nouveaux » amis qui ont rendu mon parcours « strasbourgeois » plus facile et plus agréable, alors merci à Mira, Joëlle, Patrick, Ranine, Batoul, Mireille... Et merci à mes nouveaux deux petits « frères » Mohamad et Radwan.

*A la mémoire de ma mère*

*A mon père*

*A mes sœurs et frères*

*A mon ami Ahmad*

# Tables des matières

<b>1.</b>	<b><u>Chapitre 1 : Introduction</u></b>	
<b>1.1.</b>	<b>Le virus d'immunodéficience humaine</b>	<b>01</b>
1.1.1.	Diversité du VIH	02
1.1.2.	Structure de la particule virale	04
1.1.3.	Génome viral	05
1.1.4.	Cycle viral	06
1.1.5.	L'enveloppe virale	10
1.1.5.A.	Membrane lipidique	10
1.1.5.B.	Glycoprotéines	12
<b>1.2.</b>	<b>Le VIH dans l'organisme</b>	<b>23</b>
1.2.1.	Transmission du VIH	23
1.2.2.	Evolution naturelle de l'infection virale	23
1.2.3.	Réservoirs et latence	26
1.2.4.	Réponse immunitaire développée contre virus	27
1.2.5.	Traitements	33
1.2.6.	Vaccination	34
<b>1.3.</b>	<b>Objectifs et contribution de travail de thèse</b>	<b>37</b>
<b>2.</b>	<b><u>Chapitre 2 : Aspect théorique</u></b>	
<b>2.1.</b>	<b>Stratégie d'expression des protéines membranaires hétérologues chez E.coli</b>	<b>39</b>
<b>2.2.</b>	<b>Dichroïsme circulaire</b>	<b>48</b>
<b>2.3.</b>	<b>Résonance magnétique nucléaire</b>	<b>52</b>
<b>3.</b>	<b><u>Chapitre 3 : Surexpression des polypeptides issus de gp41, fp et gp120</u></b>	
<b>3.1.</b>	<b>Les constructions peptidiques</b>	<b>61</b>
<b>3.2.</b>	<b>Méthodes et matériels</b>	<b>63</b>
<b>3.3.</b>	<b>Résultats et discussion</b>	<b>69</b>
<b>3.4.</b>	<b>Conclusion et perspectives</b>	<b>86</b>

<b>4.</b>	<b><u>Chapitre 4 : étude biophysiques de gp120TM, gp41TM et fpTM</u></b>	
<b>4.1.</b>	<b>Introduction</b>	<b>88</b>
<b>4.2.</b>	<b>Etude par dichroïsme circulaire</b>	<b>90</b>
<b>4.3.</b>	<b>Etude préliminaire par RMN</b>	<b>105</b>
<b>5.</b>	<b><u>Chapitre 5 : Etude de l'interaction du CRAC avec le cholestérol</u></b>	
<b>5.1.</b>	<b>Introduction</b>	<b>108</b>
<b>5.2.</b>	<b>Matériels et méthodes</b>	<b>110</b>
<b>5.3.</b>	<b>Résultats et discussion</b>	<b>114</b>
<b>5.4.</b>	<b>Conclusion et perspectives</b>	<b>140</b>
<b>6.</b>	<b><u>Chapitre 6 : Conclusion générale</u></b>	
<b>6.1.</b>	<b>Expression et étude structurale des polypeptides contenant des épitopes issus des glycoprotéines d'enveloppe du VIH-1</b>	<b>142</b>
<b>6.2.</b>	<b>Etude de l'interaction entre le cholestérol et le CRAC de la gp41</b>	<b>145</b>
	<b><u>Bibliographie</u></b>	<b>147</b>

# Liste des figures

Figure 1 : Diversité géographique du VIH-1. ....	3
Figure 2 : Diversité génétique du VIH et du SIV.....	3
Figure 3: Structure du VIH (NIAID, 2013).....	4
Figure 4: Structure génomique du provirus VIH-1.....	5
Figure 5 : Cycle réplcatif du VIH (NIAID, 2013). ....	6
Figure 6 : Modèle d'entrée du VIH et de la fusion des membranes virale et cellulaire. ....	8
Figure 7 : Composition lipidique de la membrane plasmique. ....	11
Figure 8 : Composition lipidique des particules virales. ....	11
Figure 9 : Schéma de la gp120 montrant la topologie des régions variables et constantes. ....	13
Figure 10 : Comparaison des structures de la gp120 issues du VIS et VIH-1 .....	15
Figure 11 : Structure quaternaire de la gp120.....	17
Figure 12 : Représentation schématique des domaines de la gp41.....	18
Figure 13 : Structure post-fusionnelle de la gp41.....	19
Figure 14 : Structure du peptide de fusion du VIH-1. ....	21
Figure 15 : Topologie de la MPER contenant le motif CRAC. ....	22
Figure 16 : Évolution des différentes phases de l'infection à VIH et variation de paramètres biologiques. ....	24
Figure 17 : Différents mécanismes d'action des anticorps anti-VIH. ....	30
Figure 18 : Schéma du « spike » du VIH-1 montrant les épitopes et leurs anticorps neutralisants correspondants. ....	32
Figure 19 : Représentation schématique des éléments essentiels d'un vecteur d'expression procaryote.....	40
Figure 20 : Carte du plasmide pTIPX-4 .....	42
Figure 21 : Système d'expression dans pET.....	45
Figure 22 : Diagramme montrant l'équilibre les différentes formes de la protéine exprimée. ....	46
Figure 23 : Diagramme montrant l'orientation des transitions dipolaires.....	49
Figure 24 : Spectres dichroïsme circulaire montrant les différentes structures secondaires des protéines .....	49
Figure 25 : Eclatement des niveaux d'énergie pour un spin $\frac{1}{2}$ par application d'un champ magnétique statique $B_0$ : Effet Zeeman. ....	53
Figure 26 : Représentation du tenseur du déplacement chimique dans le référentiel PAS. ....	55
Figure 27 : Paramètres d'asymétrie influençant la forme de spectre. ....	55
Figure 28 : Représentation du couplage dipolaire entre deux spins I et S. ....	57
Figure 29 : Comparaison de la distribution de la charge entre deux noyaux ayant un spin égale ou supérieure à $1/2$ . ....	58
Figure 30 : Conception des séquences polypeptidiques à surexprimer. ....	62
Figure 31 : Cinétique de la croissance bactérienne au cours de l'expression du TAF12-gp120TM. ....	69
Figure 32 : Influence de la température et la DO d'induction sur la cinétique de la croissance bactérienne au cours de l'expression du TAF12-gp41TM .....	70
Figure 33 : Migration des différentes fractions d'extraction protéique pour TAF-gp120TM.....	71
Figure 34 : Western Blot anti-His tag sur les différentes fractions de l'extraction protéique des bactéries exprimant TAF12-gp120TM .....	72
Figure 35 : Migration de quantités équivalentes de la fraction insoluble des bactéries exprimant TAF12-gp120TM .....	73
Figure 36 : Western Blot anti-His tag sur les différentes fractions de l'extraction protéique des bactéries exprimant TAF12-gp41TM.....	74
Figure 37 : Western Blot anti-His tag sur les différentes fractions de l'extraction protéique des bactéries exprimant TAF12-fpTM .....	74
Figure 38 : SDS-PAGE 16% sur les différentes fractions de la purification TAF12-gp120TM par résine de cobalt.....	75
Figure 39 : Western blot anti-His tag sur les différentes fractions de la purification TAF12-gp120TM par résine de cobalt. ....	76
Figure 40 : Chromatogramme de TAF12-gp120TM à t=0 du clivage .....	78
Figure 41 : Western blot anti-His tag sur les fractions collectées à t=0 de clivage.....	78

Figure 42 : Spectre de masse ESI Q-TOF de la fraction 16 à t=0 de clivage. ....	80
Figure 43 : Spectre de masse MALDI-TOF des fractions collectés à t=0 de clivage .....	80
Figure 44 : Chromatogramme de TAF12 / gp120TM à t=24h du clivage .....	81
Figure 45 : Identification de la fraction d'éluion contenant la gp120TM par spectrométrie de masse et SDS-PAGE .....	83
Figure 46 : Spectre du dichroïsme circulaire de la TM insérée dans des SUV de POPC/POPS (3/1). ....	93
Figure 47 : Spectre du dichroïsme circulaire de la gp120TM insérée dans des SUV POPC/POPS .....	94
Figure 48 : Spectre du dichroïsme circulaire de la gp120TM .....	96
Figure 49 : Spectre du dichroïsme circulaire de la gp41TM .....	98
Figure 50 : Spectre dichroïsme circulaire du fpTM. ....	100
Figure 51 : Reconstitution des polypeptides dans des liposomes.....	104
Figure 52 : Spectre 1H-15N HSQC du fpTM dans un milieu micellaire .....	106
Figure 53 : Oligomérisation des peptides CRAC_1, son contrôle et CRAC_2 .....	115
Figure 54 : Spectre de dichroïsme circulaire de CRAC_1.....	117
Figure 55 : Analyse des structures secondaires du CRAC_1 en fonction de [cholestérol] .....	117
Figure 57 : Analyse des structures secondaires du CRAC_2 en fonction de [cholestérol] .....	118
Figure 56 : Spectre de dichroïsme circulaire de CRAC_2.....	118
Figure 58 : Spectres FTIR.....	120
Figure 59 : modèle montrant les orientations potentielles du motif CRAC par rapport à la séquence transmembranaire. ....	123
Figure 60 : Spectres RMN 15N du CRAC_1. ....	125
Figure 61 : Spectres RMN 15N du CRAC_1 Contrôle.....	126
Figure 62 : Spectres 15N-RMN du CRAC_2. ....	129
Figure 63 : Spectres RMN de deutérium. ....	132
Figure 64 : Structure du cholestérol marqué en deutérium. ....	133
Figure 65 : Spectres RMN de deutérium du cholestérol-d6.....	135
Figure 66 : Influence de la température sur les éclatements quadripolaires. ....	136
Figure 67 : Spectres RMN de deutérium du cholestérol-d6.....	138
Figure 68 : Spectres RMN de deutérium du cholestérol-d7.....	139

# Liste des tableaux

Tableau 1 : Composition lipidique majeure du VIH-1. _____	11
Tableau 2 : Les différents catégories et fonctions des anticorps neutralisants dirigés contre le VIH-1. __	31
Tableau 3 : Les partenaires de fusion les plus utilisés. _____	46
Tableau 4 : Différentes méthodes de clivage des partenaires de fusion. _____	46
Tableau 5 : Récapitulatif des conditions d'expression et de purification des trois séquences. _____	87
Tableau 6 : Analyse des structures secondaires de la gp120TM. _____	97
Tableau 7 : Analyse des structures secondaires de la gp41TM. _____	99
Tableau 8 : Analyse des structures secondaires du fpTM. _____	101
Tableau 9 : Liste des peptides conçus et préparés contenant le motif CRAC. _____	112
Tableau 10 : Tableau récapitulatif des spectres azote de la catégorie 1 des peptides. _____	128
Tableau 11 : Tableau récapitulatif des spectres azote de la catégorie 2 des peptides. _____	131
Tableau 12 : Tableau récapitulatif des spectres deutérium des catégories 1 et 2 des peptides. _____	134
Tableau 13: Tableau récapitulatif des éclatements quadripolaires de cholestérol-d6 10%. _____	139
Tableau 14 : Tableau récapitulatif des éclatements quadripolaires de cholestérol-d7 10%. _____	140

# Liste des abréviations

ADN: Acide Désoxyribonucléique

ARN: Acide Ribonucléique

CD4: Cluster de Différenciation 4

CI : Corps d'Inclusion

DC: Dichroïsme Circulaire

DO : Densité Optique

DHSM: Dihydrosphingomyeline

DPC: n-dodécylephospholine

*E. coli* : *Escherichia coli*

EDTA : EthyleneDiamineTetraacetic Acid

ET : Echantillon Total

ESI-QTOF : ElectroSpray Ionization-Quadrupole Time Of Flight

Fmoc : FluorenylMethylOxyCarbonyl

gp: glycoprotéines

HPLC: High Performance (or Pressure) Liquid Chromatography

HSQC: Heteronuclear single quantum coherence spectroscopy

IMAC: Immobilized Metal Affinity Chromatography

IPTG: Isopropyl  $\beta$ -D-1-thiogalactopyranoside

LB: Lorenz Broadening

LBYE : Luria Bertani Yeast Extract

LTR: Long Terminal Repetition

MALDI-TOF: Matrix-assisted laser desorption/ionization – Time Of Flight

Mbe : membranes

PE: Phosphatidylethanolamine

POPC: 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine

POPS: 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-L-serine

pTIPX-4: Taf12-mediated Insoluble Peptide eXpression 4

RMN: Résonance Magnétique Nucléaire

SDS: Sodium Dodecyl Sulfate

SDS-PAGE: SDS- Polyacrylamide gel electrophoresis

SIDA: Syndrome d'Immunodéficience Acquis

SIV: Virus de l'Immunodéficience Simienne

SM: Sphingomyéline

TAF12: TATA box binding protein (TBP)-Associated Factor 12 Histone fold Domain

TFA: Acide Tri-fluoro-acétique

TFE: 2, 2, 2-trifluoroethanol

TM : Transmembranaire

VIH: Virus d'Immunodéficience Humaine

VIH-1: VIH de type 1

# 1. Chapitre 1 : Introduction

## 1.1. Le virus d'immunodéficience humaine

Le Syndrome d'ImmunoDéficience Acquis (SIDA) est une pandémie qui touche près de 34 millions de personnes dont 2,5 millions d'enfants, et a tué environ 1,8 millions de personnes en 2010 (UNAIDS, 2011). Le SIDA est provoqué par le virus d'immunodéficience humaine (VIH) qui est un rétrovirus transmis par contact direct des muqueuses ou voies sanguines. Ce virus a comme cible les cellules présentant la protéine CD4 (Cluster de Différenciation 4) entraînant la destruction du système immunitaire, ce qui rend la personne atteinte vulnérable à d'autres infections.

Actuellement, le principal traitement reste la trithérapie composée d'antirétroviraux permettant de diminuer la charge virale mais pas l'éradication de la maladie. Ce type de traitement à vie, extrêmement coûteux, n'est pas accessible à tous. De plus, de nombreuses résistances ont été développées par le virus (Leonard *et al.*, 1990; Lorenzi *et al.*, 1999; Miller *et al.*, 2005; Tossonian *et al.*, 2009). C'est pour ces raisons que la prévention contre l'infection reste toujours une priorité des organisations de la santé. Cependant, jusqu'à aujourd'hui, aucun vaccin anti-VIH ni aucun traitement curatif du SIDA n'a encore été mis au point.

En effet, le développement d'un vaccin contre le VIH/SIDA se heurte à des difficultés considérables qui sont bien connues: l'hypervariabilité génétique du virus et la génération de mutants d'échappement aux lymphocytes T cytotoxiques ou aux anticorps neutralisants (error-prone reverse transcriptase), l'intégration du génome viral sous forme de provirus potentiellement latent à l'abri du système immunitaire (Girard and Sanders, 2013), d'où l'incapacité où nous sommes aujourd'hui encore d'induire par la vaccination des anticorps neutralisants de large spectre d'activité qui soient actifs contre les isolats sauvages du virus, etc.

### 1.1.1. Diversité du VIH

Le VIH est originaire du virus de l'immunodéficience simienne (SIV) et a été transmis des primates non humains aux humains en Afrique occidentale et centrale (Hemelaar, 2012). Cette transmission entre les espèces a probablement eu lieu par le processus de la chasse, la boucherie de la viande, ou la conservation des singes comme animaux de compagnie (Hahn *et al.*, 2000)... Plus de quarante espèces différentes de primates non humains développent des infections causées par un SIV spécifique à chacune d'elles.

Indépendamment de la transmission inter-espèces, le VIH, lui-même, existe sous forme de lignées distinctes qui diffèrent par leur localisation géographique et l'organisation de leur génome (voir Figure 1 et Figure 2): le VIH de type 1 (VIH-1), le plus répandu au niveau mondial, et le VIH de type 2 (VIH-2) qui se concentre principalement en Afrique Occidentale. Chaque lignée est divisée en plusieurs groupes (VIH-1 : sous-types M, N, O ... ; VIH-2 : sous-types A, B, C...). Puis, au sein même de ces groupes, il existe des sous-groupes (A1, B1, BC...) (Locatelli and Peeters, 2012).



## 1.1.2. Structure de la particule virale

Les particules virales ou virions sont constituées d'une enveloppe d'origine cellulaire (Earl *et al.*, 2013) formée par une membrane lipidique dans laquelle est insérée la glycoprotéine 41 (gp41) sous forme de trimère. Cette dernière interagit avec une autre glycoprotéine (gp120) qui occupe une position plus périphérique. Elles forment un « spike » ou un complexe hétéro-trimérique. L'enveloppe du VIH est tapissée à l'intérieur par une matrice formée par la protéine p17. Au cœur de la particule virale, se trouve la capside constituée par la protéine p24. Cette capside protège le génome viral et contient les enzymes nécessaires à la réplication du virus (transcriptase inverse, intégrase...) (Sundquist and Kräusslich, 2012).

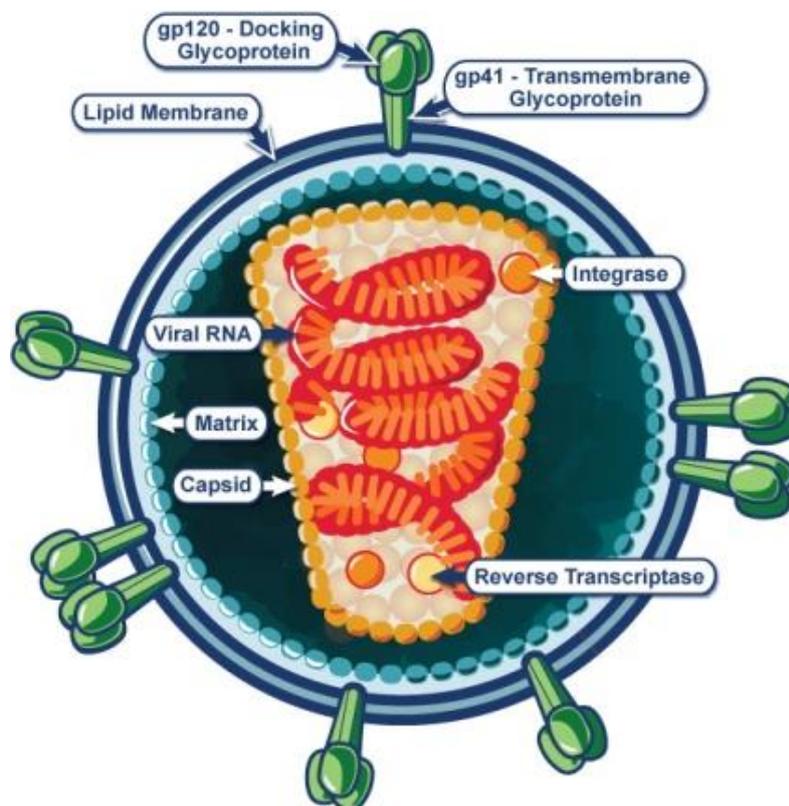


Figure 3: Structure du VIH (NIAID, 2013).

### 1.1.3. Génome viral

Le virus porte son matériel génétique sur deux molécules d'ARN simple brin d'environ 9 kilobases (kb) renfermant 9 gènes qui codent pour 15 protéines différentes. Chaque molécule d'ARN est flanquée par des séquences spécifiques, des longues répétitions terminales (LTR). Entre ces LTR, le génome comporte trois gènes principaux (Gag pour Group-specific antigen, Pol pour polymérase, et Env pour envelope) dans le sens 5'-3' (voir Figure 4) :

- le gène Gag code pour la poly-protéine précurseur qui se clive en plusieurs protéines : p17 et p24 pour les capsides ; p7 et p9 qui sont des protéines de la nucléocapside
- le gène Pol code pour la protéine précurseur p180 qui se clive en p66 et p51 (reverse transcriptase), p32 (intégrase) et p10 (protéase).
- le gène env code pour le précurseur des glycoprotéines gp160 qui se clive en gp120 et gp41

Ce génome code aussi pour des protéines régulatrices (Vif, Vpr, Tat...). Elles interviennent dans la régulation de l'expression des protéines virales et dans la multiplication du virus. Il semble même modifier l'expression de certains gènes cellulaires entraînant leur altérations d'où la destruction du système immunitaire hôte (Harris *et al.*, 2012).

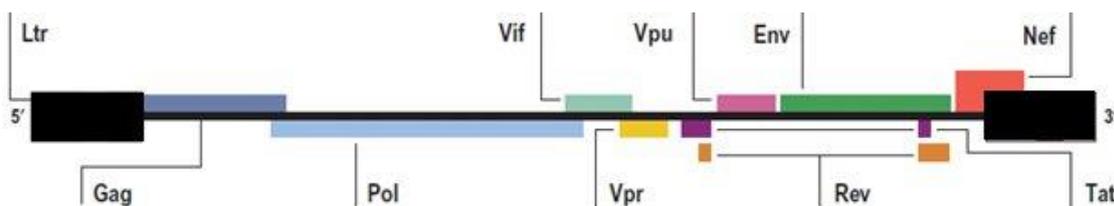


Figure 4: Structure génomique du provirus VIH-1 (Greene and Peterlin, 2008).

## 1.1.4. Cycle viral

La Figure 5 résume le cycle répliatif du VIH qui a besoin d'utiliser la machinerie de la cellule hôte pour pouvoir se répliquer.

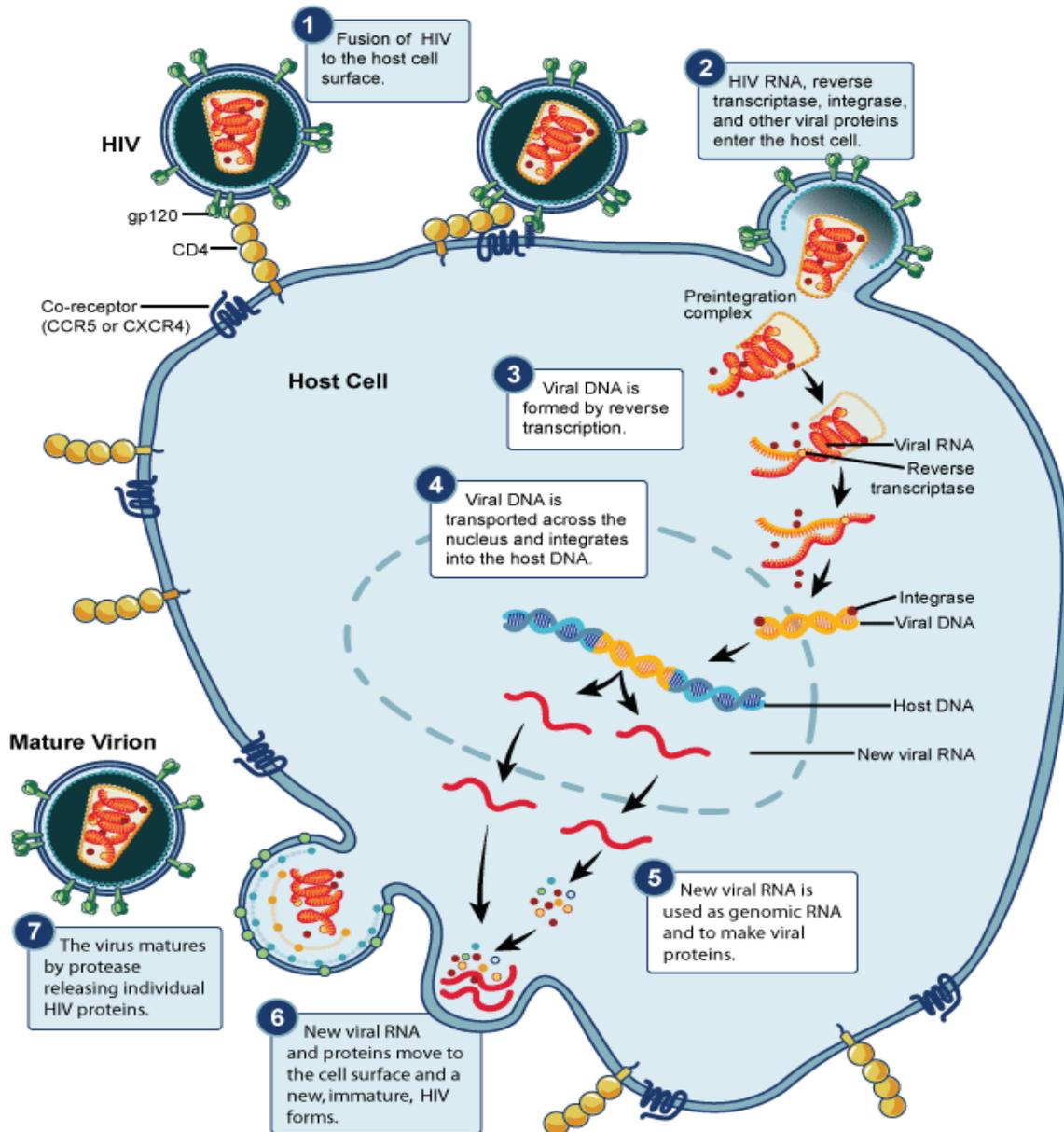


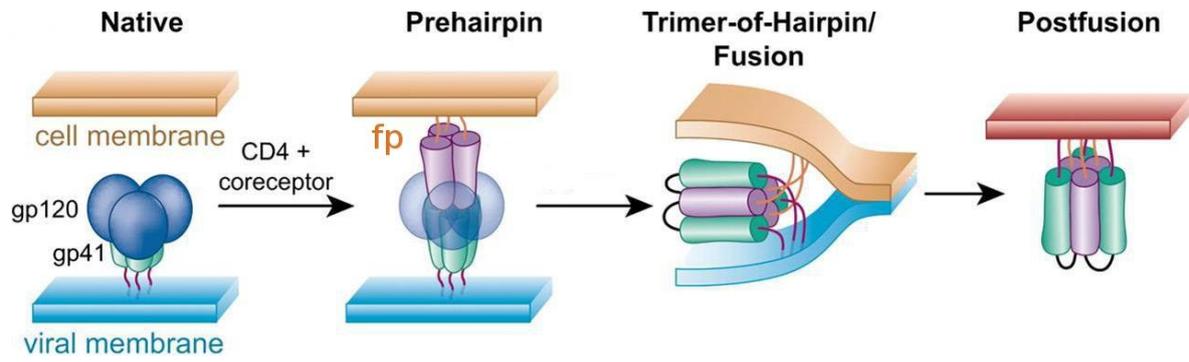
Figure 5 : Cycle répliatif du VIH (NIAID, 2013).

### 1.1.4.A. Entrée du virus

Les glycoprotéines (gp120 et gp41) d'enveloppe du VIH sont initialement produites sous la forme d'un précurseur, la glycoprotéine 160 (gp160) qui est clivée par une convertase en deux sous-unités liées par des liaisons non covalentes :

- gp120, sous-unité de surface, responsable de la reconnaissance de la cellule cible par interaction avec le récepteur CD4 et avec l'un des corécepteurs de la famille des récepteurs aux chimiokines (CCR5 ou CXCR4) ;
- gp41, sous-unité transmembranaire, responsable de la fusion de la membrane virale avec celle de la cellule cible. Cette sous-unité a deux conformations : native et celle qui permet la fusion entre les deux cellules.

Après la phase de contamination, le virus se retrouve dans la circulation sanguine. La Figure 6. montre d'une façon schématisée le mécanisme d'entrée du virus qui reconnaît la cellule hôte et se fixe à elle grâce à l'interaction entre la gp120 et le récepteur CD4 de la cellule (Sattentau and Weiss, 1988) (lymphocyte T CD4+, macrophage...). La seule présence du CD4, si elle est nécessaire, n'est pas suffisante. Des récepteurs supplémentaires sont impliqués dans cette reconnaissance. Il s'agit l'un des corécepteurs de la famille des récepteurs aux chimiokines (CCR5 ou CXCR4) (Alkhatib *et al.*, 1996). A l'issue de cette reconnaissance, le « spike » subit un changement de conformation, de la forme pré-fusionnante à la forme fusionnante. En effet, un intermédiaire « pre-hairpin » se forme tout en exposant une séquence cachée, le peptide de fusion (fp en orange), qui s'insère dans la membrane cellulaire. Le passage de cet intermédiaire à l'état *hairpin* est une réaction lente qui consiste en l'association des parties N-terminale et C-terminale de la gp41 et l'adoption d'une conformation hélicoïdale qui correspond à l'intermédiaire *hairpin*. Finalement, les membranes virale et cellulaire fusionneront (Chan *et al.*, 1997a; Miller *et al.*, 2005).



**Figure 6 : Modèle d'entrée du VIH et de la fusion des membranes virale et cellulaire (Chan et al., 1997; Miller et al., 2005).**

### 1.1.4.B. Évènements post-fusionnels

Une fois les membranes virale et hôte fusionnées, le contenu du virion est libéré dans la cellule hôte. Plusieurs étapes vont alors s'enchaîner pour mener à la production de nouvelles particules virales.

- La décapsidation : la capsid est détruite par des protéines cellulaires (cyclophilin A...) et virales (Nef, Vif...) ce qui permet la libération du génome et des enzymes virales dans le cytoplasme de la cellule hôte.
- La réplication du génome viral par la transcriptase inverse. Cette enzyme qui a un taux de mutations élevé rétrotranscrit les deux molécules ARN virale en ADN simple brin. Puis, elle les hydrolyse et duplique l'ADN néosynthétisé pour former un ADN bicaténaire.
- L'intégration : l'ADN viral double brin est transporté dans le noyau de la cellule cible avec l'intégrase virale qui coupe l'ADN de la cellule et insère l'ADN viral. A cette étape, il s'agit du provirus qui peut rester latent pendant plusieurs années dans la cellule hôte.
- La transcription et expression des gènes viraux : l'activation de la cellule hôte infectée entraîne la transcription du génome viral par l'ARN polymérase II cellulaire. Les ARN messagers viraux sont transportés dans le cytoplasme pour être traduit en protéines virales.
- L'assemblage et le bourgeonnement des virions : les particules virales s'assemblent au niveau de la membrane plasmique cellulaire tout en formant des virions immatures. Leur maturation et leur bourgeonnement s'effectuent après clivage des poly-protéines Gag-Pol, Gap et Env en protéines structurales du VIH et en ses enzymes fonctionnelles (Greene and Peterlin, 2008; Teixeiraa *et al.*, 2011).

## 1.1.5. L'enveloppe virale

L'enveloppe virale est formée par les glycoprotéines (gp120 et gp41) ancrées dans une membrane lipidique. Celle-ci est tapissée à l'intérieur par la matrice. Dans ce qui suit, nous nous concentreront uniquement sur les composants qui nous intéressent.

### 1.1.5.A. Membrane lipidique

L'enveloppe des virions néosynthétisés est d'origine cellulaire. Pourtant, la composition en lipides de la membrane virale est différente que celle de la membrane de la cellule productrice du virion. En effet, une étude menée par l'équipe de Hans-Georg Krausslich et de Britta Brugger (Brugger *et al.*, 2006) a montré que la membrane virale est fortement enrichie en cholestérol (Chol) par rapport à la membrane cellulaire, de même, dans une moindre mesure, pour la sphingomyéline (SM) et le dihydrosphingomyéline (DHSM), le phosphatidylserine (PS), plasményléthanolamine (pL-PE). Par contre, le taux de phosphatidylcholine (PC) et celui du phosphatidylethanolamine (PE), composants majeurs de la membrane des cellules de mammifère, sont plus bas dans la membrane virale (voir Figure 8 et Figure 7).

Lipide	Nb/virion
PC	26000
SM	37000
DHSM	17000
PE	13000
p1-PE	44000
PS	25000
Chol	134000

Tableau 1 : Composition lipidique majeure du VIH-1 (D'après Brügger et al., 2006)

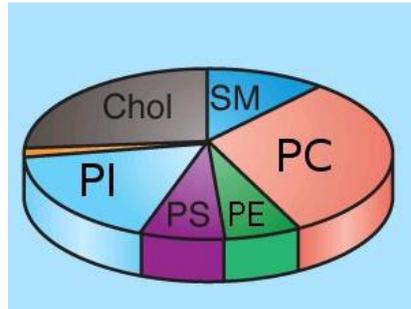


Figure 7 : Composition lipidique de la membrane plasmique (van Meer and de Kroon, 2011).

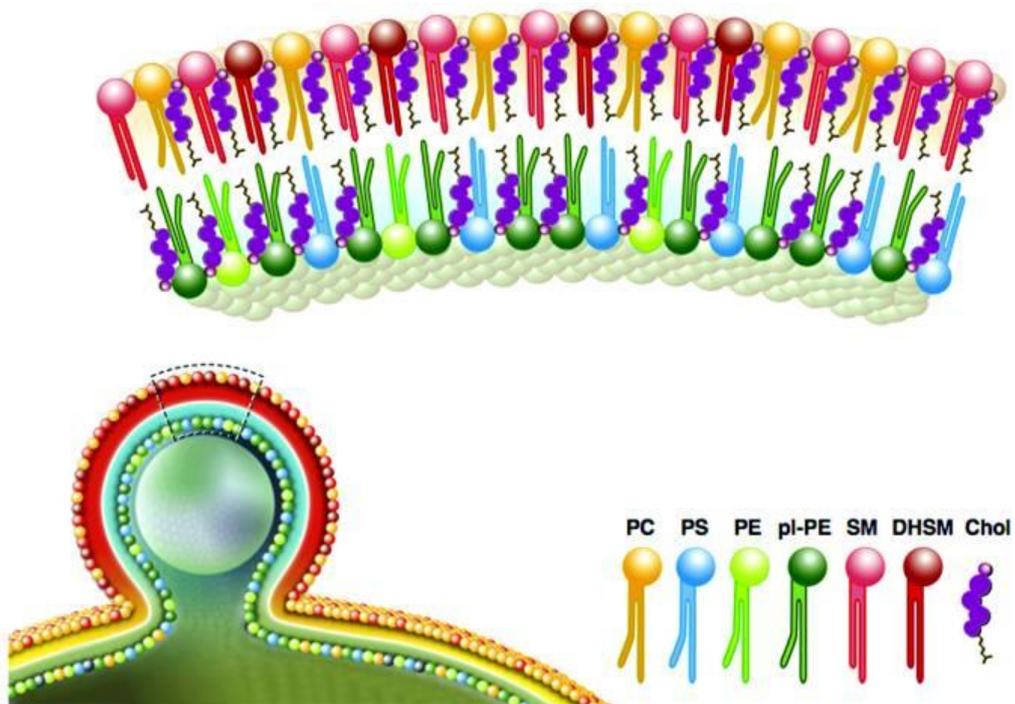


Figure 8 : Composition lipidique des particules virales (Brügger et al., 2006).

## 1.1.5.B. Glycoprotéines

La glycoprotéine d'enveloppe est une protéine transmembranaire de type I, c.-à-d. que sa partie transmembranaire est une hélice  $\alpha$  flanquée d'une extrémité N-terminale dirigée vers le milieu extracellulaire ou le lumen et par une extrémité C-terminale qui est dirigée vers le cytoplasme de la cellule. Cette protéine précurseur (gp160) est synthétisée dans le réticulum endoplasmique. La gp160 se structure en un trimère associé d'une façon non-covalente, hautement glycosylé et comprenant plusieurs ponts disulfures. Puis, elle est transportée dans l'appareil de Golgi où elle subit un clivage par une protéase pour donner naissance aux gp120 et gp41.

### 1.1.5.B.a. gp120

#### 1.1.5.B.a.i. Topologie de la gp120

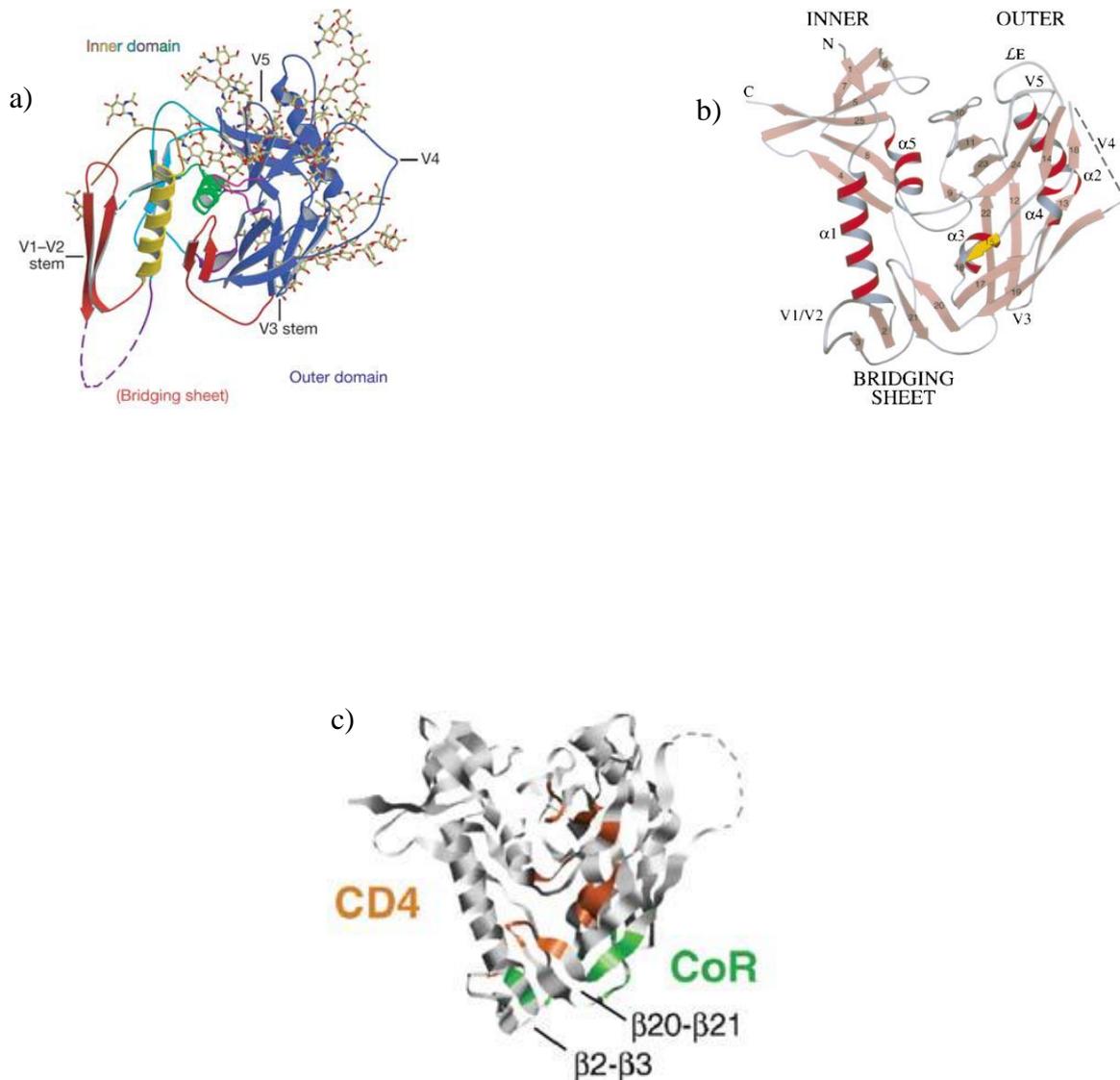
L'alignement des différentes séquences du gène *env* appartenant à différents virus d'immunodéficience infectant les primates montre que la gp120 est formée de cinq régions variables (V1-V5) exposés à la périphérie de celle-ci. La variabilité de ces régions est due à la recombinaison, les mutations, les insertions et les délétions au niveau de l'ARN viral. Elles sont disposées en alternance avec des régions conservées (C1-C5), situées au cœur de la gp120 (Figure 9). Ces dernières sont indispensables à l'interaction avec l'ectodomaine de la gp41 et les récepteurs de la cellule hôte. Toutes ces régions sont hautement glycosylées (Kwong *et al.*, 1998; Leonard *et al.*, 1990). En effet, la moitié de la masse de cette gp provient principalement des N-glycosylations avec une petite contribution des O-glycosylations (Wyatt *et al.*, 1998). De plus, plusieurs ponts disulfures sont fortement conservés permettant la formation de la structure tertiaire de l'Env.



### 1.1.5.B.a.ii. Structure atomique de la gp120

A ce jour, il n'existe pas une structure tridimensionnelle à haute résolution de la gp120 du VIH non liée, entière et glycosylée, autrement dit la structure native de la gp120 du VIH. Cela est dû à la difficulté de cristalliser cette glycoprotéine en absence de ligand (CD4 ou anticorps) pour la stabiliser (Guttman *et al.*, 2012; Huang *et al.*, 2005). Cependant, il existe une structure de la gp120 du VIS non liée et glycosylée. Les deux séquences de la gp120, celles de VIH-1 et VIS, sont identiques à 30% et ont une similarité de 70% (Chen *et al.*, 2005; Wyatt *et al.*, 1998) ceci contribue à mieux comprendre la structure de la gp120 du VIH-1.

La Figure 10 (a et b) montre que la gp120 est composée de trois domaines : « inner domain », « outer domain » et « bridging sheet » selon leur orientation dans le complexe trimérique. Qu'elle soit liée ou non, le domaine externe est hautement glycosylé, ce qui permet au virus de s'échapper au système immunitaire. Le domaine interne formé essentiellement des régions conservées C1 et C5 est dépourvu de glycosylation. Ce domaine interagit probablement avec la gp41 et les extrémités N et C terminales de la gp120. Ces deux domaines sont reliés par le « bridging sheet ». La gp120 dans sa forme complexée est composée par 25 feuillets  $\beta$ , 5 hélices  $\alpha$  et 10 boucles.



**Figure 10 : Comparaison des structures de la gp120 issues du VIS et VIH-1. a) Structure de la gp120 glycosylée du VIS non liée (Chen et al., 2005). b) et c) Structure de la gp120 non glycosylée du VIH-1 liée à CD4 et au fragment Fab d'un anticorps monoclonal (Kwong et al., 1998; Pantophlet and Burton, 2006; Wyatt et al., 1998). V : région variable,  $\alpha$  : région hélicale conservée, L : boucle.**

Les structures de la gp120 du VIH-1 lié à CD4 et du VIS représentent une grande similarité au niveau du domaine extérieur. Tandis que la conformation du domaine interne de la gp120 diffère entre elles. Cela veut dire que ce domaine a une flexibilité vis à vis du CD4. De même, la comparaison des sous-structures du domaine interne dans le complexe lié ou non, suggère que ces sous-structures se réarrangent d'une manière indépendante entre elles. Les changements conformationnels qui ont lieu dans le domaine interne influencent aussi la formation du « bridging sheet » qui se replie sous forme de quatre brins antiparallèles ( $\beta 2$ - $\beta 3$  et  $\beta 20$ - $\beta 21$ ) dans le complexe lié à CD4. Les deux brins  $\beta 2$  et  $\beta 3$  qui flanquent la boucle variable V1/V2 sont situés à proximité du domaine interne (ces brins oscilleraient entre deux conformations, celle du complexe non lié et celle du complexe lié). Alors que les deux autres brins sont proches du domaine externe et conservent leur position que le complexe soit lié ou non à CD4.

#### 1.1.5.B.a.ii.i. Le site de fixation sur CD4

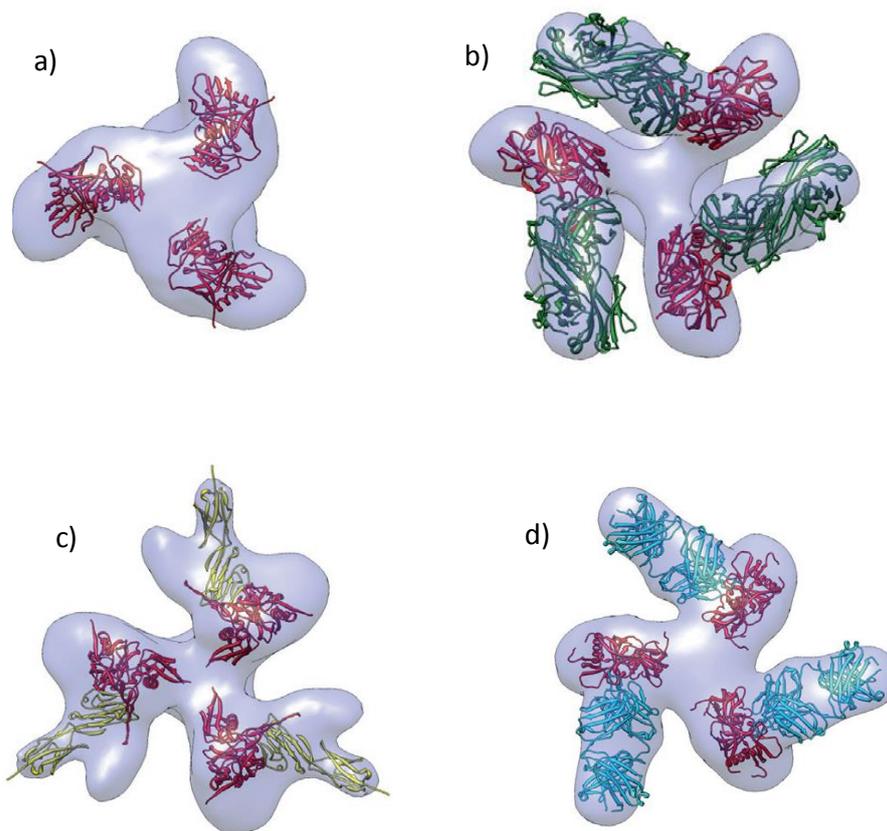
Le site de fixation à CD4 (Figure 10.c, en orange) est formé par l'interface entre le domaine interne, le domaine externe et le segment  $\beta 20$ - $\beta 21$  du « bridging sheet ». Au centre de cette interface, il y a une cavité hydrophobe appelée cavité Phe43. Cependant, la majorité des résidus qui entre en contact avec CD4 sont localisés dans le domaine externe du complexe lié à CD4. Cette cavité est absente dans la gp120 non liée du VIS.

#### 1.1.5.B.a.ii.ii. Le site de fixation sur le corécepteur

La région primordiale pour l'interaction avec le corécepteur (CCR5 ou CXCR4) est composée par des résidus présents dans le « bridging sheet » et proche de la boucle V3 (Figure 10.c, en vert). Ces résidus se regroupent ensemble dans le complexe gp120 lié à CD4, par contre ceux du VIS sont séparés en deux régions. Ceci signifie que la fixation à CD4 est indispensable pour rendre le site de fixation sur le corécepteur accessible à la gp120.

### 1.1.5.B.a.iii. Structure quaternaire de la gp120

La combinaison de la tomographie cryo-électronique avec les structures tertiaires obtenues par cristallisation aux rayons X contribue à déterminer les structures quaternaires de la gp120 sur les virus intacts (Figure 11). De plus, la localisation des boucles variables a été bien définie au sein du trimère. Par exemple, la boucle V1/V2 est localisée à proximité de l'apex du « spike » (Hu *et al.*, 2011; White *et al.*, 2010). En fait, lorsque le trimère de la gp120 se lie au récepteur CD4 ou à l'anticorps monoclonal b17, il subit une transition importante d'une conformation « fermée » à une autre plus ouverte. Cela est dû à la rotation de chacun des trois monomères de la gp120 et à l'exposition concomitante des boucles V1/V2 à la périphérie du « spike ». Par contre, cette transition est moins prononcée dans le cas où le trimère se lie à un autre anticorps tel que le b12.



**Figure 11 : Structure quaternaire de la gp120 (Merk and Subramaniam, 2013). Vue de dessus des cartes de densité de l'Env native : a) non lié, b) lié à l'anticorps monoclonal 17b, c) lié au CD4 soluble et d) lié à l'anticorps monoclonal b12. Monomère gp120 en rouge, CD4 soluble en jaune, 17b en vert et b12 en bleu.**

### 1.1.5.B.b. gp41

#### 1.1.5.B.b.i. Structure primaire de la gp41

La gp41 est composée de trois domaines (Figure 12): un domaine extracellulaire ou ectodomaine, un domaine transmembranaire (TM) et un domaine cytoplasmique (CT). Le domaine extracellulaire contient les séquences indispensables à la fusion des membranes virale et cellulaire : le peptide de fusion (fp), une région hautement conservée et riche en glycine, est située à l'extrémité N-terminale (Nt). Une région polaire (PR) qui sépare le fp des heptades répétées (4-3 répétitions hydrophobes) : NHR et CHR (ou HR1 et HR2) sont situées respectivement aux extrémités Nt et Ct; elles sont connectées par une boucle glycosylée de 30-40 résidus. La région externe pré-transmembranaire ou MPER (membrane proximal external region) est une région riche en tryptophane.

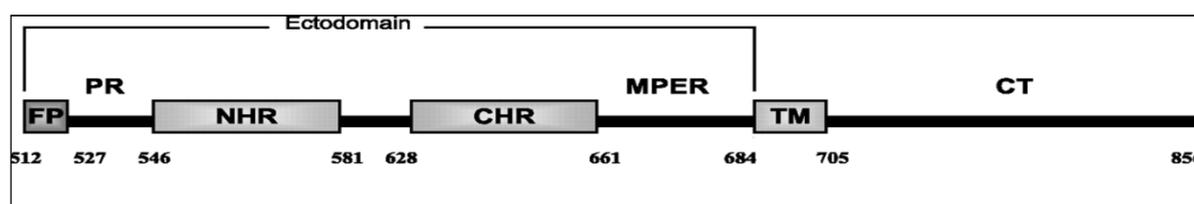
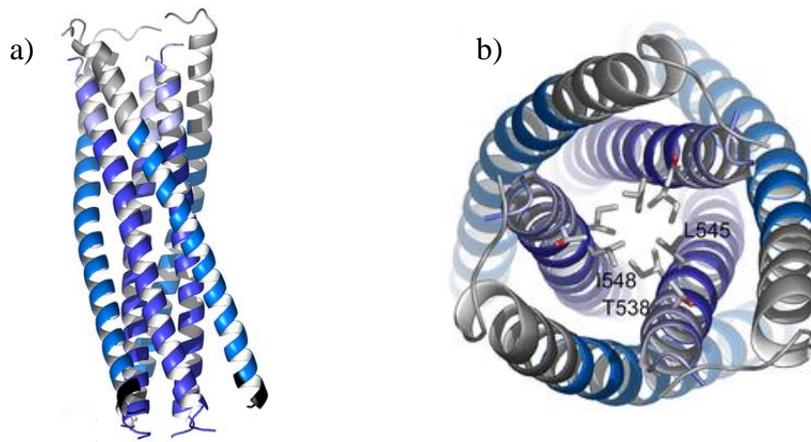


Figure 12 : Représentation schématique des domaines de la gp41 (Montero *et al.*, 2008). La numérotation est basée sur le sous-type HIV-1 HXB2 servant de référence. FP : peptide de fusion, NHR : Nt Heptades Repeat, CHR : Ct Heptades Repeats, MPER : Membrane Proximal External Region, TM : Transmembranaire et CT : Cytoplasmique.

#### 1.1.5.B.b.ii. Structure moléculaire de la gp41

La structure tridimensionnelle de la gp41 native n'est pas encore connue. Pourtant, quelques structures ont été résolues pour la gp41 tronquée ou pour l'un de ses domaines. Le premier aperçu était obtenu en 1997 lorsque plusieurs équipes ont résolu la structure de l'ectodomaine de la gp41 dans sa conformation post-fusionnelle qui consiste en six brins hélicoïdaux (six-helix-bundle: 6HB). Au centre de cette conformation, trois domaines NHR empaquetés en une super hélice gauche forment un trimère coiled-coil qui est entouré par trois domaines CHR empaquetés d'une manière antiparallèle par rapport aux NHR Figure 12 (Buzon *et al.*, 2010; Chan *et al.*, 1997b; Tan *et al.*, 1997; Weissenhorn *et al.*, 1997).



**Figure 13 : Structure post-fusionnelle de la gp41 (Buzon *et al.*, 2010; Chan *et al.*, 1997b). Représentation en ruban de la conformation 6HB. a) vue de côté: la structure déterminée par Chan et al. (NHR, en bleu sombre et CHR, en bleu ciel), la boucle qui relie NHR et CHR est représentée en noir, la région polaire en Ct du fp est représentée en bleu clair et la MPER en gris. b) vue de dessus : les NHR au centre du 6HB entourés par les CHR.**

Malgré le fait que la gp41 ait été cristallisée dans différentes conditions, liée ou non à un ligand (inhibiteur de la fusion des membranes virale et cellulaire ou anticorps) ou bien sous une forme tronquée avec ou sans mutations, toutes les structures de la gp41 connues jusqu'à présent sont quasiment identiques (Merk and Subramaniam, 2013), une structure bien définie et conservée.

#### 1.1.5.B.b.ii.i. Le rôle du peptide de fusion

Le fp correspond aux premiers 15 acides aminés (aa) de la gp41 qui sont suivis par une région polaire. La séquence primaire de ce peptide montre un motif qui ressemble à celui des polypeptides amphiphiles, un heptade hydrophobe répété (HPPHPPP)<sub>n</sub>. Les H et P représentent respectivement des résidus hydrophobes et polaires (Volkov and Bonn, 2013). Ce peptide caché dans le complexe gp120/gp41 s'expose transitoirement après la reconnaissance de la cellule cible et les changements de conformation subis par les deux gp. Il s'insère dans la membrane cellulaire pour la déstabiliser tout en formant des pores de fusion. Il permet également le rapprochement des membranes cellulaire et virale.

Sa structure dans la gp41 native est toujours inconnue. En effet, plusieurs études menées sur des peptides issus du fp montrent des résultats contradictoires. Par exemple, des études de Fourier Transform Infrared (FTIR) sur le fp en contact avec des membranes de phospholipides (Gordon *et al.*, 2004) et de RMN dans des micelles de phospholipides (Li and Tamm, 2007) ou de dodécylsulfate de sodium (SDS) (Jaroniec *et al.*, 2005) montrent que ce peptide, à faible concentration, s'insèrent dans les membranes sous forme d'hélice  $\alpha$ . Par contre, à concentration élevée, le FTIR et la RMN à basse température (Moblely *et al.*, 1999; Schmick and Weliky, 2010) montre que le peptide s'assemble comme les peptides amyloïdes (feuillet  $\beta$ ). De plus, la RMN montre que le peptide se structure en majoritairement en feuillet  $\beta$  antiparallèle lorsqu'il est congelé dans des membranes.

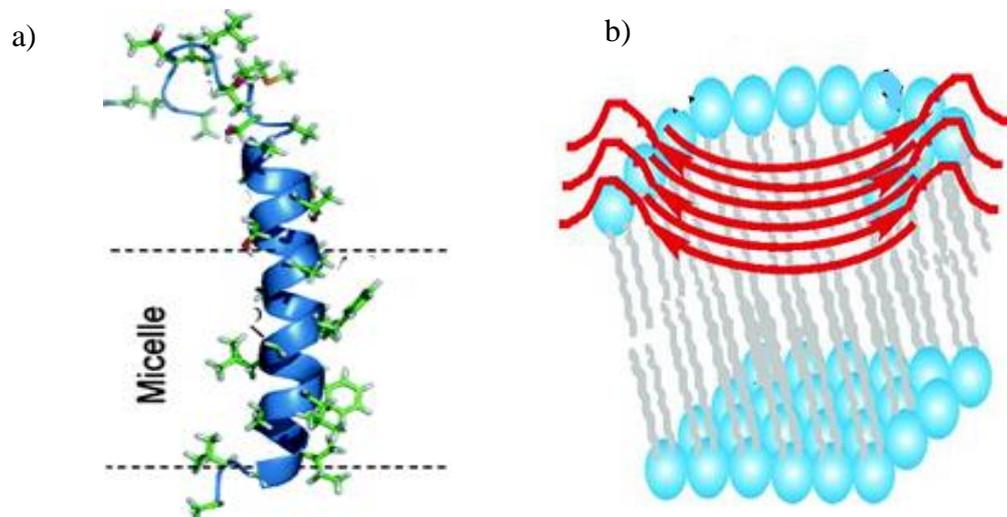


Figure 14 : Structure du peptide de fusion du VIH-1. a) Conformation en hélice  $\alpha$  dans des micelles de SDS (Jaroniec et al., 2005). b) Conformation en feuillet  $\beta$  dans des membranes modèles (Schmick and Weliky, 2010).

### 1.1.5.B.b.ii.ii. Le rôle de la MPER

La gp41 est la plus conservée parmi les deux gp du VIH et notamment au niveau de la MPER (Coutant *et al.*, 2008). Cette région constituée par 24 aa situés en Ct de l'ectodomaine de la gp41 contient des résidus hydrophobes hautement conservés et notamment du tryptophane. Des études de mutagenèse montrent que cette région joue un rôle crucial dans la fusion des membranes cellulaire et virale (Muñoz-Barroso *et al.*, 1999; Salzwedel *et al.*, 1999). En fait, elle participe à l'empaquetage de la gp41 avec l'enveloppe du VIH et à la déstabilisation de la bicouche lipidique durant la fusion des deux membranes. A l'extrémité Ct de cette région on trouve le motif CRAC (cholesterol recognition/interaction amino acid consensus) composé des acides aminés : LWYIK, qui joue un rôle important dans les phases précoce et tardive du cycle viral du VIH. Il se comporte comme un déterminant structural modulant la fusion des membranes et les événements post-fusion (Abad *et al.*, 2009).

La structure de la MPER, comme pour le fp, reste toujours controversée. En fait, les prédictions de conformation pour cette région montrent une structure en hélice  $\alpha$ . Cela est confirmé par plusieurs études structurales par cristallisation aux rayons X (Buzon *et al.*, 2010; Liu *et al.*, 2009; Weissenhorn *et al.*, 1997) et par RMN (Coutant *et al.*, 2008; Schibli *et al.*, 2001). Par contre, d'autres études montrent que la MPER se structure en  $\beta$ -turn (Bryson *et al.*, 2008; Ofek *et al.*, 2004a).

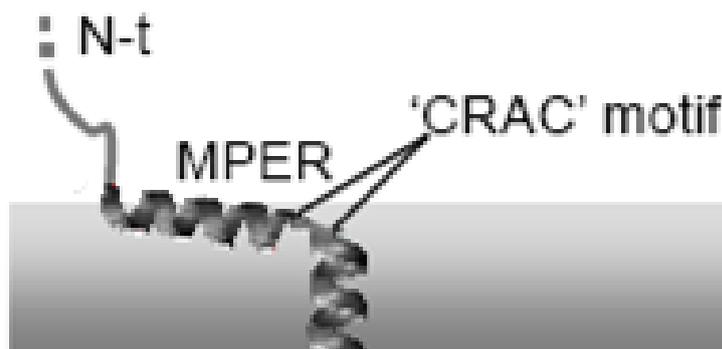


Figure 15 : Topologie de la MPER contenant le motif CRAC (Abad *et al.*, 2009).

## 1.2. Le VIH dans l'organisme

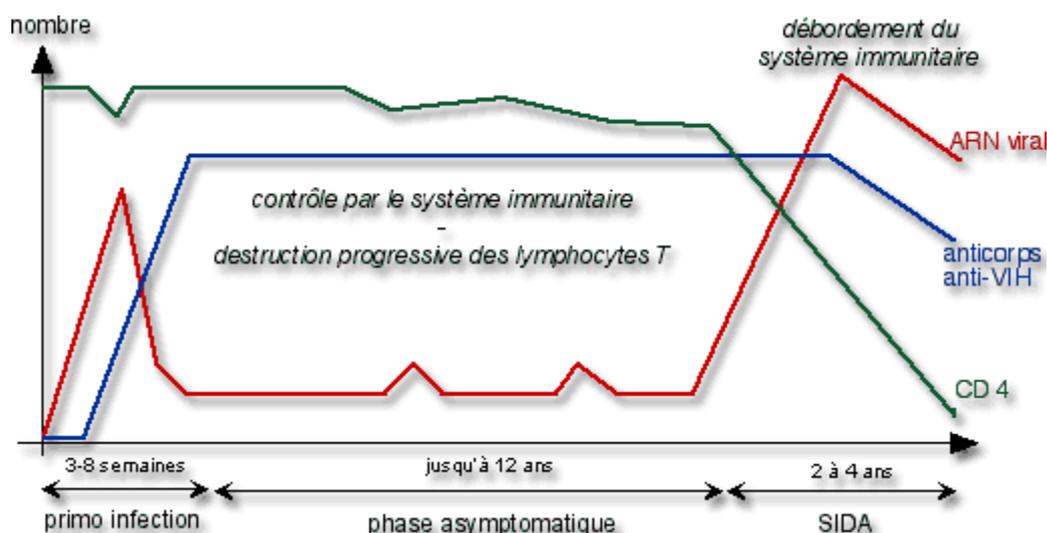
### 1.2.1. Transmission du VIH

Le virus est présent dans les liquides biologiques de l'organisme des personnes infectées tel que le sang, le sperme, le liquide séminal, les sécrétions anales, les sécrétions vaginales et le lait maternel. Les trois principaux modes de transmission sont donc:

- la transmission par voie sexuelle (l'efficacité varie entre 0,3 à 30%) : elle est d'autant plus forte que les muqueuses sont lésées facilitant la contamination par le VIH et plus généralement par les maladies sexuellement transmissibles.
- la transmission par voie sanguine (efficacité de 80 à 90%) : transfusion par du sang contaminé, échange des seringues souillées...
- la transmission verticale, de la mère à l'enfant (efficacité de 25%) : elle peut avoir lieu *in utero*, lors de l'accouchement ou lors de l'allaitement.

### 1.2.2. Évolution naturelle de l'infection virale

La chronologie de l'infection par le VIH suit une évolution en trois phases distinctes : la primo-infection ou phase aigüe, la phase asymptomatique ou latence clinique et la phase SIDA (Genetet and Ruffault, 2002). La figure 15 résume cette évolution au niveau de trois facteurs : la charge virale (en rouge), la quantité des anticorps produits par le système immunitaire contre ce virus (en bleu) et le nombre des cellules cibles CD4+ (en vert).



**Figure 16 : Évolution des différentes phases de l'infection à VIH et variation de paramètres biologiques.**

### 1.2.2.A. Primo-infection

C'est la phase précoce de l'infection qui dure de 3 à 8 semaines. Après l'entrée du virus, ce dernier se multiplie de façon intensive et diffuse dans l'ensemble de l'organisme (voir Figure 16, courbe rouge). Durant cette phase, 50% à 70% des personnes présentent des symptômes qui ressemblent à ceux de la grippe ou de la mononucléose. Elle peut aussi passer inaperçue ou s'accompagner de signes cliniques (présence de ganglions, fièvre, malaise général, maux de tête, courbatures et douleurs articulaires, éruption cutanée, ulcérations des muqueuses). Ces symptômes durent généralement environ une ou deux semaines, puis disparaissent. En parallèle, le système immunitaire commence à réagir en produisant des anticorps dirigés contre ce virus (Figure 16, courbe bleue).

### 1.2.2.B. Phase chronique

La phase chronique est asymptomatique où seule la présence d'adénopathies persistantes témoigne d'une stimulation active du système immunitaire. La durée de cette phase est très variable d'un sujet à un autre (en moyenne 8-12 ans). Durant cette phase, la répllication du virus dans le sang diminue et se stabilise à un niveau qui varie selon les personnes. Dans certains cas, la quantité de virus dans le sang reste faible. Les personnes dont le système immunitaire reste à peu près intact après 10 ans représentent environ 10% des personnes atteintes par le VIH. Plus souvent, la quantité

de virus augmente dans le sang et le nombre de lymphocytes T-CD4 diminue sur des périodes pouvant s'étendre de 3 à 12 ans. La phase de séropositivité sans symptôme clinique correspond à la période durant laquelle les effets toxiques du virus semblent apparemment contrôlés par l'organisme, notamment par le système immunitaire. Lorsque les personnes atteintes par le VIH ne présentent aucun signe physique de maladie, elles sont dites asymptomatiques.

### 1.2.2.C. SIDA

Le sida correspond au stade avancé de l'infection par le VIH, c'est-à-dire de la forme la plus grave de l'infection par le VIH, lorsqu'une personne séropositive est atteinte par l'une des vingt-cinq maladies répertoriées dans la liste des pathologies définissant le sida (candidose bronchique, trachéale ou pulmonaire ; cancer du col invasif, encéphalopathie due au VIH...). La déplétion lymphocytaire (Figure 16, courbe noire) est tout d'abord compensée par la production de nouvelles cellules T CD4 jusqu'à ce que le processus de compensation s'effondre sous la poussée de la multiplication virale, apparaît alors le stade SIDA. Cette évolution de l'infection se traduit par la survenue de pathologies opportunistes plus ou moins graves. Certains symptômes d'allure banale peuvent apparaître (dermite séborrhéique, zona, herpès, candidoses oropharyngée). D'autres lésions sont plus spécifiques des infections opportunistes liées au VIH (leucoplasie chevelue de la langue). Le système immunitaire est maintenant en état d'insuffisance grave. Il se trouve alors dans l'incapacité de défendre correctement l'organisme contre la survenue de certaines infections.

## 1.2.3. Réservoirs et latence

### 1.2.3.A. Réservoirs cellulaires

Le VIH infecte plusieurs types cellulaires tels que les cellules de Langerhans, les macrophages, les monocytes, les cellules dendritiques... Mais, il cible principalement les lymphocytes T CD4+. On distingue les lymphocytes T CD4+ (LT4) activés et les cellules mémoires. L'infection des LT4 activés est cytopathique et réduit considérablement leur durée de vie. En fait, la demi-vie de ces cellules est en général de quelques jours.

Quant aux LT4 mémoires, on peut en distinguer deux types : les cellules naïves et les cellules mémoires. Dans les cellules naïves, l'infection va bien avoir lieu, mais ces cellules sont réfractaires à la réplication. On parle de latence pré-intégrative, qui induit le blocage du cycle viral au niveau de la transcription inverse ou juste après celle-ci. La protéine virale *nef* permettrait l'infection et la réplication du VIH dans des LT quiescents. Ce type de latence, n'est pas cliniquement conséquente, car elle ne peut pas expliquer la résurgence du virus à l'arrêt des traitements et car la demi-vie de ces cellules est relativement courte.

Les cellules mémoires, qui résultent de cellules activées passant à un état quiescent, montrent une latence post-intégrative. Elles expriment également plus de CCR5 que les cellules circulant dans la voie sanguine. Ce type de latence est très important car elle explique la résurgence du virus à l'arrêt du traitement.

### 1.2.3.B. Réservoirs anatomiques

Les réservoirs anatomiques correspondent à des zones à immunologies privilégiées où le virus aura une cinétique de réplication plus stable que celle du virus actif dans le reste de l'organisme. Ces réservoirs sont également appelés sanctuaires viraux.

On distingue trois principaux réservoirs : le tractus génital, le système nerveux central et les organes lymphoïdes (principalement le tractus digestif).

## 1.2.4. Réponse immunitaire développée contre le virus

Face au VIH, le corps humain met en place tous les composants d'une réponse immunitaire spécifique et dirigée contre les différents constituants de la particule virale. Elle semble assurer un certain degré de protection pour un temps limité. Mais toute l'ambiguïté de la réponse immunitaire vient de ce que les effecteurs spécifiques, anticorps et lymphocytes T sont détournés de leur fonction de protection. En fait, leurs propriétés mises en jeu pour contrôler l'infection se voient utilisées par le VIH pour assurer sa survie et sa dissémination.

### 1.2.4.A. Réponse cellulaire

La réponse immunitaire cellulaire est assurée par les lymphocytes T cytotoxiques ou CD8<sup>+</sup> (CTL). Ces derniers ayant acquis la capacité de détruire des cellules infectées par le VIH sont observées très rapidement après le contact contaminant. Avec leur apparition, coïncide la décroissance de la charge virale, dans les semaines qui suivent la contamination. Une des caractéristiques de ce type de réponse est le large spectre de réactivité des effecteurs. De plus, elle persiste tout au long de l'infection chronique chez la plupart des patients asymptomatiques à un niveau élevé. En effet, les CTL sont présents dans le sang avec une proportion élevée parmi les lymphocytes. Cependant, la réponse cellulaire se détériore à mesure de la progression vers le SIDA. Ce déclin de l'activité cytotoxique accompagné par l'augmentation de la charge virale peut être expliqué par plusieurs mécanismes :

- Epuisement de la réponse toxique : le nombre des CTL circulants et dans les organes lymphoïdes périphériques est considérablement augmenté durant la phase chronique. Une proportion élevée d'entre eux exprime un phénotype des cellules activées parmi lesquelles des cellules entrent dans un processus d'apoptose. Néanmoins, le défaut de renouvellement ne se manifestera que tardivement par une chute rapide des cellules CD8, qui survient aux stades avancés de l'infection.
- Anergie : le déclin de la réponse cytotoxique peut s'expliquer par une inactivation fonctionnelle, due à la stimulation persistante des CTL spécifiques, mais dans des conditions qui ne permettront par leur différenciation en effecteurs. En effet, cette réponse inadaptée qui aboutit à une anergie des cellules réactives. Cela est dû à l'absence de facteurs critiques tels que les interleukines 2 ou 12 secrétées normalement par les lymphocytes T4, les macrophages...

- Apoptose : les lymphocytes T4 infectés expriment le ligand Fas et deviennent capables de délivrer un signal de mort cellulaire aux CTL activés qui expriment le Fas.

### 1.2.4.B. Réponse humorale

Les anticorps anti-VIH produits par les lymphocytes B apparaissent très tôt après la contamination. Ils sont détectables dans les semaines qui suivent le contact infectant et sont utiles au diagnostic de l'infection : ils définissent la séropositivité. La réponse humorale se diversifie tout en produisant des anticorps dirigés contre les différents composants du VIH : les glycoprotéines d'enveloppe, les enzymes virales... Pourtant cette réponse est dominée par les anticorps anti-Env. Ils appartiennent à la classe des immunoglobulines G et A (respectivement IgG et IgA) et sont présents tout au long de l'infection. Leur titre ne décroît qu'aux périodes avancées de la progression de la maladie.

- Pouvoir neutralisant

Les anticorps neutralisants doivent reconnaître des épitopes des glycoprotéines d'enveloppe du VIH et notamment ceux des régions impliquées dans l'interaction avec les récepteurs de la cellule hôte, tel que la boucle V3 de la gp120... Cette interaction ne nécessite que le paratope situé au niveau du fragment Fab (Fragment antibody binding) de l'anticorps et l'épitope (voir Figure 17, i. Neutralisation). Le recouvrement du virus par les anticorps inhibe sa fixation sur la cellule cible.

- Pouvoirs lytiques

Les anticorps anti-VIH ne sont pas capables de lyser le virus en présence du complément. Par contre, ils sont susceptibles d'agir par des mécanismes de cytotoxicité dépendant des anticorps (ADCC). Ils sont responsables de la destruction de la cellule infectée, mais également de la cellule non infectée en reconnaissant des glycoprotéines sous leur forme soluble mais associées à sa membrane par liaison avec le récepteur CD4.

- Propriétés opsonisantes

Une des propriétés des IgG dans les défenses infectieuses est de favoriser l'internalisation des microorganismes par les cellules phagocytaires. Celles-ci expriment des récepteurs reconnaissant le fragment cristallisable (Fc) des IgG. Dans le cas du VIH, ils sont soupçonnés de contribuer à la dissémination de l'infection en favorisant les contacts qui permettront aux virions d'être pris en charge par les cellules (macrophages) capables d'assurer leur réplication ou les cellules capables d'assurer leur survie extracellulaire (cellule dendritique) et de les mettre en contact avec de nouvelles cellules cibles.

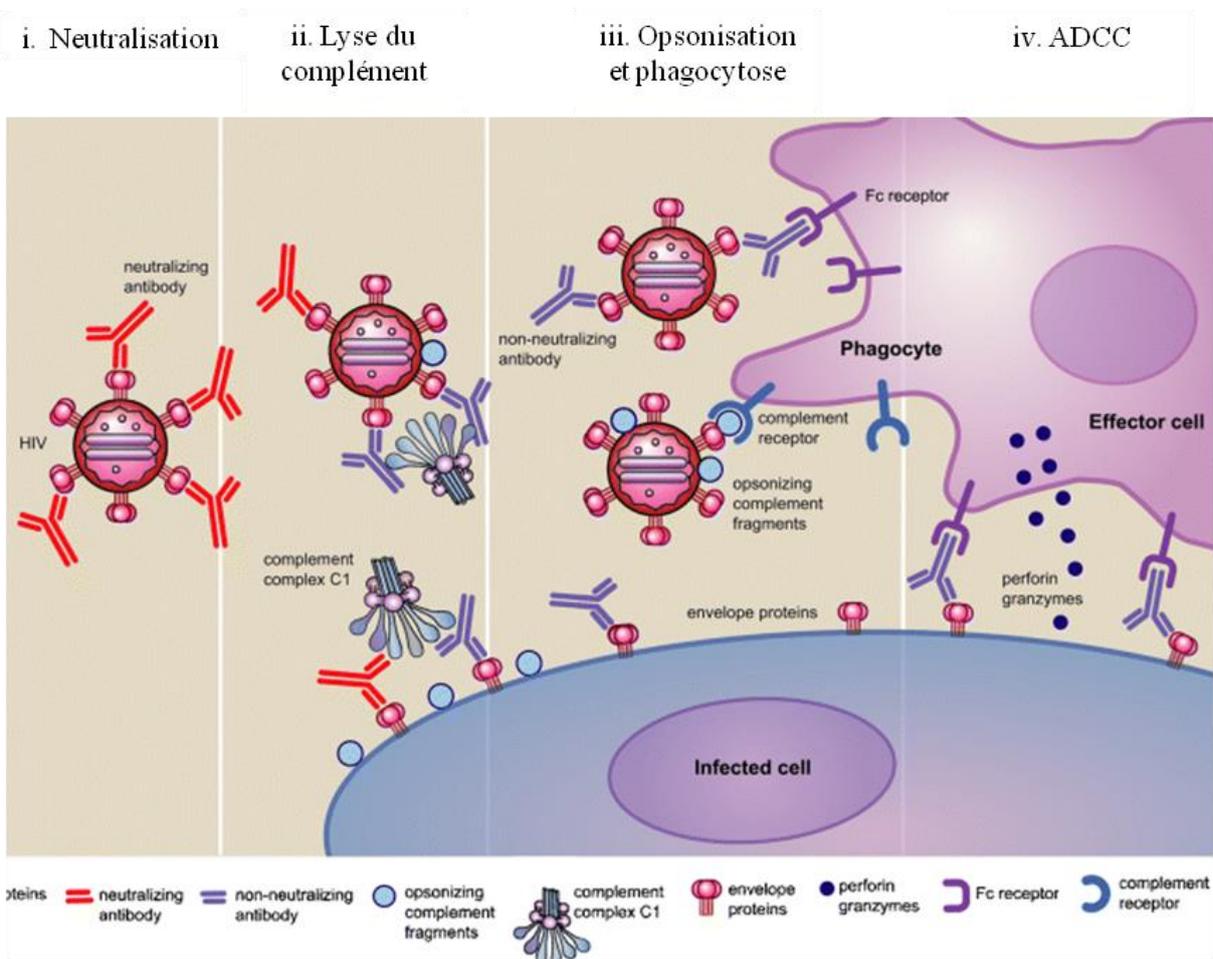


Figure 17 : Différents mécanismes d'action des anticorps anti-VIH (Huber et al., 2007).

### 1.2.4.B.a. Anticorps neutralisants dirigés contre l'Env du VIH-1

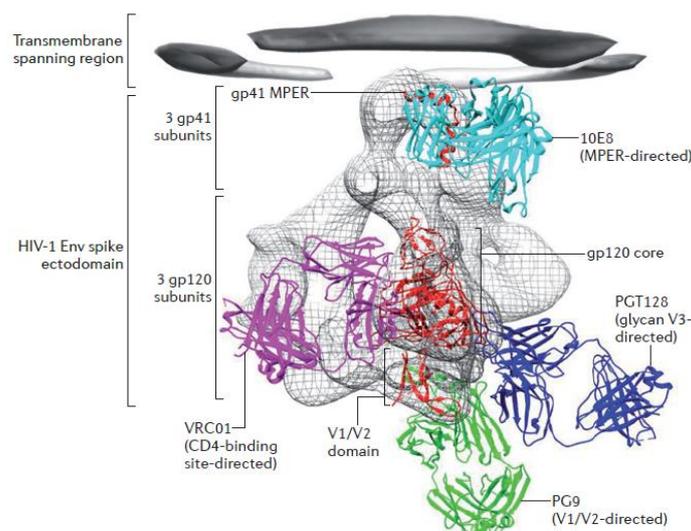
Durant les premiers mois de l'infection, les patients produisent des anticorps non neutralisant dirigés contre l'Env ou spécifiques à une souche du VIH bien déterminée. En fait, l'Env du VIH est protégée par une forte glycosylation, des boucles hypervariables et le masquage des épitopes qui s'exposent d'une façon transitoire après changement de conformation. Cependant, environ 20% des personnes infectées par le VIH-1 ont pu produire des anticorps neutralisant un large panel des souches virales.

Le tableau 2 et la Figure 18 récapitulent les différentes catégories et fonctions des anticorps neutralisant le VIH-1. Ces anticorps ont pour cible soit la gp120, soit la MPER de la gp41 (Kwong *et al.*, 2013; Pantophlet and Burton, 2006; Walker and Burton, 2008) :

- b12 : cet anticorps reconnaît un épitope discontinu entre la gp120 et CD4 et empêche ainsi leur interaction. En effet, l'élément clé de la liaison de la gp120 au CD4 est la présence d'une boucle de la gp120 au niveau du site de liaison au CD4 qui interagit avec la Phe-43, un acide aminé saillant du CD4. En absence de cette interaction, la gp120 se lie faiblement au CD4. La structure cristallographique du b12 a montré la présence d'une saillie, le CDR3 (complementarity-determining region 3), qui rentre dans le site de fixation du CD4.
- 2G12 : cet anticorps reconnaît les clusters de résidus d'oligomannose sur la gp120 grâce à sa structure unique. En effet, le 2G12 ne possède pas la forme « Y » traditionnelle des Ac mais une structure différente selon laquelle les deux bras sont adjacents et pointés verticalement.
- 2F5, 4E10 et Z13 : ces anticorps reconnaissent les épitopes sur la région proximale de la membrane de la gp41 tout en bloquant la fusion des membranes virales et cellulaires.

Prototypic antibody	Breadth*	IC <sub>80</sub> (µg per ml)†	Unique features	Similar monoclonal antibodies
<b>CD4-binding site</b>				
b12	33%	2.70	Derived from phage display	NA
HJ16	30%	0.77	Extensive affinity maturation	NA
VRC01	87%	0.98	CD4 mimicry by V <sub>H</sub> 1-2-derived heavy chain	VRC02, VRC03, NIH45-46, 3BNC60, BNC62, 3BNC117, 12A12, 12A21, 12A30, VRC-PG04 and VRC-CH31
8ANC131	57%	4.02	CD4 mimicry by V <sub>H</sub> 1-46-derived heavy chain	8ANC37 and 8ANC134
CH103	34%	8.00	CDR H3 mode of recognition and reasonable affinity maturation	NR
<b>V1/V2</b>				
PG9	70%	0.31	Extended CDR H3, which is often tyrosine-sulphated	PG16 and CH01-04
PGT145	60%	0.31	Discontinuous conformational epitopes	PGT141 to PGT144
<b>Glycan V3</b>				
2G12	18%	4.85	Glycan-only recognition	NR
PGT121	53%	0.08	Recognizes V1/V2 and V3 glycan	PGT122 and PGT123
PGT128	56%	0.11	Recognizes V3 glycan	PGT125 to PGT127, PGT130 and PGT131
PGT135	<30%	NR	Recognizes V3 and V4 glycans	PGT136 and PGT137
<b>MPER</b>				
2F5	48%	9.42	ELDKWAS peptide recognition	m66
4E10	88%	8.98	Sequence before transmembrane domain	NR
10E8	97%	2.05	No autoreactivity	7H6
Z13	<20%	NR	WNWFDITN peptide recognition derived from phage display	NR

**Tableau 2 : Les différents catégories et fonctions des anticorps neutralisants dirigés contre le VIH-1 (Kwong et al., 2013). CDR H3 : troisième région déterminant de la complémentarité de la chaîne lourde ; NA : non applicable ; NR : non reporté dans la littérature ; V<sub>h</sub> : segment variable de la chaîne lourde de l'immunoglobuline ; Breadth (ou spectre) est défini comme étant le pourcentage de la neutralisation de 80% des virus, au moins, par la concentration maximale d'inhibition de 50µg/ml d'anticorps.**



**Figure 18 : Schéma du « spike » du VIH-1 montrant les épitopes et leurs anticorps neutralisants correspondants (Kwong et al., 2013).**

## 1.2.5. Traitements

Jusqu'à aujourd'hui, il n'existe pas de traitement curatif contre le VIH. Par contre, il existe des traitements dirigés contre ce virus et qui contribuent à réduire la charge virale dans le sang, c'est-à-dire réduire le nombre des particules virales jusqu'à atteindre une charge virale indétectable par les techniques actuelles de dépistage (<50 particules virales ou copies / mL du sang). Ces antirétroviraux sont répartis en plusieurs classes principales (<http://www.doctissimo.fr/>, 2013) et ciblent différentes étapes du cycle d'infection/réplication virale :

- les inhibiteurs d'entrée qui empêchent le VIH d'entrer dans les cellules hôtes. Il en existe deux types : les inhibiteurs de fusion et les inhibiteurs du co-récepteur CCR5 (Enfuvirtide...);
- les inhibiteurs nucléosidiques de la transcriptase inverse (INTI) et les inhibiteurs nucléotidiques de la transcriptase inverse (INtTI) qui ont pour cible la transcriptase inverse : ces molécules interviennent dans la cellule pour bloquer l'action de cette enzyme virale, et ainsi empêchent le virus de transcrire son ARN en ADN (Emtricitabine, Abacavir...);
- les inhibiteurs non-nucléosidiques de la transcriptase inverse (INNTI), qui visent également la transcriptase inverse mais d'une façon différente aux INTI et INtTI. En fait, ces molécules sont des variantes chimiques des inhibiteurs de la transcriptase inverse. Ces molécules se fixent sur cette enzyme tout en bloquant son activité ;
- les inhibiteurs de l'intégrase qui visent l'intégrase et empêchent le virus d'intégrer son ARN dans l'ADN des cellules humaines ;
- les inhibiteurs de la protéase qui sont dirigés contre la protéase. Ces composés agissent à un autre stade du processus de la réplication virale, en s'attaquant à la protéase arrêtant ainsi la maturation de nouveaux virus créés par la cellule infectée (Amprénavir, Fosamprénavir, Atazanavir...);
- les Immuno-Modulateurs : ces traitements agissant sur l'immunité de la personne tout en restaurant une réponse immunitaire et stimulant la production et le renouvellement des cellules naïves (Proleukin : Il2)

Plusieurs stratégies de traitement sont actuellement disponibles en associant plusieurs médicaments. En général, ils sont associés par trois pour assurer une attaque puissante contre le VIH, d'où l'appellation trithérapie antirétrovirale qui combine deux INTI avec une anti-protéase ou un INTI...

Malgré l'existence et la combinaison de tous ces médicaments, ces derniers n'arrivent pas à guérir de l'infection ni à éradiquer le VIH. De plus, ils entraînent beaucoup d'effets secondaires : ils sont contraignants, souvent mal tolérés et à long terme générateurs d'effets indésirables sévères (digestifs, hématologiques, neurologiques...) qui altèrent la qualité de vie et peuvent conduire les patients à interrompre leur traitement. Il faut signaler aussi que ces traitements deviennent parfois inutiles. En effet, le virus qui mute naturellement devient insensible face à ces molécules thérapeutiques, d'où la nécessité de trouver un vaccin préventif anti-VIH.

## **1.2.6. Vaccination**

Un vaccin anti-VIH semble être la seule solution pour lutter contre le VIH et son épidémie, le SIDA. Mais sa mise au point impose d'énormes contraintes qui rendent hasardeuse toute prévision dans ce domaine. Les obstacles sont à la fois biologiques et techniques : biologiques du fait de la nature inconstante du virus et techniques du fait qu'il est impossible d'envisager l'utilisation d'un vaccin tué ou atténué qui garderait la capacité de se recombiner avec d'autres séquences virales, chez son hôte, et de produire de nouveaux virus avec des conséquences imprévisibles à long terme. C'est pour ces raisons que, jusqu'à présent, les essais vaccinaux ont toujours échoué.

### **1.2.6.A. Essais Vaccinaux**

Depuis 1987, plus de 200 essais vaccinaux ont été effectués. Pourtant, seulement quatre d'entre eux ont montré une efficacité relativement plus importante. Ces essais sont basés sur plusieurs approches dans le but d'induire une réponse immunitaire soit en produisant des anticorps neutralisants (réponse humorale), soit en générant des CTL capables de lutter contre le virus (réponse cellulaire), soit en associant ces deux réponses immunitaires. Les deux premiers essais (VAX003 (Montefiori *et al.*, 2012) et VAX004 (Bartholow, 2005) ont évalué la capacité de la sous-unité gp120 à induire la production des anticorps spécifiques à l'Env. Malgré le fait que ce vaccin

soit immunogène, il n'est parvenu ni à prévenir les vaccinés contre l'infection par le VIH ni à diminuer la charge virale.

Les essais « Step and Phambili » (Gray *et al.*, 2010) ont pour but de générer des CTL spécifiques au VIH. L'immunogène utilisé dans ces essais est le vecteur-vaccin recombinant de l'adénovirus stéréotype 5 (S5). L'étude Step a été arrêtée en 2007 du fait qu'elle n'a pu prévenir contre l'infection virale. De plus, ce vaccin a augmenté le risque d'infection chez les hommes incirconcis ayant une immunité préalable contre l'adénovirus S5. L'étude Phambili n'a elle aussi montré aucune efficacité. Par contre, aucune association entre l'augmentation de l'infection par VIH-1 et la séropositivité en adénovirus S5 et l'incirconcision n'ont été mises en évidence.

Le quatrième essai, RV144 (de Souza *et al.*, 2012), a adopté une stratégie qui consiste à évaluer une vaccination hétérologue. Cette méthode consiste en un vecteur vaccin issu du canarypox qui exprime gag, pol, en combinaison avec la vaccination par le même immunogène utilisé dans VAX003/VAX004. Cet essai également n'a pu ni induire la production des anticorps neutralisants à large spectre viral, ni la génération de CTL spécifiques du VIH. Cependant, il a induit la cytotoxicité dépendant des anticorps (ADCC) qui neutralisent quelques souches virales.

### 1.2.6.B. Obstacles à contourner

Le développement d'un vaccin anti-VIH se heurte à des obstacles considérables qui sont bien connus (Barouch, 2008; Johnston and Fauci, 2008; Nabel, 2013; Walker and Burton, 2008). Tout d'abord, ce virus attaque principalement les lymphocytes T auxiliaires, LT4, (ou helper en anglais) qui stimulent les CTL en secrétant des cytokines en cas d'infection, autrement dit, il attaque le chef d'orchestre du système immunitaire. L'hypervariabilité génétique du virus due à la transcriptase inverse « error-prone » est également l'un des plus grands challenges. En effet, cette enzyme qui a un taux d'erreurs élevés et le fait que le virus se réplique rapidement *in vivo* contribuent à générer des variantes ayant une grande capacité d'adaptation et d'échappement au système immunitaire. De plus, la protéine Nef diminue l'expression des complexes majeurs d'histocompatibilité qui jouent un rôle important dans la reconnaissance des cellules infectées par les CTL. Ajoutons à tout cela, la latence virale dans les réservoirs cellulaires et anatomiques durant laquelle le virus est immunologiquement silencieux.

Plus spécifiquement et au niveau de la réponse humorale, de nos jours, aucun immunogène n'a pu induire la production des anticorps neutralisant un large spectre des souches virales, même dans une infection naturelle. En effet, les anticorps produits ne peuvent pas cross-neutraliser tous les variantes qui apparaissent rapidement *in vivo*. Pourtant, quelques anticorps monoclonaux neutralisant un large spectre des souches virales *in vitro* ont été découverts mais ils n'arrivent pas à assurer une immunisation contre le virus. En effet, les épitopes de ces anticorps sont localisés majoritairement sur les glycoprotéines d'enveloppe qui sont hautement glycosylées. De plus, d'autres de ces épitopes ne sont exposés que transitoirement durant l'entrée du virus à cause de la dynamique et l'instabilité du « spike » qui subit des changements de conformation énormes pour assurer la fusion des membranes virale et cellulaire.

### 1.3. Objectifs et contribution de travail de la thèse

Certains patients infectés par VIH ont développé des anticorps capables de neutraliser *in-vitro* différents sous-types de VIH, en reconnaissant des régions conservées sur les glycoprotéines d'enveloppe. Les plus prometteuses sont localisées dans la MPER mais elles ne sont probablement exposées que d'une manière transitoire durant la fusion des membranes virales et cellulaires.

Nous nous sommes focalisés sur ces épitopes pour mieux comprendre le mécanisme d'entrée du virus et pour pouvoir par la suite concevoir un vaccin anti-VIH. Ainsi, les séquences de trois épitopes ont été choisies pour leur proximité avec la membrane et leur implication dans le mécanisme de fusion (épitope gp120 et gp41 : cibles des anticorps neutralisants ; épitope fp : acteur clé de la fusion). Un autre facteur important est l'environnement lipidique. En effet, la composition lipidique de la membrane du VIH est différente de celle des cellules eucaryotes. En outre, la région MPER de la gp41 présente le motif CRAC de reconnaissance et d'interaction avec le cholestérol qui a un rôle important dans la formation d'un micro-environnement lipidique. Nous nous sommes donc basés sur ces facteurs pour concevoir les différentes constructions et préparation vaccinales qui font l'objet de ce projet.

Ces constructions prévoient donc l'ajout d'une séquence transmembranaire (TM) qui permettra un ancrage dans les membranes et la possibilité de trimérisation. Le but étant l'obtention d'hétérotrimères (gp120/gp41/fp) mimant l'état de fusion, dans des membranes de composition lipidique proche de celle de la membrane virale du VIH.

Mon travail de thèse a donc consisté, dans un premier temps, à la surexpression en système bactérien des trois protéines issues des différentes constructions et à leur purification par chromatographie d'affinité et/ou HPLC. Puis, les propriétés biophysiques ont été étudiées par les techniques de spectroscopie, résonance magnétique nucléaire (RMN) et de dichroïsme circulaire (DC). Chaque construction a été caractérisée de manière individuelle dans différents systèmes lipidiques. En parallèle, des vaccinations à partir de préparations liposomales (composition lipidique proche de celle de la membrane du VIH) ont été effectuées pour chaque polypeptide chez la souris. Les propriétés immunogéniques et la spécificité des anticorps produits sont testées par des techniques d'immunodétection (ELISA, immunoblot).

A plus long terme, l'analyse comparative entre le potentiel d'immunogénicité et les résultats structuraux obtenus seront confrontés dans le but de mieux comprendre l'entrée du VIH-1 dans la cellule hôte et par la suite améliorer la conception d'un vaccin efficace dirigé contre le VIH-1.

## 2. Chapitre 2 : Aspect théorique

### 2.1. Stratégie d'expression des protéines membranaires hétérologues chez *E. coli*

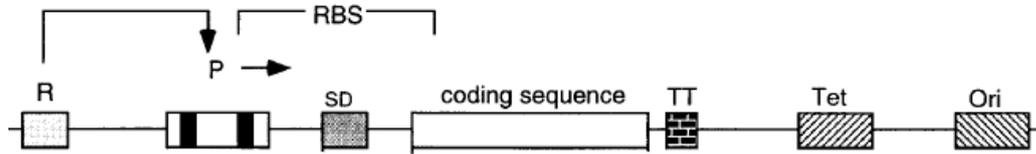
*Escherichia coli* est le système le plus utilisé pour exprimer une protéine recombinante ou hétérologue (protéine qui n'est pas synthétisée naturellement par l'organisme utilisé). Cela est dû au fait que la génétique de cette bactérie est bien connue et elle croît rapidement, avec des densités très hautes dans des milieux peu onéreux. De plus, la commercialisation d'une grande variété des vecteurs d'expression et des souches mutées d'*E. coli*, rend cette dernière de plus en plus attractive. Cependant, ce système n'assure pas de modifications post-traductionnelles ce qui est par la suite un frein à la production des protéines biologiquement actives. D'où le recours à des systèmes plus complexes (levure, cellules d'insectes ou des mammifères...). De même, ce système ne garantit pas la production des protéines en quantités suffisantes surtout dans le cas où il s'agit des protéines dites « toxiques » pour le microorganisme, par exemple les protéines ou les peptides membranaires dont leur étude par RMN nécessite une grande quantité (Samuelson, 2011). Donc, plusieurs facteurs doivent être contrôlés et testés pour assurer l'expression de telles protéines : le choix du vecteur d'expression et de son promoteur, la souche d'*E. coli* utilisée, la température d'expression, la densité optique (DO) d'induction...

#### 2.1.1. Le vecteur d'expression

Généralement, le vecteur d'expression procaryote doit contenir les éléments suivants (Makrides, 1996) (voir Figure 19) :

- l'origine de réplication (Ori) : une séquence d'ADN qui permet au plasmide d'être répliqué indépendamment du chromosome bactérien et détermine le nombre de copies du vecteur d'expression dans la bactérie.
- un gène de résistance à un antibiotique qui permet de sélectionner les clones transformés par le vecteur (ampicilline, kanamycine, chloramphénicol, tétracycline...)

- un promoteur de transcription qui permet l'initiation de la transcription par la fixation de l'ARN polymérase, ainsi qu'un terminateur et les régions d'initiation (Ribosome Binding Site) et de terminaison de la traduction (codon STOP).



**Figure 19 : Représentation schématique des éléments essentiels d'un vecteur d'expression procaryote (Makrides, 1996). Le promoteur (P) qui est positionné en amont de la séquence codante (coding sequence) est sous le contrôle d'un gène régulateur (R) qui code pour un répresseur. Ce gène peut exister sur le vecteur d'expression ou encore sur le chromosome de la cellule hôte. Le site de fixation du ribosome (RBS) contient la séquence de Shine-Dalgarno (SD). La séquence de terminaison de la transcription (TT) sert à stabiliser l'ARNm. Le marqueur de sélection (Tet) qui code un gène de résistance à la tétracycline, permet la sélection des souches bactériennes possédant le vecteur. Le nombre de copies du vecteur dans la cellule est fixé grâce à l'origine de réplication (Ori). La flèche indique le sens de la**

## 2.1.2. Régulation de l'expression de la protéine d'intérêt

La surexpression de protéines membranaires chez *E. coli* peut provoquer chez cette dernière une situation de stress ce qui entraîne le ralentissement de la croissance bactérienne, la sélection de variantes de mutation ou bien la dégradation de la protéine d'intérêt. D'où la nécessité de réguler l'expression de cette classe des protéines. Tout d'abord, il faut éviter d'utiliser les vecteurs qui se répliquent trop (plus que 40 copies/bactérie, par exemple pUC) et plutôt utiliser des vecteurs contrôlables. Ce contrôle s'effectue par un système de répression et/ou d'induction via le promoteur. L'induction peut-être thermique : pour le système  $\lambda$ pL, (bio)chimique : pour le promoteur lac inductible par l'Isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG) ou l'allolactose, ou pour le promoteur araBAD inductible par le L-arabinose (Samuelson, 2011). Cependant, dans les systèmes de surexpression bactérienne qui contiennent des promoteurs forts et inductibles, il peut exister une « fuite » de l'expression de la protéine d'intérêt en absence de l'inducteur. Il s'agit de l'expression basale qui peut entraîner surtout si la protéine est « toxique », une instabilité du vecteur d'expression et par la suite à son expulsion de la cellule hôte. Dans le cas de protéines membranaires, cette surexpression peut conduire à la mort de la bactérie avant même l'induction. Pour réguler l'expression de la protéine d'intérêt et limiter tant que possible l'expression basale de la protéine d'intérêt, nous avons utilisé la souche bactérienne BL21 (DE3) pLysS ainsi que le plasmide d'expression pTIPX-4 (Taf12-mediated Insoluble Peptide eXpression) (voir Figure 20), disponible dans notre laboratoire et développé par le Dr Werten (Vidovic *et al.*, 2009). Brièvement, le plasmide pTIPX-4 a été obtenu par l'insertion dans le vecteur pET-28b (Novagen), du fragment de restriction BglII-Bpu1102I issu du vecteur pTIPX-2. Cet insert comprend un nouveau polylinker, un site de clivage à l'entérokinase et la séquence de TAF12 (TATA box binding protein (TBP)-Associated Factor 12 Histone fold Domain), facteur de transcription utilisé ici comme protéine de fusion. Afin faciliter la purification, une mutation a été introduite dans la séquence naturelle de TAF12 au niveau du site de clivage Asp-Pro substitué par Ala-Pro. Par la suite, nous décrirons en détail le mécanisme de régulation du système d'expression pET / BL21 (DE3) pLysS.

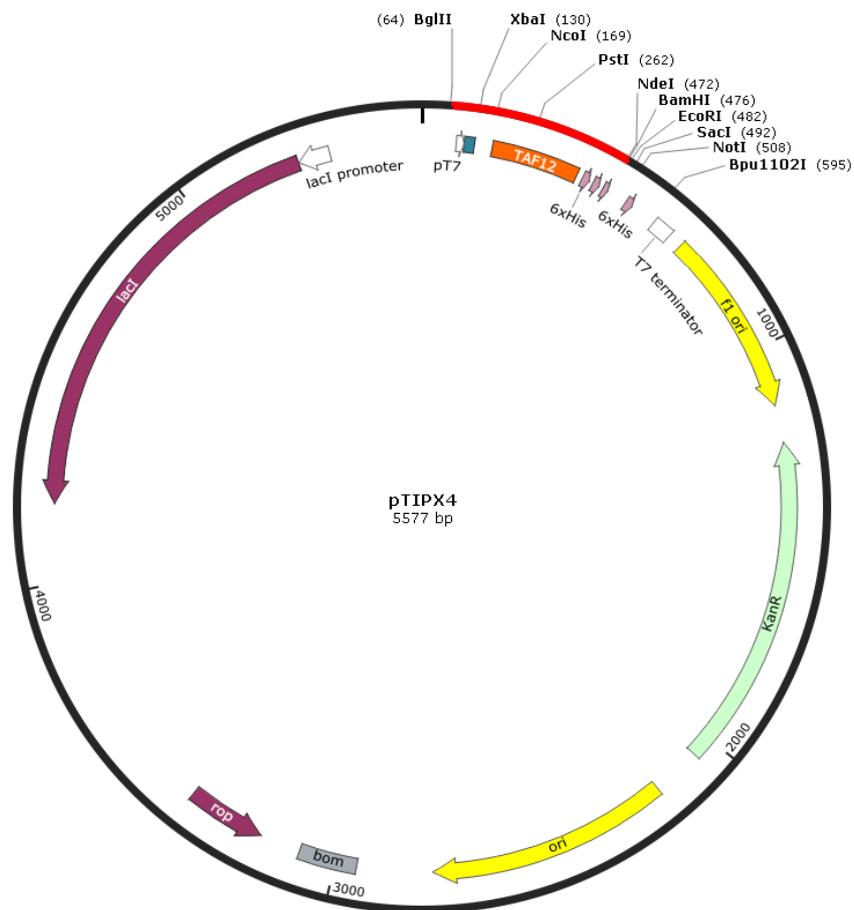


Figure 20 : Carte du plasmide pTIPX-4. Le plasmide d'expression pTIPX-4 contient les gènes de résistance à la kanamycine, la protéine de fusion TAF12HFD un site multiple de clonage comprenant les sites de restriction de SacI et BamHI, et le gène opéron lactose permettant l'induction de l'expression.

### 2.1.1.A. Système d'expression du vecteur pET

Le système pET contient différents éléments (Figure 21.a):

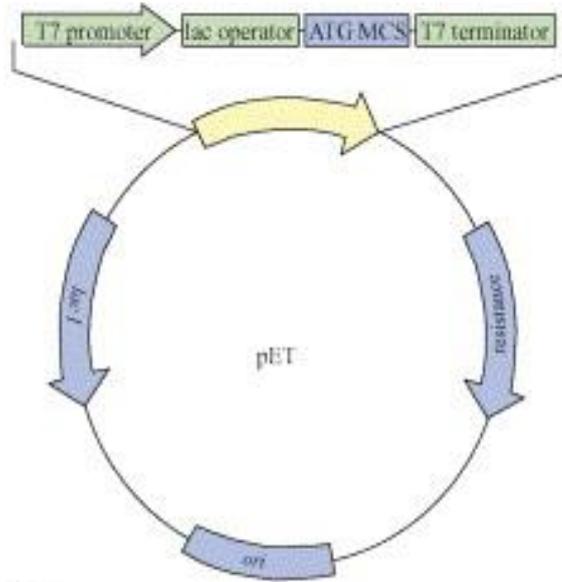
- une cassette de clonage multiple (MCS) : ensemble de sites de restriction présents en un seul exemplaire dans le vecteur, ce qui permet l'insertion du gène d'intérêt
- le promoteur T7 : c'est un promoteur dit « fort » car il permet, via la liaison de l'ARN polymérase T7, la transcription des ARN sous son contrôle en grande quantité
- un terminateur de transcription (T7 terminator) qui permet de mettre fin à la transcription du brin d'ARN
- le gène codant pour le répresseur *lacI*. Ce dernier qui régule le promoteur T7 en se fixant sur l'opérateur *lac o* (*lac operator*).
- un marqueur de résistance à un antibiotique (résistance) permettant la sélection des bactéries ayant intégré le vecteur
- une origine de répllication (*ori*) qui permet la répllication du vecteur

Dans ce système, l'expression de la protéine d'intérêt nécessite une souche bactérienne ayant intégré dans son génome le gène qui code pour l'ARN polymérase T7 (fragment  $\lambda$ DE3 du bactériophage). En effet, c'est cette enzyme qui en se liant à la séquence du promoteur T7 va permettre la transcription du gène sous son contrôle. En général, le gène codant l'ARN polymérase T7 se trouve lui-même sous le contrôle du promoteur *LacUV5* inductible par l'IPTG (régulation par l'opéron *lac*). Cette enzyme a pour avantage, contrairement à l'ARN polymérase bactérienne, d'avoir une très haute processivité et d'être cinq fois plus rapide (230 nucléotides par seconde).

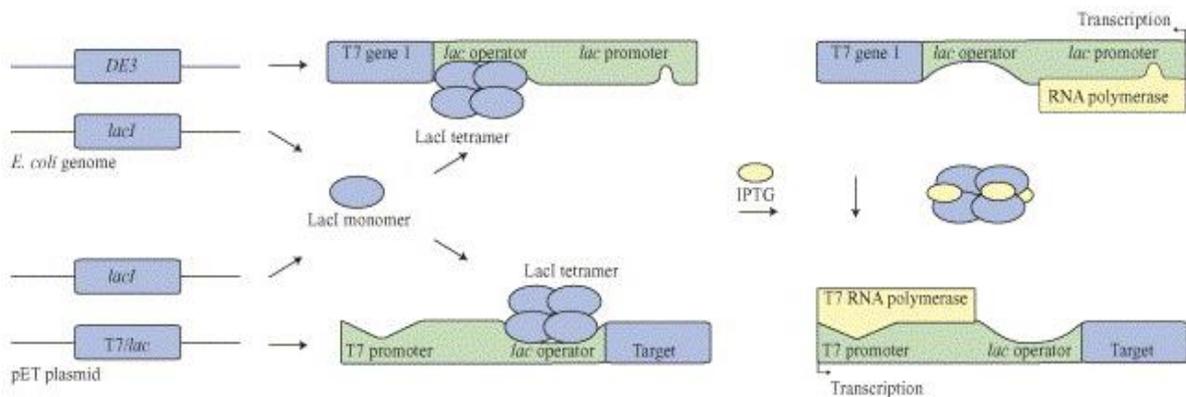
*LacI*, en se liant sous forme d'un tétramère à l'opérateur *lac o*, réprime le promoteur *lacUV5* présent sur le chromosome bactérien et le promoteur hybride T7/*lac o* présent sur le plasmide d'expression. En présence de l'IPTG, ce dernier se lie au répresseur *lacI*, libérant ainsi les opérateurs *lac*. Ceci a deux conséquences, la première étant que l'ARN polymérase de la bactérie va pouvoir se fixer sur le promoteur *lacUV5* et initier la transcription du gène codant l'ARN polymérase T7, la deuxième étant la fixation de l'ARN polymérase T7 exprimée sur le promoteur T7 du vecteur pET et par

conséquent la production de la protéine hétérologue (Baneyx, 1999; Chen, 2012; Hannig and Makrides, 1998; Studier *et al.*, 1990). Cependant, même en absence d'IPTG, l'ARN polymérase T7 peut être produite ce qui entraîne la mort des bactéries dans le cas où la protéine d'intérêt produite est toxique pour la bactérie,. Cette expression basale peut être contrôlée via la présence d'un inhibiteur de l'ARN polymérase T7 : le lysozyme T7. C'est dans ce but que les souches bactériennes BL21 (DE3) pLysS ont été ingénierées. Elles possèdent en effet un plasmide codant pour le gène *lys T7*. (Samuelson, 2011).

a)



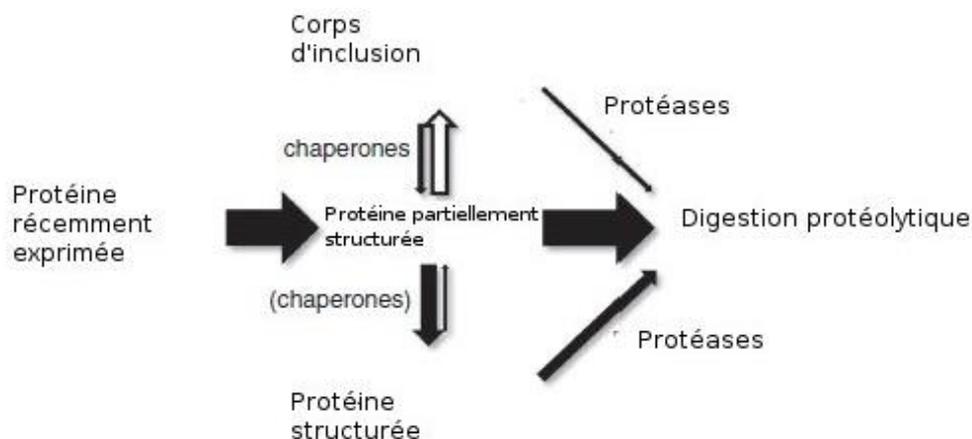
b)



**Figure 21 : Système d'expression dans pET (Sørensen and Mortensen, 2005). a) Configuration générale d'un plasmide pET montrant les éléments clés : origine de répliation (ori), marqueur de résistance (resistance), le gène codant le répresseur *lacI* qui régule le promoteur T7 en se fixant sur l'opérateur *lac* (lac operator), la cassette de clonage multiple (MCS) qui permet l'insertion du gène d'intérêt ainsi que le terminateur de transcription (T7 terminator). b) Mécanisme de répression du promoteur *lacUV5* (régulation de l'expression de l'ARN T7 polymérase) et du promoteur T7 (régulation de l'expression de la protéine d'intérêt), en absence ou en présence d'IPTG.**

### 2.1.1.B. Les partenaires de fusion

Bien que les partenaires de fusion aient été élaborés pour faciliter la purification et l'immobilisation des protéines d'intérêt, ils sont également utilisés pour améliorer leur solubilité (Baneyx, 1999). Les protéines surexprimées ont tendance à s'agréger dans le cytoplasme sous forme des corps d'inclusion qui correspondent à des particules denses de protéines agrégées d'environ 1 µm de diamètres (Lilie *et al.*, 1998). Ce processus doit être limité dans le cas où on veut avoir des protéines structurées et biologiquement actives. En générale, les protéines sont surexprimées sous forme soluble dans le cytoplasme, cette solubilité étant nettement améliorée par certains partenaires de fusion. Mais dans le cas des protéines membranaires et/ou toxiques pour la bactérie, d'autres partenaires peuvent aussi être utilisés pour réduire la toxicité. En effet, l'accumulation de la protéine d'intérêt sous forme de corps d'inclusion neutralise leur toxicité (Vidovic *et al.*, 2009). Généralement, la protéine exprimée existe toujours sous trois formes : structurée, partiellement structurée et agrégée qui existent en équilibre dynamique (voir Figure 22). Dans sa forme partiellement peu ou non structurée, la protéine d'intérêt est facilement dégradée par les protéases et les chaperonnes bactériennes. Le déplacement chimique est contrôlé bien évidemment par le partenaire de fusion et par d'autres facteurs (la température d'expression, insertion des mutations...) (Hwang *et al.*, 2013). Alors le choix du partenaire de fusion, la température d'expression, etc... se fait en fonction de la nature de la protéine d'intérêt et du type d'études pour lequel elle va être utilisée.



**Figure 22 : Diagramme montrant l'équilibre les différentes formes de la protéine exprimée (Hwang *et al.*, 2013).**

Les partenaires de fusion les plus utilisés sont : la Gluthation-S-Transférase (GST), la Maltose Binding Protein (MBP), TAF12 contenant tag His (voir Tableau 3)... Ces partenaires facilitent la purification des protéines car ils se lient spécifiquement sur des ligands bien précis. Par contre, il y a des partenaires qu'ils n'ont pas cette fonction alors il est impératif d'y introduire une étiquette ou « tag », il s'agit du tag 6-histidine, le peptide FLAG... Après la purification, ces partenaires doivent être séparés de la protéine d'intérêt par clivage chimique, enzymatique (voir Tableau 4)...

<b>Partenaire de fusion</b>	<b>Ligand/Matrice</b>
GST	Sépharose Gluthatione
MBP	Résine amylose
TAF12 contenant 6-His	Ni <sup>2+</sup> , Co <sup>2+</sup> - acide nitriloacétique
Peptide FLAG	Anticorps monoclonaux anti-FLAG

**Tableau 3 : Les partenaires de fusion les plus utilisés (Makrides, 1996).**

	<b>Méthode clivage</b>	<b>Séquence de reconnaissance</b>
Chimique	Acide formique	-Asp/Pro-
	Bromure de cyanogène	-Met/X-
Enzymatique	Entérokinase	-Asp-Asp-Asp-Asp-Lys/X-
	Thrombine	-Leu-Val-Pro-Arg/Gly-Ser-

**Tableau 4 : Différentes méthodes de clivage des partenaires de fusion (Hwang *et al.*, 2013; Vidovic *et al.*, 2009). / indique la position exacte du clivage.**

## 2.2. Dichroïsme circulaire

La spectroscopie du dichroïsme circulaire (DC) est une technique utilisée de plus en plus dans le domaine de la biologie structurale. En effet, elle procure des informations sur la structure des biomolécules, notamment la structure secondaire des protéines, d'une façon rapide et non destructive. De plus elle nécessite très peu de produits (<20 µg).

### Principe :

Le DC se base sur la capacité d'une molécule considérée comme chromophore chiral ou placée dans un environnement asymétrique ou en se liant à un ligand chiral, d'absorber différemment la lumière polarisée circulairement droite et la lumière polarisée circulairement gauche (Corrêa and Ramos, 2009; Kelly *et al.*, 2005). Le spectre dichroïque obtenu correspond à la différence d'absorbance entre ces deux types de lumière, pour chaque longueur d'onde. Il reflète alors la structure secondaire des protéines du fait qu'il est dominé, dans l'ultraviolet lointain, par les transitions  $n \rightarrow \pi^*$  et  $\pi \rightarrow \pi^*$  des groupements amide (voir Figure 23) et par les différentes géométries des squelettes polypeptidiques (Whitmore and Wallace, 2008). En effet, la contribution de la liaison peptidique prédomine aux longueurs d'onde inférieures à 240 nm et les transitions  $\pi \rightarrow \pi^*$  donnent naissance à deux bandes centrées à 190 et 208 nm par contre les transitions  $n \rightarrow \pi^*$  donnent naissance à une bande centrée à 222 nm (Niezborala, 2008). L'association de ces bandes caractérise une structure bien particulière : l'hélice  $\alpha$ . Les spectres des autres structures (feuillet  $\beta$ , coudes...) étant très différents, il est fréquent de mesurer le spectre de DC d'une protéine pour déterminer les proportions de ces différentes structures secondaires (Figure 24). Le spectre d'une hélice  $\alpha$  montre deux pics négatifs à 208 et 222 nm, et un pic positif à 193 nm. Par contre celui du feuillet  $\beta$  antiparallèle montre un seul pic négatif à 218 nm et un pic positif à 195 nm. Les protéines non structurées ont un pic négatif vers 195 nm et un faible signal au-dessus de 210 nm.

L'estimation de la contribution de chaque type de structure secondaire se fait par déconvolution des spectres par des différents algorithmes (CONTIN, SELCON, CDSSTR, K2D...) et en se basant sur différentes bases de données (Greenfield, 2007; Sreerama and Woody, 2000).

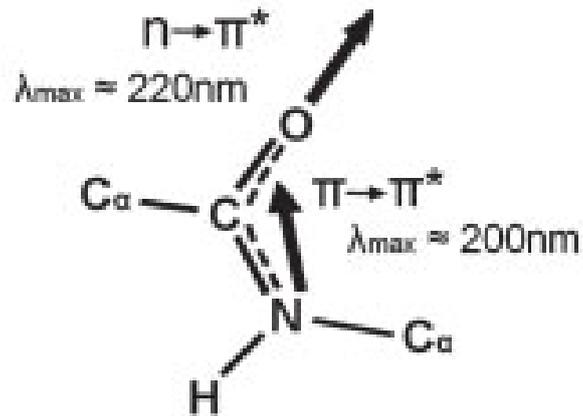


Figure 23 : Diagramme montrant l'orientation des transitions dipolaires (Whitmore and Wallace, 2008). Les flèches montrent les transitions  $n \rightarrow \pi^*$  et  $\pi \rightarrow \pi^*$  avec leurs longueur d'onde maximale correspondante.

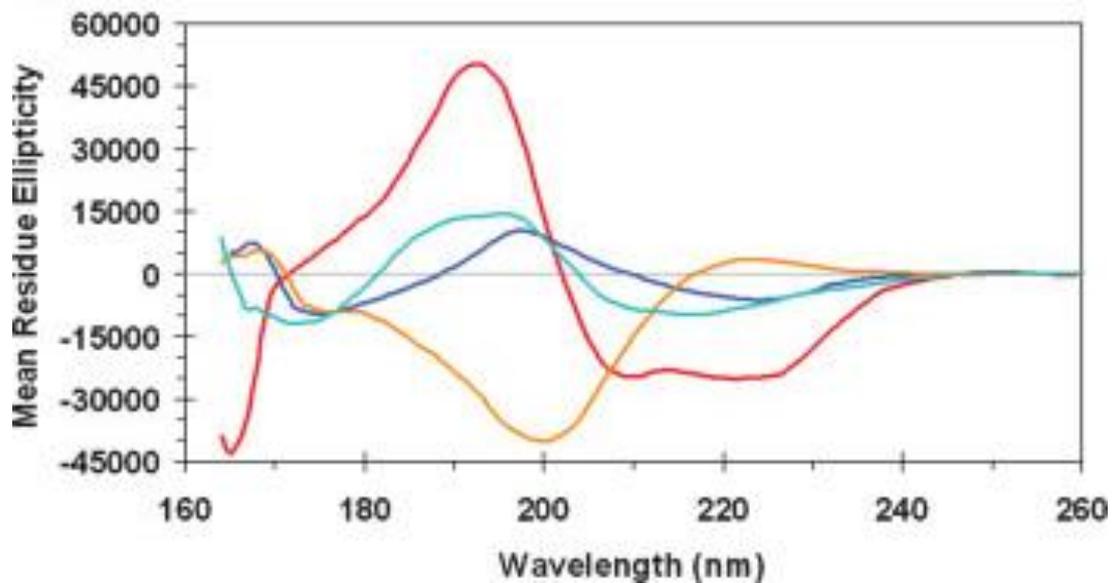


Figure 24 : Spectres dichroïsme circulaire montrant les différentes structures secondaires des protéines (Whitmore and Wallace, 2008). Spectre rouge : myoglobine structurée en hélice  $\alpha$  ; spectres bleu et cyan : respectivement concanavaline A et bêta-lactoglobuline structurés en feuillet  $\beta$  ; Spectre orange : collagène non structuré.

Quelques algorithmes exigent l'import des données sous forme d'une unité bien précise d'où la nécessité d'expliquer la conversion d'ellipticité mesuré en milli degré ( $\theta$  : mdeg) ou la différence d'absorption ( $\Delta A$ ) par le spectropolarimètre (Kelly *et al.*, 2005):

$\Delta A = A_L - A_R = \Delta \epsilon \cdot c \cdot l$  (loi de Beer-Lambert) avec :

- $A_L$  et  $A_R$  : l'absorption des composants, respectivement, gauche (L : Left) et droit (R : Right) de la lumière circulairement polarisée
- $\epsilon$  : coefficient d'extinction molaire ( $M^{-1} \cdot cm^{-1}$ )
- $c$  : concentration de la protéine (M)
- $l$  : trajet optique de la cuve (cm)

$\Delta A = \theta / 32,98 = \tan^{-1}(b/a) / 32,98$  avec  $a$  et  $b$  les axes d'ellipsoïde résultant de la différence d'absorption des différents composants de la lumière circulairement polarisée

Pour comparer différents spectres dichroïques, il faut les normaliser en termes de concentration et du nombre d'acides aminés.

La masse moyenne d'un résidu (Mean Residue Weight : MRW) :

$MRW = \frac{M}{N-1}$  avec  $M$  : masse moléculaire de la protéine (Da) et  $N$  nombres d'acides aminés

L'ellipticité molaire  $[\theta]$  :

$$[\theta] = \frac{M \cdot \theta}{c \cdot l}$$

- $[\theta]$  : ellipticité molaire  $degree \cdot cm^2 \cdot mol^{-1}$
- $M$  : masse molaire en g/mol
- $c$  : concentration en  $g/cm^{-3}$
- $l$  : trajet optique en cm

L'ellipticité molaire résiduelle (MRE) :

$$MRE = \frac{MRW.\theta}{10.l.c} \text{ avec :}$$

- MRE:  $\text{degree.cm}^2.\text{mol}^{-1}.\text{residue}^{-1}$
- q: degré
- l : cm
- c : g/ml

Il faut noter que le DC n'apporte pas des informations structurales au niveau atomique contrairement à la cristallographie aux rayons X ou la RMN. Pourtant et grâce à ses avantages, cette technique est en train de se développer rapidement. Citons, par exemple, Synchrotron Radiation Circular Dichroism (SRCD) qui permet d'avoir une meilleure ratio signal/bruit sur une acquisition partant des longueurs d'onde  $< 190$  nm. Cette extension permet la mesure du DC des échantillons contenant des composés non-chiraux hautement absorbant comme les sels, les détergents...

## 2.3. Résonance magnétique nucléaire

La spectroscopie de résonance magnétique nucléaire (RMN) consiste à étudier l'interaction d'une molécule placée dans un champ magnétique statique avec une onde électromagnétique. Il s'agit de la radiofréquence.

Dans cette partie, je rappellerai brièvement les principes de la résonance magnétique nucléaire (RMN) (Cavanagh *et al.*, 2007a, b) et les techniques utilisées dans la RMN de solide (Bechinger *et al.*, 2004; Bechinger *et al.*, 2011; Saitô *et al.*, 2006)

### 2.3.1. Phénomène de RMN

La résonance magnétique nucléaire (RMN) est basée sur une propriété de certains noyaux atomiques possédant un moment angulaire intrinsèque, un spin nucléaire :  $\vec{I}$ , (par exemple  $^1\text{H}$ ,  $^{13}\text{C}$ ,  $^{31}\text{P}$ ,  $^{15}\text{N}$ ...), et donc un moment magnétique non nul  $\vec{\mu}$ .

$$\vec{\mu} = \gamma \cdot \hbar \cdot \vec{I} \quad (1)$$

$\gamma$  est le rapport gyromagnétique exprimé en rad/s

$\hbar$  est la constante de Planck divisée par  $2\pi$  ( $\hbar = 1,054 \cdot 10^{-34}$  J.S)

En présence d'un champ magnétique statique  $B_0$  appliqué le long de l'axe Z, les moments magnétiques s'orientent en adoptant  $2I+1$  orientations. Chaque orientation correspond à un niveau d'énergie la séparation entre deux niveau est donnée par

$$\Delta E = \gamma \cdot \hbar \cdot B_0 = h \cdot \nu_0 \quad (2)$$

$\Delta E$  (Joules, J),  $B_0$  (Tesla, T)

$\nu_0$  est la fréquence de Larmor qui correspond à la fréquence de précession (Hertz, Hz)

A l'équilibre et pour un spin  $=1/2$ , la différence de population entre les deux niveaux d'énergie  $N^+$  (pour le niveau  $\alpha$ ) et  $N^-$  (pour le niveau  $\beta$ ) est caractérisée par un très faible excès des noyaux sur le niveau  $\alpha$ . En effet, les spins se répartissent sur les niveaux d'énergies selon la loi de Boltzmann :

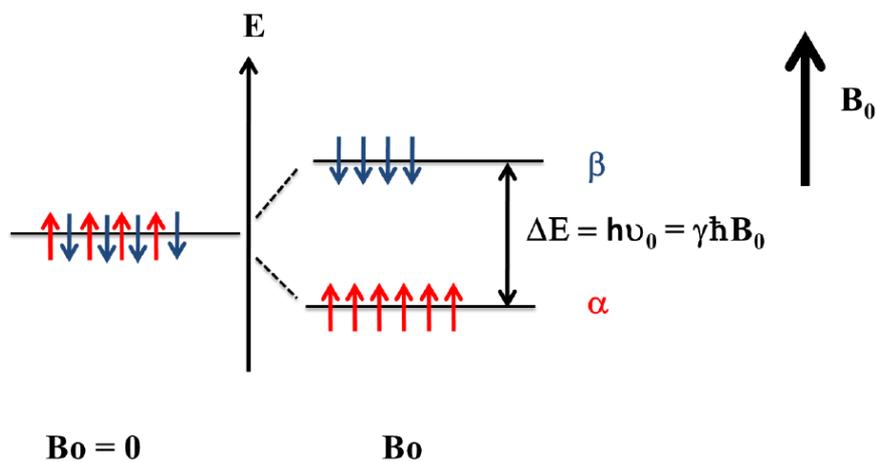
$$N^-/N^+ = e^{-\Delta E/kT} \quad (3)$$

$k$  est la constante de Boltzmann ( $1,3805 \cdot 10^{-23}$  J/Kelvin) et  $T$  la température en Kelvin

Si les noyaux sont soumis à un second champ magnétique oscillant (composant magnétique de la radiofréquence) perpendiculaire au champ  $B_0$  et ayant la valeur de la fréquence de Larmor de ces noyaux, ces derniers peuvent absorber l'énergie et sauter au niveau d'énergie le plus élevé. Puis ils la libèrent lors de la relaxation. Il s'agit de deux types de relaxation :

- Relaxation longitudinale ou spin-réseau (T1)
- Relaxation transversale ou spin-spin (T2)

Le courant induit dans la bobine excitatrice est mesuré sous forme d'un signal temporel transverse ou Free Induction Delay (FID). Le spectre RMN correspond à une transformée de Fourier de ce FID.



**Figure 25 : Eclatement des niveaux d'énergie pour un spin  $\frac{1}{2}$  par application d'un champ magnétique statique  $B_0$ : Effet Zeeman.**

## 2.3.2. Interactions en RMN

Il existe plusieurs types d'interaction en RMN : externes et internes. Il s'agit, tout d'abord, de l'interaction Zeeman et l'interaction avec la radiofréquence (externes). Mais il existe aussi d'autres interactions internes qui sont responsables de fluctuations du champ magnétique ressenti par le noyau. Toutes ces interactions sont décrites par un Hamiltonien, noté par  $\hat{H}$ .

$$\hat{H}_{\text{total}} = \hat{H}_{\text{externe}} + \hat{H}_{\text{interne}} = \hat{H}_Z + \hat{H}_{\text{rf}} + \hat{H}_{\text{DC}} + \hat{H}_{\text{CD}} + \hat{H}_{\text{CJ}} + \hat{H}_Q \quad (4)$$

Z : Zeeman, rf : radiofréquence, DC : déplacement chimique, CD : couplage dipolaire, CJ : couplage J ou couplage scalaire, Q : couplage quadripolaire.

Les interactions internes sont plus faibles que celle de Zeeman. C'est pour cela qu'elles peuvent être considérées comme des perturbations de l'Hamiltonien de Zeeman

### 2.3.2.A. Interaction Zeeman

L'interaction Zeeman est la plus forte interaction (centaines de MHz) décrite par  $\hat{H}_Z$  (entre le champ magnétique  $B_0$  et les spins nucléaires)

$$\hat{H}_Z = -\mu \cdot B_0 \quad (5)$$

### 2.3.2.B. Déplacement chimique

Le déplacement chimique est une interaction due au couplage magnétique entre le spin nucléaire et le champ statique  $B_0$  perturbé par les électrons entourant le noyau. En effet, ces électrons induisent un champ magnétique local (conformément à loi de Lenz) et le noyau ne ressent plus le  $B_0$  mais un champ effectif :

$$B_{\text{eff}} = B_0 - B_{\text{local}} \text{ avec } B_{\text{local}} = \sigma \cdot B_0 \rightarrow$$

$$B_{\text{eff}} = B_0 (1 - \sigma) \quad (6)$$

$$\nu_0 = \frac{(1 - \sigma)\gamma \cdot B_0}{2\pi} \quad (7)$$

$\sigma$  est la constante d'écran

La fréquence de résonance du spin dépend alors du  $B_0$  et du  $\sigma$ . De plus, le champ local  $B_{\text{local}}$  est très petit devant  $B_0$ . On est donc contraint de passer par une mesure relative dans laquelle le pic de résonance d'une molécule standard sert de référence. Pour le proton il s'agit du tétraméthylsilane (TMS)... Le spectre RMN est ainsi caractérisé par le déplacement chimique ( $\delta$ ) exprimé en partie par millions (ppm), qui dépend seulement de l'environnement chimique entourant le noyau :

$$\delta = (v - v_{\text{ref}}) \times 10^6 / v_{\text{ref}} \quad (8)$$

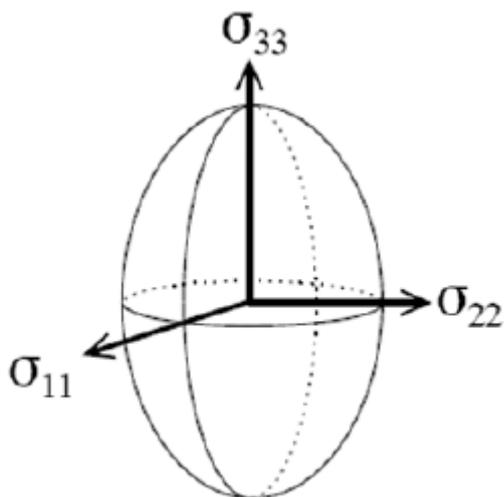


Figure 26 : Représentation du tenseur du déplacement chimique dans le référentiel PAS.

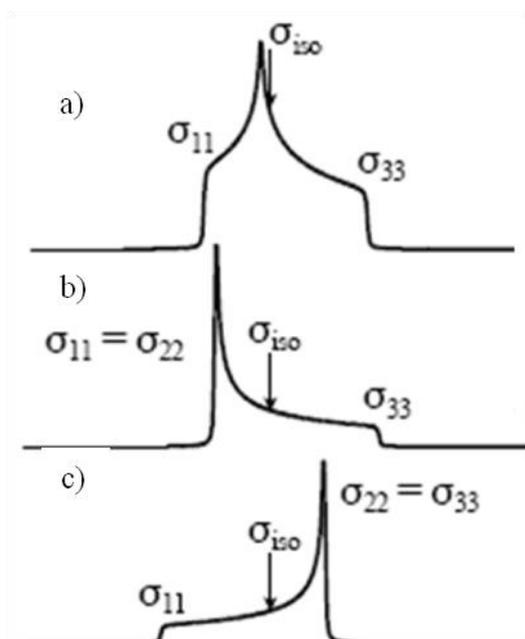


Figure 27 : Paramètres d'asymétrie influençant la forme de spectre. a) Le tenseur d'ADC avec asymétrie. b) et c) Le tenseur d'ADC avec symétrie axiale pour respectivement  $\sigma_{11} = \sigma_{22}$  et  $\sigma_{22} = \sigma_{33}$ .

Il faut noter que si la distribution des électrons autour du noyau n'est pas sphérique alors la constante d'écran dépend de l'orientation de ces électrons vis-à-vis du champ magnétique. Ceci est reconnu par l'anisotropie du déplacement chimique (ADC) dont  $\sigma$  correspond à un tenseur (matrice 3x3) :

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{zy} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \quad (9)$$

Ce tenseur peut être diagonalisé en trois composantes principales  $\sigma_{xx}$ ,  $\sigma_{yy}$  et  $\sigma_{zz}$  dans le système d'axe principal (PAS) :

$$\sigma = \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \quad (10)$$

Les tenseurs sont généralement illustrés par des ellipses où les axes correspondent à ceux de l'ellipse.

### 2.3.2.C. Couplage dipolaire

Le couplage dipolaire entre deux spins (S et I) est la conséquence d'une interaction entre leurs moments magnétiques,  $\vec{\mu}_I$  et  $\vec{\mu}_S$ , voisins dans l'espace. Ce couplage peut être homonucléaire ou hétéronucléaire, c.à.d. il s'agit respectivement deux spins S ou deux spins différents, S et I. L'Hamiltonien de cette interaction est proportionnelle à la constante du couplage directe, notée D. Cette constante dépend de la distance r entre ces deux noyaux, et de l'orientation relative du vecteur reliant les deux spins par rapport à la direction du champ magnétique ( $\theta$ ).

$$D = (\mu_0 / 4\pi) \gamma_1 \gamma_2 \hbar \cdot r^3 (3\cos^2\theta - 1) \quad (11)$$

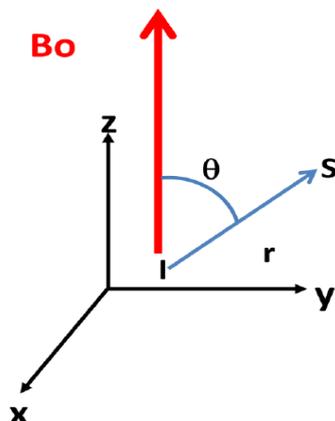


Figure 28 : Représentation du couplage dipolaire entre deux spins I et S.

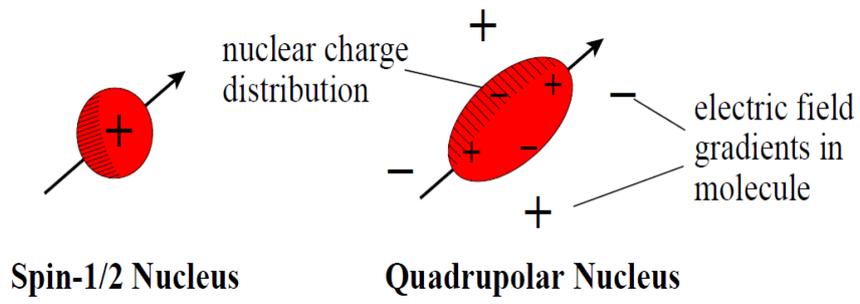
### 2.3.2.D. Couplage indirect ou couplage j

Le couplage J se fait entre des spins liés d'une façon covalente via les électrons liants. Evidemment ce couplage diminue avec l'augmentation du nombre des liaisons qui séparent ces spins. Il est désigné par la constante du couplage J,  ${}^nJ_{x-y}$ , avec n le nombre des liaisons séparant les deux spins x et y. Cette constante est le facteur principal agissant sur l'Hamiltonien du couplage J ( $\hat{H}_{CJ}$ ).

### 2.3.2.E. Interaction quadripolaire

Tous les noyaux ayant des spins nucléaires supérieurs à  $\frac{1}{2}$  ont un moment électrique quadripolaire qui interagit avec le champ électrique traversant le noyau (Figure 29). Ce moment électrique quadripolaire est la conséquence d'une distribution de la charge dans le noyau.

L'interaction résultant est reconnue comme étant l'interaction quadripolaire qui perturbe les niveaux d'énergie Zeeman. L'Hamiltonien de cette interaction dépend, entre autre, du moment quadripolaire spécifique à chaque isotope du noyau (Q).



**Figure 29 : Comparaison de la distribution de la charge entre deux noyaux ayant un spin égale ou supérieure à 1/2.**

### 2.3.3. RMN en phase solide

La RMN en phase solide (ssRMN) aboutit, contrairement à la RMN en solution, à des spectres compliqués avec des pics trop larges. Cela est dû à l'anisotropie des interactions internes surtout le déplacement chimique et le couplage dipolaire qui sont moyennés en solution par les fluctuations rapides des molécules (mouvements browniens). Or les pics trop larges obtenus par ssRMN portent des informations importantes sur la structure et la dynamique des molécules comme les distances interatomiques, les angles dièdres, vitesse de fluctuations... D'où c'était la nécessité de développer des techniques en RMN de solide pour augmenter la résolution de ces spectres tout en gardant les informations qu'ils contiennent. Il s'agit principalement de la rotation à l'angle magique, la polarisation croisée et découplage dipolaire.

#### 2.3.3.A. Rotation à l'angle magique

Cette technique se base sur la partie spatiale des Hamiltoniens, plus précisément  $(3\cos^2\theta-1)$ . Ce facteur disparaît en solution par moyennement mais il peut être aussi annulé lorsque  $\cos^2\theta = 1$ , ceci donne  $\theta = 54,7^\circ$  appelé l'angle magique. Alors la rotation autour d'un axe formant un angle de  $54,7^\circ$  par rapport au  $B_0$  moyenne les interactions dipolaire et de déplacement chimique à leurs valeurs isotropes.

#### 2.3.3.B. Polarisation croisée

La polarisation croisée ou en anglais Cross Polarization (CP) permet un gain en sensibilité en mesurant le signal d'un noyau peu abondant (typiquement  $^{13}\text{C}$  ou  $^{15}\text{N}$ ) par l'intermédiaire d'un noyau abondant voisin (typiquement  $^1\text{H}$ ). Cette technique consiste à exciter le noyau abondant et de faire un transfert de polarisation vers le noyau rare que l'on veut observer. Ce transfert se fait via le couplage dipolaire et obéissant à la condition Hartmann-Hahn. Il est d'autant plus efficace que les distances sont courtes et que la liaison est rigide entre les deux noyaux.

### 2.3.3.C. Découplage dipolaire

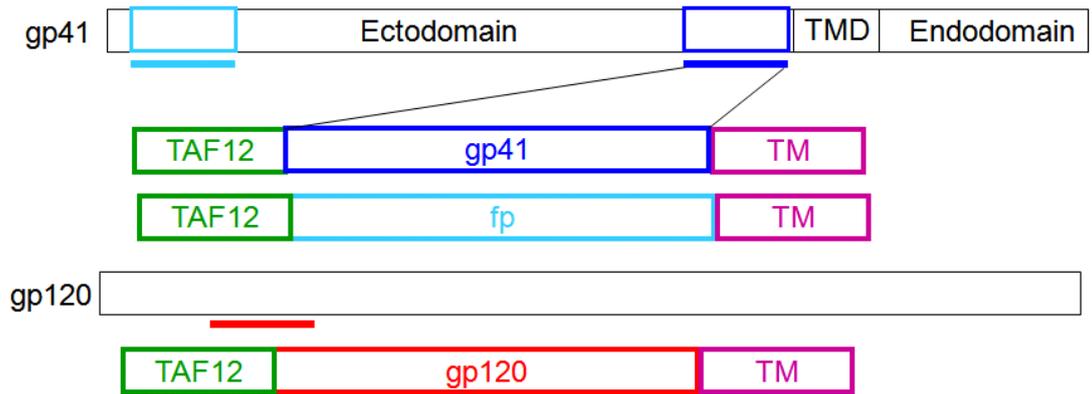
Le signal RMN des spins rares ( $^{13}\text{C}$  et  $^{15}\text{N}$ ) sont larges à cause du couplage dipolaire avec les protons existant à proximité de ces spins, ces couplages ne sont pas supprimés efficacement par le MAS. C'est pour cela qu'il faut supprimer ce type de couplage en irradiant par une radiofréquence à la fréquence de la résonance du proton, c'est le découplage.

# 3. Chapitre 3 : Surexpression des polypeptides issus de gp41, fp et gp120

## 3.1. Les constructions peptidiques

Les trois polypeptides à surexprimer en système bactérien sont issus des glycoprotéines d'enveloppe du VIH-1 (deux de la gp41 : gp41 et fp, et le troisième de la gp120). Chaque construction comporte :

- en N-terminale, un partenaire de fusion comprenant la protéine TAF12, qui permet de diminuer la toxicité du produit d'expression, suivi d'un tag histidine 6x pour faciliter les étapes de purification,
- le site (-Asp-Pro-), qui permet, suite à clivage chimique à l'acide formique, de se débarrasser du partenaire de fusion,
- puis, la séquence de l'épitope d'intérêt. Le choix des séquences épitopiques a porté sur le fait qu'elles soient reconnues par des anticorps neutralisants. Ces épitopes ont également été choisis pour leur localisation à proximité de la membrane lors de l'étape de fusion et pour le fait que l'anticorps 4E10 se lie à ces trois épitopes (Hager-Braun *et al.*, 2006). En fait, le polypeptide gp41<sup>TM</sup> contient les deux épitopes des anticorps 4E10 (ELDKWA) et 2F5 (NWFDIT), le polypeptide gp120<sup>TM</sup> renferme la séquence 33 à 45 de la gp120 et fp<sup>TM</sup> correspond au peptide de fusion du VIH-1.
- et enfin, en C-terminale, une séquence transmembranaire (TM) construite sur la base du peptide MS1. Le choix de ce dernier s'est fait sur sa capacité à former des dimères et des trimères en présence des micelles. Lorsqu'il est muté en position valine 13 par une sérine, il présente uniquement des trimères (Choma *et al.*, 2000; Tatko *et al.*, 2006) ce qui permet aux épitopes de mimer le trimère des gp.



**Figure 30 : Conception des séquences polypeptidiques à surexprimer.**

## 3.2. Matériels et méthodes

### 3.2.1. Clonage des polypeptides

Les clonages des trois protéines (TAF12-gp41<sup>TM</sup>, TAF12-fp<sup>TM</sup> et TAF12-gp120) ont été faits par le Dr. Marc PRUDHON (Thèse soutenue à l'Université de Strasbourg, septembre 2011).

Les plasmides de clonage contenant les séquences nucléotidiques codant pour les épitopes-TM (gp120<sup>TM</sup>, gp41<sup>TM</sup> et fp<sup>TM</sup>) ont été commandés chez GeneArt AG, Regensburg. Les fragments codant pour les polypeptides d'intérêt sont amplifiés par Polymer Chain Reaction (PCR). Les produits PCR purifiés et le vecteur d'expression pTIPX-4 sont digérés par les deux enzymes de restriction BamH1 et SacI. Ensuite, une réaction de ligation est lancée pour insérer les fragments d'intérêt dans le pTIPX-4 en aval et en phase de lecture avec la séquence codant pour le TAF12. Le tout est amplifié *in vivo* dans les bactéries de clonage (XL1 Blue Competent Cells, 200249, Stratagene). L'ADN plasmidique est ensuite purifié et gardé à -20°C. La conformité du vecteur d'expression ayant inséré le gène d'intérêt est vérifié par séquençage nucléotidique (envoi à Eurofins MWG Operon, Allemagne).

### 3.2.2. Transformation des bactéries d'expression

Un  $\mu\text{L}$  d'ADN plasmidique (+/- 100 ng/ $\mu\text{L}$ ) est mélangé à 100  $\mu\text{L}$  des bactéries compétentes décongelées (BL21 Star (DE3) pLysS One Shot®, C6020-03, Invitrogen), puis incubés pendant 30 min dans la glace. Un choc thermique à 42°C pendant 30 sec est alors effectué suivi d'une incubation dans la glace pendant au moins 2 min. 250  $\mu\text{L}$  de milieu LB (Luria Bertani) sont ajoutés. Après incubation sous agitation (250 rpm) pendant 1h à 37°C, différents volumes (20, 40 et 80  $\mu\text{L}$ ) de la transformation sont étalés sur des boîtes de gélose contenant les antibiotiques permettant la sélection des transformants : kanamycine (kan) 50  $\mu\text{g}/\text{mL}$  pour le plasmide PTIPX4 et chloramphénicol (CAM) 60  $\mu\text{g}/\text{mL}$  pour le plasmide pLysS. Les géloses sont alors incubées à 37°C à l'étuve (Memmert,) sur la nuit.

### 3.2.3. Surexpression des polypeptides

Une préculture de bactéries est obtenue par ensemencement d'une colonie des transformants précédemment obtenus dans du milieu LB liquide pendant une nuit à 37°C sous agitation (250 rpm). Des milieux de culture (LBYE : Luria Bertani Yeast Extract, glucose 1%, kan 50 µg/mL et CAM 34 µg/mL), respectivement de 500 mL (dans un erlenmeyer de 2 litres) ou de 4 L (dans le fermenteur : Biostat® B) sont ensemencées au 1/100 par les volumes appropriés de précultures. Un prélèvement de 1 mL est effectué chaque 30 min pour suivre la cinétique de la croissance bactérienne par mesure de la densité optique (DO) à 600 nm (Uvikon Spectrophotometer). Lorsque la DO atteint la valeur optimale d'induction, préalablement déterminée pour chacune des constructions, une partie est prélevée pour constituer la condition non induite (respectivement 100 mL et 500 mL), puis l'inducteur (IPTG) est ajouté à une concentration finale de 1mM dans le volume restant (respectivement 400 mL et 3.5 L) qui constitue la condition induite. L'expression est stoppée après 4h ou lorsque la DO chute. 10 mL de suspension bactérienne sont alors prélevés pour chaque condition afin de déterminer la présence et la localisation de la protéine surexprimée dans les différentes fractions de purification par migration électrophorétique sur gel. Les bactéries sont centrifugées à 5000 g pendant 30 min à 4°C puis les culots sont congelés à -20°C.

### 3.2.4. Extraction protéique

Les culots bactériens correspondant soit aux 10 mL prélevés à la fin de l'expression ou soit au reste de l'expression sont traités de la même manière avec des volumes adaptés. Le protocole ci-dessous donne les volumes pour les 10 mL de culture.

Les culots bactériens décongelés sont repris dans 2 mL de tampon de lyse (Tris 50 mM ; NaCl 300 mM ; Triton 0,5% à pH 8) puis soniqués deux fois pendant 5 min à 50% du cycle temps. 80 µL du lysat obtenu sont prélevés et constituent l'échantillon lysat total (ET) à déposer sur gel. Après centrifugation (11000 g ; 30 min et 4°C), 80 µL du surnageant est prélevé et constitue la fraction soluble. Le culot bactérien lysé correspondant à la fraction insoluble, est lavé deux fois par le tampon de lavage (Tris 50 mM ; NaCl 300 mM à pH 8) pour éliminer le Triton : le premier lavage se fait avec le ratio lyse/lavage, V/V, et le deuxième avec V/0,5V. Le surnageant est éliminé après chaque centrifugation, le culot est finalement repris dans le tampon de dénaturation (Tris 50 mM ; NaCl 300 mM ; urée 8 M à pH 8) puis la suspension est incubée sur la roue (rotation à 20 rpm) à température ambiante sur la nuit. La fraction membranes et corps d'inclusion (mbe et CI) est alors

prélevée. Une centrifugation (11000 g ; 20 min à 20°C) permet de culoter les membranes qui sont conservés à -20°C. La fraction correspondant au surnageant (corps d'inclusion) et dans laquelle les protéines d'intérêt sont attendues, est incubée durant 1h30min à température ambiante sur la roue (rotation 20 rpm) avec de la résine de Cobalt afin de fixer les protéines présentant le tag Histidine (tag His). Une centrifugation (11000g, 20min) à température ambiante permet de séparer la fraction FT (Flow Through) constituée des protéines non retenues sur la résine. Le culot résine est lavé, puis des éluions successives avec la solution d'éluion (Imidazole 150 mM ; Tris 50mM ; NaCl 300 mM à pH7) permettent de récupérer les protéines d'intérêt fixées grâce à leur tag His. L'acide éthylène diamine tétra acétique (EDTA) est utilisé dans la dernière éluion pour décrocher tout ce qui est retenu sur la résine.

### 3.2.5. Clivage par l'acide formique

Le clivage se fait sur la fraction insoluble (gp41TM, gp120TM et fpTM) par ajout d'acide formique à une concentration finale de 75% et incubation à 50°C dans un bain-marie durant 24h. L'acide formique permet le clivage des polypeptides au niveau du site Asp-Pro, en les libérant du partenaire de fusion. La cinétique du clivage est suivie par des prélèvements toute les 12h et est visualisée par migration de SDS-PAGE (Sodium Dodecyl Sodium-PolyAcrylamide Gel Electrophoresis).

### 3.2.6. Purification des polypeptides

#### 3.2.6.A. HPLC en phase inversée

L'HPLC (High Performance Liquid Chromatography) est une technique de chromatographie utilisant deux phases : une phase stationnaire (billes de silice greffées) et une phase mobile (solvants) permettant la séparation des protéines en fonction de leurs propriétés physicochimiques. Dans notre cas, les protéines à purifier sont des protéines membranaires et donc à caractère hydrophobe, la séparation des protéines se fera donc en fonction de leur hydrophobicité. Il s'agit alors d'une HPLC à phase inversée (Reverse phase-HPLC). L'échantillon protéique est injecté sur une colonne HPLC équilibrée avec un solvant composé d'eau 90% ; acide tri fluoro acétique (TFA) 0.1%. Les protéines adsorbées sur la colonne, seront progressivement éluées par un gradient

croissant d'isopropanol 90% ; eau 10% ; TFA 0.1%. La phase stationnaire est composée de billes de silice greffées avec des chaînes de 4 carbones (C4). Le choix d'une chaîne courte s'explique par le caractère très hydrophobe des protéines membranaires à purifier. En effet, des chaînes carbonées plus longues entraîneraient des interactions trop fortes avec les protéines et ne permettraient pas la dissociation et donc l'élution de celles-ci par le solvant. Deux types d'HPLC ont été utilisés : la première est une HPLC semi-préparative (Bischoff) pour mettre au point les conditions d'élutions et faire les premiers tests sur chaque préparation protéique ; la seconde est une HPLC préparative (Gilson) pour purifier de grandes quantités de protéines. Deux types de colonnes ont respectivement été utilisées pour ces HPLC : une colonne analytique : Grace Vydac Protein C4 (214TP5415; S/N 030311-1-1; Particule 300Å, 5µm, Colonne 4.6mm x 150mm) et une colonne préparative (Phenomenex Jupiter C4 Particule 300Å, 5µm, Colonne 30mm x 150mm).

La purification finale des polypeptides se fait en une ou deux étapes (respectivement pour gp41TM et gp120TM, et fpTM) par HPLC en phase inversée. L'élution se fait suivant deux gradients d'isopropanol/eau selon le polypeptide à purifier.

### 3.2.6.B. IMAC

Après extraction protéique, la dernière fraction récupérée est la fraction insoluble qui contient nos protéines d'intérêt. Cette fraction comprend les corps d'inclusion et les membranes qui sont resolubilisés dans de l'urée 8M à température ambiante. Après centrifugation, le surnageant contenant les corps d'inclusion est collecté et les protéines sont purifiées par chromatographie d'affinité IMAC (Ion Metal Affinity Chromatography) qui est basée sur l'affinité des métaux de transition (Ni<sup>2+</sup>, Co<sup>2+</sup>...) à quelques acides aminés considérés comme donneur d'électrons particulièrement l'histidine (Porath *et al.*, 1975).

Les corps d'inclusion sont incubés avec la résine de cobalt à température ambiante durant 1h30. Après centrifugation, une fraction du surnageant correspondant au matériel non retenu (Flowthrough, FT) est conservée pour être déposée sur gel. Le culot de résine est lavé et des éluions successives à l'imidazole 150 mM permettent d'éluer les protéines adsorbées. En fait, l'imidazole entre en compétition avec les protéines adsorbées à la résine ce qui permet leur élution.

## 3.2.7. Techniques de caractérisation

### 3.2.7.A. Migration sur gel

Les protéines sont séparées par électrophorèse en conditions dénaturantes en présence de 8% de SDS selon la technique « SDS-PAGE ». Elles migrent successivement à travers deux gels. Le premier gel est appelé gel de concentration ou d'alignement, faiblement concentré en acrylamide 4%. Il sert de concentrer ces protéines à la limite du second gel. Ce dernier, dit de séparation est plus concentré en acrylamide 16% et il permet de séparer les protéines selon leur taille. La migration a lieu pendant 1h à 150V. Les gels sont colorés par le bleu de Coomassie (Roti<sup>®</sup> Blue 5X, A15.1, Roth) puis décolorés avec une solution de méthanol 25% et d'eau 75%. Les photos des gels sont prises avec le lecteur de gel El Logic 200, Imaging System de Kodak.

### 3.2.7.B. Western blot

Les extraits protéiques qui sont séparés par migration électrophorétique sont transférés sur une membrane de nitrocellulose durant 1h à 4°C dans le tampon de transfert conservé à 4°C (3 g de Tris-HCl, 14,4 g glycine dans 20% de méthanol et 80% du tampon TBST (Tris-HCl 10 mM, NaCl 150 mM et Tween 0,0625% à pH=8). La saturation de la membrane est faite dans le tampon TBST avec 5% lait à 4°C sous agitation durant la nuit. Le lendemain, la membrane est incubée avec l'anticorps primaire (Anti-Histidine-Tagged Protein Mouse mAb, OB05) dilué à 1/1000 dans le tampon de saturation, durant 2h (1h à 4°C puis 1h à RT). Après six lavages (5 min/lavage sous agitation) dans du tampon TBST, la membrane est incubée 1h avec l'anticorps secondaire (Goat anti-Mouse IgG Antibody, (H+L) HRP conjugate, AP308P) dilué au 1/2000 dans le tampon de saturation de la membrane. La membrane est ensuite lavée comme la première fois.

La révélation est faite par chimiluminescence (Enhanced Chemiluminescent : ECL). Deux mélanges réactionnels (R1 et R2) sont préparés à partir du luminol (A-8511, Sigma) et l'acide p-coumarique (C9008, Sigma) :

- R1 : 44 µL acide p-coumarique (90 mM dans DMSO) + 100 µL luminol (250 mM dans DMSO) compléter le volume à 10 mL Tris-HCl 0,1 M à pH8
- R2 : 5,5 µL H<sub>2</sub>O<sub>2</sub> compléter le volume à 10 mL Tris-HCl 0,1 M à pH8

Au moment de la révélation, la membrane est incubée durant 1 min sous agitation dans R1 et R2. Les signaux sont détectés par le spectromètre de luminescence (SLM AMICO BOWMAN series 2).

### 3.2.7.C. Spectrométrie de masse

La matrice (acide sinapique, SA, ou  $\alpha$ -cyano-4-hydroxycinnamique,  $\alpha$ -cyano) est préparée à une concentration finale de 10 mg/mL dans acétonirile 30% et TFA 0,1%. Un  $\mu$ L de la matrice utilisée est déposé sur une plaque (05174, Bruker) et mélangé avec 1  $\mu$ L de l'échantillon à tester. Après séchage à l'air libre, la plaque est passée dans le spectromètre MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization - Time Of Flight ; BRUKER Daltonics model Autoflex).

Les spectres de masse ESI Q-TOF (Electro Spray Ionisation Quadrupole-TOF) ont été faits au laboratoire de « Spectrométrie de Masse des Interactions et des Systèmes », dirigé par Dr. Emmanuelle LEIZE-WAGNER.

### 3.2.7.D. Quantification des polypeptides

Le kit Micro BCA<sup>TM</sup> Protein Assay (23235) était utilisé pour quantifier les polypeptides purifiés. Cette méthode est basée sur la réduction des ions  $\text{Cu}^{2+}$  par les protéines en formant des ions  $\text{Cu}^+$ . Ces derniers sont chelatés par deux molécules d'acide bicinchoninique tout en donnant à un complexe violet qui absorbe à 562 nm. Cette absorption est proportionnelle à la concentration des protéines présentes.

## 3.2.8. Immunisation des souris

Une série d'immunisation des souris est faite par un polypeptide. Chaque série comprends trois groupes en fonction de la solution injectée (polypeptide insérée dans des liposomes, polypeptide seule et un contrôle avec du PBS). Chaque groupe était composé de cinq souris, de deux différentes souches (3 BALB/c et 2 C57BL/6). Les injections intrapéritonéales ont été effectuées tous les 15 jours avec un volume de 200 $\mu$ L (100  $\mu$ L d'antigène + 100  $\mu$ L d'adjuvant). En parallèle, des prélèvements de sérum ont été effectués pour vérifier la présence d'anticorps par Enzyme-Linked ImmunoSorbent Assay (ELISA).

### 3.3. Résultats et discussion

#### 3.3.1. Expression des polypeptides

Une préculture de 50 mL (glucose 1%, kan 50 $\mu$ g/mL et CAM 34 $\mu$ g/mL) estensemencée avec une colonie de BL21 DE3 pLysS transformées par le pTIPX4-gp41TM. Après une nuit de culture à 37°C à l'incubateur, un fermenteur (4L Biostat® B) de 4 L du milieu LBYE (glucose 1%, kan 50  $\mu$ g/mL et CAM 34  $\mu$ g/mL) estensemencé au 1/100 avec 40 mL de cette préculture. La cinétique d'expression qui se fait à 37°C est suivie par des prélèvements toutes les 30 min et la lecture de la densité optique (DO) à 600nm par un spectrophotomètre.

L'induction est faite par ajout d'IPTG à une concentration finale de 1 mM lorsque les bactéries sont en phase exponentielle de leur croissance (DO=1,4 pour gp120TM, voir Figure 31). Il est à noter que 500 mL de culture ont été prélevés avant induction et constitue la condition non induite, contrôle d'expression. La culture est arrêtée lorsque la DO dans la condition d'induction chute ou après 4h d'induction. Les milieux de cultures sont alors centrifugés et les culots bactériens sont pesés et conservés à -20°C.

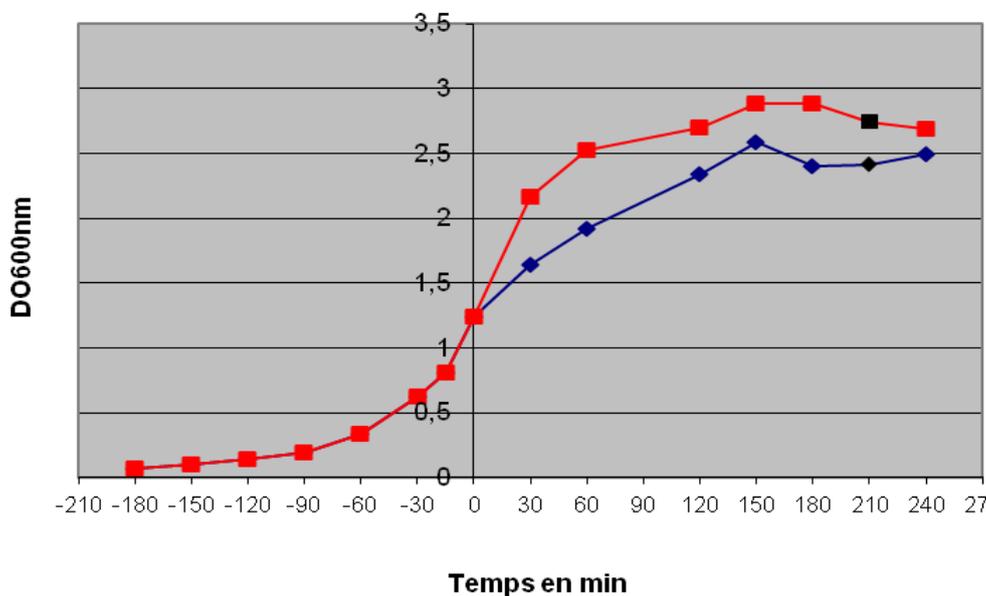
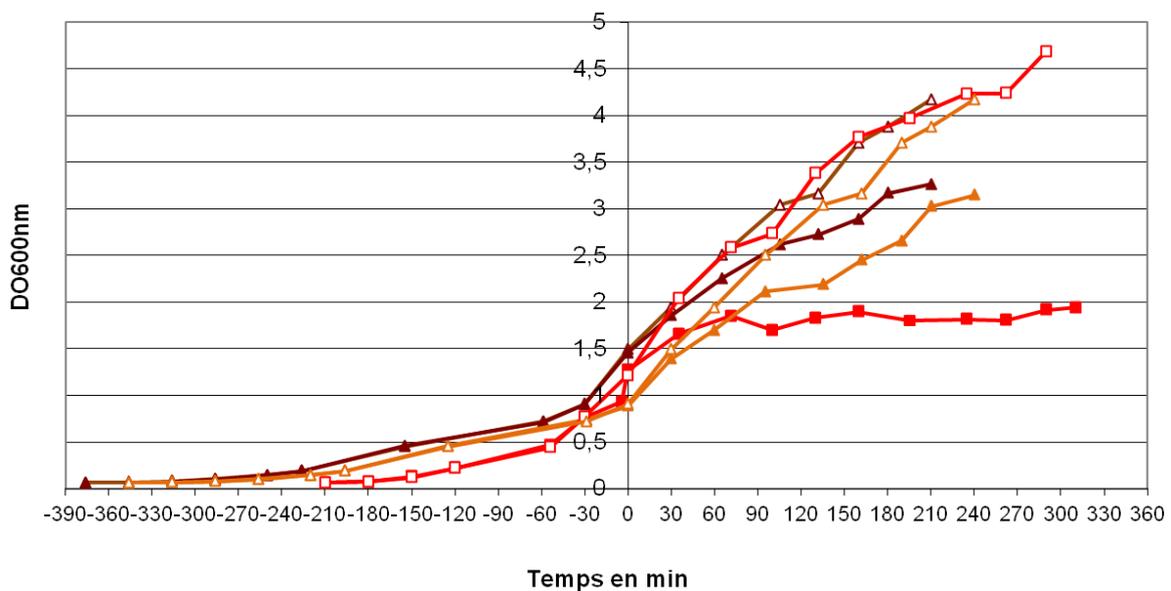


Figure 31 : Cinétique de la croissance bactérienne au cours de l'expression du TAF12-gp120TM. La croissance bactérienne est suivie par la mesure de la DO<sub>600nm</sub> en fonction du temps. Le temps zéro correspond à l'induction par l'ajout de l'IPTG (concentration finale 1 mM) effectué à DO=1.4. La courbe rouge correspond à la condition induite et la bleue à celle non induite

Plusieurs DO d'induction et plusieurs températures ont été testées afin de déterminer la DO d'induction et la température optimales pour l'expression de chacun des trois polypeptides d'intérêt. Comme le montre la Figure 32, l'expression bactérienne du TAF12-gp41TM à 37°C (carré plein rouge) entraîne un net ralentissement de la croissance bactérienne après induction en comparaison au contrôle non induit (carré vides rouge), et ce quel que soit la DO d'induction (pour plus de clarté, seule la DO d'induction =1,2 est présentée sur le graphe). Tandis que l'expression du TAF12-gp41TM à 28°C (triangles pleins orange et marron) montre une croissance bactérienne plus importante qui est plus prononcée lorsque l'induction est effectuée à DO=1,3 (triangles pleins marron).

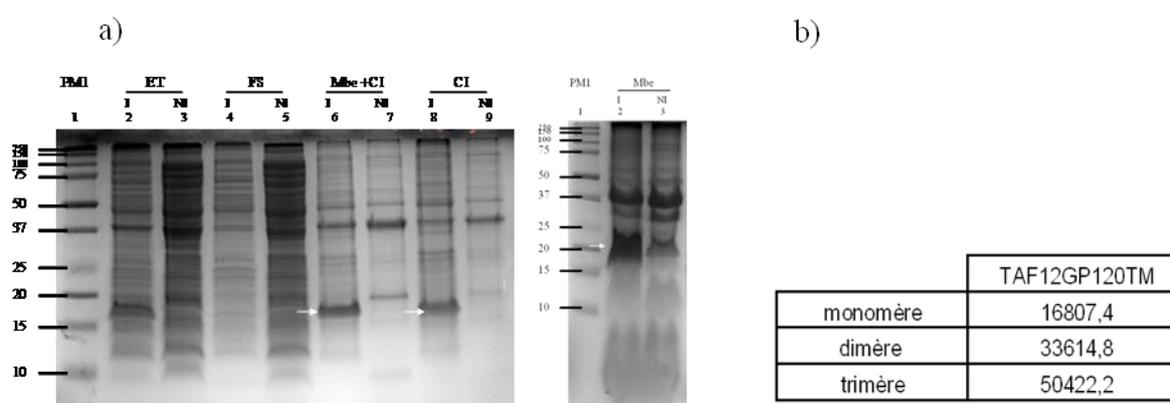


**Figure 32 : Influence de la température et la DO d'induction sur la cinétique de la croissance bactérienne au cours de l'expression du TAF12-gp41TM. Formes pleines : conditions induites (I). Formes vides : conditions non induites (NI). Carré rouge : température d'expression de 37°C et DO d'induction = 1,2. Triangles orange et marron : température d'expression de 28°C et respectivement DO d'induction = 0,8 et 1,3.**

## 3.3.2. Purification des polypeptides

### 3.3.2.A. Localisation des protéines d'intérêt

Les culots bactériens subissent ensuite une extraction protéique. Durant cette extraction, les différentes fractions sont collectées à chaque étape puis déposées sur gel polyacrylamide en conditions dénaturantes (SDS-PAGE) pour vérifier la localisation de la protéine d'intérêt. En fait, cette technique permet de séparer les protéines en fonction de leur taille (le dodécylsulfate de sodium (SDS) dénature les protéines et leur donne une charge globale négative).

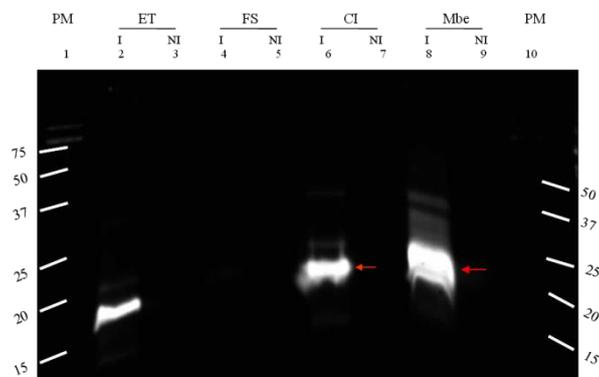


**Figure 33 :** a) Migration des différentes fractions d'extraction protéique pour TAF-gp120TM. Les différentes fractions issues de l'extraction protéique des conditions induite (I) et non induite (NI) ont été déposées sur gel SDS-PAGE 16 %. Les volumes de dépôts sont les suivants : 45µL des fractions ET, FS, Mbe+CI et CI (équivalent à 144 µL de culture), 10 µL Mbe (équivalent à 1,6 mL de culture). Après migration électrophorétique 1h à 150V, les gels sont colorés au bleu de Coomassie. ET : Echantillon Total, FS : Fraction Soluble, Mbe+CI : membrane+Corps d'Inclusion, Mbe : Membranes, CI : Corps d'Inclusion; PM : marqueur de poids moléculaire (Precision Plus Protein™ Dual Color Standards #161-0374). b) Masses théoriques des différentes formes de TAF12-gp120TM. Elles sont obtenues grâce au logiciel ProtParam du site ExPASy (<http://expasy.org/tools/protparam.html>) qui calcule la masse

D'après Figure 33, plusieurs bandes dans différentes fractions ont une taille proche de TAF12-gp120TM (non clivé). En comparant, sur les deux gels de la, le profil de migration dans les deux conditions induite et non induite, nous remarquons qu'une bande d'un poids moléculaire apparent se situant entre 15 et 20 kDa est présente uniquement en condition induite. Elle apparaît dans

l'échantillon total, les fractions membranes + corps d'inclusion, corps d'inclusion et membranes. Cette bande peut correspondre au monomère de TAF12-gp120TM (16807,4 Da ; voir le tableau présenté dans la Figure 33.b). Mais dans les membranes, une autre bande dans la condition NI présente approximativement la même taille que TAF12-gp120TM. La technique SDS-PAGE permet de définir un poids moléculaire apparent mais ne donne pas une information suffisamment précise sur la masse réelle des protéines visualisées pour permettre de conclure sur l'identité de celles-ci.

Pour confirmer qu'il s'agit bien du TAF12-gp120TM, un Western blot a été réalisé, tout d'abord en un SDS-PAGE puis le transfert des protéines sur une membrane et leur détection par un anticorps spécifique. Dans notre cas, un anticorps anti-his tag a été utilisé puisque l'épitope (His6)

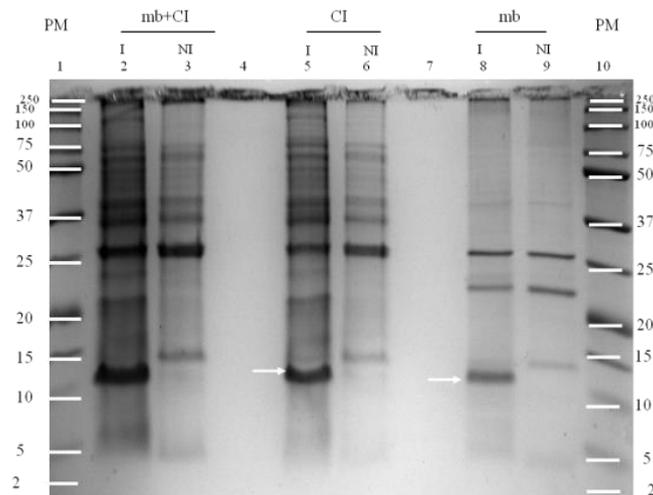


**Figure 34 : Western Blot anti-His tag sur les différentes fractions de l'extraction protéique des bactéries exprimant TAF12-gp120TM. Les volumes de dépôts sont les suivants : 45µL des fractions ET, FS et CI (équivalent à 144 µL de culture), 10 µL Mbe (équivalent à 1,6 mL de culture). Après migration électrophorétique 1h à 150V, les protéines sont transférées sur une membrane nitrocellulose durant 1h à 100V et à 4°C. La révélation est faite suivant le protocole d'ECL (Enhanced Chemiluminescent).**

correspondant est présent en aval du TAF12.

D'après la Figure 34, nous observons un signal correspondant à la taille d'un monomère de TAF12-gp120TM (16807,4 Da) dans les fractions ET (échantillon totale), CI (corps d'inclusion) et Mbe (membranes). Ceci confirme que cette protéine exprimée est localisée dans la fraction insoluble.

Dans le but d'estimer la proportion de cette protéine de façon plus précise dans chacune des composants de la fraction insoluble (CI et Mbe), des quantités équivalentes en termes de litres de culture bactérienne ont été déposées sur gel SDS-PAGE. La Figure 35 montre que la protéine d'intérêt est majoritairement localisée dans les corps d'inclusion. C'est donc cette fraction qui sera utilisée pour purifier le TAF12-gp120TM.



**Figure 35 : Migration de quantités équivalentes de la fraction insoluble des bactéries exprimant TAF12-gp120TM. Les volumes de dépôts sont les suivants : 45µL des fractions Mbe+CI et CI ; et 17 µL de Mbe dilué 20 fois (équivalents à 144 µL de culture).**

Les mêmes démarches ont été effectuées pour les autres constructions. Les autres protéines, le TAF12-gp41TM et TAF12-fpTM sont également localisés dans la fraction insoluble et répartie entre les corps d'inclusion et les membranes (voir Figure 36 et Figure 37).

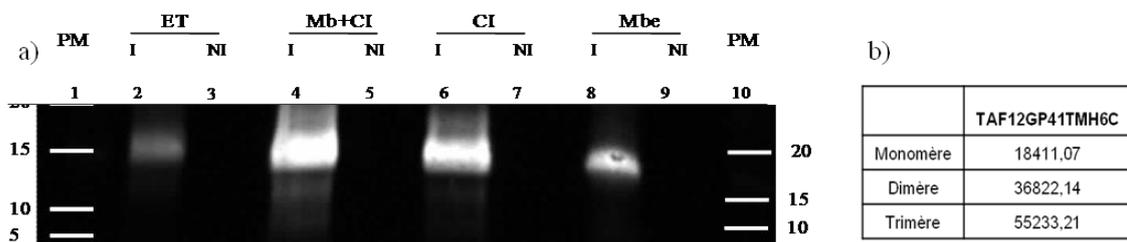


Figure 36 : a) Western Blot anti-His tag sur les différentes fractions de l'extraction protéique des bactéries exprimant TAF12-gp41TM. b) Masses théoriques des différentes formes de TAF12-gp41TM.

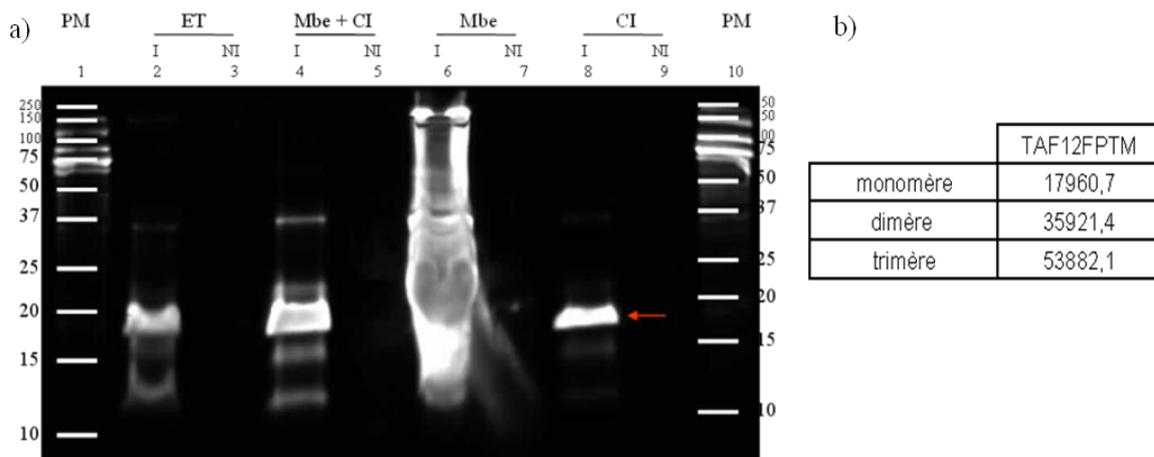
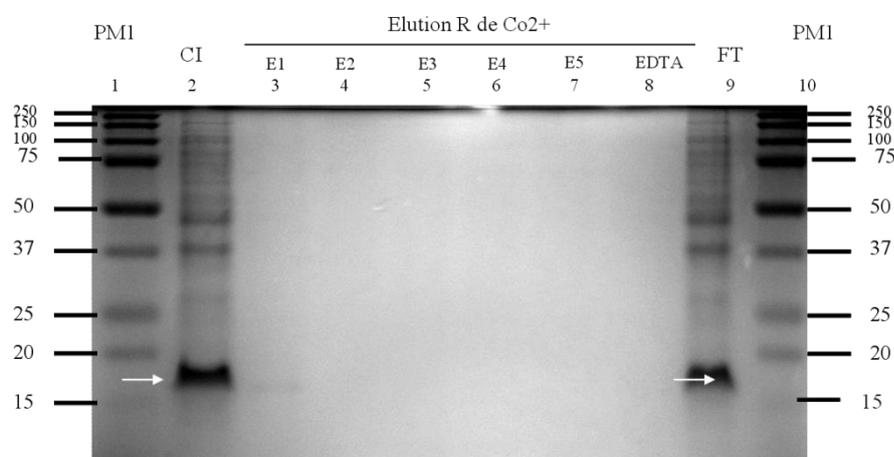


Figure 37 : a) Western Blot anti-His tag sur les différentes fractions de l'extraction protéique des bactéries exprimant TAF12-fpTM. b) Masses théoriques des différentes formes de TAF12-

### 3.3.2.B. Purification à partir des corps d'inclusion

Après extraction protéique, la dernière fraction récupérée est la fraction insoluble qui contient nos protéines d'intérêt. Cette fraction comprend les corps d'inclusion et les membranes qui sont resolubilisés dans de l'urée 8M à température ambiante. Après centrifugation, le surnageant contenant les CI est collecté et les protéines sont purifiées par chromatographie d'affinité IMAC (résine de cobalt). Différentes fractions sont collectées à chaque étape de cette purification et déposées sur gel SDS-PAGE : une fraction du surnageant correspondant au matériel non retenu (Flowthrough, FT), les éluions successives à l'imidazole 150 mM et la dernière fraction correspondant à l'éluion avec l'EDTA.

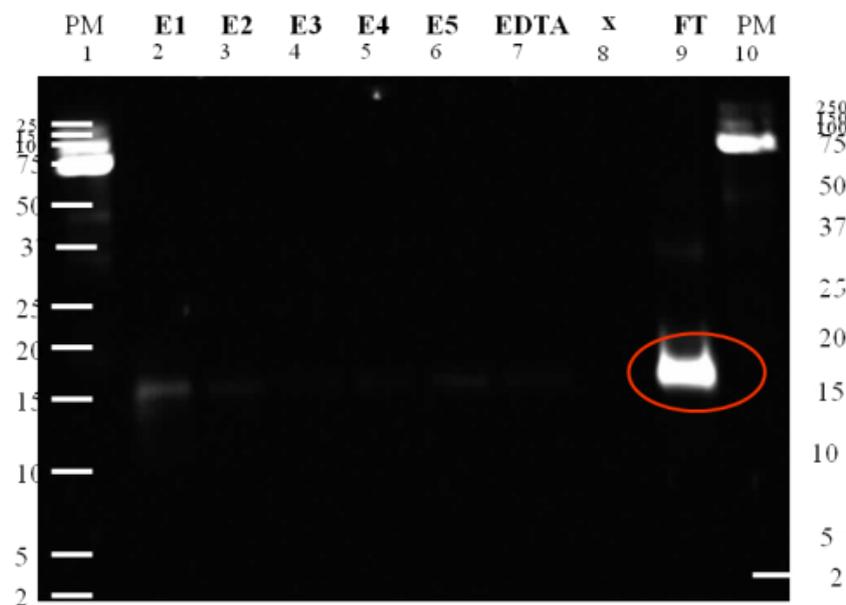


**Figure 38 : SDS-PAGE 16% sur les différentes fractions de la purification TAF12-gp120<sup>TM</sup> par résine de cobalt. La fraction insoluble est dénaturée dans urée 8M. Les volumes de dépôts sont les suivants : 37 $\mu$ L des fractions de CI et FT, Flow through, (équivalent à 200  $\mu$ L de culture), 45  $\mu$ L d'éluions avec imidazole de 1 à 5 et d'éluion avec EDTA (équivalent à 10 mL de culture). Après migration électrophorétique 1h à 150V, les gels sont colorés au bleu de Coomassie**

D'après la Figure 38, le puits 2 correspondant aux CI et servant de contrôle avant incubation, présente une bande de forte intensité et majoritaire migrant à la taille attendue pour le TAF12-gp120<sup>TM</sup> mais également de nombreuses autres protéines de plus grande taille. Les mêmes bandes sont retrouvées dans la fraction FT après incubation avec la résine, et présente les mêmes intensités. Dans l'éluion E1, seule la bande correspondant à la protéine d'intérêt est présente mais de très faible intensité. Rien n'a été détecté dans les autres fractions d'éluion. Ceci indique que le TAF12-gp120<sup>TM</sup> n'a pas été retenu sur la résine cobalt.

D'autres conditions ont alors été testées : diminution de la concentration de l'urée de 8M à 6M. En effet, une concentration de 8M est limite dans ce type de chromatographie (Cheung *et al.*, 2012). Mais là encore, la protéine d'intérêt n'a pas été retenue sur la résine cobalt (voir Figure 39). Une autre solution aurait été de dénaturer la fraction insoluble dans guanidine-HCl 6M mais ce dernier précipite avec le SDS présent dans le tampon de charge utilisé dans le SDS-PAGE.

Probablement, le tag His n'est pas accessible à la résine malgré la présence d'un dénaturant (urée 8M). Il faut alors soit changer l'emplacement de ce tag ou bien trouver une autre méthode de purification (Biotech, 2013).



**Figure 39 :** Western blot anti-His tag sur les différentes fractions de la purification TAF12-gp120<sup>TM</sup> par résine de cobalt. La fraction insoluble est dénaturée dans de l'urée 6M. Les volumes de dépôts sont les suivants : 37 $\mu$ L des fractions de FT (équivalent à 200  $\mu$ L de culture), 45  $\mu$ L d'élutions avec imidazole de 1 à 5 et d'élution avec EDTA (équivalent à 10 mL de culture). Après migration électrophorétique 1h à 150V, les protéines sont transférées sur une membrane nitrocellulose durant 1h à 100V et 4°C. La révélation est faite suivant le protocole d'ECL.

Finalement, nous avons décidé de tenter une nouvelle méthode de purification par RP-HPLC à partir de la fraction insoluble toute entière pour ne pas perdre la protéine exprimée de la fraction membranaire.

### 3.3.2.C. Clivage et purification par RP-HPLC

La fraction insoluble obtenue après extraction protéique est solubilisée dans l'acide formique puis l'eau est ajoutée pour avoir une concentration finale en acide formique de 75%. En effet, l'acide formique permet le clivage du partenaire de fusion, le TAF12, au niveau du site -Asp-Pro-. Le clivage se fait à 50°C durant 24h. La cinétique du clivage est suivie par prélèvement à différents temps et dépôt sur SDS-PAGE (ou western blot anti-His tag) et par injection sur RP-HPLC, Bischoff. Seuls les résultats obtenus pour le TAF12-gp120TM seront présentés dans les paragraphes suivants mais la même méthodologie a été appliquée aux trois polypeptides.

**A t=0 du clivage :** immédiatement après solubilisation du culot dans l'acide formique 75%, l'équivalent de 10 mL de culture est injecté sur la colonne analytique et ce pour les deux conditions induite et non induite). Les deux chromatogrammes (Voir Figure 40) présentent un profil similaire. Cependant, certains pics sont beaucoup plus intenses dans la condition induite. Les fractions correspondantes sont collectées afin de caractériser leur contenu par western blot et spectrométrie de masse (SM).

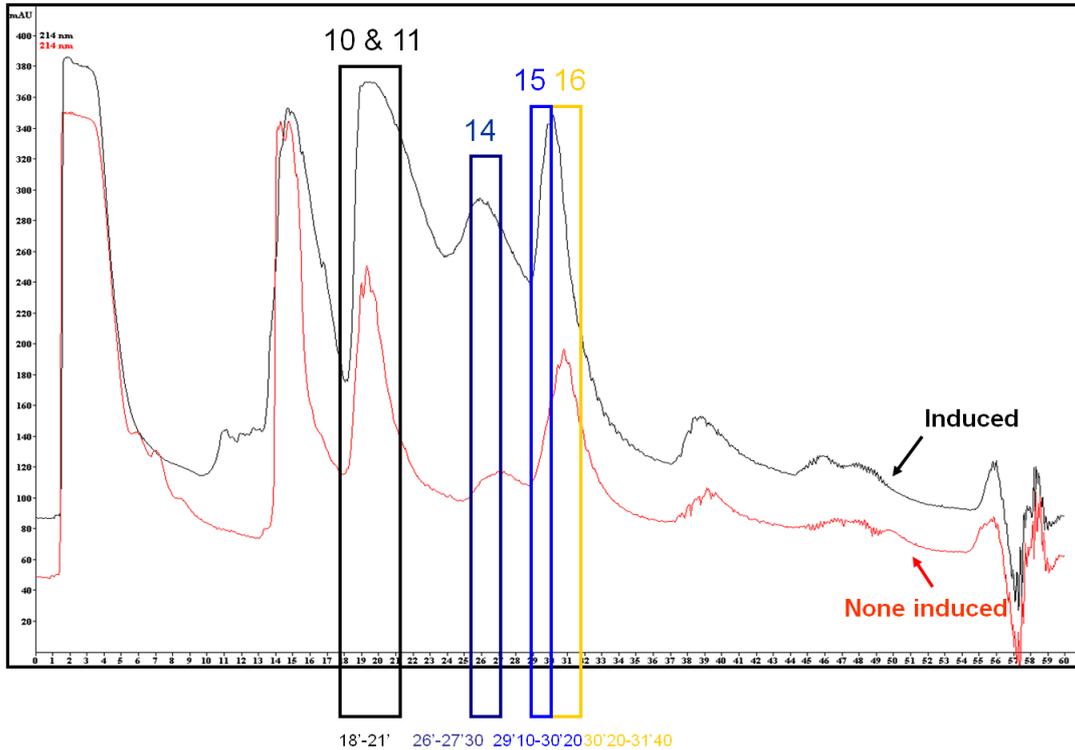
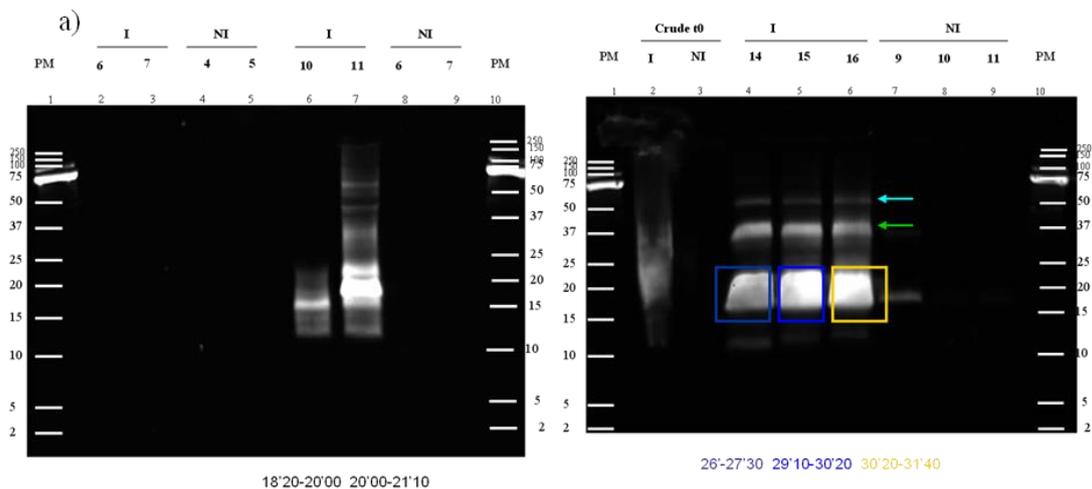


Figure 40 : Chromatogramme de TAF12-gp120TM à t=0 du clivage. Il montre le profil de l'éluion des protéines : absorbance à 214 nm en fonction du temps dans deux conditions induite (noir) et non induite (rouge). Les fractions collectées 10, 11, 14, 15 et 16 sont utilisées pour confirmer leur correspondance au TAF12-gp120TM par gels et spectrométrie de masse.



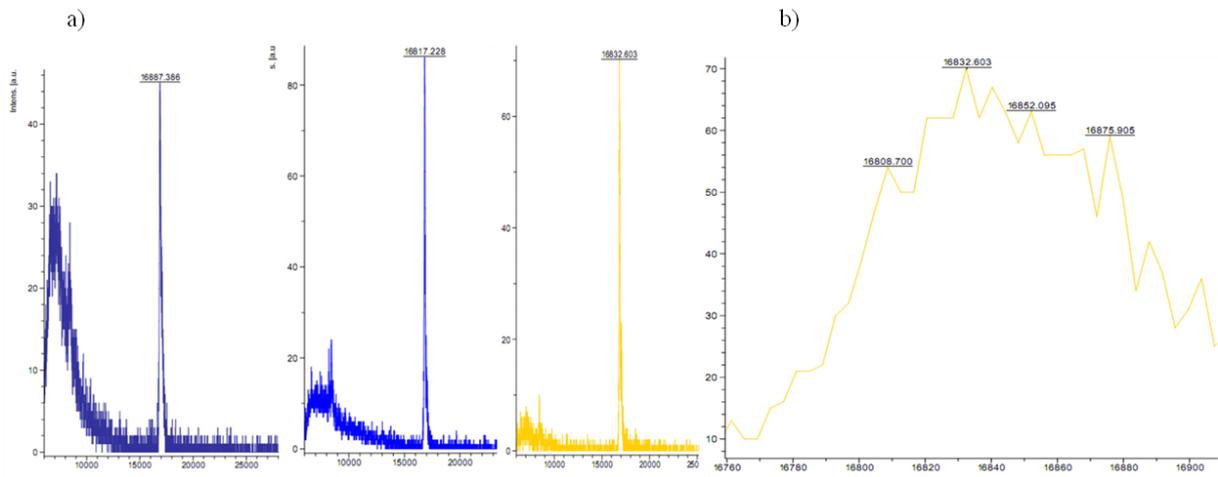
b)

	TAF12GP120TM
monomère	16807.4
dimère	33614.8
trimère	50422.2

Figure 41 : a) Western blot anti-His tag sur les fractions collectées à t=0 de clivage. Fraction 14 en bleu marine, fraction 15 bleu claire, fraction 16 en jaune. Les temps d'éluion des ces fractions sont présentés au-dessous. b) Masses théoriques des différentes formes de TAF12-

Le Western blot anti-His tag effectué sur les fractions collectées montre que les fractions 15 et 16 présentent un signal intense à la taille attendue pour le TAF12-gp120TM. Cette bande est également présente dans les fractions 10 et 11. Ceci veut dire que cette protéine peut s'adsorber sur la colonne de différentes manières. En outre, d'autres bandes migrant à 37 kDa et à 50 kDa sont visibles dans les fractions 14, 15 et 16. La protéine d'intérêt se trouve donc sous plusieurs formes. Le signal qui sort vers 37 kDa correspondant au dimère (en vert) et celui qui sort vers 50 kDa correspond au trimère (en cyan) du TAF12-gp120TM (voir Figure 41).

La confirmation par spectrométrie de masse a tout d'abord été faite par MALDI-TOF (Matrix-Assisted Laser Desorption/Ionization - Time Of Flight) puis par ESI Q-TOF (Electro Spray Ionisation Quadrupole-TOF) (voir respectivement Figure 43 et Figure 42). Pour le MALDI-TOF, plusieurs matrices ont été essayées afin de trouver celle qui convient aux nos protéines d'intérêt. Tout d'abord, l'acide  $\alpha$ -cyano-4-hydroxycinnamique, normalement utilisé pour les peptides, s'est révélé inadaptée aux tailles de ces protéines. Puis, l'acide sinapique a été testé et a donné de meilleurs résultats. Les résultats obtenus avec le MALDI-TOF présente des pics très larges ne permettant pas une détermination précise de la masse de l'échantillon. Cette technique est effectivement limitée par la taille des biomolécules testées (<3000 masse-to-charge, m/z) (Trauger *et al.*, 2002). Le recours à l'ESI Q-TOF a donc été nécessaire. Le calcul des masses exactes est présenté dans la Figure 42.

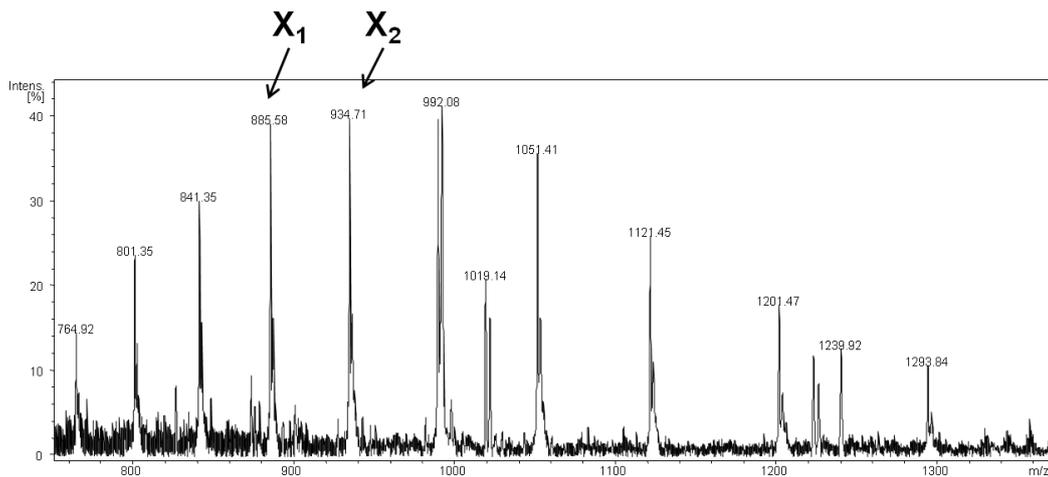


**Figure 43 : Spectre de masse MALDI-TOF des fractions collectés à t=0 de clivage. a) Fraction 14 en bleu marine, fraction 15 bleu clair, fraction 16 en jaune. b) L'agrandissement du spectre de masse de la fraction 16.**

$$X_1 = \frac{M + z_1}{z_1} = 885,58$$

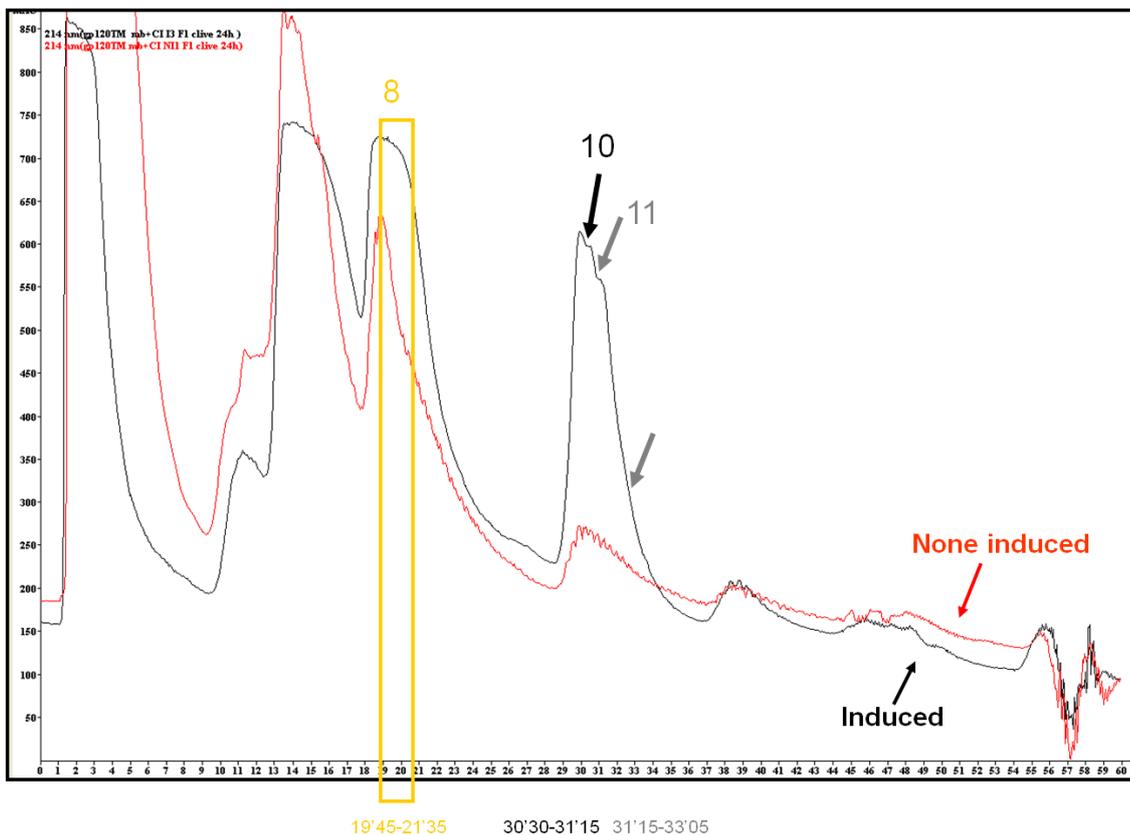
$$X_2 = \frac{M + z_2}{z_2} = 934,71$$

$$z_1 = \frac{X_2 - 1}{X_2 - X_1} = +19 \rightarrow M = 16807,2 \text{ Da}$$



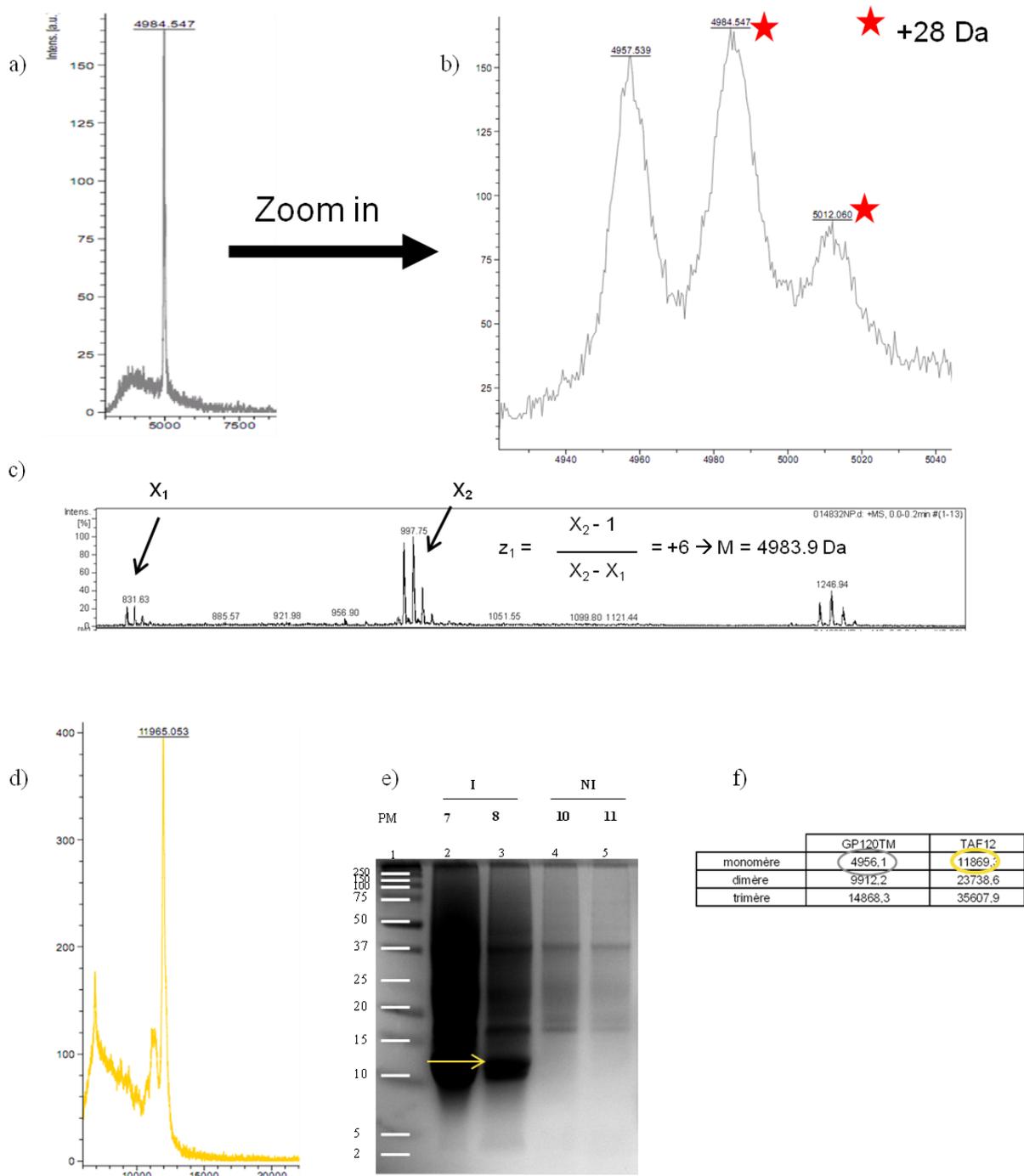
**Figure 42 : Spectre de masse ESI Q-TOF de la fraction 16 à t=0 de clivage. La masse moléculaire (M, en Daltons) de la protéine est déterminée par les formules détaillées au-dessus du spectre de la masse. Les X correspondent aux rapports (m/z).**

**A t=24h du clivage :** après 24h d'incubation à 50°C de la solution de clivage, l'équivalent de 10 mL de culture est injecté sur la colonne analytique et ce pour les deux conditions, induite et non induite (voir Figure 44). En comparant, les deux chromatogrammes nous remarquons que des pics supplémentaires, absents dans la condition non induite, apparaissent dans la condition induite, (fraction 10 et 11) Les fractions correspondantes ont donc été collectées pour caractériser leur contenu par gels et SM.



**Figure 44 :** Chromatogramme de TAF12 / gp120TM à t=24h du clivage. Il montre le profil de l'éluion des protéines : absorbance à 214 nm en fonction du temps dans deux conditions induite (noir) et non induite (rouge). Les fractions collectées 8, 10 et 11 sont utilisées pour confirmer leur correspondance au TAF12 ou gp120TM par gels et spectrométrie de masse.

La Figure 45 (a et b) qui correspond au spectre de masse de la fraction 11 montre une masse proche de celle de la gp120<sup>TM</sup> (polypeptide clivé, voir Figure 45.f) mais en agrandissant ce spectre nous distinguons que ce polypeptide se trouve sous la forme de trois populations : une présentant la masse exacte attendue, et deux autres présentant respectivement un ajout de 28 Da et 56 Da. Le spectre MALDI-TOF du polypeptide clivé montre des masses plus précises que celui du non clivé, ce qui s'explique par la différence de masse. Nous avons tout de même passé ces fractions sur l'ESI Q-TOF, qui a montré des résultats identiques à ceux obtenus par MALDI-TOF (Figure 45.c). L'analyse de la fraction 8 montre la présence de TAF12 (Figure 45. d, e et f). Comme attendue, cette fraction est éluée avant celle contenant la gp120<sup>TM</sup> puisque le TAF12 est moins hydrophobe que la gp120<sup>TM</sup>. Ces résultats ont été également confirmés par MALDI-TOF et SDS-PAGE qui montre une bande intense aux alentours de 11000 Da). Comme la gp120<sup>TM</sup>, le TAF12 présente aussi des formylations.



**Figure 45 : Identification de la fraction d'éluion contenant la gp120TM par spectrométrie de masse et SDS-PAGE. a) Spectre de MALDI-TOF de la fraction 11 (en gris). b) l'agrandissement de ce spectre de masse, les étoiles rouges correspondent aux ajouts de 28 Da. c) Spectre ESI Q-TOF de la fraction 11. d) Spectre de MALDI-TOF de la fraction 8 (en jaune). e) SDS-PAGE montrant le profil de migration de la fraction 8. f) Tableau montrant les masses théoriques de la gp120TM et du TAF12 sous leurs trois formes.**

La réaction de formylation consiste en un ajout d'un groupement formyle (-C(O)-H). Cette réaction peut avoir lieu naturellement chez les cellules procaryotes. Elle est assurée par une enzyme, la transformylase, qui transfère un groupement formyl à la méthionine durant l'initialisation de la biosynthèse des protéines (Kozak, 1983). Il s'agit alors d'une réaction de N-formylation qui peut aussi atteindre la lysine (Wisniewski *et al.*, 2008). Il faut ajouter aussi l'acide formique, lui-même, peut provoquer une telle réaction. Il s'agit d'O-formylation de la sérine et de la thréonine (Hwang *et al.*, 2013). En observant les spectres de masses obtenus au temps t=0 du clivage, nous remarquons que les produits présentent des formylations, alors qu'ils n'ont été en contact que très peu de temps avec l'acide formique : le temps de la solubilisation des fractions insolubles suivi de l'injection sur HPLC est inférieur à 30 min. De plus, lors de cette étape les échantillons ont été maintenus à 4°C. La formylation est donc une réaction rapide même à basse température. Pour éviter cette réaction indésirable, nous avons tenté de cliver la protéine d'intérêt par une solution concentrée d'acide chlorhydrique (HCl) (Gavit and Better, 2000; Wang *et al.*, 2002) . Mais ce dernier n'est pas parvenu à solubiliser la fraction insoluble.

Concernant la purification du TAF12-gp41TM, elle se fait de la même manière que le TAF12-gp120TM. Par contre, celle de TAF12-fpTM nécessite l'ajout d'une étape supplémentaire. La fraction insoluble doit être purifiée par RP-HPLC avant d'être clivée. Ainsi, le culot (membrane + corps d'inclusion) est solubilisé dans l'acide formique puis directement injecté sur HPLC. Les fractions collectées contenant le peptide de fusion TAF12-fpTM sont poolées, évaporées sur le Rotavapor et resolubilisées dans la solution de clivage acide formique 75% suivie d'une incubation à 50°C pendant 24h. Le polypeptide clivé est alors à nouveau purifié par l'HPLC.

Le protocole de surexpression et de purification pour les trois polypeptides pour assurer des quantités suffisantes pour étudier leur structure et leur immunogénicité était établi. En effet, deux fermenteurs de 4 L, au moins, sont lancés pour produire un seul polypeptide. Chaque fermenteur donne une concentration finale en polypeptides de 3 mg/L de culture bactérienne.

### 3.3.3. Test d'immunogénicité des polypeptides

Deux séries d'immunisation sur des souris ont été faites, l'une pour gp120<sup>TM</sup> et l'autre pour fp<sup>TM</sup>. La quantification de la réponse immunitaire humorale (anticorps) était faite par ELISA. Les résultats obtenus en utilisant le fp<sup>TM</sup> comme antigène inséré dans les liposomes ont donné des résultats contradictoires. En fait, l'adjuvant utilisé peut être à la base de ces résultats, il pourrait détruire les liposomes grâce à sa nature huileuse. De plus, durant la préparation des échantillons, nous avons observé des agrégats ce qui nous a amené à retrouver une autre méthode de reconstitution.

### 3.4. Conclusion et perspectives

Le but du travail décrit dans ce chapitre était d'établir un protocole de production, en quantités suffisantes pour les études structurales et immunogéniques, des trois polypeptides contenant des épitopes ayant la capacité d'induire la production des anticorps neutralisants chez quelques patients infectés par le VIH-1. Ces séquences sont exprimées sous forme de protéine de fusion dont le partenaire est le TAF12 en extrémité N-terminale. Ce dernier réduit la cytotoxicité des protéines d'intérêt en les dirigeant dans la fraction insoluble. De même, il facilite leur purification grâce à son His tag. En extrémité C-terminale des épitopes, nous trouvons une séquence transmembranaire issue du peptide MS1 qui permet la trimérisation de ces séquences pour bien mimer la conformation native des glycoprotéines virales.

Les conditions d'expression et de purification des trois séquences sont bien établies (Tableau 5). Elles sont majoritairement localisées dans la fraction insoluble (membranes et corps d'inclusion). Cette fraction est solubilisée directement dans l'acide formique 75%, pour les peptides TAF-gp120<sup>TM</sup> et TAF12-gp41<sup>TM</sup>, afin de cliver la protéine de fusion TAF12 au niveau du site de clivage Asp/Pro situé juste en amont du polypeptide d'intérêt. Pour le peptide TAF12-fp<sup>TM</sup>, une étape de purification de la fraction insoluble par RP-HPLC, préalablement au clivage est nécessaire. Les fractions collectées comportant le peptide TAF12-fp<sup>TM</sup> sont alors clivées dans l'acide formique. Les polypeptides clivés sont purifiés par RP-HPLC. Le contrôle par spectrométrie de masse des fractions d'intérêt montrent pour quelques pics un ajout d'un multiple de 28 Da à la masse moléculaire attendue; ceci indique que une partie des polypeptides purifiés sont formylés (plusieurs états de formylation). Afin d'éviter cette réaction, un clivage avec l'acide chlorhydrique a été testé mais sans succès, l'acide chlorhydrique n'a pas permis de solubiliser correctement la fraction insoluble et le clivage n'était pas spécifique. La quantité du polypeptide obtenu est estimée 3 mg/L de culture, par micro BCA (Bicinchoninic Acid Assay).

Les perspectives à envisager consistent à améliorer le clivage des protéines d'intérêt pour avoir des protéines non formylées. D'autres conditions plus douces et permettant de conserver un taux de clivage raisonnable peuvent être envisagées, par exemple, une diminution de la concentration finale en acide formique, ou bien le remplacement de l'acide formique par l'acide trifluoroacétique (Hwang *et al.*, 2013). Cependant, il faut éviter le clivage enzymatique qui nécessite un milieu aqueux.

	<b>Température d'induction (°C)</b>	<b>DO d'induction</b>	<b>Purification</b>	<b>Concentration finale (mg/L de culture)</b>	<b>Formylations</b>
<b>gp120TM</b>	37	1	1x RP-HPLC	3	0, 1 et 2x
<b>gp41TM</b>	28	1,7	1x RP-HPLC	3	0, 1 et 2x
<b>fpTM</b>	37	1,2	2x RP-HPLC	3	0, 1 et 2x

**Tableau 5 : Récapitulatif des conditions d'expression et de purification des trois séquences.**

# 4. Chapitre 4 : Étude biophysique de gp120TM, gp41TM et fpTM

## 4.1. Introduction

De nos jours, les études se focalisent de plus en plus sur la structure native des glycoprotéines virales. Récemment, des structures des gp tronquées ont été élucidées. Il s'agit des trimères de gp140 formés par la gp120 et l'ectodomaine de la gp41. L'étude de la MPER nécessite d'utiliser des formes tronquées car du fait de son rôle dans la trimérisation et de sa position juxta-transmembranaire, elle se présente sous sa forme entière, est tronquée car, grâce à son rôle, elle rend les trimères insolubles et provoque leurs agrégation (Khayat *et al.*, 2013; Klasse *et al.*, 2013). De plus, cette région a une plasticité structurale vis-à-vis des anticorps qui s'y lient. Par exemple, l'épitope de l'anticorps neutralisant 2F5 se structure en hélice  $3_{10}$  dans un milieu aqueux (Biron *et al.*, 2002). Par contre, en se liant à cet anticorps, la MPER se structure en feuillet  $\beta$  (Ofek *et al.*, 2004b). Elle forme une hélice  $\alpha$  avec l'anticorps non neutralisant 13H11 (Nicely *et al.*, 2010). Dans un milieu micellaire, la MPER adopte une structure en « L » formée par deux hélices  $\alpha$  liées par une coude (Sun *et al.*, 2008).

Cette région interagit avec le peptide de fusion dont la structure reste toujours sous controverse. En fait, cette structure dépend de plusieurs facteurs : le ratio peptide/lipide, les lipides utilisés, la présence du cholestérol (Gabrys *et al.*, 2013). De plus, ce peptide a une tendance à s'oligomériser au contact des membranes modèles (Yang *et al.*, 2004). Probablement, ce peptide forme des oligomères en feuillet  $\beta$  antiparallèles mais en absence du cholestérol il y a apparition de monomères en hélice  $\alpha$ .

La gp120, elle aussi, participe à cette interaction au niveau de 1<sup>er</sup> domaine constant (C1) qui se trouve à l'extrémité N-terminale de cette glycoprotéine. Ce domaine se structure principalement en feuillet  $\beta$  (Pancera *et al.*, 2010).

C'est pour ces raisons que nous nous sommes intéressés à ces régions. De fait, le polypeptide gp41TM est issu de la séquence de MPER contenant les épitopes induisant la production des anticorps (4E10 et 2F5) qui neutralisent un large spectre de souches virales. Le polypeptide gp120TM, lui, contient le domaine C1 de la gp120 et enfin, le polypeptide fpTM correspond au peptide de fusion du VIH.

Au vue de ce polymorphisme structural, nous avons voulu voir comment ces polypeptides se structurent dans différents environnements lipidiques. Or, leur caractère hydrophobe et leur capacité à former des multimères limitent leur solubilité, ce qui nous a amené à tester plusieurs méthodes de reconstitution. Afin d'avoir un contrôle dans cette reconstitution, la séquence TM a été synthétisée chimiquement.

## 4.2. Étude par dichroïsme circulaire

### 4.2.1. Matériels et méthodes

Les solvants organiques proviennent du Sigma Aldrich (Saint Louis, MO) avec une pureté de 99% ; POPC, POPS et cholestérol de Avanti Polar Lipids ; Birmingham, AL, les acides aminés de protégés par F-moc de Bachem, Heidelberg, Allemagne ou d'Applied Biosystems, Weiterstadt, Allemagne.

#### 4.2.1.A. Synthèse peptidique

La séquence TM (QKLIAAVLLLLIA<sup>(4H)</sup>SNL<sup>(15N)</sup>ILLIANARLRYL-CO-NH<sub>2</sub>) a été conçue par synthèse peptidique en phase solide sur un synthétiseur automatique Millipore 9050, et assemblée sur une résine tentagel-R-RAM (RAPP Polymere, GmbH, Tubingen, Allemagne) en utilisant des cycles standards de synthèse avec un excès de quatre fois en acides aminés (A, L, K et H) protégés par un groupement Fmoc (Merck, Darmstadt, Allemagne). Une fois la synthèse terminée, la résine a été placée sur un verre frité et lavée avec de l'éthanol, de l'acide acétique et de l'éther. Le clivage de la TM a été réalisé dans une solution contenant 28 mL de TFA, 1,5 mL d'eau et 300 µL de triéthylsilane sous agitation pendant 4 heures. Le peptide a été ensuite repris dans une solution d'acide tri-fluoro-acétique (TFA)/toluène (50/50) (v/v). Le solvant a été évaporé à l'aide d'un évaporateur rotatif (Büchi R-200) à 37°C. Le peptide a ensuite été solubilisé dans de l'acide acétique 100% et précipité dans l'éther. Les tubes ont été centrifugés à 11000 g pendant 30 minutes à 4°C. Le culot contenant le peptide a été finalement solubilisé dans de l'acide acétique 4 %, congelé dans de l'azote liquide et mis à lyophiliser. Le produit clivé a été analysé par MALDI-TOF pour vérifier la présence de la TM. Elle a été purifiée par RP-HPLC à l'aide d'un gradient d'acétonitrile/eau sur une colonne C18 (Luna 5u c18(2) 100A ; 150x30 mm ; Phenomenex, 547606-2).

### 4.2.1.B. Reconstitution des polypeptides

La membrane plasmique sépare les organites cellulaires du milieu extracellulaire. Elle est formée par des lipides (phospholipides, cholestérol, glycolipides) et par des protéines. Pour que ces dernières accomplissent leur fonction, elles doivent adopter une structure bien déterminée qui est modulée par les interactions protéines-lipides. Or, l'étude des protéines membranaires *in vivo* est difficile car elles sont associées à des mélanges complexes d'autres protéines et elles ont la tendance à s'agréger en solution. C'est pourquoi nous avons préféré des études *in vitro* avec des modèles de membranes parmi lesquels, on peut citer les bicouches lipidiques, les liposomes (Shen *et al.*, 2013)...

#### 4.2.1.B.a. Par le biais des solvants organiques

A titre d'exemple, nous prenons la reconstitution de la TM. Ce polypeptide n'est pas soluble dans beaucoup de solvants qui sont souvent utilisés pour les polypeptides membranaires (2,2,2-trifluoro-éthanol (TFE), Hexafluoro-2-propanol...) mais un système formé de chloroforme/méthanol (V/V = 2/1) chauffé à ~40°C, a pu être identifié pour solubiliser 1 mg/mL de TM. Les lipides (POPC : 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine et POPS : 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phospho-L-serine ; chez Avanti Polar Lipids) dont leur quantité est calculée en fonction du ratio molaire peptide/lipide (mol/mol) sont solubilisés en parallèle dans le mélange chloroforme/méthanol (V/V = 2/1). La TM et les lipides solubilisés sont ensuite mélangés. Les solvants sont évaporés sous un flux doux d'azote. Les dernières traces de solvants sont éliminées par lyophilisation durant la nuit. Le film obtenu est hydraté, pendant une heure, sous forte agitation à température ambiante, avec le volume approprié de tampon phosphate 10 mM pH7 afin d'obtenir la concentration peptide/lipide désirée. Pour le DC, la concentration idéale est de 0,1 – 0,2 mg/mL de peptide. Pour avoir des « Small Unilamellar Vesicles » (SUV), une étape de sonication dans un bain de glace est nécessaire jusqu'à l'obtention d'une solution translucide. Les agrégats et les morceaux du titane qui peuvent provenir de la sonde de sonication sont éliminés par une centrifugation durant 30 min à 11000 g à température ambiante.

#### 4.2.1.B.b. Par le biais des détergents

Les polypeptides exprimés (gp120<sup>TM</sup>, gp41<sup>TM</sup> et fp<sup>TM</sup>) ne sont pas solubles dans le SDS mais il a été possible de les solubiliser à 1 mg/mL dans le n-dodecylphosphocholine (DPC) 5-10 mM qui est supérieure à sa concentration micellaire critique (CMC = 1 mM). La quantité des lipides appropriée est solubilisée dans du chloroforme qui est par la suite évaporé sous un flux d'azote puis par lyophilisation durant la nuit (pour le DC la concentration idéale est de 0,1 – 0,2 mg/mL de peptide). Le film lipidique obtenu est hydraté par le volume approprié de la solution protéique, pendant une heure sous forte agitation à température ambiante. Le détergent est éliminé par dialyse dans des cassettes (Slide-A-Lyzer Dialysis Cassettes, 3.5K MWCO et 0,5mL, Thermo Scientific). Le volume du tampon phosphate utilisé est équivalent à 1000 fois le volume de l'échantillon. Les agrégats sont éliminés par une centrifugation durant 30 min à 11000 g à RT.

#### 4.2.1.C. Dichroïsme circulaire

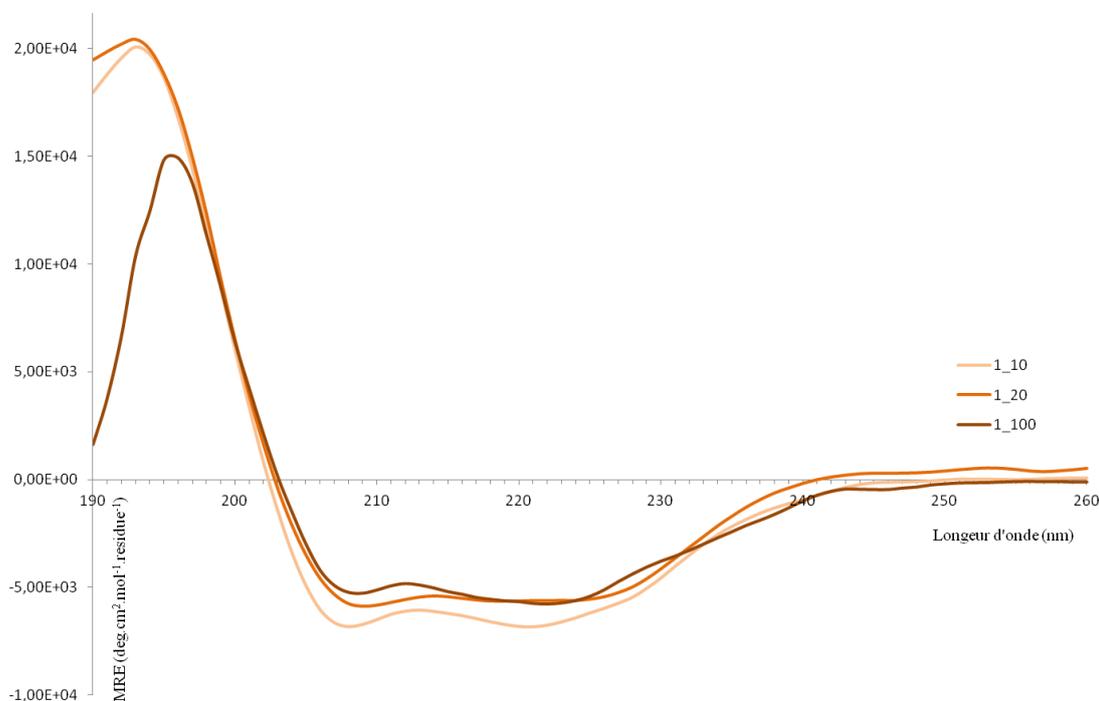
Les spectres du dichroïsme circulaire ont été enregistrés sur un spectrophotomètre J-510 de Jasco à 20°C avec une fenêtre spectrale de 190-260 nm, une bande spectrale (bandwidth) de 3 nm, vitesse de scan 1 nm/min et 10 scans. Les traitements de ces spectres consistent en une soustraction de la contribution des vésicules présentes dans le tampon utilisé. Les spectres obtenus et analysés correspondent alors à la seule contribution des polypeptides. L'estimation des contributions des différentes structures secondaires est faite par Dichroweb (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>). Deux algorithmes ont été utilisés CONTIN/LL (une variante de CONTIN) et CDSSTR avec le set référence SP175 (Whitmore and Wallace, 2004, 2008).

## 4.2.2. Résultats

### 4.2.2.A. Structure secondaire de la TM

La TM est reconstituée dans des liposomes de POCP/POPS (3/1) avec une concentration finale de 0,1 mg/mL et avec plusieurs ratio molaires (peptide/lipides, P/L = 1/10, 1/20 et 1/100).

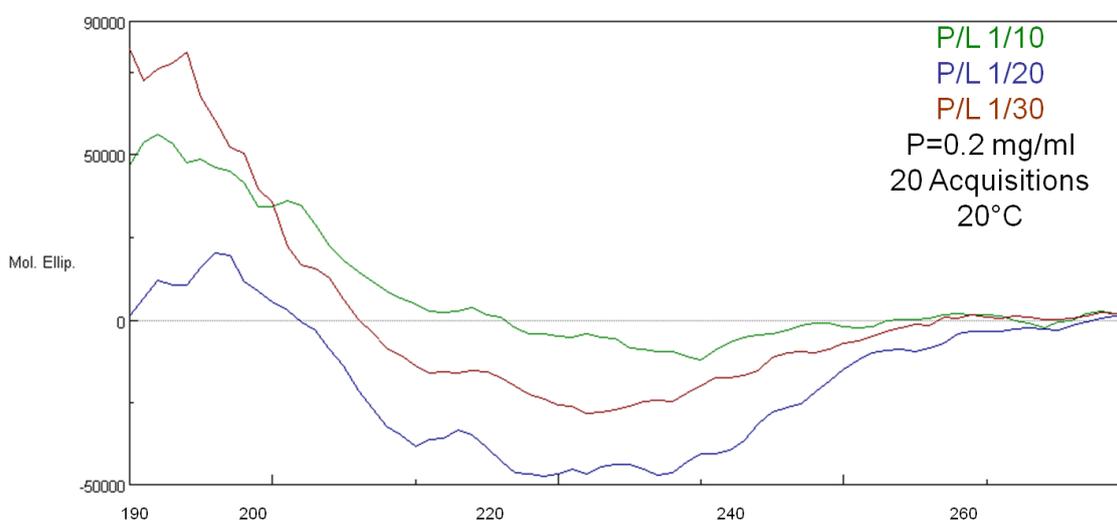
La Figure 46 montre que le profil des spectres du DC de la TM est similaire quel que soit le ratio molaire peptide/lipide utilisé. Il présente un pic positif vers 195 nm et deux minima, vers 208 et 222 nm, qui ont approximativement les mêmes intensités. Ces valeurs de longueur d'onde pour les maxima et minima sont caractéristiques d'une hélice  $\alpha$ . Par contre, avec un ratio élevé de P/L (1/100), l'intensité du pic positif à 195 nm est inférieure à celle des ratios 1/10 et 1/20. Ceci peut s'expliquer par le fait que les lipides diffusent fortement la lumière dans cette zone surtout à une concentration élevée alors ils perturbent le signal de la TM.



**Figure 46 : Spectre du dichroïsme circulaire de la TM insérée dans des SUV de POCP/POPS (3/1). [TM] = 0,1 mg/mL dans un tampon phosphate 10 mM pH =7. P/L (mol/mol) = 1/10 en orange, 1/20 en orange foncé et 1/100 en marron.**

#### 4.2.2.B. Structure secondaire des polypeptides

Les échantillons des polypeptides (gp120<sup>TM</sup>, gp41<sup>TM</sup> et fp<sup>TM</sup>) préparés par le biais des solvants organiques ont donné des spectres DC inexploitable (par exemple le spectre de la gp120<sup>TM</sup> est montré dans la Figure 47). Bien que les trois polypeptides se soient bien solubilisés dans le mélange isopropanol/eau (50/50), ils n'ont pas été probablement capables de s'insérer correctement dans vésicules lors de l'ajout des lipides. Ce qui nous a amené à modifier la préparation des échantillons en utilisant l'approche de la dialyse. Le premier détergent utilisé a été SDS. Mais ce dernier n'a pas permis la solubilisation de la gp120<sup>TM</sup>. Un autre détergent a donc été utilisé, c'est le DPC qui a permis de solubiliser les polypeptides. En fait, ce détergent zwitterionique, contrairement au SDS qui est chargé négativement, a une tête polaire similaire à celle des phosphatidylcholines, les composants majeurs de la membrane plasmique des cellules eucaryotes (Manzo *et al.*, 2013). Et n'oublions pas le fait que l'enveloppe virale est issue de la membrane plasmique de la cellule infectée avec une forte contribution de phosphatidylcholines.

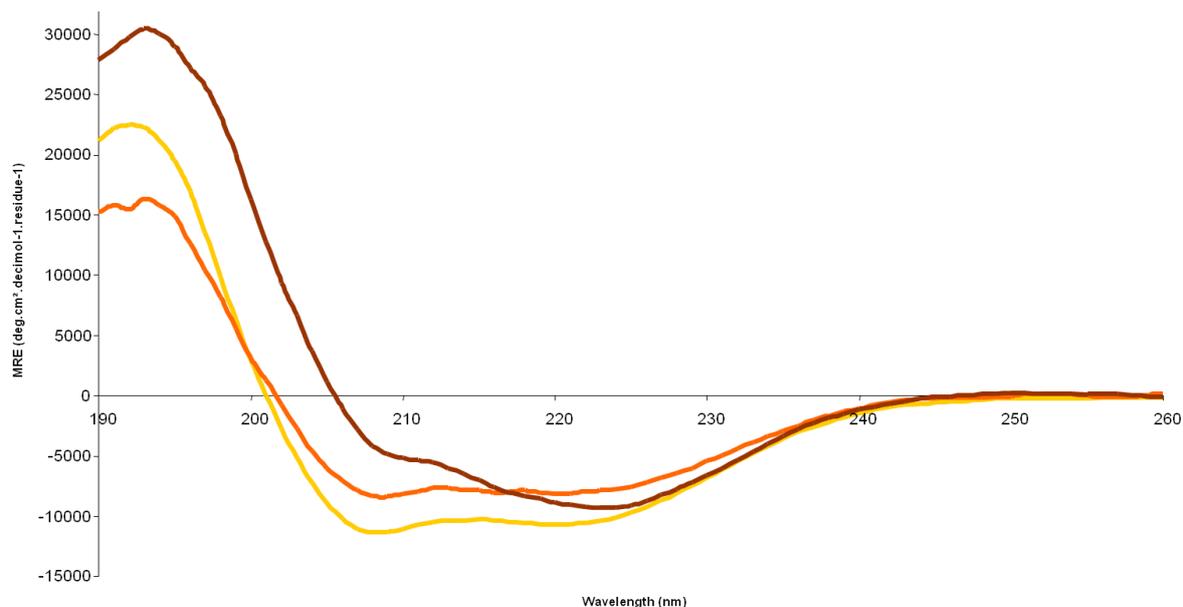


#### 4.2.2.B.a. Structure secondaire de la gp120<sup>TM</sup>

La structure secondaire de la gp120<sup>TM</sup> a été étudiée, tout d'abord, dans un milieu micellaire (DPC 5mM) avec une concentration finale de 0,05mg/mL ; puis, insérée, par voie d'échange avec le détergent, dans deux types de liposomes formés par:

- POPC/POPS (3/1) avec [gp120<sup>TM</sup>] = 0,1 mg/mL
- POPC/ SM/ DHSM/ PE-pLPE/ PS/ Chol (9/12,5/5,75/19,25/8,45/45), avec [gp120<sup>TM</sup>] = 0,1 mg/mL (une composition similaire à celle de la membrane virale (Brügger *et al.*, 2006))

Le spectre DC de la gp120<sup>TM</sup> (Figure 48) solubilisée dans du DPC 5 mM possède un pic positif vers 195 nm et deux minima ayant à peu près la même intensité vers 208 et 222 nm. Ceci indique que ce polypeptide se structure majoritairement en hélice  $\alpha$  dans le DPC. Par contre, dans les liposomes, surtout ceux qui ont une composition similaire à celle de la membrane virale, le spectre DC de la gp120<sup>TM</sup>, montre un profil différent. Ceci indique que ce polypeptide se structure autrement que dans le DPC (Figure 48). D'après les calculs des contributions des structures secondaires (Tableau 6), nous remarquons que la gp120<sup>TM</sup> perd son caractère majoritaire d'hélice au détriment du caractère en feuillet  $\beta$  en contact avec les lipides.



**Figure 48 :** Spectre du dichroïsme circulaire de la gp120TM. Les données spectrales ont été normalisées en termes de concentration de la gp120TM et du nombre d'acides aminés (MRE : Molar Residue Ellipticity). La gp120TM dans le DPC 5 mM est en jaune, insérée dans POPC/POPS (3/1) avec P/L = 1/70 en orange et dans des membranes ayant une composition similaire à la membrane virale avec P/L = 1/100 en marron.

<b>gp120TM</b>			
	<b>DPC 5mM</b>	<b>PCPS</b>	<b>Vrl mb</b>
<b>CONTIN (set 7: sp175)</b>			
NRMSD	0,021	0,14	0,1
alpha	97	12	23
beta		38	40
T	3	12	9
U		37	28
<b>CDSSTR (set 7 : sp175)</b>			
NRMSD	0,001	0,03	0,015
alpha	92	36	19
beta	3	21	39
T	3	11	9
U	2	32	33

**Tableau 6 :** Analyse des structures secondaires de la gp120TM. Les contributions des structures secondaires de la gp120TM (dans le DPC 5 mM ou insérée dans des liposomes de POPC/POPS (PCPS) ou Viral-Like membrane (Vrl mb) sont calculées sur Dichroweb par deux algorithmes CONTIN et CDSSTR en utilisant le set de référence sp175. NRMSD: Normalized Root-Mean-Square Deviation, alpha: hélice  $\alpha$ , beta: Feuillet  $\beta$ , T: Turn (coude),

#### 4.2.2.B.b. Structure secondaire de la gp41<sup>TM</sup>

La structure secondaire de la gp41<sup>TM</sup> a été étudiée, tout d'abord, dans un milieu micellaire (DPC 5mM) avec une concentration finale de 0,27 mg/mL ; puis, insérée, par voie d'échange avec le détergent, dans trois types de liposomes formés par :

- POPC/POPS (3/1) avec [gp41<sup>TM</sup>] = 0,05 mg/mL
- POPC/POPS (3/1)/Chol 30%, avec [gp41<sup>TM</sup>] = 0,1 mg/mL
- POPC/ SM/ DHSM/ PE-pLPE/ PS/ Chol (9/12,5/5,75/19,25/8,45/45), avec [gp41<sup>TM</sup>] = 0,05 mg/mL (une composition similaire à celle des membranes virale)

La Figure 49 montre que le spectre DC de la gp41<sup>TM</sup> solubilisée dans du DPC 5 mM présente un pic positif vers 195 nm et deux minima vers 208 et 222 nm. Par contre, le minimum à 208 nm est plus intense que celui du 222 nm. Ce qui signifie que la structure de la gp41<sup>TM</sup> tend vers une hélice  $\alpha$  dans DPC mais il se peut que d'autres structures secondaires contribuent. Une fois insérée dans les liposomes, les spectres DC conservent la même allure sauf que les minima, à 208 et 222 nm, changent d'intensité indiquant que le caractère d'hélice change d'un environnement lipidique à un autre. D'après les calculs des contributions des structures secondaires, nous remarquons que la gp41<sup>TM</sup> se structure majoritairement en hélice  $\alpha$  quel que soit son environnement mais que ce caractère est renforcé dans les liposomes ayant une composition similaire à celle de la membrane virale (Tableau 7).

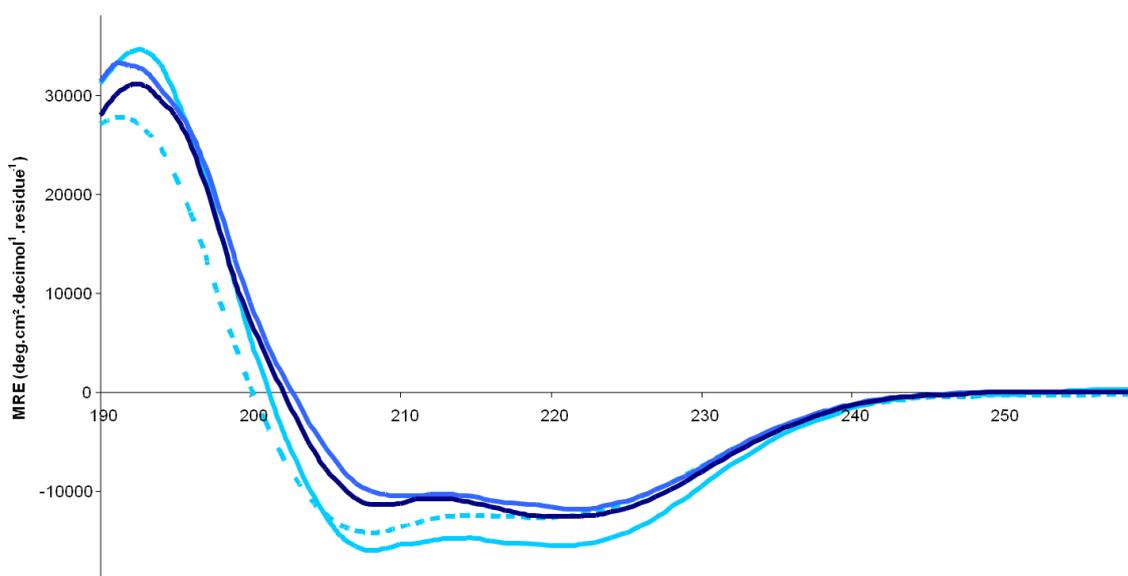


Figure 49 : Spectre du dichroïsme circulaire de la gp41TM. Les données spectrales ont été normalisées en termes de concentration de la gp41TM et du nombre d'acides aminés. La gp41TM dans le DPC 5 mM est en bleu pointillé, insérée dans POPC/POPS (3/1) avec P/L = 1/70 en bleu, dans POPC/POPS (3/1)/Chol 30% avec P/L = 1/70 en bleu foncé et dans des membranes ayant une composition similaire à la membrane virale avec P/L = 1/100 en bleu

gp41TM				
	DPC 5mM	PCPS	PCPSchol	Vrl mb
<b>CONTIN (set 7: sp175)</b>				
NRMSD	0,094	0,1	0,137	0,07
alpha	39	43	40	60
beta	20	16,3	22	8
T	15	10	10	10
U	24	30	27	21
<b>CDSSTR (set 7 : sp175)</b>				
NRMSD	0,01	0,021	0,01	0,01
alpha	38	43	35	54
beta	18	19	27	15
T	12	10	9	8
U	31	28	29	22

Tableau 7 : Analyse des structures secondaires de la gp41TM. Les contributions des structures secondaires de la gp41TM (dans le DPC 5 mM ou insérée dans des liposomes de POPC/POPS (3/1) (PCPS), POPC/POPS (3/1)Chol 30% (PCPSchol) ou Viral-Like membrane (Vrl)) sont calculées sur Dichroweb par deux algorithmes CONTIN et CDSSTR en utilisant le set de référence sp175. NRMSD: Normalized Root-Mean-Square Deviation, alpha: hélice  $\alpha$ , beta: Feuillet  $\beta$ , T: Turn (coude), U: Unordered (non structuré).

#### 4.2.2.B.c. Structure secondaire du fpTM

La structure secondaire du fpTM a été étudiée, tout d'abord, dans un milieu micellaire (DPC 10 mM) avec une concentration finale de 0,06 mg/mL ; puis insérée, par voie d'échange avec le détergent, dans deux types de liposomes formés par :

- POPC/POPS (3/1) avec [fpTM] = 0,05 mg/mL
- POPC/ SM/ DHSM/ PE-pLPE/ PS/ Chol (9/12,5/5,75/19,25/8,45/45), avec [fpTM] = 0,1 mg/mL (une composition similaire à celle de la membrane virale)

L'allure du spectre DC du fpTM solubilisée dans du DPC 10 mM montre un maxima vers 195 nm et deux minima vers 208 et 222 nm. Le minimum à 208 nm est plus intense que celui du 222 nm. Ce qui signifie que la structure du fpTM tend aussi vers une hélice  $\alpha$  dans le DPC et que d'autres structures secondaires peuvent également contribuer. Par contre, l'insertion de ce polypeptide dans les liposomes provoque des changements drastiques au niveau des spectres DC. Ces changements correspondent, selon les calculs des structures secondaires (Tableau 8), à une déstructuration du fpTM d'une part et de l'apparition des feuillets  $\beta$  d'autre part.

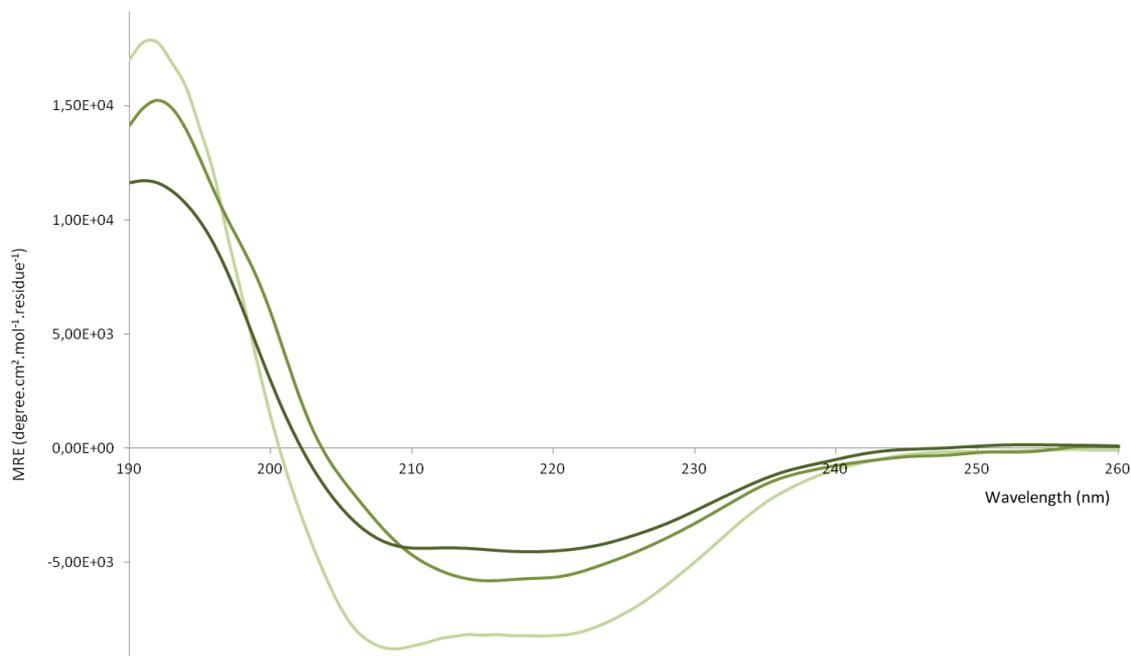


Figure 50 : Spectre dichroïsme circulaire du fpTM. Les données spectrales ont été normalisées en termes de concentration du fpTM et du nombre d'acides aminés. Le fpTM dans du DPC 10 mM est en vert clair, insérée dans POPC/POPS (3/1) avec P/L = 1/70 en vert et dans des membranes ayant une composition similaire à la membrane virale avec P/L = 1/100 en vert

fpTM	CONTIN		
	DPC	POPS/POPS	VM
NRMSD	0,032	0,134	0,097
A	100	12	12
B		42	40
T		11	12
U		35	36
	CDSSTR		
	DPC	POPS/POPS	VM
NRMSD	0,001	0,048	0,053
A	88	4	4
B	5	46	44
T	4	10	10
U	4	40	41

Tableau 8 : Analyse des structures secondaires du fpTM. Les contributions des structures secondaires du fpTM DPC 10 mM sont calculées sur Dichroweb par deux algorithmes CONTIN et CDSSTR en utilisant le set de référence sp175. NRMSD: Normalized Root-Mean-Square Deviation, alpha: hélice  $\alpha$ , beta: Feuillet  $\beta$ , T: Turn (coude), U: Unordered (non

### 4.2.3. Discussion

Les trois polypeptides se structurent en hélice  $\alpha$  principalement dans le DPC. La gp41TM conserve et renforce cette structure dans les liposomes qui ressemblent à la membrane virale. Par contre, la structure des gp120TM et fpTM tend à se changer dans les liposomes. Cette différence de comportement entre les différents polypeptides peut être basée simplement sur le fait que la MPER qui fait partie de la gp41TM est naturellement en contact avec la membrane lipidique et sa structure dépend d'autres facteurs (anticorps, gp120...). Cependant, le domaine C1 (gp120TM) et le peptide de fusion (fpTM) qui sont à la base localisés respectivement à la surface de la membrane virale ou caché dans le spike deviennent de plus en plus proches de cette membrane. Ce rapprochement est dû au changement de conformation en cascade subi par le trimère.

Les analyses des spectres DC donnent une idée générale sur la structure secondaire et ne permettent pas de savoir quels sont les résidus qui contribuent à telle ou telle structure. Leur fiabilité est reflétée par le « Normalized Root-Mean-Square Deviation » (NRMSD). Une bonne superposition des spectres expérimental et calculé aboutit à un faible NRMSD qui doit être, généralement, inférieur à 0,1. Pourtant, le fait que ce facteur doit être faible est nécessaire mais ce n'est pas suffisant pour avoir des analyses précises (Wallace, 2009). En effet, un autre paramètre à considérer est la concentration des protéines dans l'échantillon doit être connue de manière précise (Kelly *et al.*, 2005), ce qui est particulièrement difficile dans l'approche de dialyse utilisée pour préparer nos échantillons. De plus, le choix du set de référence doit être compatible avec la nature de la protéine étudiée et avec la manière dont elle pourrait se structurer. Dans notre cas, nous aurions dû choisir le set de référence SMP180 qui inclut le set SP175 ainsi que plusieurs protéines membranaires (Abdul-Gader *et al.*, 2011). Mais l'utilisation de ce set de référence nécessite une fenêtre spectrale qui commence de 180 nm. Or, dans cette zone le signal dichroïque est faussé par la forte absorption des lipides existant dans nos échantillons, surtout lorsque ceux-ci forment des grosses vésicules. C'est pourquoi nous avons choisi le SP175 car il contient beaucoup plus des protéines en comparaison aux autres sets couvrant les différentes structures secondaires.

Il faut signaler que les protéines membranes sont extrêmement difficiles à étudier par DC du fait de leur insolubilité dans les tampons aqueux. Elles sont souvent analysées dans des milieux micellaires ou insérées dans des couches lipidiques. Cet environnement hydrophobe et anisotrope est à l'origine de quelques phénomènes qui expliquent les difficultés rencontrées dans l'analyse des spectres CD (Wallace, 2009):

- Un aplatissement du signal dichroïque dû à la distribution non uniforme des chromophores au sein des vésicules et de ces vésicules dans l'ensemble de l'échantillon. Ceci, en provoquant la diminution des pics d'absorbance (en comparaison à leur intensité en milieu isotrope) interfère avec la loi de Beer-Lambert qui n'est dès lors plus applicable
- La diffusion différentielle qui dépend directement de la taille des vésicules et apparaît lorsque la taille de celles-ci dépasse  $1/20$  de la longueur d'onde de la lumière polarisée. Les vésicules diffusent alors la lumière dans toutes les directions et donc en dehors de l'angle du détecteur.
- Le déplacement des longueurs d'onde causé par la différence des constantes diélectriques entre une solution aqueuse et les queues apolaires des lipides ou les chaînes amphipatiques des détergents. La différence d'énergie des transitions va alors être modifiée et les pics du spectre dichroïque vont être déplacés.

## 4.2.4. Conclusion et perspectives

Les trois polypeptides (gp41<sup>TM</sup>, gp120<sup>TM</sup> et fp<sup>TM</sup>) ont été reconstitués dans les liposomes par l'approche de dialyse (voir Figure 51). Leur structure secondaire a, tout d'abord, été étudiée par spectroscopie CD. Le changement de conformation des trois polypeptides a été détecté en comparant la structure des polypeptides dans un milieu micellaire à celle obtenue dans une membrane modèle. Tous les trois peptides se structurent majoritairement en hélice  $\alpha$  dans le DPC qui, ayant une structure proche de celle des phospholipides, permet de bien les solubiliser. Cette structure en hélice  $\alpha$  est bien conservée par la gp41<sup>TM</sup> lorsqu'elle est insérée dans les liposomes quel que soit leur composition tandis que fp<sup>TM</sup> et gp120<sup>TM</sup> subissent un changement de conformation lorsque ces deux polypeptides sont reconstitués dans des liposomes. Ce changement est beaucoup plus prononcé pour le fp<sup>TM</sup>.

Pour améliorer l'analyse des structures secondaires, plusieurs facteurs doivent être contrôlés. Tout d'abord les concentrations des polypeptides et des lipides doivent être calculés après dialyse. La taille des vésicules doit être testée par DLS (Dynamic Light Scattering). En effet, les vésicules de grandes tailles faussent le signal DC. Une autre technique peut être utilisée en parallèle, c'est la spectroscopie d'infrarouge à transformée de Fourier qui permet de contourner les problèmes liés à la tailles des vésicules et la précision de la concentration.

Si le DC nous donne des indications sur les structures secondaires de nos polypeptides, elle n'apporte aucune information sur la structure atomique, ce qui nous amène à utiliser la technique de RMN pour résoudre les structures de nos polypeptides.

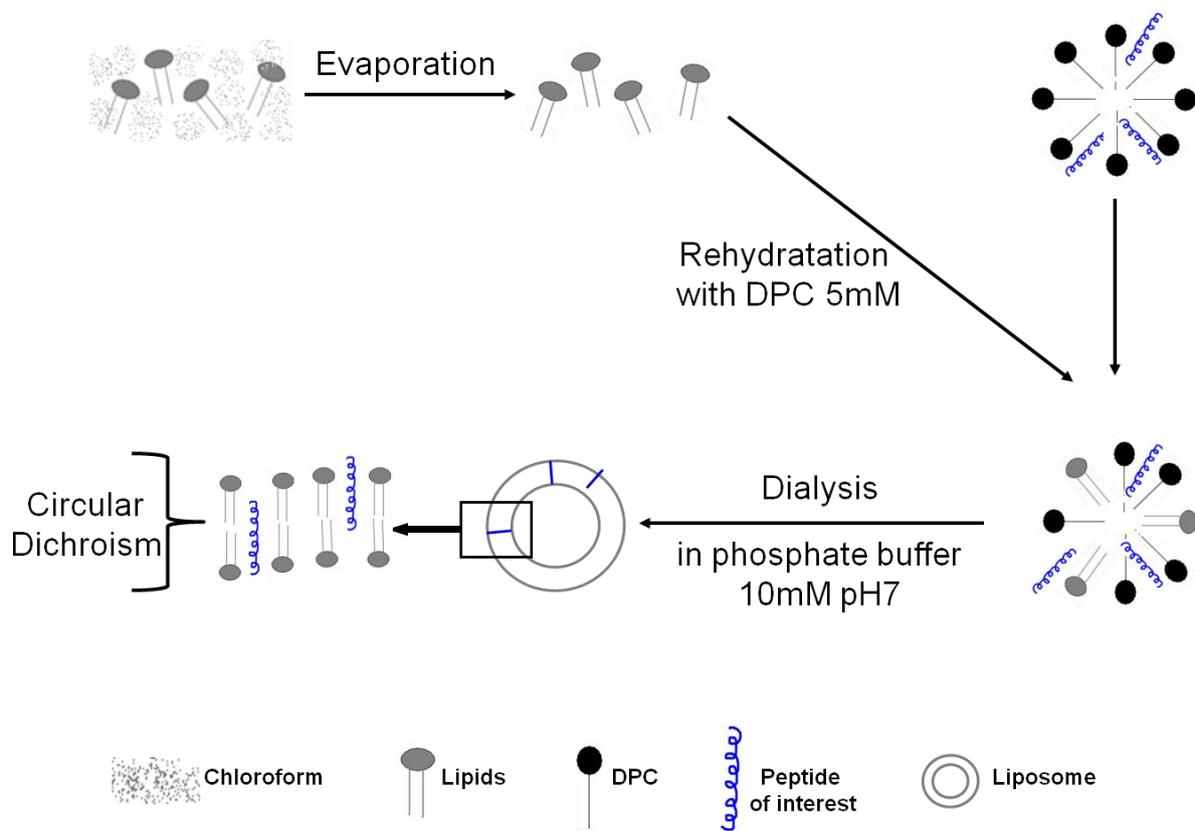


Figure 51 : Reconstitution des polypeptides dans des liposomes.

## 4.3. Étude préliminaire par RMN

### 4.3.1. Matériels et méthodes

#### 4.3.1.A. Surexpression du polypeptide fpTM marqué

Le polypeptide fpTM uniformément marqué en  $^{15}\text{N}/^{13}\text{C}$  a été produit de la même manière que le peptide non marqué (chapitre 3) à la seule différence que le milieu LBYE a été remplacé par le milieu minimum M9 (voir annexe) qui contient 0,5 g  $^{15}\text{NH}_4\text{Cl}$  et 2 g de  $^{13}\text{C}$  glucose par litre de culture comme seule source d'azote et de carbone. Tous les produits proviennent de Sigma-Aldrich (St-Louis, MO) sauf le  $^{13}\text{C}$  glucose qui vient d'EURISOTOP (Saint Aubin, France).

#### 4.3.1.B. Préparation des échantillons

L'échantillon RMN du polypeptide fpTM marqué uniformément en  $^{15}\text{N}$  et  $^{13}\text{C}$  a été préparé de la façon suivante : fpTM dans DPC-d38 (160 mM, Avanti Polar Lipids, Alabaster, AL) avec une concentration finale de 0,6 mM dans  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (80/20).

#### 4.3.1.C. Expérience RMN

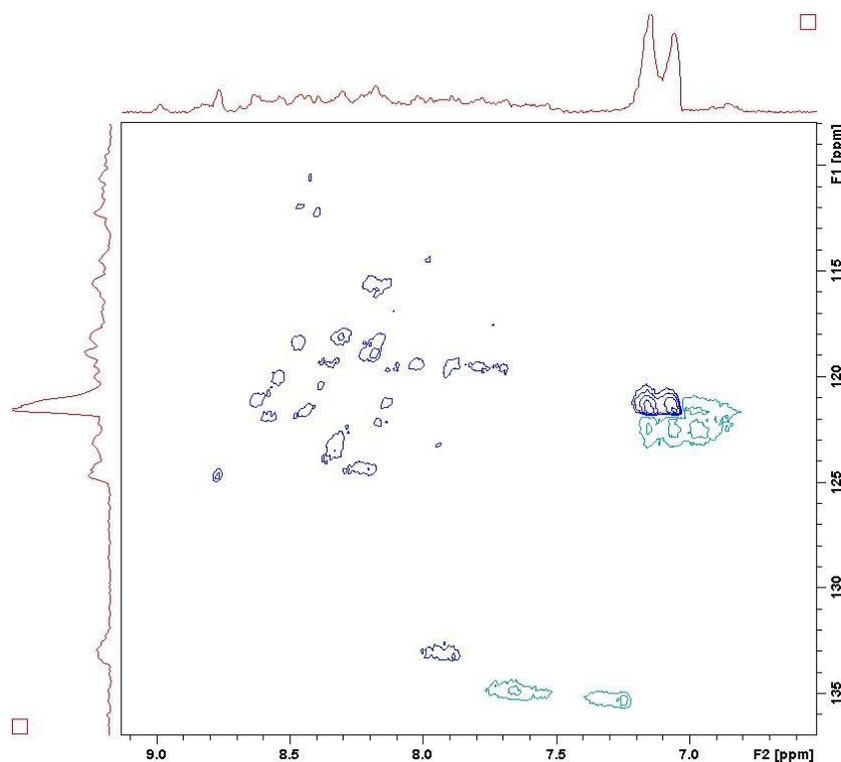
Le spectre en deux dimensions  $^1\text{H}-^{15}\text{N}$  HSQC a été enregistré sur un spectromètre de 600 MHz, la fenêtre spectrale en proton est de 2400 Hz. La fenêtre spectrale et le nombre d'incrément dans la dimension  $^{15}\text{N}$  sont respectivement 3040 Hz et 328 incréments. Le nombre de scans est de 16. Le délai de recyclage est de 1 s.

## 4.3.2. Résultats et discussion

### HSQC

Les spectres  $^1\text{H}$ - $^{15}\text{N}$  HSQC montrent la corrélation entre les protons et les noyaux azote directement liés. Ainsi pour une protéine, on observe des taches de corrélation entre le proton amide et l'azote pour chaque acide aminé sauf pour le tout premier en N-ter. Théoriquement sur notre spectre nous devrions compter 60 taches.

Ce type de spectre permet de confirmer l'homogénéité de notre échantillon mais aussi de voir rapidement si notre peptide est présent de façon structurée. Un peptide bien structuré, est caractérisé par une bonne dispersion des résonances sur les deux dimensions ce qui est le cas dans la Figure 52 qui correspond au spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC du fpTM contenu dans 160 mM de micelles de DPC. Cet échantillon semble donc être prometteur pour l'étude structurale du peptide dans des membranes mimétiques.



**Figure 52 : Spectre  $^1\text{H}$ - $^{15}\text{N}$  HSQC du fpTM dans un milieu micellaire. fpTM dans DPC-d38 avec une concentration finale de 0,6 mM dans  $\text{H}_2\text{O}/\text{D}_2\text{O}$  (80/20).**

### 4.3.3. Conclusions et perspectives

Le DC ne donne pas une la structure atomique, ce qui nous amène à résoudre les structures de nos polypeptides par RMN. Par contre, cette technique nécessite le marquage complet ( $^{13}\text{C}$  et  $^{15}\text{N}$ ) des polypeptides dans un milieu minimum (milieu M9 (Marley *et al.*, 2001)) ce qui impose un nouvel obstacle à contourner, surtout que la surexpression de ces polypeptides dans ce milieu devrait être plus faible par rapport à son surexpression dans un milieu complet (LBYE). Or, nous avons pu surexprimer et purifier le fpTM complètement marqué avec une quantité suffisante pour lancer les spectres RMN (en phases solide et liquide). Plusieurs spectres ont été enregistrés pour fpTM en poudre (lyophilisé), dans DPC 160 mM et inséré dans des vésicules de POPC/POPS (3/1) et les attributions sont en cours.

# 5. Chapitre 5 : Étude de l'interaction du CRAC avec le cholestérol

## 5.1. Introduction

Le cholestérol est le constituant majeur de la membrane plasmique des cellules eucaryotes. Il contrôle la fluidité de cette membrane, augmente son épaisseur et forme une barrière de perméabilité. En effet, il est distribué d'une façon non uniforme dans la membrane et y contribue à la formation des micro-domaines latéraux. De plus, le cholestérol régule l'activité de plusieurs protéines membranaires. Cette régulation peut se faire par des séquences d'acides aminés bien précises (Fantini and Barrantes, 2013). Il s'agit principalement du motif CRAC (Cholesterol Recognition/interaction Amino acid Consensus) identifié par Li et Papadopoulos chez le récepteur périphérique de la benzodiazépine. Ce motif est constitué par la succession des acides aminés suivants : -L/V-X<sub>1-5</sub>-Y-X<sub>1-5</sub>-R/K- (Li and Papadopoulos, 1998). Deux axes de recherche ont été abordés pour mettre en évidence son interaction avec le cholestérol soit des études de mutagenèse au niveau de la protéine d'intérêt visant à déterminer les mutations qui altèrent la fonction vis-à-vis des domaines riches en cholestérol, soit des études menées sur des peptides contenant le motif CRAC visant à tester si ces peptides interagissent préférentiellement avec le cholestérol. Des études ont mis en évidence l'existence de ce motif au sein de nombreuses protéines parmi lesquelles plusieurs membres des récepteurs couplés à la protéine G (par exemple le récepteur de la sérotonine <sub>1A</sub>...) et ont prouvé la capacité de ce motif à interagir avec le cholestérol (Jafurulla et al., 2011). Une démarche similaire a montré l'existence du motif CRAC (-L-W-Y-I-K-) au sein de la gp41. Vincent *et al.* ont mis en évidence sa localisation au niveau de sa région pré-transmembranaire. En effet, cette équipe a conçu plusieurs constructions protéiques correspondant à différents domaines de glycoprotéines virales et a étudié leur affinité pour le cholestérol hemisuccinate lié à l'agarose (CHS). Seule, la construction correspondant à la gp41 a montré une affinité pour le CHS. Ils ont alors ciblé différents domaines de la gp41 afin de déterminer lesquels présentaient l'affinité au CHS. Celle-ci a principalement été détectée pour les peptides pré-transmembranaires. Ces derniers contiennent un motif qui correspond à celui identifié par Li et Papadopoulos (Vincent et al., 2002).

D'autre part, des études ont montré l'importance du cholestérol dans l'infection virale et dans la fusion des membranes virale et cellulaire (Liao et al., 2001; Viard et al., 2002). En parallèle, des mutations au niveau du motif CRAC de cette glycoprotéine ont montré un ralentissement du cycle viral aussi bien qu'une importante réduction de l'infection par les particules virales libres et par la fusion des cellules infectées ; alors que la synthèse, l'oligomérisation, l'expression en surface de la cellule et l'assemblage des protéines virales n'ont pas été affectés (Chen et al., 2009). Il est à noter que ce motif est hautement conservé chez les différentes souches virales (VIH-1 et 2, et SIV) (Vincent et al., 2002). Ce lien entre la fusion catalysée par la gp41 contenant ce motif conservé et le cholestérol a amené les chercheurs à chercher à mieux comprendre le mécanisme de cette interaction.

A cet égard, il reste de nombreuses questions à élucider: comment le cholestérol agit-il sur le mécanisme de l'entrée du virus ? Le cholestérol interagit-il directement avec le motif CRAC ou en changeant les propriétés physico-chimiques des membranes pour permettre à la gp41 d'accomplir son rôle dans la fusion des membranes? Ou bien par les deux mécanismes?

Afin d'y répondre, une des techniques envisagée a été la RMN de l'état solide sur des échantillons orientés. Cette technique apporte des informations sur la topologie et la dynamique des protéines ou des lipides. Pour cela, deux peptides ont été conçus de la façon suivante : les deux peptides contiennent le motif CRAC (L-W-Y-I-K) flanqué en N-terminale de quelques acides aminés de la gp41 (Voir Tableau 9). Ils diffèrent l'un de l'autre de par la séquence transmembranaire ajoutée en C-terminal de ce motif. Cette séquence correspond soit à la séquence transmembranaire naturelle de la gp41, soit à une succession d'A-L (alanine-leucine) pour rendre ce peptide moins hydrophobe que le premier. Les deux constructions sont enfin flanquées à chaque extrémités (C- et N-terminales) de deux résidus lysine afin d'améliorer leur solubilité. Des peptides contrôles ont été conçus sur le même modèle mais présentent un réarrangement des acides aminés constituant le motif CRAC (L-K-W-I-Y). Une variante a été également conçue en remplaçant l'isoleucine du motif CRAC par la leucine tout en conservant une séquence qui correspond à ce motif.

## 5.2. Méthodes et matériels

Les solvants organiques proviennent du Sigma Aldrich (Saint Louis, MO) avec une pureté de 99%, POPC, POPS et cholestérol non marqué de Avanti Polar Lipids ; Birmingham, AL, le cholestérol marqué en positions 2,2,3,4,4,6 par  $^2\text{H}$  de Eurisotop, Paris , France ; les acides aminés de protégés par F-moc de Bachem, Heidelberg, Allemagne ou d'Applied Biosystems, Weiterstadt, Allemagne.

### 5.2.1. Synthèse peptidique

Les peptides conçus (voir Tableau 9) ont été produits par synthèse peptidique en phase solide sur un synthétiseur automatique Millipore 9050, et assemblés sur une résine tentagel-R-RAM (RAPP Polymere, GmbH, Tubingen, Allemagne) en utilisant des cycles standards de synthèse de l'appareil avec un excès de quatre fois en acides aminés (A, L, K et H) protégés par un groupement Fmoc (Merck, Darmstadt, Allemagne). Une fois la synthèse terminée, la résine a été placée sur un verre frité et lavée avec de l'éthanol, de l'acide acétique et de l'éther. Le clivage des peptides a été réalisé dans une solution contenant 28 mL TFA, 1,5 mL d'eau et 300  $\mu\text{L}$  de triéthylsilane sous agitation pendant 4 heures. Les peptides ont été ensuite repris dans une solution TFA/toluène (50/50) (v/v). Le solvant a été évaporé à l'aide d'un évaporateur rotatif (Büchi R-200) à 37°C. Les peptides ont été solubilisés dans de l'acide acétique 100% et précipités dans l'éther. Les tubes ont été centrifugés à 11000 g pendant 30 minutes à 4°C. Les culots contenant les peptides ont été finalement solubilisés dans de l'acide acétique 4 %, congelés dans de l'azote liquide et mis à lyophiliser. Les produits clivés sont analysés par MALDI-TOF pour vérifier la présence des peptides d'intérêt qui ont été purifiés par RP-HPLC par un gradient d'acétonitrile/eau/TFA 0,1% sur une colonne C18 (Luna 5u c18(2) 100A ; 150x30 mm ; Phenomenex, 547606-2).



traitements de ces spectres consistent en une soustraction de la contribution des vésicules présentes dans le tampon utilisé. Les spectres obtenus et analysés correspondent alors à la seule contribution des polypeptides. L'estimation des contributions des différentes structures secondaires est faite à l'aide du logiciel Dichroweb (<http://dichroweb.cryst.bbk.ac.uk/html/home.shtml>). Deux algorithmes ont été utilisés CONTIN/LL (une variante de CONTIN) et CDSSTR avec le set référence SP175 (Whitmore and Wallace, 2004, 2008).

#### 5.2.4. SDS-PAGE

Les protéines sont séparées par électrophorèse en conditions dénaturantes en présence de 8% de SDS selon la technique « SDS-PAGE ». Elles migrent successivement à travers deux gels. Le premier gel est appelé gel de concentration ou d'alignement, faiblement concentré en acrylamide 4%. Il sert de concentrer ces protéines à la limite du second gel. Ce dernier, dit de séparation est plus concentré en acrylamide 20% et il permet de séparer les protéines selon leur taille. La migration a lieu pendant 1h à 150V. Les gels sont colorés par le bleu de coomassie (Roti® Blue 5X, A15.1, Roth) puis décolorés avec une solution de méthanol 25% et d'eau 75%. Les gels sont colorés aussi par le nitrate d'argent (protocole dans annexe). Les photos des gels sont prises avec le lecteur de gel El Logic 200, Imaging System, Kodak.

#### 5.2.5. Préparations des échantillons orientés

La quantité désirée du peptide (CRAC\_1, CRAC\_2 ou leurs contrôles) est dissoute dans un mélange de solvants organiques (chloroforme/méthanol (V/V = 2/1) pour CRAC\_1 et son contrôle ; TFE/eau (15%) pour CRAC\_2 et son contrôle. Les quantités appropriées des lipides sont dissous dans chloroforme/méthanol (V/V = 2/1) s'ils sont utilisés pour la reconstitution du CRAC\_1 ou dans TFE/Chloroforme (V/V = 1/1) s'ils sont utilisés pour reconstituer le CRAC\_2. Le tout est mélangé pour avoir le ratio molaire (peptide/lipides) = 2%. Les solvants sont évaporés sous un flux d'azote jusqu'à obtention d'un film lipidique (~200 µL). Celui-ci est déposé sur 22 plaques de verre (8x22 ou 6x11 ; Marienfield, Lauda-Königshofen, Allemagne). Les traces de solvants sont évaporées sous vide. L'échantillon est mis à hydrater à 93% d'humidité relative, à température ambiante pendant une nuit. Les lamelles sont empilées et emballées dans du téflon et un film plastique pour éviter la déshydratation de l'échantillon.

## 5.2.6. Acquisition des spectres RMN

### 5.2.6.A. Spectres phosphore ( $^{31}\text{P}$ )

Les spectres  $^{31}\text{P}$  découplés en  $^1\text{H}$  ont été enregistrés à 161,937 MHz sur un spectromètre RMN Bruker Avance 400 MHz à l'aide d'une sonde à double résonance équipée d'une bobine plate. L'enregistrement s'est fait à température ambiante en utilisant une séquence d'impulsions écho de Hahn (Rance and Byrd, 1983) avec un pulse  $\pi/2$  de 3,5  $\mu\text{s}$ . La largeur spectrale est de 75 kHz. Le délai de l'écho et du recyclage sont de 40  $\mu\text{s}$  et 3 s, respectivement. Un spectre  $^{31}\text{P}$  de référence de l'acide phosphorique 85 % est enregistré et le phosphore est calibré à 0 ppm. Avant transformation de Fourier une multiplication exponentielle correspondant à un « Lorenz Broadening » (LB) de 30 Hz est appliquée.

### 5.2.6.B. Spectres azote ( $^{15}\text{N}$ )

Les spectres  $^{15}\text{N}$  découplés su  $^1\text{H}$  ont été enregistrés à 40,54 MHz sur un spectromètre RMN Bruker Avance 400 MHz à l'aide d'une sonde à double résonance équipée d'une bobine plate. L'enregistrement s'est fait à température ambiante en utilisant une séquence d'impulsions de polarisation croisée avec un fort découplage en  $^1\text{H}$  (Pines et al., 1973). Les paramètres utilisés : une fenêtre spectrale de 37,5 kHz, un temps d'acquisition de 3,5 ms, temps de contact de la polarisation croisée 800  $\mu\text{s}$ , délai de recyclage 3 s, le champ B1 pendant l'impulsion  $90^\circ$  est 42 kHz, Les spectres sont accumulés avec 40 k scans. Avant transformation de Fourier une multiplication exponentielle correspondant à un LB de 100 Hz est appliquée. Les spectres sont calibrés sur  $^{15}\text{NH}_4\text{Cl}$  à 40,0 ppm.

### 5.2.6.C. Spectres deutérium ( $^2\text{H}$ )

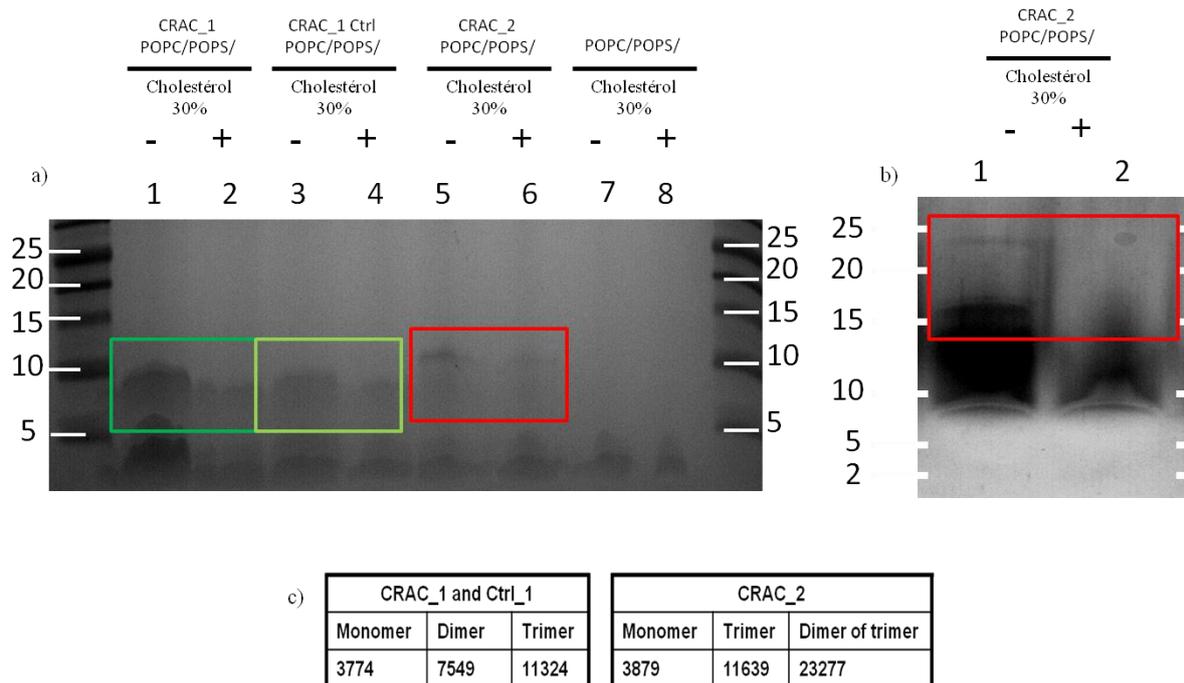
Les spectres  $^2\text{H}$  ont été enregistrés à 61,4 MHz sur un spectromètre RMN Bruker Avance 400 MHz à l'aide d'une sonde à double résonance équipée d'une bobine plate. L'enregistrement s'est fait à une température bien précisée en utilisant une séquence d'écho quadripolaire (Davis et al., 1976). La fenêtre spectrale, le temps d'acquisition et le délai de recyclage valent 125 kHz, 6,8 ms, 1 ms et 1,5 s respectivement. Les spectres sont accumulés avec 40 k scans. Avant transformation de Fourier une multiplication exponentielle correspondant à un LB de 300 Hz est appliquée. Les spectres sont calibrés sur  $\text{D}_2\text{O}$  à 0 ppm.

## 5.3. Résultats et discussion

### 5.3.1. Oligomérisation des peptides

Les gels SDS-PAGE sont utilisés couramment pour accéder à l'état d'oligomérisation de diverses protéines membranaires telles que le domaine transmembranaire de phospholambane (Chu et al., 2010), la glycophorine A (Lemmon et al., 1992), le peptide MS1 (Choma et al., 2000).

La Figure 53 montre le profil de migration de nos peptides d'intérêt qui donne une idée sur leur état d'oligomérisation. Dans le gel a, les puits 1 et 2 correspondent au dépôt de 2 µg de CRAC\_1 reconstitué dans des liposomes (POPC/POPS, 3/1) contenant ou non 30% du cholestérol, nous y observons des bandes entre 5 et 10 kDa. Ces bandes sont absentes dans les puits 7 et 8 correspondant au dépôt des liposomes tous seuls sans peptide. Les puits 3 et 4 contenant 1,3 µg de CRAC\_1 contrôle inséré dans des liposomes contenant ou non 30% du cholestérol montrent également des bandes ayant la même mobilité électrophorétique que celles du CRAC\_1. Ceci indique le CRAC\_1 et son contrôle forment un dimère (7549 Da) dans les liposomes contenant ou non cholestérol, cette dimérisation est due à la séquence transmembranaire modèle qui tend à se dimériser lorsque sa concentration dépasse le 1% (Sparr et al., 2005; Yano et al., 2002). Dans les puits 5 et 6 contenant 1,3 µg de CRAC\_2 reconstitué dans liposomes, nous observons des bandes entre 10 et 15 kDa ce qui correspond à des trimères. La présence d'un motif GxxxG et d'un résidu arginine hautement conservés au sein de la séquence transmembranaire de la gp41 pourrait être à l'origine de cette trimérisation (Curran and Engelman, 2003; Kim et al., 2009; Long et al., 2011; Senes et al., 2004). De plus, dans le gel b, nous voyons que le CRAC\_2 reconstitué dans des liposomes formés de POPC/POPS (3/1) a une bande faible entre 20 et 25 kDa (puits 1) ce qui correspond à un dimère d'un trimère (26277 Da). Cependant aucune bande n'est observée dans le puits 2 où le CRAC\_2 est reconstitué dans POPC/POPS 3/1 en présence de 30% mole% cholestérol. Ceci peut être expliqué par le fait que la quantité effective du CRAC\_2 déposée peut être non détectable. En effet, lors de leur évaporation sous le flux d'azote, les solvants organiques commencent, à un moment donné, à former deux phases. Alors il est possible que les composants du mélange peptide-lipides s'agrègent au fur et à mesure. Ces agrégats sont ensuite éliminés par centrifugation. Ce phénomène est beaucoup plus prononcé dans le cas de CRAC\_2 à cause de sa solubilité limitée et qui nécessite un certain pourcentage d'eau.



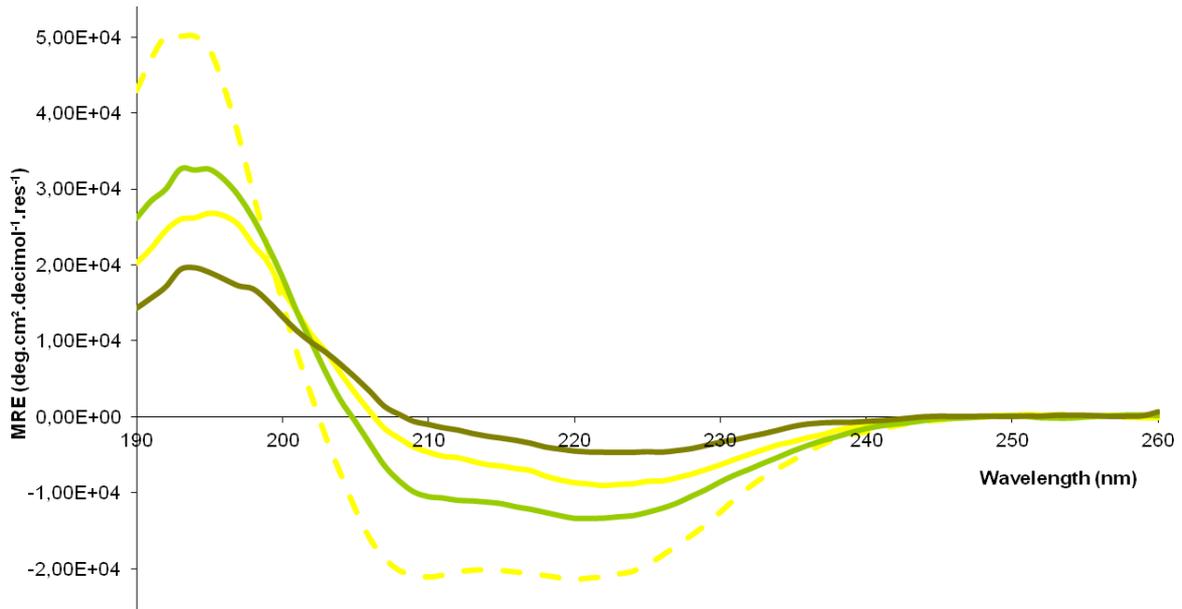
**Figure 53 : Oligomérisation des peptides CRAC\_1, son contrôle et CRAC\_2.** Les échantillons utilisés correspondent aux peptides reconstitués dans POPC/POPS (3/1) sans (-) ou avec cholestérol 30% (+). Ils sont déposés sur des gels SDS-PAGE 20%. Après migration électrophorétique 1h à 150V, les gels sont colorés soit a) au bleu de Coomassie (2 µg de CRAC\_1, 1,3µg de CRAC\_1 Contrôle (Ctrl) et de CRAC\_2) ; soit b) au nitrate d'argent (0,5 µg de CRAC\_2). Le marqueur de poids moléculaire utilisé est (Precision Plus Protein™ Dual Color Standards #161-0374). c) Masses théoriques des différentes formes des peptides. Elles sont obtenues grâce au logiciel ProtParam du site ExPASy (<http://expasy.org/tools/protparam.html>) qui calcule la masse d'une protéine d'après sa séquence en acide aminé.

### 5.3.2. Structure secondaire de CRAC\_1 et CRAC\_2

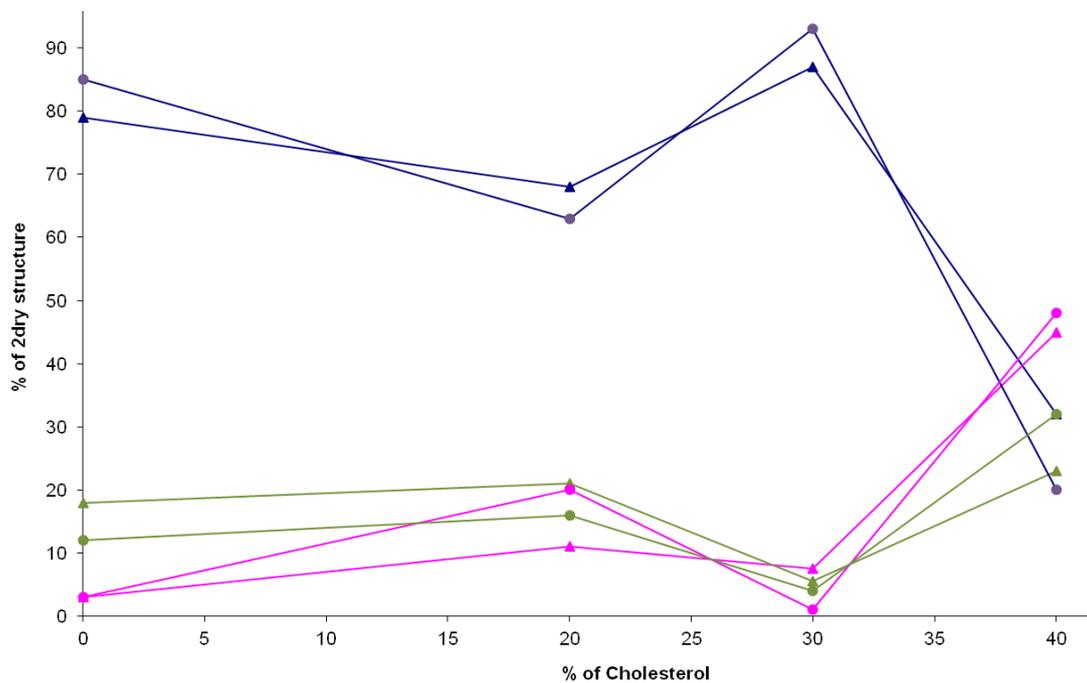
Dans le but de voir l'influence du cholestérol sur la structure secondaire des peptides CRAC\_1 et CRAC\_2, des spectres DC de ces deux peptides insérés dans des liposomes POPC/POPS (3/1) ont été enregistrés en fonction d'une concentration croissante de cholestérol. En parallèle, les structures secondaires ont été étudiées par spectroscopie d'infrarouge à transformée de Fourier (FTIR).

Les spectres DC du CRAC\_1 (Figure 55) montrent un pic positif vers 195nm et deux pics négatifs vers 208 et 222 nm lorsque la concentration du cholestérol est entre 0 et 30 mole% ce qui correspond à une signature hélice  $\alpha$ . L'analyse de ces spectres (Figure 54) montre une contribution d'une hélice  $\alpha > 70\%$ . Lorsque la concentration du cholestérol dépasse 30%, l'allure du spectre DC change et montre un seul pic négatif vers 220 nm. Ce changement est dû à l'apparition d'une structure en feuillet  $\beta$  (~50%).

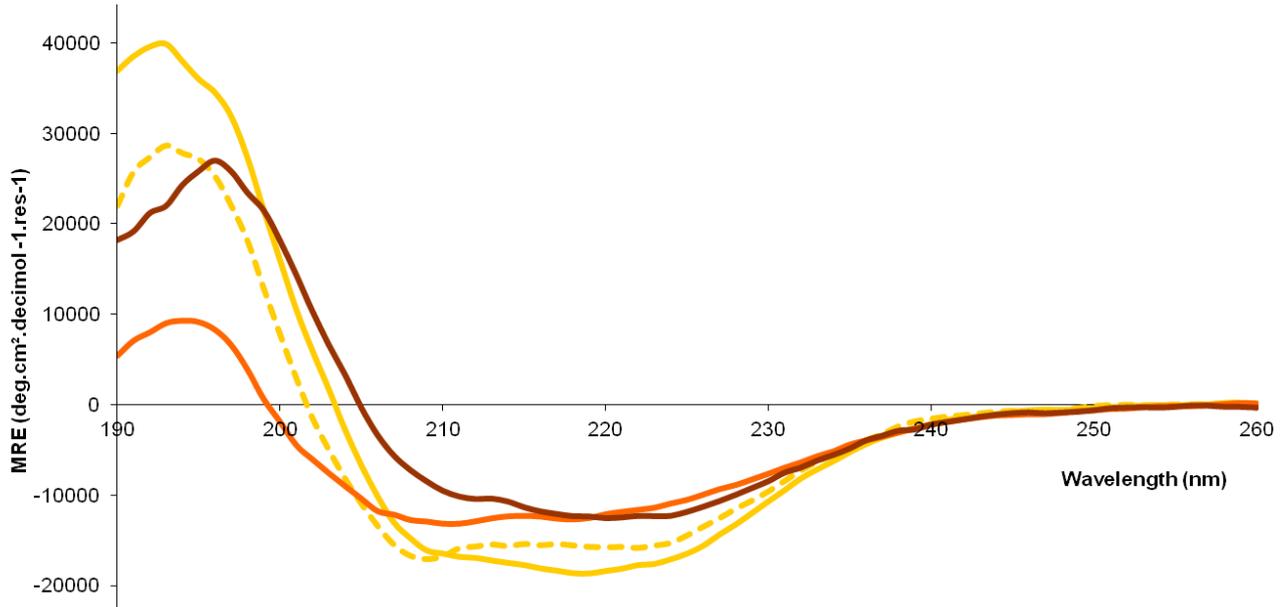
Par contre le spectre DC du CRAC\_2 (Figure 57) montre un maximum vers 195nm et deux minima vers 208 et 222 nm, ce qui est caractéristique d'une hélice  $\alpha$ , lorsqu'il est inséré dans des liposomes dépourvus du cholestérol. Lorsque ce dernier est rajouté, le spectre DC de CRAC\_2 subit des changements drastiques surtout quand la concentration du cholestérol dépasse 30%, ceci est dû à une diminution du taux d'hélice  $\alpha$  (~50%, voir Figure 56). De même ces changements sont dus à une déstructuration partielle de CRAC\_2 (~40%, voir Figure 56).



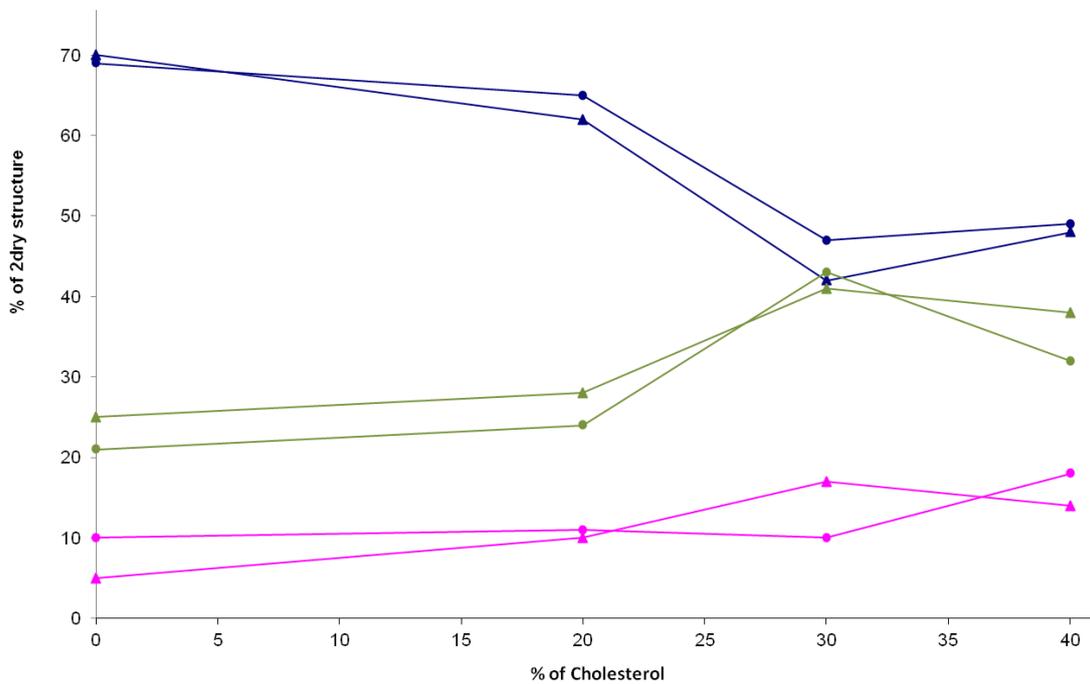
**Figure 55 : Spectre de dichroïsme circulaire de CRAC\_1. Ce peptide est reconstitué dans des liposomes P/L = 1/50) formés de POPC/POPS (3/1) avec des concentrations croissantes en cholestérol 0% (jaune pointillé), 20% (jaune), 30% (vert clair) et 40% (vert foncé). La concentration finale du CRAC\_1 = 0,1 mg/mL dans un tampon phosphate 10mM pH7. MRE:**



**Figure 54 : Analyse des structures secondaires du CRAC\_1 en fonction de [cholestérol]. Elles sont faites à l'aide du logiciel Dichroweb en utilisant deux algorithmes CONTIN/LL (triangles) et CDSSTR (cercles). Il s'agit trois types de contributions : hélice  $\alpha$  (bleu), feuillet  $\beta$  (rose) et**



**Figure 57 : Spectre de dichroïsme circulaire de CRAC\_2.** Ce peptide est reconstitué dans des liposomes (P/L = 1/50) formés de POPC/POPS (3/1) avec des concentrations croissantes en cholestérol 0% (jaune pointillé), 20% (jaune), 30% (orange) et 40% (marron). La concentration finale du CRAC\_2 = 0,05 mg/mL dans un tampon phosphate 10mM pH7. MRE:



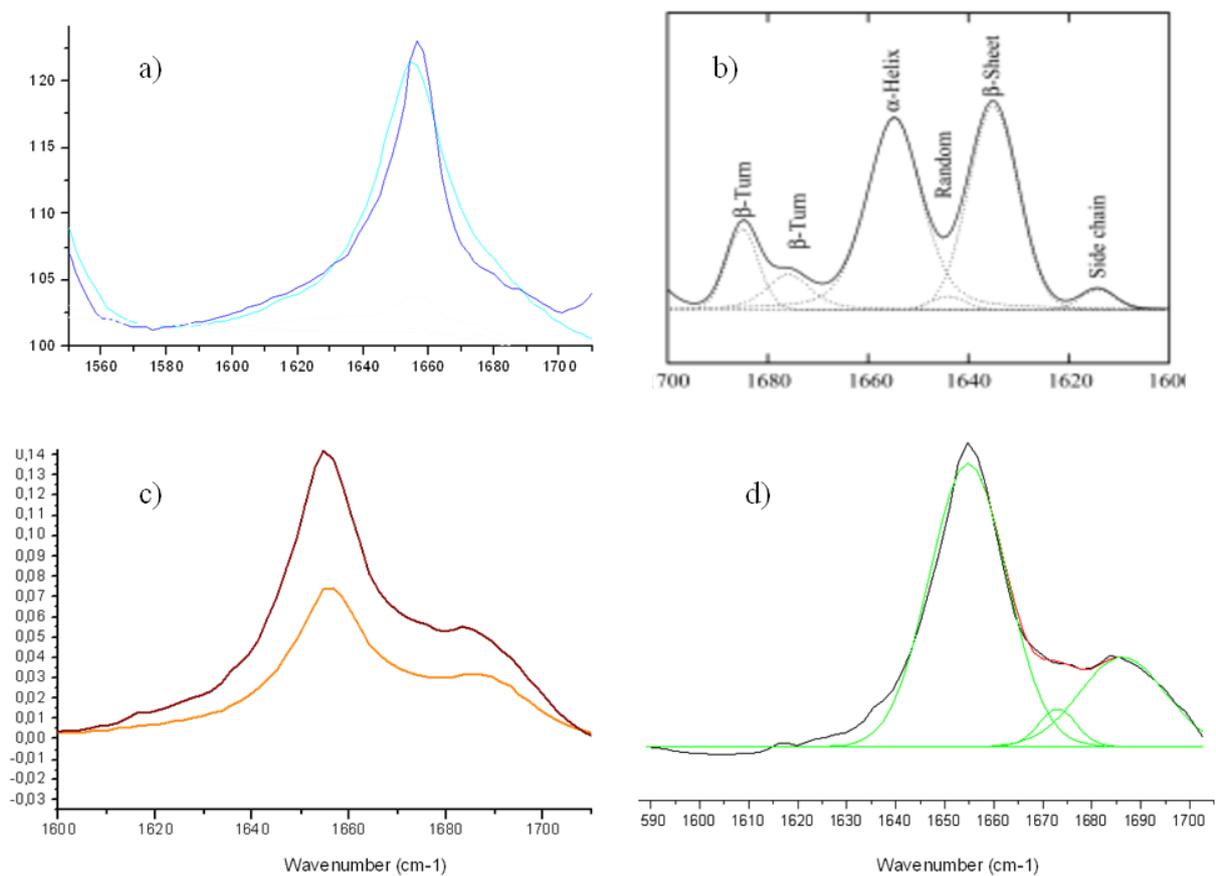
**Figure 56 : Analyse des structures secondaires du CRAC\_2 en fonction de [cholestérol].** Elles sont faites à l'aide du logiciel Dichroweb en utilisant deux algorithmes CONTIN/LL (triangles) et CDSTTR (cercles). Il s'agit trois types de contributions : hélice  $\alpha$  (bleu), feuillet  $\beta$  (rose) et

Vu les problèmes discutés dans le chapitre 4 du dichroïsme circulaire, nous avons voulu d'étudier les structures secondaires des peptides avec une autre technique, le FTIR. En fait, cette technique présente un avantage par rapport au DC qu'elle ne nécessite pas que les échantillons soient dans un tampon. Les lipides et les peptides mélangés dans des solvants organiques sont déposés sur le cristal puis séchés. Ce qui permet ainsi de contourner quelques obstacles : taille des vésicules, précision de la concentration...

Le spectre FTIR du CRAC\_1 reconstitué dans POPC ou POPC/cholestérol 30% montre une seule bande entre 1640 et 1670  $\text{cm}^{-1}$  (Figure 58.a). C'est pour cela qu'il n'a pas subi une déconvolution. En le comparant à la Figure 58.b, nous remarquons que cette bande correspond à une hélice  $\alpha$ . Le spectre du CRAC\_2 reconstitué dans POPC ou POPC/cholestérol 30% montre une seule bande entre 1640 et 1670  $\text{cm}^{-1}$  avec un épaulement entre 1670-1690  $\text{cm}^{-1}$  (Figure 58.c). La déconvolution de ce spectre montre trois contributions, une qui correspond à une hélice  $\alpha$  entre 1640-1670  $\text{cm}^{-1}$  et deux contributions entre 1665-1685 et 1665-1700  $\text{cm}^{-1}$  qui correspondent à des coudes  $\beta$ .

Ces résultats de CRAC\_1 sont en accord avec ceux du DC qui donne, en présence du cholestérol 30%, ~90% d'hélice ce qui n'est pas le cas de CRAC\_2. Cela est peut-être dû au fait que le DC nécessite la détermination de la concentration de peptide avec précision. Or comme nous avons déjà vu le CRAC\_2 a une solubilité limitée (Voir la partie 5.3.1 Oligomérisation des peptides).

Comme les résultats obtenus, surtout ceux du FTIR, montrent que la structure secondaire ne dépend pas trop de la présence du cholestérol. Ceci nous a amené à étudier l'orientation des peptides dans des bicouches et étudier l'influence du cholestérol sur cette orientation.



**Figure 58 : Spectres FTIR. a) CRAC\_1 reconstitué dans une bicouche de POPC (bleu clair) ou POPC/cholestérol 30% (bleu foncé). b) Contributions des différentes structures secondaires. c) CRAC\_2 reconstitué dans une bicouche de POPC (orange) ou POPC/cholestérol 30% (marron). d) Déconvolution du spectre c).**

### 5.3.3. Topologie et dynamique des peptides

L'orientation d'un peptide dans des bicouches lipidiques peut être déterminée par spectroscopie de RMN de solide sur le noyau azote ( $^{15}\text{N}$ ). Elle indique directement l'angle tilt entre le vecteur NH de la liaison peptidique et le champ magnétique extérieure  $B_0$  (Bechinger et al., 1999). Si ce peptide se structure complètement en hélice  $\alpha$ , cette technique permet de déterminer son orientation par rapport à la normale de la bicouche (si le déplacement chimique de  $^{15}\text{N}$   $>200$  ppm, il s'agit d'une hélice transmembranaire ; si ce déplacement chimique  $<100$  ppm, il s'agit d'une hélice orientée parallèlement à la membrane (Bechinger and Sizun, 2003). C'est pour cela nous avons marqué nos peptides par  $^{15}\text{N}$  au niveau d'un des cinq acides aminés du motif CRAC.

Une autre technique de RMN de solide est utilisée pour déterminer l'orientation d'un peptide, il s'agit de la RMN de solide sur le noyau  $^2\text{H}$  qui assure aussi des informations sur la dynamique de ce peptide (Aisenbrey and Bechinger, 2004; Aisenbrey et al., 2006; Bechinger et al., 2004). Cette technique nécessite le marquage d'un acide aminé par  $^2\text{H}$ , c'est pour cela que la majorité de nos peptides ont été marqués par  $^2\text{H}$  au niveau d'un résidu alanine de la séquence transmembranaire. En effet, cet acide aminé est le plus utilisé dans cette technique car il présente plusieurs avantages. Tout d'abord le groupement méthyle possède trois deutériums qui s'échangent rapidement entre eux à température ambiante ce qui améliore la sensibilité et augmente le signal. De plus ce groupement est directement attaché au squelette peptidique et la liaison  $\text{C}\alpha\text{-C}\beta$  qui adopte un alignement bien précis vis-à-vis de la normale de la membrane détermine directement l'orientation du peptide. Cet alignement de la liaison  $\text{C}\alpha\text{-C}\beta$  est complètement différent du NH dont l'orientation par rapport au champs magnétique détermine majoritairement le déplacement chimique de l'azote ce qui assure que les deux noyaux apportent des informations complémentaires (Aisenbrey and Bechinger, 2004)...

Tous les peptides sont reconstitués dans une membrane orientée avec un ratio molaire de 1/50 (P/L). Cette membrane est formée par POPC puis par POPC/Cholestérol (30mol% de cholestérol sont utilisés dans le cas de CRAC\_1 et des concentrations variables, 2 jusqu'à 30 mol%, avec CRAC\_2). L'acquisition des spectres RMN se fait soit dans une membrane dont sa normale est alignée parallèlement au champ magnétique extérieur pour le noyau azote 15, soit dans une membrane dont sa normale est alignée parallèlement ou perpendiculairement au  $B_0$  pour le noyau  $^2\text{H}$ . L'alignement des lipides est mesuré par spectroscopie de RMN de solide sur le noyau  $^{31}\text{P}$ .

### 5.3.3.A. Spectroscopie de RMN sur l'azote

#### 5.3.3.A.a. CRAC\_1 (ayant une séquence transmembranaire modèle)

Trois peptides ont été conçus avec une séquence transmembranaire modèle (succession d'-Ala-Leu-). Le premier contient le motif CRAC (-LWYIK-) marqué par  $^{15}\text{N}$ -Leucine en position 8, le deuxième contient une variante de ce motif (Ile est substitué par Leu, -LWYLK) marqué aussi au niveau de la même leucine et le troisième correspond à un contrôle négatif (les acides aminés du CRAC sont randomisés : - LKWIY- ce qui ne correspond pas à la séquence consensus du motif CRAC) ayant le même marquage ( $^{15}\text{N}$ ). Ces trois peptides ont été également marqués par  $^2\text{H}$  au niveau de l'alanine 24 dont la partie transmembranaire.

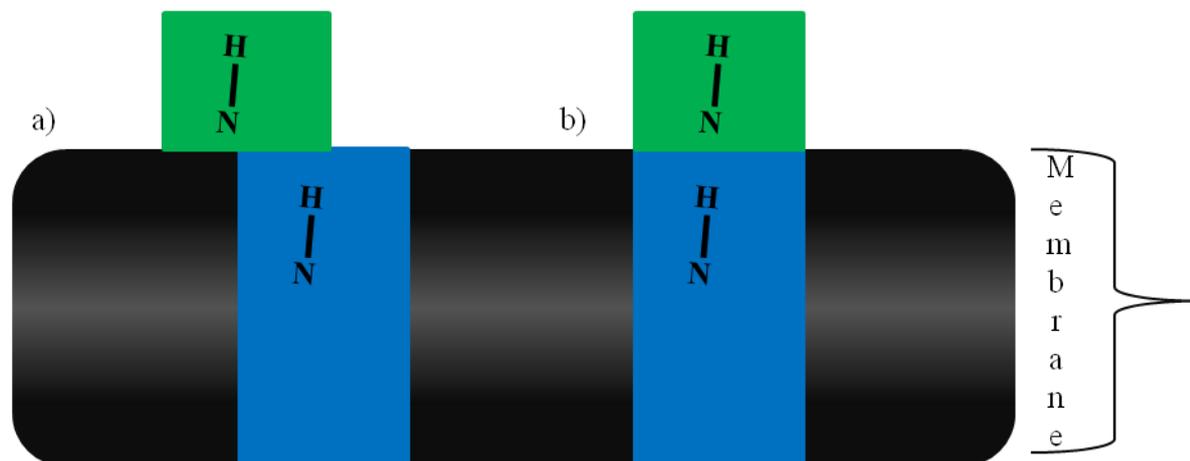
Lorsque le peptide CRAC\_1 est reconstitué dans des bicouches formées soit par POPC soit par POPC/cholestérol 30% équilibrées à 93% d'humidité relative à température ambiante, les spectres RMN  $^{31}\text{P}$  découplé  $^1\text{H}$  (Figure 60, a et b, à droite) montrent un pic majoritaire vers 30 ppm et une petite contribution entre 30 et 15 ppm. Ceci indique que les membranes sont dans une phase liquide cristalline et qu'elles sont relativement bien alignées avec la normale de la bicouche parallèle au champ magnétique (Harzer and Bechinger, 2000). Les spectres  $^{31}\text{P}$  découplé  $^1\text{H}$  du CRAC\_1 contrôle présentés dans la Figure 61 (a et b, à droite) montrent également les mêmes allures que les spectres de CRAC\_1 et par la suite les bicouches contenant le CRAC\_1 contrôle sont bien alignés aussi par rapport au champ magnétique extérieur. Tous les spectres  $^{31}\text{P}$  découplé  $^1\text{H}$  montrent un bon alignement des lipides pour tous les échantillons qui vont être présentés au fur et à mesure dans ce chapitre alors ils ne vont pas y être présentés.

Le spectre  $^{15}\text{N}$  découplé  $^1\text{H}$  du CRAC\_1, marqué en  $^{15}\text{N}$  au niveau de la Leu8 et reconstitué dans des bicouches orientées de POPC (Figure 60.a, à gauche), montre un seul pic fin ayant un déplacement chimique de  $196,5 \pm 2,2$  ppm. Lorsqu'il est reconstitué dans POPC/Cholestérol 30% (Figure 60.b, à gauche et Tableau 10), son spectre azote montre un pic vers  $208 \pm 6,5$  ppm. Cette différence du déplacement chimique de l'azote, sans ou avec 30% du cholestérol, est observée également chez la variante du CRAC\_1 marquée au niveau du même acide aminé. Son spectre azote montre un pic à  $197,5 \pm 2,5$  ppm qui se déplace vers  $208 \pm 6$  ppm en présence du 30% du cholestérol (Tableau 10). Par contre, le contrôle CRAC\_1 qui porte aussi le même marquage ne montre presque aucune différence au niveau du déplacement chimique de l'azote 15 étant

reconstitué dans des membranes contenant ou pas 30% de cholestérol (Voir Figure 61 et Tableau 10).

Dans POPC, le CRAC\_1 et sa variante marqués sur une seule position en  $^{15}\text{N}$  adoptent une seule orientation dont le vecteur NH de la liaison peptidique est approximativement parallèle à la normale de la membrane.

L'acide aminé marqué  $^{15}\text{N}$  (Leu8) est localisé au sein du motif CRAC qui a une localisation pré-transmembranaire (Abad et al., 2009) alors si ce motif se structure en hélice  $\alpha$  comme la séquence transmembranaire modèle (déplacement chimique de  $^{15}\text{N} > 200$  ppm (Bechinger and Sizun, 2003)) cela veut dire qu'il est orienté de la même façon que cette séquence même il se peut qu'il soit une extension de la TM (Voir Figure 59, respectivement a et b). D'après nos études de structure secondaire, surtout le FTIR, le CRAC\_1 est majoritairement structuré en hélice  $\alpha$  ce qui peut renforcer le modèle présenté dans la Figure 59.



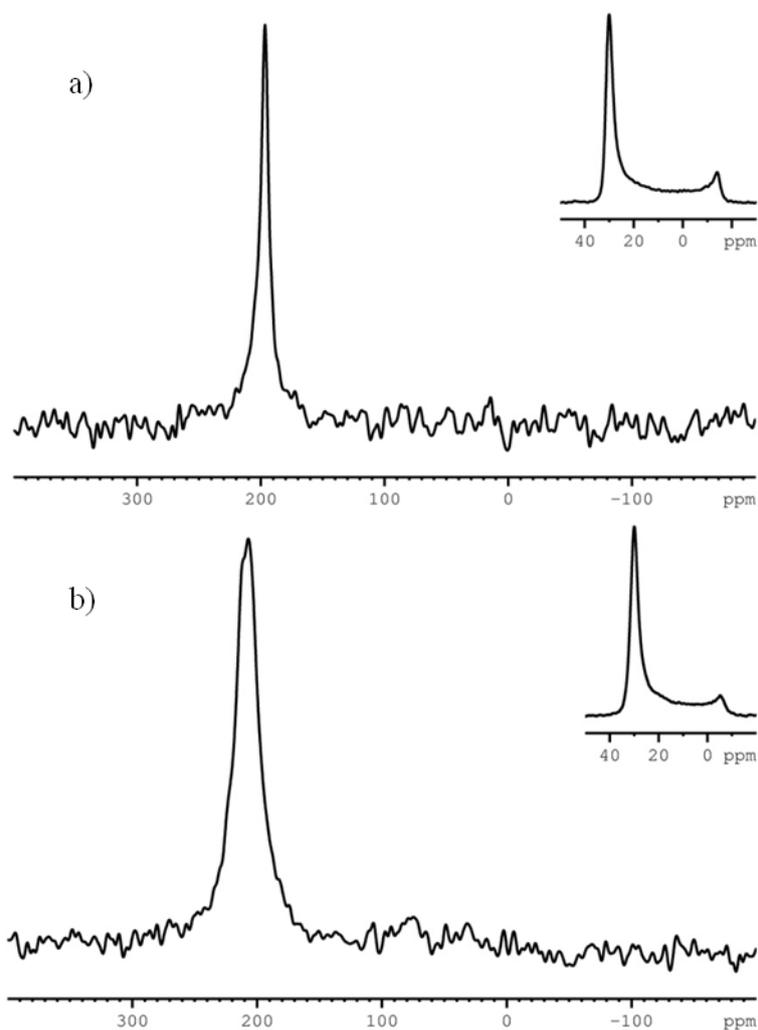
**Figure 59** : modèle montrant les orientations potentielles du motif CRAC par rapport à la séquence transmembranaire. La membrane est présentée en noir, les hélices sont présentées sous formes des rectangles : la séquence transmembranaire en bleu et le motif CRAC en vert. a) La liaison N-H de Leu 8 du motif CRAC est orientée à peu près de la même façon que les N-H de la TM modèle. b) Le motif CRAC forme une extension de la TM modèle.

L'orientation du CRAC\_1 contrôlé reconstitué dans POPC montre une similarité par rapport à celle du CRAC\_1 et sa variante mais avec un petit changement au niveau des angles entre la liaison NH peptidique et la normale de la membrane (différentes valeurs du déplacement chimique mais qui sont toutes aux alentours de 200 ppm).

En présence de 30% de cholestérol, les pics des trois peptides deviennent plus larges par rapport à ceux obtenus dans POPC (voir Tableau 10). Cela veut dire que les trois peptides ont une distribution plus large d'orientations. Par contre, vu les valeurs du déplacement chimique, nous

pouvons dire que CRAC\_1 et sa variante subissent une réorientation au niveau des angles dû à la présence du cholestérol, ce qui n'est pas le cas du CRAC\_1 contrôle qui garde toujours la même valeur.

Ces résultats prometteurs nous ont poussés à étudier la deuxième catégorie des peptides contenant le domaine transmembranaire naturelle de la gp41.



**Figure 60 : Spectres RMN 15N du CRAC\_1. 10mg du peptide sont reconstitués avec un ratio P/L = 1/50 dans : a) POPC, b) POPC/Cholestérol 30%. A gauche spectre <sup>15</sup>N découpé <sup>1</sup>H et à droite spectre <sup>31</sup>P découpé <sup>1</sup>H.**

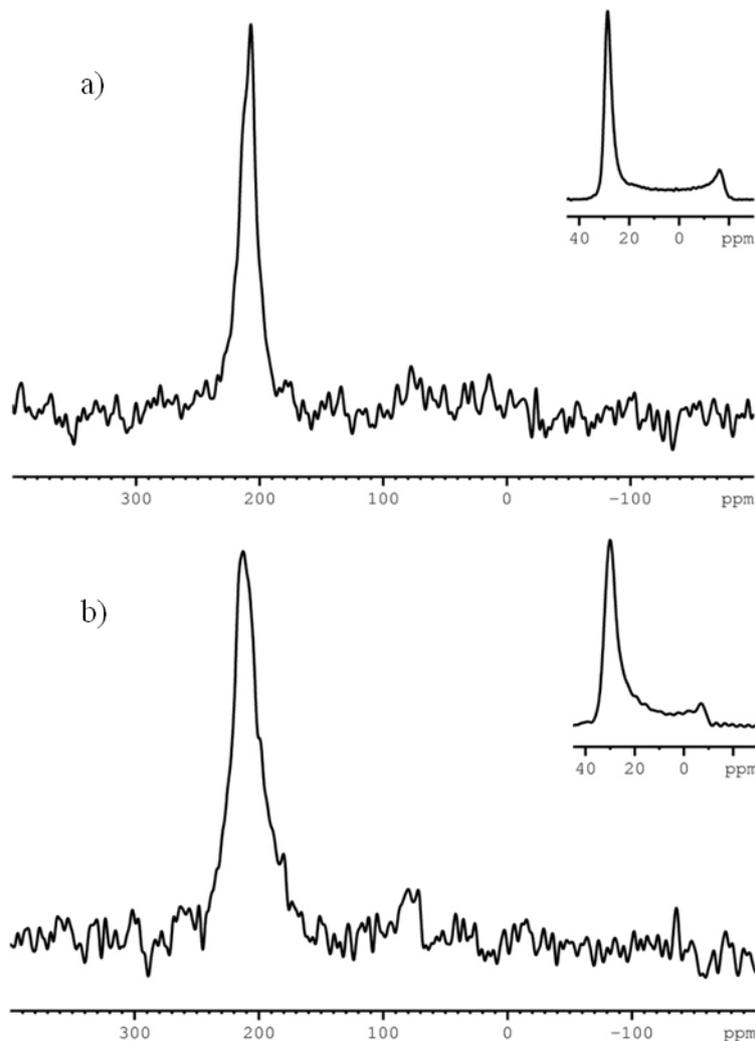


Figure 61 : Spectres RMN 15N du CRAC\_1 Contrôle. 10mg du peptide sont reconstitués avec un ratio P/L = 1/50 dans : a) POPC, b) POPC/Cholestérol 30%. A gauche spectre <sup>15</sup>N découplé <sup>1</sup>H et à droite spectre <sup>31</sup>P découplé <sup>1</sup>H.

Déplacement chimique de 15N (ppm)					
CRAC_1		CRAC_1 Variante		CRAC_1 Ctrl	
POPC	POPC/Chol 30%	POPC	POPC/Chol 30%	POPC	POPC/Chol 30%
196,5 ± 2,2	208 ± 6,5	197,5 ± 2,5	208 ± 6	209 ± 4,5	211 ± 7
Largeur du pic à mi-hauteur (ppm)					
12,5	25	15	30	20	30

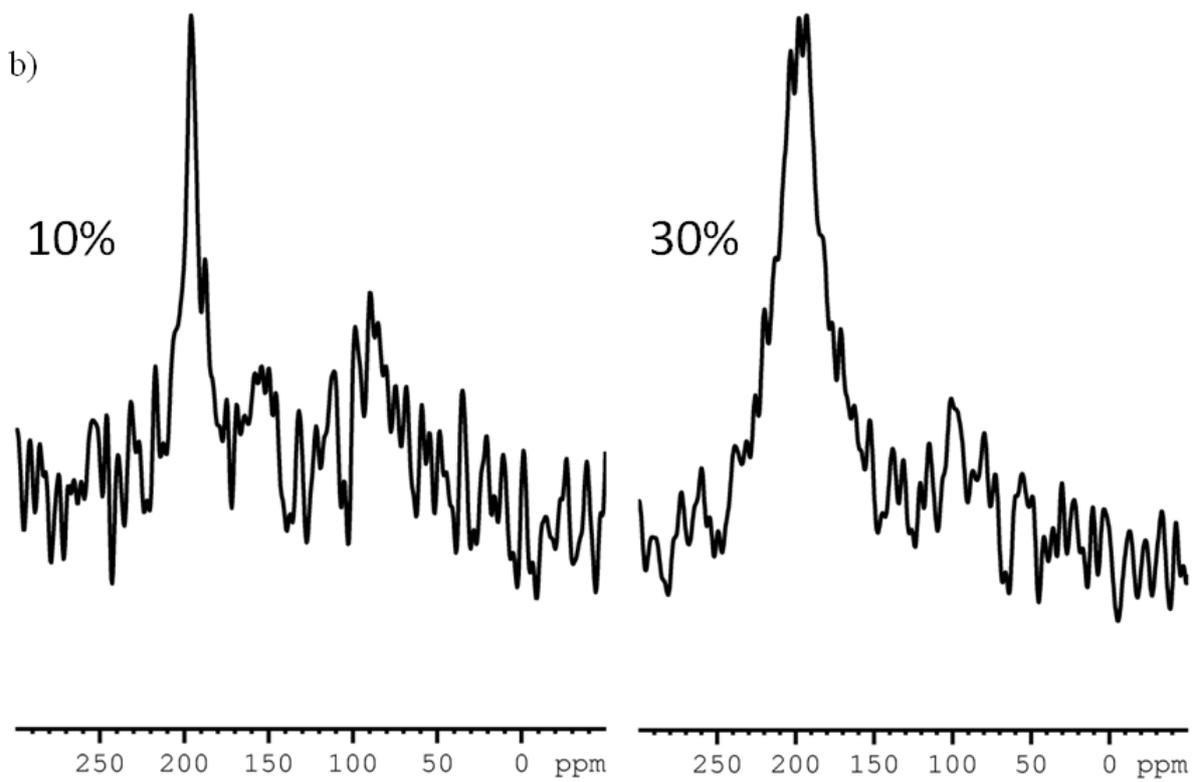
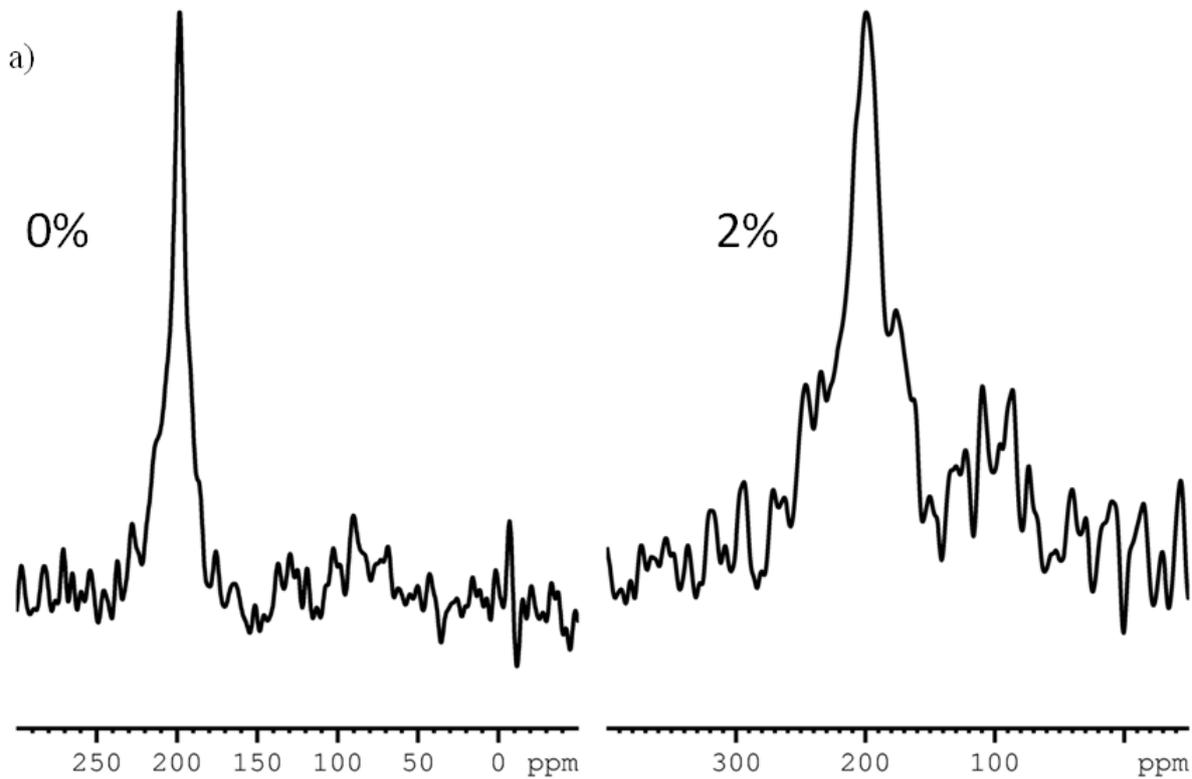
Tableau 10 : Tableau récapitulatif des spectres azote de la catégorie 1 des peptides.

### 5.3.3.A.b. CRAC\_2 (avec la séquence naturelle de la gp41)

Deux peptides ont été conçus avec la séquence transmembranaire naturelle de la gp41. Le premier contient le motif CRAC (-LWYIK-) marqué soit au niveau de  $^{15}\text{N}$ -leucine 8 et  $^2\text{H}$ -Alanine 29 soit au niveau de  $^{15}\text{N}$ -Isoleucine 11 et le deuxième correspond à un contrôle négatif (les acides aminés du CRAC sont randomisés : -LKWYI-) marqué au niveau de  $^{15}\text{N}$ -Leucine 8.

D'après la Figure 62 (a et b) et Tableau 11, nous observons que le spectre  $^{15}\text{N}$  découplé  $^1\text{H}$  du CRAC\_2, marqué en  $^{15}\text{N}$  au niveau de la Leu8 et reconstitué dans des bicouches orientées de POPC, montre un seul pic fin autour de  $199,5 \pm 1,5$  ppm. Les autres spectres correspondant au CRAC\_2 reconstitué dans POPC/cholestérol (concentrations croissantes) montrent également le même pic ayant presque le même déplacement chimique. Le marquage avec  $^{15}\text{N}$ -Leu8 montre que le CRAC\_2 adopte la même orientation et donc les mêmes angles tilt et pitch quel que soit la concentration du cholestérol. La largeur de pic à la mi-hauteur conserve la même valeur pour une concentration de cholestérol <10% mais elle devient deux fois plus importante lorsque cette concentration vaut 30%.

Le marquage de ce peptide au niveau d'Ile11 donne les mêmes résultats du premier marquage sauf que les pics sont plus fins (Voir Figure 62.b et Tableau 11)



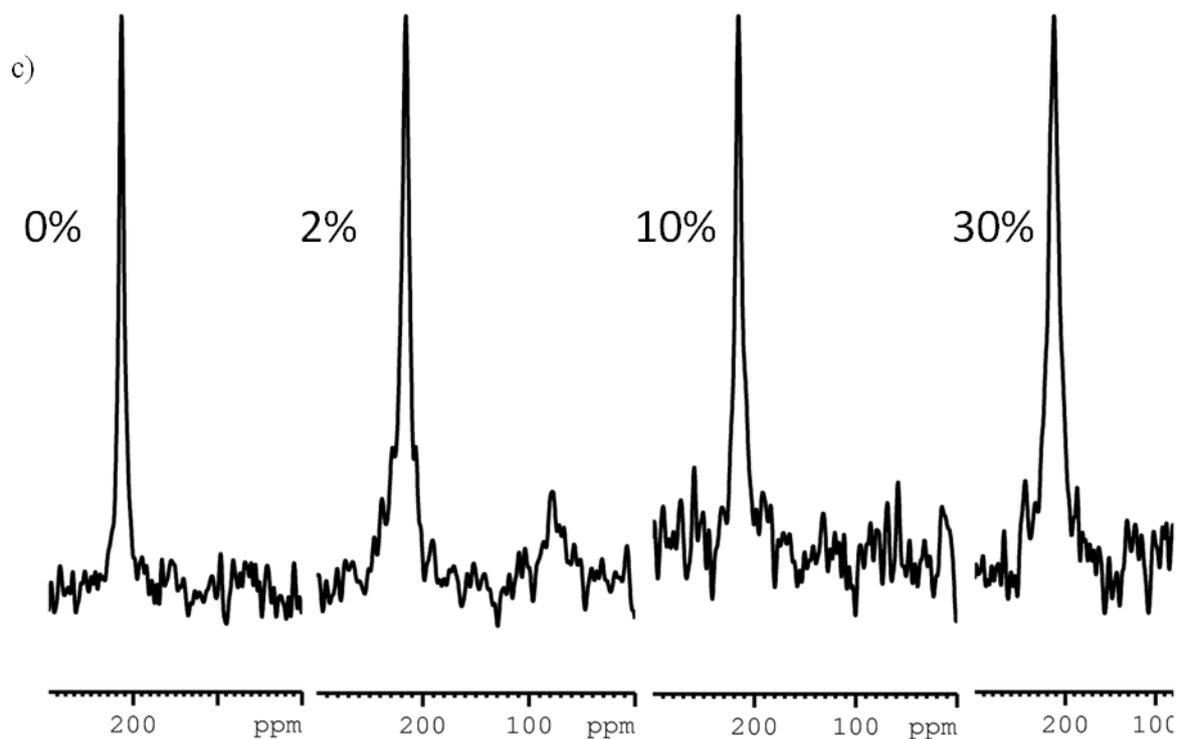


Figure 62 : Spectres 15N-RMN du CRAC\_2. Le CRAC\_2 est marqué au niveau a) et b) Leu8 (15N) ou c) Ile11 (15N). 10mg des peptides sont reconstitués dans avec un ratio P/L = 1/50 dans POPC et 0%, 2%, 10% et 30% du cholestérol.

Déplacement chimique 15N (ppm)								
CRAC_2 (15N-Leu8)					CRAC_2 (15N-Ile11)			
0%	2%	10%	30%		0%	2%	10%	30%
199,5 ± 1,5	197,6 ± 5	195,5 ± 2,5	197,5 ± 9		213,5 ± 2,2	216 ± 2,2	215,4 ± 2,3	213 ± 4,3
Largeur du pic à mi-hauteur								
15	20	17,5	40		7	9,5	8	14

Tableau 11 : Tableau récapitulatif des spectres azote de la catégorie 2 des peptides.

### 5.3.3.B. Spectroscopie de RMN sur deutérium

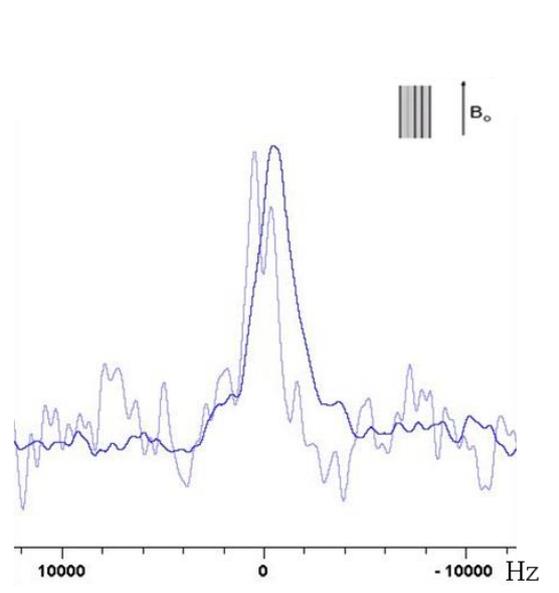
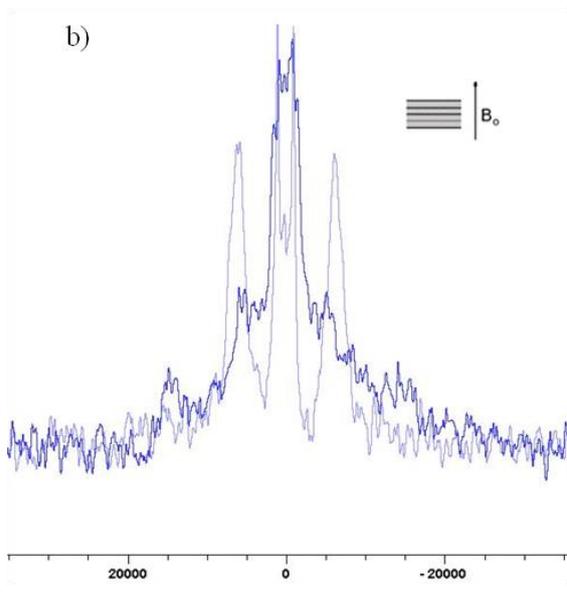
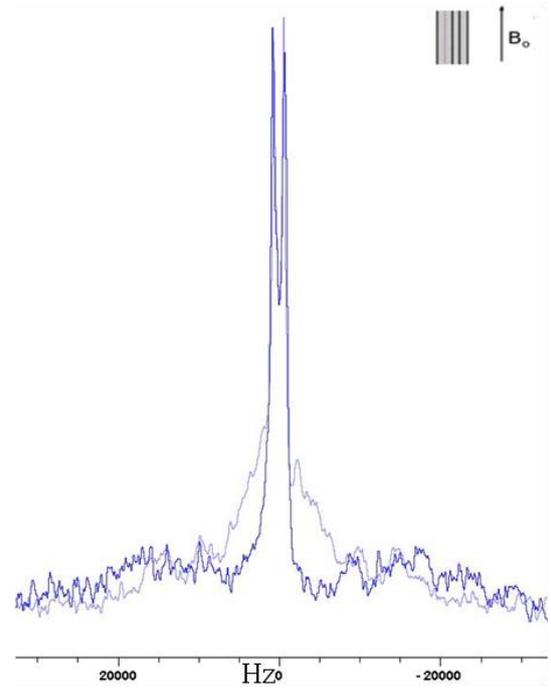
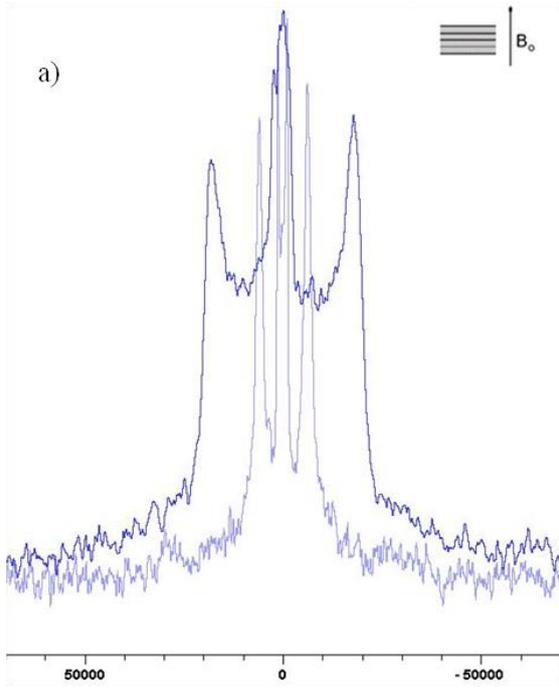
#### 5.3.3.B.a. CRAC\_1

Dans la position où la normale de la bicouche est parallèle au champ magnétique, le spectre deutérium du CRAC\_1 (Figure 63.a, à gauche), marqué au niveau d'alanine 24 est reconstitué dans des bicouches orientées de POPC, montre un éclatement quadripolaire bien défini de 12 kHz à température ambiante. Au contraire des spectres orientés 15N, l'orientation de l'hélice transmembranaire ne peut pas être déduite d'un seul spectre 2H, mais l'allure du spectre montre que le peptide est bien aligné par rapport au champ magnétique extérieur (Aisenbrey et al., 2007). Tandis que lorsque ce peptide est reconstitué dans des bicouches formées par POPC/Cholestérol 30%, son spectre montre une allure qui ressemble plus à celle de la poudre (Bloom et al., 1981) dont l'éclatement quadripolaire est compris entre 0 et 30 kHz. Ceci correspond à une distribution d'orientations des liaisons C $\alpha$ -C $\beta$  reflétant le fait que le peptide n'est pas aussi bien aligné de façon homogène par rapport au champ magnétique extérieur. Cette différence de la dynamique du CRAC\_1 au niveau de son Ala24 peut être due au changement des propriétés de la bicouche. En effet, en présence de 30% cholestérol et à température ambiante (>température de transition) la membrane devient plus épaisse (Khelashvili and Harries, 2013).

Après inclinaison par 90° de la bicouche (POPC ou POPC/cholestérol) (Figure 63.a, à droite), les spectres des deux échantillons montrent un aplatissement des pics. Alors le peptide n'a pas une diffusion rotationnelle rapide autour la normale (échelle de temps de 20  $\mu$ s (50 kHz<sup>-1</sup>)). Le pic au centre de tous les spectres qui montre un éclatement quadripolaire étroit (gamme de 2 kHz) correspond principalement aux deutérons de l'eau (Aisenbrey and Bechinger, 2004). Les deux autres peptides (variante et contrôle de CRAC\_1) montrent des spectres deutérium similaires à ceux obtenus par CRAC\_1 (Voir Figure 63. b et c, et Tableau 10).

#### 5.3.3.B.b. CRAC\_2

Le peptide CRAC\_2 marqué au niveau <sup>2</sup>H-Ala 29 donne des spectres semblables à ceux obtenus dans le cas de CRAC\_1. Sauf que dans la position inclinée de la bicouche et en présence de 30% de cholestérol, le spectre montre un éclatement quadripolaire qui est égal à celui obtenu dans la position normale de la bicouche (Voir Figure 63.d, et Tableau 12). Alors CRAC\_2 reconstitué dans une bicouche contenant 30% de cholestérol a une diffusion rotationnelle latérale autour de l'axe de l'hélice relativement lente.



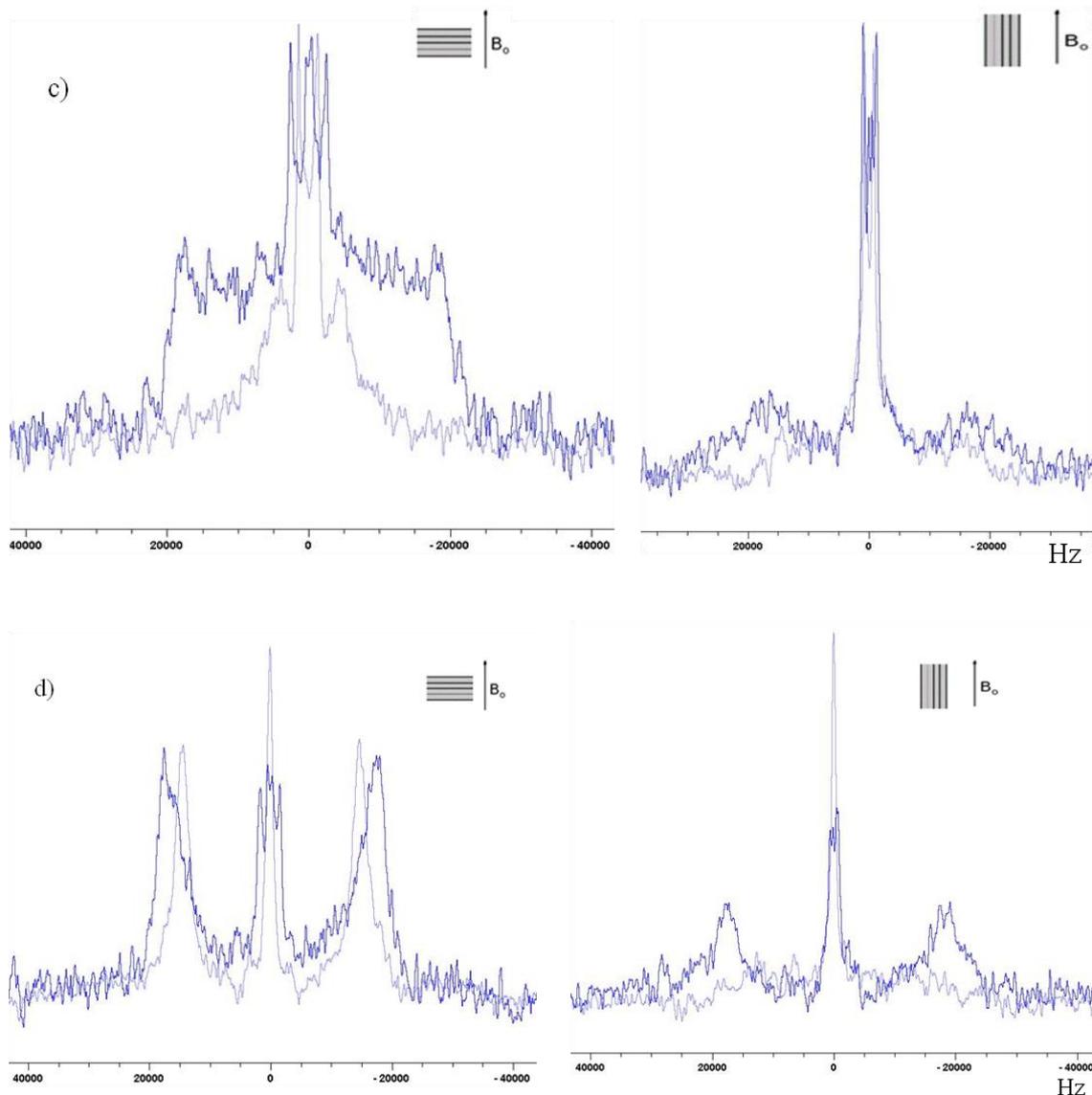


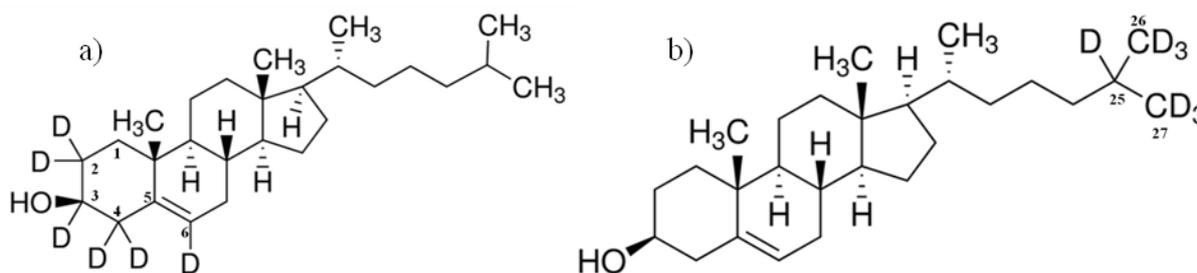
Figure 63 : Spectres RMN de deutérium. a) CRAC\_1 marqué [3,3,3- $^2\text{H}_3$ ]alanine-24. b) CRAC\_1 variante marqué [3,3,3- $^2\text{H}_3$ ]alanine-24. c) CRAC\_1 contrôle marqué [3,3,3- $^2\text{H}_3$ ]alanine-24. d) CRAC\_2 marqué [3,3,3- $^2\text{H}_3$ ]alanine-29. Les peptides sont reconstitués dans des bicouches formées par POPC (bleu claire) ou par POPC/Cholestérol 30% (bleu marine). Les spectres à gauche correspondent à l'orientation perpendiculaire au champ magnétique de la bobine et ceux à droite l'orientation parallèle au champ magnétique.

	Splitting de $2\text{H}$ (KHz)							
	CRAC_1		CRAC_1 Variante		CRAC_1 Ctrl		CRAC_2	
	POPC	POPC/Chol 30%	POPC	POPC/Chol 30%	POPC	POPC/Chol 30%	POPC	POPC/Chol 30%
Normal	12	0-30	12	0-30	9	0-35	30	35
Tilted	-	-	-	-	-	-	-	35

Tableau 12 : Tableau récapitulatif des spectres deutérium des catégories 1 et 2 des peptides.

### 5.3.4. L'effet de CRAC\_2 sur le cholestérol

Les deux catégories des peptides n'ont pas montré aucun changement de dynamique (spectres deutérium) spécifique vis-à-vis du cholestérol. De plus, CRAC\_2 n'a pas montré aucun changement au niveau de son orientation au vis-à-vis du cholestérol (spectres azote). Ceci nous a amené à tester l'effet des peptides sur le cholestérol. Pour cela nous avons eu recours à la même technique (RMN de solide sur le deutérium) mais en utilisant du cholestérol marqué soit en positions 2, 2, 3, 4, 4 et 6 (cholestérol-d6) ; soit en positions 25, méthyle-26 et méthyle-27 (cholestérol-d7). Le cholestérol-d6 a été choisi pour cette étude car le site marqué en 2H est au niveau du groupement hydroxyle qui est probablement localisé à proximité des têtes polaires des phospholipides (Khelashvili and Harries, 2013). L'autre type du cholestérol a été choisi pour tester si le changement, s'il y en a, atteint toute la molécule du cholestérol. En effet, les groupements méthyle sont probablement localisés au cœur des membranes (Khelashvili and Harries, 2013).



**Figure 64 : Structure du cholestérol marqué en deutérium. a) en positions 2,2,3,4,4,6 et en positions b) 25,26,27.**

Cette étude n'était menée que sur la deuxième catégorie des peptides (CRAC\_2). Tout d'abord, le CRAC\_2 est reconstitué dans des bicouches formées par POPC/Cholestérol-d6 30% pour mimer le plus possible la membrane virale. La comparaison des spectres  $^2\text{H}$  du cholestérol-d6, enregistrés à température ambiante, présentée dans la Figure 65 ne montre aucune différence en absence ou en présence du CRAC\_2. En effet, la bicouche formée par POPC/cholestérol-d6 30%, contenant ou pas 2% du CRAC\_2, présente des séries des éclatements quadripolaires (Marsan et al., 1999; Tiburu et al., 2004). La première qui est la plus externe correspond à l'éclatement quadripolaire des deutériums 3 et, 2 et 4 axiaux (96 kHz). La deuxième présentant une valeur de 67 kHz correspond à éclatement quadripolaire des deutériums 2 et 4 équatoriaux. Le pic au centre correspond à l'éclatement des deutériums de l'eau.

Ensuite, nous avons voulu déterminer l'influence de la température sur la transition de phase de la bicouche formée par POPC/Cholestérol-d6 30% contenant ou non le peptide CRAC\_2. En effet, c'est possible que l'interaction entre le CRAC et le cholestérol dépende de la phase des lipides.

Des spectres  $^2\text{H}$  ont été enregistrés sur les mêmes échantillons en changeant la température de 268K jusqu'à 318K. Tous les spectres (sans ou avec le peptide) montrent la même allure que celui présenté dans la Figure 65 sauf que les éclatements quadripolaires des deutériums 2, 3 et 4 diminuent en fonction de la température (voir Figure 66). Ceci indique qu'il n'y pas une transition de phase ce qui est en accord avec le diagramme des phases publié par l'équipe de E. Melo (de Almeida et al., 2003). En comparant, la variation des éclatements de la bicouche contenant le CRAC (2%) à celle qui ne le contient pas, nous n'observons aucune différence au niveau des deutériums 2 et 4 équatoriaux. Ce qui n'est pas le cas des deutériums 3 et, 2 et 4 axiaux. En fait, nous trouvons une différence de 2 kHz à basse température (<288 K) et de 3,5 kHz à haute température (>298 K). Or ces valeurs proviennent d'un seul échantillon par condition (sans ou avec le peptide) alors nous ne sommes pas sûrs si cette différence est réelle et significative. Cette différence peut être due à une hydratation qui n'est pas la même dans les deux types de bicouches. En effet, l'hydratation est une étape importante dans la préparation des échantillons orientés (Bechinger and Weik, 2003).

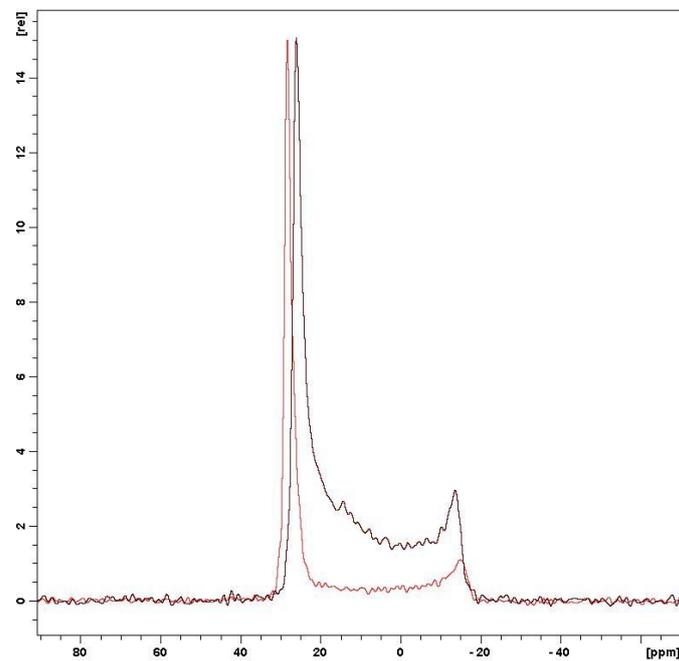
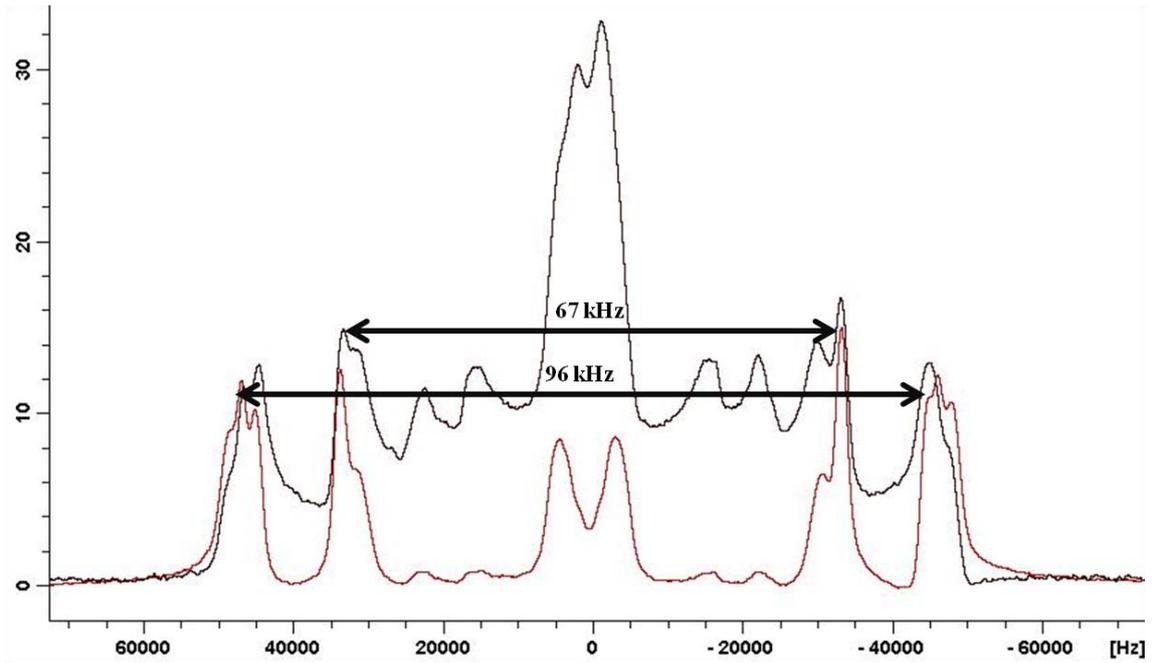
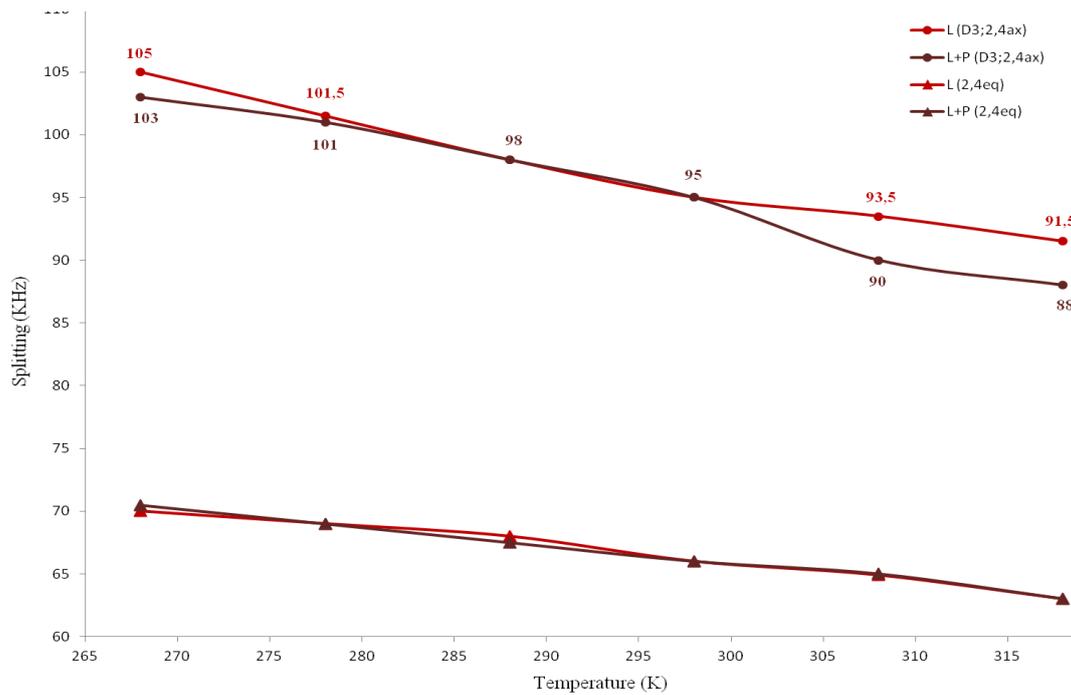


Figure 65 : Spectres RMN de solide : deutérium du cholestérol-d6, en haut, et  $^{31}\text{P}$  découplé  $^1\text{H}$ , en bas. En marron, CRAC\_2 marqué  $[3,3,3\text{-}^2\text{H}_3]$ alanine-29 reconstitué une bicouche formée par POPC/Cholestérol-[2,2,3,4,4,5,6] 30%. En rouge, bicouche formée par POPC/Cholestérol-[2,2,3,4,4,5,6] 30%. Les spectres correspondent à l'orientation de la membrane où sa normale est parallèle au champ magnétique.



**Figure 66 : Influence de la température sur les éclatements quadripolaires. En marron (L+P), CRAC\_2 reconstitué dans une bicouche formée par POPC/Cholestérol-[2,2,3,4,4,5,6] 30%. En rouge (L), une bicouche formée par POPC/Cholestérol-[2,2,3,4,4,5,6] 30%. Les éclatements des deutériums 3 et 2,4 axiaux sont présentés par les cercles. Les éclatements des deutériums 2,4 équatoriaux sont présentés par les triangles.**

Vu ces différentes minimales, nous nous sommes décidés de diminuer la concentration du cholestérol dans la bicouche. Cela nous permet d'observer plus facilement l'interaction potentielle entre le motif CRAC et le cholestérol. En effet, si une proportion faible des molécules de cholestérol interagit avec le peptide leur signal RMN (ou le changement de ce signal en cas d'échange) sera trop faible par rapport au signal des autres molécules n'interagissant pas avec le cholestérol. De plus, une concentration faible en cholestérol assure une homogénéité de distribution des composants au sein de l'échantillon étudié. Autrement dit une concentration faible en cholestérol évite la formation des de plusieurs phases co-existantes dans la bicouche (Silvius, 2003).

Le peptide CRAC\_2 et son contrôle sont reconstitués dans des bicouches formées par POPC/Cholestérol-d6 10%, toujours dans le même ratio P/L = 1/50. La Figure 67 montre le spectre deutérium du cholestérol-d6 présent dans les bicouches formée par POPC avec (respectivement bleu clair et foncé) ou sans peptide (vert). La présence du peptide dans, quel que soit sa nature (le CRAC\_2 ou son contrôle), entraîne une légère augmentation des éclatements quadripolaires par rapport à ceux obtenus par la bicouche sans peptide (Voir Tableau 13). Cette augmentation n'est pas due à une interaction spécifique entre le motif CRAC et le cholestérol parce que le peptide contrôle se comporte de la même manière que le peptide contenant le motif CRAC vis-à-vis de la bicouche contenant du cholestérol. Mais elle peut être due simplement à un changement d'organisation des lipides provoqué par le peptide.

Un autre cholestérol marqué a été utilisé également, le cholestérol-d7, les spectres ne montrent pas une grosse différence avec ou sans CRAC\_2. Le peptide CRAC\_2 et son contrôle sont reconstitués dans des bicouches formées par POPC/Cholestérol-d7 10% (P/L = 1/50). La bicouche formée par POPC/cholestérol-d7 10% montre un éclatement quadripolaire central (4 kHz) qui correspond à celui des deutériums du carbone 25 (Engelbrecht et al., 2012), en rajoutant le CRAC\_2 cet éclatement diminue légèrement (3,5 kHz). Les éclatements externes correspondent aux deutériums des carbones 26 et 27 (Engelbrecht et al., 2012), ils valent 48,5 et 51,5 kHz respectivement dans les bicouches contenant ou pas le CRAC\_2. Cette différence peut être significative mais cette expérience n'était pas menée sur le CRAC\_2 contrôle.

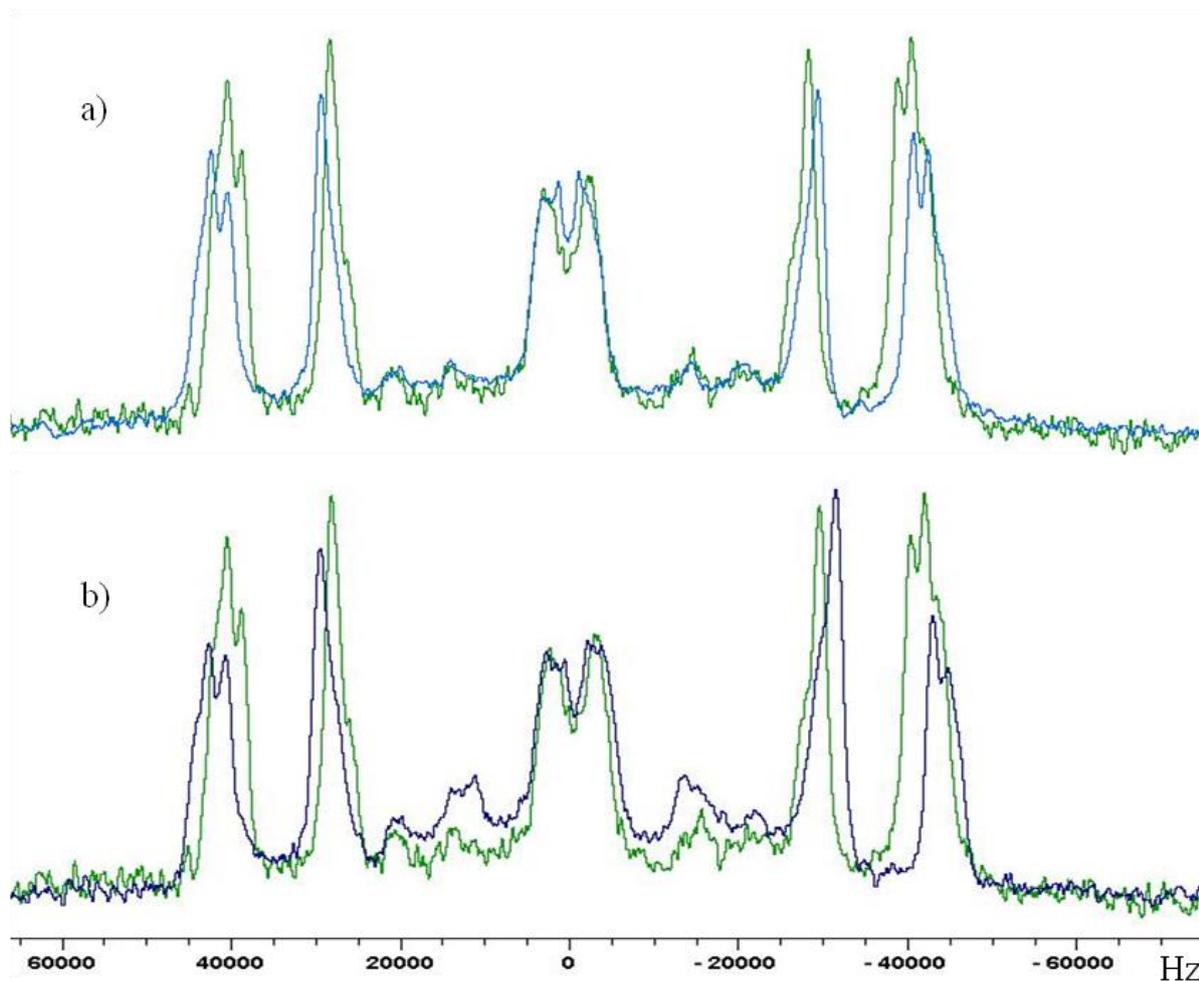
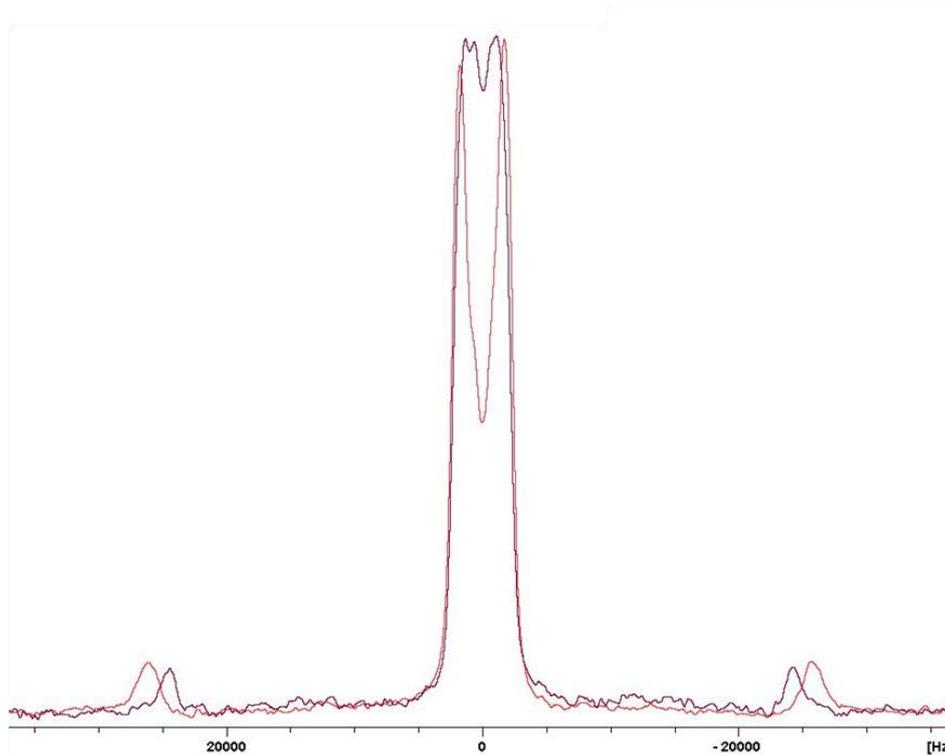


Figure 67 : Spectres RMN de deutérium du cholestérol-d6. a) CRAC\_2 (bleu clair) et b) Contrôle (bleu foncé) reconstitués dans une bicouche formée par POPC/Cholestérol-[2,2,3,4,4,5,6] 10%. En vert, une bicouche formée par POPC/Cholestérol-[2,2,3,4,4,5,6] 10%. Les spectres correspondent à l'orientation de la membrane où sa normale est parallèle au champ magnétique.

Chol_d6 10%	CRAC_2	CRAC_2 Ctrl	L
$^2\text{H} : 3;2,4_{ax}$	86,5	87	85,5
$^2\text{H} : 2,4_{eq}$	60	60	57,5

Tableau 13: Tableau récapitulatif des éclatements quadripolaires de cholestérol-d6 10%.



**Figure 68 : Spectres RMN de deutérium du cholestérol-d7. En marron, CRAC\_2 reconstitué dans une bicouche formée par POPC/Cholestérol-[25, 26, 27] 10%. En rouge, bicouche formée par POPC/Cholestérol-[25, 26, 27] 10%. Les spectres correspondent à l'orientation de la membrane où sa normale est parallèle au champ magnétique de la bobine.**

Chol_d7 10%	CRAC_2	L
C25	3,5	4
C26 et C27	48,5	51,5

**Tableau 14 : Tableau récapitulatif des éclatements quadripolaires de cholestérol-d7 10%.**

### **Comparaison entre CRAC 1 et CRAC 2 :**

D'après les résultats obtenus concernant la topologie des deux peptides, nous remarquons, tout d'abord, que le CRAC\_1 se réoriente spécifiquement vis-à-vis du cholestérol dans les conditions utilisées dans nos expériences. En effet le CRAC\_1 et sa variante montre un changement du déplacement chimique de l'azote (Leu8) en rajoutant du 30% cholestérol dans la bicouche tandis que le contrôle négatif ne montre aucun changement.

Par contre le CRAC\_2 ne présente aucun changement d'orientation, son déplacement chimique d'azote (Leu8 et Ile11) ne varie pas. En effet, d'après nos gels, nous avons remarqué que le CRAC\_2 est potentiellement sous sa forme trimérique voire des dimères de trimères (Figure 53). Cette oligomérisation est due au fait que le domaine transmembranaire de la gp41 contient des motifs hautement conservés d'oligomérisation comme le GxxxG (Senes et al., 2004) et un acide aminé polaire au centre de ce domaine (l'arginine) (Curran and Engelman, 2003). L'association de ces hélices transmembranaires peut entrer en compétition avec leur interaction avec le cholestérol, c'est le cas du domaine C99 du précurseur amyloïde (Song et al., 2013). C'est possible qu'il se trouve d'autres facteurs qui peuvent moduler cette interaction surtout que le CRAC\_2 contient la TM naturelle de la gp41, par exemple cette interaction nécessite que le cholestérol soit figé dans des microdomaines contenant de la sphingomyéline (Lorizate et al., 2008). Toutes ces raisons peuvent expliquer en même temps pourquoi la dynamique du cholestérol marqué par deutérium n'était pas influencée par l'ajout du peptide CRAC\_2.

La dynamique de CRAC\_1 et CRAC\_2 était similaire vis-à-vis du cholestérol. En fait, tous les peptides ont montré des éclatements quadripolaires similaires et bien résolus dans des bicouches formées par POPC mais ces éclatements deviennent plus larges, moins résolus et qui ressemblent au spectre d'une poudre dans des bicouches contenant 30% du cholestérol. Ceci est simplement dû à la réduction de la fluidité des bicouches formées par POPC.

## 5.4. Conclusion et perspectives

Pour mieux comprendre l'interaction entre le cholestérol et le motif CRAC localisé dans la séquence pré-transmembranaire de la gp41, nous avons eu recours à la RMN de solide sur des échantillons orientés. Cette technique apporte des informations sur la topologie et la dynamique des protéines ou des lipides. Pour cela, deux catégories de peptides ont été conçues de la façon suivante : la première catégorie correspond au motif CRAC avec une séquence transmembranaire modèle et la deuxième catégorie correspond au CRAC avec la TM naturelle de la gp41.

Tout d'abord, nous avons étudié, par SDS-PAGE, l'oligomérisation des peptides. Nous avons trouvé que les CRAC\_1 sont probablement sous forme de dimères tandis que le CRAC\_2 est sous forme d'un trimère même nous avons trouvé une bande qui correspondrait à un dimère de trimère.

Ensuite la structure secondaire était le sujet d'une étude par DC et FTIR. Les peptides reconstitués dans des liposomes de POPC/POPS (3/1) contenant un taux croissant du cholestérol présentent majoritairement une structure en hélice  $\alpha$  par dichroïsme circulaire. Ces résultats sont confirmés par FTIR qui présente quelques avantages par rapport au DC surtout dans le cas des peptides membranaires. Elle n'est pas quantitative alors ne nécessitant pas une concentration précise de peptide, dépendante des tailles des vésicules...

L'influence du cholestérol sur la topologie des peptides était étudiée par spectroscopie de RMN ( $^{15}\text{N}$ ). Le CRAC\_1 a montré une réorientation vis-à-vis spécifique du cholestérol dans les bicouches lipidiques par contre le CRAC\_2 n'a rien montré même en changeant la concentration du cholestérol. Ceci peut être expliqué par le fait que la TM naturelle de la gp41 s'oligomérisent d'une manière différente et plus importante que la TM modèle favorisant l'association des hélices et non pas l'interaction du cholestérol.

Cette différence (spécificité d'interaction avec le cholestérol) n'était pas observée par spectroscopie de RMN ( $^2\text{H}$ ) qui a donné des spectres ayant des allures similaires entre CRAC\_1 et CRAC\_2. De même, cette technique a donné des résultats similaires pour le cholestérol marqué ( $^2\text{H}$ ) dans le cas du CRAC\_2. En effet, ce cholestérol ne change pas d'orientation en rajoutant ce peptide même en changeant la concentration du cholestérol ou la température d'acquisition des spectres RMN.

Vu cette différence de comportement entre CRAC\_1 et CRAC\_2 notamment au niveau de la topologie, l'influence de CRAC\_1 sur le cholestérol sera l'objet d'une étude par spectroscopie de RMN ( $^2\text{H}$ ) semblable à celle faite avec CRAC\_2. De plus, la sphingomyéline va être introduite dans la bicouche pour pouvoir éventuellement renforcer l'interaction des peptides avec le motif

CRAC. En effet, le cholestérol et la sphingomyéline interagissent pour former des micro-domaines qui ont des propriétés différentes à celles de lipides qui les entourent.

## 6. Chapitre 6 : Conclusion générale

### 6.1. Expression et étude structurale des polypeptides contenant des épitopes issus des glycoprotéines d'enveloppe du VIH-1

Quelques régions stables ont été découvertes sur les glycoprotéines d'enveloppe (gp120 et gp41) du VIH-1 contre lesquelles des patients produisent des anticorps neutralisants. Les épitopes les plus prometteurs se trouvent dans la région proche de la membrane et sont probablement exposés durant la fusion des membranes cellulaire et virale. Alors que les peptides isolés à partir de cette région n'ont pas permis d'induire de réaction immunogène neutralisante, des études suggèrent que la membrane lipidique joue un rôle dans la structuration des antigènes et dans la réponse immunogénique retrouvée chez ces patients.

C'est pourquoi, nous nous proposons d'étudier ces régions dans un environnement lipidique par des approches structurales. Trois séquences ont été choisies pour leur proximité à la membrane lors de l'étape de fusion : une issue de la gp120 et deux issues de la gp41, nommées respectivement gp120, gp41 et fp auxquelles est ajoutée en C- terminale une séquence transmembranaire (TM). L'étude de ces polypeptides requiert leur production par surexpression bactérienne en fusion avec la protéine TAF12, leur purification via des méthodes biochimiques (extraction, purification HPLC et clivage) et finalement la reconstitution des polypeptides exprimés dans des liposomes. Leur repliement et leur dynamique seront étudiés dans différents environnements lipidiques en utilisant des techniques de spectroscopie.

Les études qui sont menées sur ces polypeptides, exigeaient des quantités relativement importantes, d'où la nécessité de leur production par voie bactérienne qui peut assurer ces quantités à moindre coût.

Les conditions d'expression et de purification des trois séquences ont été établies. Elles sont majoritairement localisées dans la fraction insoluble (membranes et corps d'inclusion). Cette fraction est solubilisée dans l'acide formique 75% afin de cliver la protéine de fusion TAF12 au niveau du site de clivage Acide aspartique-Proline situé juste en amont du peptide d'intérêt. Les

polypeptides clivés sont purifiés en une (gp120<sup>TM</sup> et gp41<sup>TM</sup>) ou deux étapes (fp<sup>TM</sup>) par RP-HPLC. Le contrôle par spectrométrie de masse des fractions d'intérêt montrent un ajout d'un multiple de 28 Da à la masse moléculaire attendue; ceci indique que les polypeptides purifiés sont formylés (plusieurs états de formylation). Afin d'éviter cette réaction, un clivage avec l'acide chlorhydrique a été testé mais sans succès, l'acide chlorhydrique n'a pas permis de solubiliser correctement la fraction insoluble et le clivage n'était pas spécifique.

Pour tester l'immunogénicité des polypeptides, deux séries d'immunisation sur des souris ont été faites, l'une pour gp120<sup>TM</sup> et l'autre pour gp41<sup>bTM</sup>. Chaque série comprenait trois groupes en fonction de la solution injectée (polypeptide insérée dans des liposomes, polypeptide seule et un contrôle avec du tampon PBS). Chaque groupe était composé de cinq souris, de deux différentes souches (3 BALB/c et 2 C57BL/6). Les injections intrapéritonéales ont été effectuées tous les 15 jours avec un volume de 200 µl (100 µl d'antigène + 100 µl d'adjuvant). En parallèle, des prélèvements de sérum ont été effectués pour vérifier la présence d'anticorps (tests d'ELISA). Les premiers résultats d'ELISA ont montré des résultats contradictoires. Ceci peut être expliqué par l'utilisation d'un adjuvant qui, grâce à sa nature huileuse, peut détruire les liposomes. De plus, lors de la préparation des échantillons nous avons observé des agrégats ce qui nous a amené à essayer une autre technique de préparation.

Pour leur reconstitution dans des liposomes, les polypeptides sont solubilisés dans un détergent, en parallèle, le film lipidique obtenu après évaporation des solvants organiques est réhydraté dans le même détergent. Ensuite, les polypeptides et les lipides sont mélangés avec différents ratios (1/30, 1/70 et 1/100). Pour éliminer le détergent, une étape de dialyse dans un tampon phosphate à pH 7 s'est effectuée. Le dialysat est centrifugé durant 20 min à 11000 g pour éliminer tous les agrégats.

La structure secondaire de ces polypeptides insérés dans des liposomes de différentes compositions a été étudiée par dichroïsme circulaire. Ils adoptent majoritairement une structure en hélice  $\alpha$  dans un milieu micellaire. La gp41<sup>TM</sup> conserve cette structure une fois reconstituée dans des liposomes quel que soit leur composition, par contre la gp120<sup>TM</sup> et gp41<sup>bTM</sup> deviennent moins structurées surtout le fp<sup>TM</sup>. Les calculs des structures secondaires par des outils bioinformatiques pour ce type d'échantillon (polypeptides membranaires insérés dans des liposomes) n'est que qualitatif. En effet, la distribution non uniforme des chromophores dans les membranes, la taille des liposomes, la différence entre les constantes diélectriques des lipides et du milieu aqueux et la ration lipide/ratio sont des facteurs qui faussent le signal du dichroïsme circulaire.

Ensuite nous avons voulu étudier les structures atomiques des polypeptides par RMN qui nécessite leur marquage ( $^{13}\text{C}$  et  $^{15}\text{N}$ ) dans un milieu minimum. Le fpTM uniformément marquée ( $^{13}\text{C}/^{15}\text{N}$ ) était la première à être surexprimée et purifiée. Les premières d'expériences de RMN solide MAS ont déjà été effectuées sur le polypeptide lyophilisé (poudre), la résolution n'est pas assez bonne pour faire une attribution des signaux  $^{13}\text{C}$  et  $^{15}\text{N}$ . En parallèle, la structure de la fpTM dans un milieu micellaire (DPC : n-dodecylphosphocholine, 160 mM) ou inséré dans des vésicules formées par POPC/POPS (3 /1) est étudiée par respectivement RMN en solution et RMN de solide avec rotation à l'angle magique. L'attribution des signaux  $^{13}\text{C}$ ,  $^1\text{H}$  et  $^{15}\text{N}$  est en cours.

A plus long terme, l'analyse comparative entre le potentiel d'immunogénicité et les résultats structuraux obtenus seront confrontés dans le but de mieux comprendre l'entrée du VIH-1 dans la cellule hôte et par la suite améliorer la conception d'un vaccin efficace dirigé contre le VIH-1.

## 6.2. Etude de l'interaction entre le cholestérol et le CRAC de la gp41

L'entrée du VIH dans la cellule hôte s'effectue grâce aux glycoprotéines d'enveloppe du VIH. La gp120 est responsable de la reconnaissance de la cellule cible et la fixation sur ses récepteurs CD4. Puis cette glycoprotéine subit un changement de conformation et se lie à un corécepteur existant à proximité de la molécule CD4 (CXCR4 pour les lymphocytes T CD4<sup>+</sup> et CCR5 pour les macrophages). Ceci provoque la libération de la gp41, qui se fixe sur la membrane cytoplasmique. Par repli sur elle-même, la gp41 attire l'enveloppe virale vers la membrane cytoplasmique. Puis le peptide de fusion, déjà caché, s'insère dans la membrane plasmique et entraîne la fusion des membranes cellulaire et virale.

Ces membranes sont hautement riches en cholestérol (30% pour la membrane cellulaire et 45% pour la membrane virale). Plusieurs études ont montré que le cholestérol est indispensable l'infection des cellules par le virus ou par la fusion des cellules infectées. Pourtant, le taux d'expression des glycoprotéines virales est intact. En effet, la gp41 renferme une séquence de cinq acides aminés (LWIYK) qui correspond au motif CRAC (Cholesterol Recognition/interaction Amino acid Consensus) identifié par Li et Papadopoulos. D'où les questions qui s'imposent à cet égard, est-ce que le cholestérol interagit directement avec le CRAC ou en changeant les propriétés physico-chimiques des membranes? Ou bien par les deux voies (directe et indirecte) ?

Pour répondre à cette question, nous avons eu recours à la RMN de solide sur des échantillons orientés. Cette technique apporte des informations sur la topologie et la dynamique des protéines ou des lipides. Pour cela, deux peptides ont été conçus de la façon suivante : les deux peptides contiennent le motif CRAC flanqué par quelques acides aminés de la gp41. En C terminal de ce motif, une séquence transmembranaire était rajoutée. Cette séquence correspond soit à la séquence transmembranaire naturelle de la gp41, soit à une succession d'A-L (alanine-leucine) pour rendre ce peptide moins hydrophobe que le premier. Ces constructions sont flanquées de part et d'autre de deux résidus lysine afin d'améliorer leur solubilité. Des peptides contrôles et variantes ont été également conçus.

Les peptides reconstitués dans des liposomes de POPC contenant un taux croissant du cholestérol présentent majoritairement une structure en hélice  $\alpha$  par dichroïsme circulaire et par spectroscopie d'infrarouge à transformée de Fourier.

La topologie des peptides a été étudiée par spectroscopie de RMN de solide sur le noyau  $^{15}\text{N}$ . Les spectres  $^{15}\text{N}$  de RMN montrent que les peptides marqués sur une seule position en  $^{15}\text{N}$  adoptent, dans le POPC, une orientation dont leur vecteur NH est approximativement parallèle à la normale de la membrane. En présence de 30% du cholestérol, les peptides adoptent une distribution d'orientations. Or seulement le CRAC\_1 et sa variante (avec la TM modèle) ont montré une réorientation spécifique vis-à-vis du cholestérol (changement de leur déplacement chimique d'azote dans les bicouches formées par POPC ou par POPC/cholestérol 30% ; par contre le déplacement chimique du CRAC\_1 contrôle n'a pas changé). Cette réorientation n'était pas observée chez le CRAC\_2 (avec la TM naturelle de la gp41) même en changeant la concentration du cholestérol dans les bicouches. Ceci peut être expliqué par les différents états d'oligomérisation de CRAC\_1 et CRAC\_2.

En parallèle, nous avons étudié l'alignement et la dynamique des peptides par autre une méthode, la spectroscopie de RMN de solide sur le noyau  $^2\text{H}$ . Tous les peptides (CRAC\_1, son contrôle et CRAC\_2) ont montré un comportement semblable vis-à-vis à l'ajout du cholestérol dans les bicouches. En effet, les peptides reconstitués dans les bicouches formées par POPC possèdent un bon alignement (des éclatements quadripolaires bien définis) mais dans les bicouches contenant du cholestérol tous les peptides possèdent une distribution d'orientations (une continuité des éclatements quadripolaires). Alors cette différence (sans ou avec cholestérol) n'est pas spécifique et elle est due au changement des propriétés physico-chimiques des membranes.

Nous avons utilisé aussi cette technique pour étudier si le cholestérol agit indirectement sur le motif CRAC, alors nous avons utilisé du cholestérol deutéré (soit au niveau de son groupement hydroxyle, soit au niveau de ses groupements méthyle). Les spectres  $^2\text{H}$  de RMN ne montrent aucune différence sur des membranes de POPC contenant 30% de cholestérol- $\text{d}_6$  sans ou avec CRAC\_2 même en changeant la température d'acquisition. En diminuant la concentration du cholestérol jusqu'à 10%, nous observons une petite différence des éclatements quadripolaires du cholestérol en présence ou en absence du CRAC\_2. Par contre, cette différence était observée également chez le CRAC\_2 contrôle alors ce n'est pas dû à une interaction spécifique avec le cholestérol mais peut être au changement de l'organisation des bicouches.

# Bibliographie

Abad, C., Martínez-Gil, L., Tamborero, S., and Mingarro, I. (2009). Membrane topology of gp41 and amyloid precursor protein: Interfering transmembrane interactions as potential targets for HIV and Alzheimer treatment. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1788, 2132-2141.

Abdul-Gader, A., Miles, A.J., and Wallace, B.A. (2011). A reference dataset for the analyses of membrane protein secondary structures and transmembrane residues using circular dichroism spectroscopy. *Bioinformatics* 27, 1630-1636.

Aisenbrey, C., and Bechinger, B. (2004). Tilt and rotational pitch angle of membrane-inserted polypeptides from combined <sup>15</sup>N and <sup>2</sup>H solid-state NMR spectroscopy. *Biochemistry* 43, 10502-10512.

Aisenbrey, C., Bertani, P., Henklein, P., and Bechinger, B. (2007). Structure, dynamics and topology of membrane polypeptides by oriented <sup>2</sup>H solid-state NMR spectroscopy. *Eur Biophys J* 36, 451-460.

Aisenbrey, C., Sizun, C., Koch, J., Herget, M., Abele, R., Bechinger, B., and Tampe, R. (2006). Structure and dynamics of membrane-associated ICP47, a viral inhibitor of the MHC I antigen-processing machinery. *J Biol Chem* 281, 30365-30372.

Alkhatib, G., Broder, C.C., and Berger, E.A. (1996). Cell type-specific fusion cofactors determine human immunodeficiency virus type 1 tropism for T-cell lines versus primary macrophages. *J Virol* 70, 5487-5494.

Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. *Current Opinion in Biotechnology* 10, 411-421.

Barouch, D.H. (2008). Challenges in the development of an HIV-1 vaccine. *Nature* 455, 613-619.

Bartholow (2005). *JAIDS Journal of Acquired Immune Deficiency Syndromes*.

Bechinger, B., Aisenbrey, C., and Bertani, P. (2004). The alignment, structure and dynamics of membrane-associated polypeptides by solid-state NMR spectroscopy. *Biochim Biophys Acta* 1666, 190-204.

Bechinger, B., Kinder, R., Helmle, M., Vogt, T.C., Harzer, U., and Schinzel, S. (1999). Peptide structural analysis by solid-state NMR spectroscopy. *Biopolymers* 51, 174-190.

Bechinger, B., Resende, J.M., and Aisenbrey, C. (2011). The structural and topological analysis of membrane-associated polypeptides by oriented solid-state NMR spectroscopy: established concepts and novel developments. *Biophys Chem* 153, 115-125.

Bechinger, B., and Sizun, C. (2003). Alignment and structural analysis of membrane polypeptides by <sup>15</sup>N and <sup>31</sup>P solid-state NMR spectroscopy. *Concepts in Magnetic Resonance Part A* 18A, 130-145.

Bechinger, B., and Weik, M. (2003). Deuterium solid-state NMR investigations of exchange labeled oriented purple membranes at different hydration levels. *Biophys J* 85, 361-369.

Biotech, A.P. (2013). *AMERSHAM\_TheRecombProtHandbook.pdf*.

Biron, Z., Khare, S., Samson, A.O., Hayek, Y., Naider, F., and Anglister, J. (2002). A monomeric 3(10)-helix is formed in water by a 13-residue peptide representing the neutralizing determinant of HIV-1 on gp41. *Biochemistry* 41, 12687-12696.

- Bloom, M., Davis, J.H., and Mackay, A.L. (1981). Direct determination of the oriented sample nmr spectrum from the powder spectrum for systems with local axial symmetry. *Chemical Physics Letters* 80, 198-202.
- Brugger, B., Glass, B., Haberkant, P., Leibrecht, I., Wieland, F.T., and Krausslich, H.G. (2006). The HIV lipidome: a raft with an unusual composition. *Proc Natl Acad Sci U S A* 103, 2641-2646.
- Bryson, S., Julien, J.-P., Isenman, D.E., Kunert, R., Katinger, H., and Pai, E.F. (2008). Crystal Structure of the Complex between the Fab' Fragment of the Cross-Neutralizing Anti-HIV-1 Antibody 2F5 and the Fab Fragment of Its Antidiotypic Antibody 3H6. *Journal of Molecular Biology* 382, 910-919.
- Buzon, V., Natrajan, G., Schibli, D., Campelo, F., Kozlov, M.M., and Weissenhorn, W. (2010). Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS Pathog* 6, e1000880.
- Cavanagh, J., Fairbrother, W.J., Palmer Iii, A.G., Rance, M., and Skelton, N.J. (2007a). Chapter 1 - Classical NMR spectroscopy. In *Protein NMR Spectroscopy (Second Edition)*, J. Cavanagh, W.J. Fairbrother, A.G. Palmer, M. Rance, and N.J. Skelton, eds. (Burlington: Academic Press), pp. 1-28.
- Cavanagh, J., Fairbrother, W.J., Palmer Iii, A.G., Rance, M., and Skelton, N.J. (2007b). Chapter 2 - Theoretical description of NMR spectroscopy. In *Protein NMR Spectroscopy (Second Edition)*, J. Cavanagh, W.J. Fairbrother, A.G. Palmer, M. Rance, and N.J. Skelton, eds. (Burlington: Academic Press), pp. 29-113.
- Chan, D.C., Fass, D., Berger, J.M., and Kim, P.S. (1997a). Core structure of gp41 from the HIV envelope glycoprotein. *Cell* 89, 263-273.
- Chen, B., Vogan, E.M., Gong, H., Skehel, J.J., Wiley, D.C., and Harrison, S.C. (2005). Structure of an unliganded simian immunodeficiency virus gp120 core. *Nature* 433, 834-841.
- Chen, R. (2012). Bacterial expression systems for recombinant protein production: E. coli and beyond. *Biotechnology Advances* 30, 1102-1107.
- Chen, S.S., Yang, P., Ke, P.Y., Li, H.F., Chan, W.E., Chang, D.K., Chuang, C.K., Tsai, Y., and Huang, S.C. (2009). Identification of the LWYIK motif located in the human immunodeficiency virus type 1 transmembrane gp41 protein as a distinct determinant for viral infection. *J Virol* 83, 870-883.
- Cheung, R.C., Wong, J.H., and Ng, T.B. (2012). Immobilized metal ion affinity chromatography: a review on its applications. *Appl Microbiol Biotechnol* 96, 1411-1420.
- Choma, C., Gratkowski, H., Lear, J.D., and DeGrado, W.F. (2000). Asparagine-mediated self-association of a model transmembrane helix. *Nat Struct Biol* 7, 161-166.
- Chu, S., Abu-Baker, S., Lu, J., and Lorigan, G.A. (2010). (15)N Solid-state NMR spectroscopic studies on phospholamban at its phosphorylated form at ser-16 in aligned phospholipid bilayers. *Biochim Biophys Acta* 1798, 312-317.
- Corrêa, D.H.A., and Ramos, C., H. I. (2009). the use of circular dichroism spectroscopy to study protein folding, form and function. *African Journal of Biochemistry Research*.
- Coutant, J., Yu, H., Clément, M.J., Alfsen, A., Toma, F., Curmi, P.A., and Bomsel, M. (2008). Both lipid environment and pH are critical for determining physiological solution structure of 3-D-conserved epitopes of the HIV-1 gp41-MPER peptide P1. *FASEB J* 22, 4338-4351.
- Curran, A.R., and Engelman, D.M. (2003). Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Current Opinion in Structural Biology* 13, 412-417.

- Davis, J.H., Jeffrey, K.R., Bloom, M., Valic, M.I., and Higgs, T.P. (1976). Quadrupolar echo deuteron magnetic resonance spectroscopy in ordered hydrocarbon chains. *Chemical Physics Letters* *42*, 390-394.
- de Almeida, R.F.M., Fedorov, A., and Prieto, M. (2003). Sphingomyelin/Phosphatidylcholine/Cholesterol Phase Diagram: Boundaries and Composition of Lipid Rafts. *Biophys J* *85*, 2406-2416.
- de Souza, M.S., Ratto-Kim, S., Chuenarom, W., Schuetz, A., Chantakulkij, S., Nuntapinit, B., Valencia-Micolta, A., Thelian, D., Nitayaphan, S., Pitisuttithum, P., *et al.* (2012). The Thai Phase III Trial (RV144) Vaccine Regimen Induces T Cell Responses that Preferentially Target Epitopes within the V2 Region of HIV-1 Envelope. *J Immunol* *188*, 5166-5176.
- Earl, L.A., Lifson, J.D., and Subramaniam, S. (2013). Catching HIV 'in the act' with 3D electron microscopy. *Trends Microbiol* *21*, 397-404.
- Engelbrecht, T.N., Schroeter, A., Hau, Deme, B., Scheidt, H.A., Huster, D., and Neubert, R.H.H. (2012). The impact of ceramides NP and AP on the nanostructure of stratum corneum lipid bilayer. Part I: neutron diffraction and <sup>2</sup>H NMR studies on multilamellar models based on ceramides with symmetric alkyl chain length distribution. *Soft Matter* *8*, 6599-6607.
- Fantini, J., and Barrantes, F.J. (2013). How cholesterol interacts with membrane proteins: an exploration of cholesterol-binding sites including CRAC, CARC, and tilted domains. *Front Physiol* *4*.
- Gabrys, C.M., Qiang, W., Sun, Y., Xie, L., Schmick, S.D., and Weliky, D.P. (2013). Solid-State Nuclear Magnetic Resonance Measurements of HIV Fusion Peptide <sup>13</sup>CO to Lipid <sup>31</sup>P Proximities Support Similar Partially Inserted Membrane Locations of the  $\alpha$  Helical and  $\beta$  Sheet Peptide Structures. *The Journal of Physical Chemistry A* *117*, 9848-9859.
- Gavit, P., and Better, M. (2000). Production of antifungal recombinant peptides in *Escherichia coli*. *J Biotechnol* *79*, 127-136.
- Genetet, N., and Ruffault, A. (2002). *Immunologie*. 4<sup>ème</sup> édition - Noëlle Genetet, Collectif.
- Girard, M., and Sanders, R. (2013). HIV vaccine development at the turn of the 21st century : Current Opinion in HIV and AIDS.
- Gordon, L.M., Mobley, P.W., Lee, W., Eskandari, S., Kaznessis, Y.N., Sherman, M.A., and Waring, A.J. (2004). Conformational mapping of the N-terminal peptide of HIV-1 gp41 in lipid detergent and aqueous environments using <sup>13</sup>C-enhanced Fourier transform infrared spectroscopy. *Protein Sci* *13*, 1012-1030.
- Gray, G., Buchbinder, S., and Duerr, A. (2010). Overview of STEP and Phambili trial results: two phase IIb test of concept studies investigating the efficacy of MRK ad5 gag/pol/nef sub-type B HIV vaccine. *Curr Opin HIV AIDS* *5*, 357-361.
- Greene, W.C., and Peterlin, M.B. (2008). CHAPTER 3 – Molecular Biology of HIV: Implications for New Therapies. 23–38.
- Greenfield, N.J. (2007). Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* *1*, 2876-2890.
- Guttman, M., Kahn, M., Garcia, N.K., Hu, S.-L., and Lee, K.K. (2012). Solution Structure, Conformational Dynamics, and CD4-Induced Activation in Full-Length, Glycosylated, Monomeric HIV gp120. *Journal of Virology* *86*, 8750-8764.
- Hager-Braun, C., Katinger, H., and Tomer, K.B. (2006). The HIV-neutralizing monoclonal antibody 4E10 recognizes N-terminal sequences on the native antigen. *J Immunol* *176*, 7471-7481.

Hahn, B.H., Shaw, G.M., De, K.M., Cock, and Sharp, P.M. (2000). AIDS as a Zoonosis: Scientific and Public Health Implications. *Science* 287, 607-614.

Hannig, G., and Makrides, S.C. (1998). Strategies for optimizing heterologous protein expression in *Escherichia coli*. *Trends in Biotechnology* 16, 54-60.

Harris, R.S., Hultquist, J.F., and Evans, D.T. (2012). The Restriction Factors of Human Immunodeficiency Virus. *The Journal of Biological Chemistry* 287, 40875-40883.

Harzer, U., and Bechinger, B. (2000). Alignment of Lysine-Anchored Membrane Peptides under Conditions of Hydrophobic Mismatch: A CD, 15N and 31P Solid-State NMR Spectroscopy Investigation.

Hemelaar, J. (2012). The origin and diversity of the HIV-1 pandemic. *Trends Mol Med* 18, 182-192.

<http://www.doctissimo.fr/> (2013). Sida : Le point sur les traitements - Doctissimo.

Hu, G., Liu, J., Taylor, K.A., and Roux, K.H. (2011). Structural Comparison of HIV-1 Envelope Spikes with and without the V1/V2 Loop.

Huang, C.-c., Tang, M., Zhang, M.-Y., Majeed, S., Montabana, E., Stanfield, R.L., Dimitrov, D.S., Korber, B., Sodroski, J., Wilson, I.A., *et al.* (2005). Structure of a V3-Containing HIV-1 gp120 Core.

Huber, M., From the Division of Infectious Diseases, U.H.Z., Zurich, Switzerland, Trkola, A., and From the Division of Infectious Diseases, U.H.Z., Zurich, Switzerland (2007). Humoral immunity to HIV-1: neutralization and beyond. *Journal of Internal Medicine* 262, 5-25.

Hwang, P.M., Pan, J.S., and Sykes, B.D. (2013). Targeted expression, purification, and cleavage of fusion proteins from inclusion bodies in *Escherichia coli*. *FEBS Letters*.

Jafurulla, M., Tiwari, S., and Chattopadhyay, A. (2011). Identification of cholesterol recognition amino acid consensus (CRAC) motif in G-protein coupled receptors. *Biochem Biophys Res Commun* 404, 569-573.

Jaroniec, C.P., Joshua D. Kaufman, Stephen J. Stahl, Mathias Viard, Robert Blumenthal, Paul T. Wingfield, a., and Bax, A. (2005). Structure and Dynamics of Micelle-Associated Human Immunodeficiency Virus gp41 Fusion Domain.

Johnston, M.I., and Fauci, A.S. (2008). An HIV Vaccine — Challenges and Prospects — *NEJM. N Engl J Med*.

Kanki, P.x.J.x., and Meloni, S.x.T.x. (2009). Biology and Variation in HIV-2 and HIV-1.

Kelly, S.M., Jess, T.J., and Price, N.C. (2005). How to study proteins by circular dichroism. *Biochim Biophys Acta* 1751, 119-139.

Khayat, R., Lee, J.H., Julien, J.-P., Cupo, A., Klasse, P.J., Sanders, R.W., Moore, J.P., Wilson, I.A., and Ward, A.B. (2013). Structural Characterization of Cleaved, Soluble HIV-1 Envelope Glycoprotein Trimers.

Khelashvili, G., and Harries, D. (2013). How sterol tilt regulates properties and organization of lipid membranes and membrane insertions. *Chem Phys Lipids* 169, 113-123.

Kim, J.H., Hartley, T.L., Curran, A.R., and Engelman, D.M. (2009). Molecular dynamics studies of the transmembrane domain of gp41 from HIV-1. *Biochim Biophys Acta* 1788, 1804-1812.

Klasse, P.J., Depetris, R.S., Pejchal, R., Julien, J.-P., Khayat, R., Lee, J.H., Marozsan, A.J., Cupo, A., Cocco, N., Korzun, J., *et al.* (2013). Influences on Trimerization and Aggregation of Soluble, Cleaved HIV-1 SOSIP Envelope Glycoprotein.

- Kozak, M. (1983). Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol Rev* 47, 1-45.
- Kwong, P.D., Mascola, J.R., and Nabel, G.J. (2013). Broadly neutralizing antibodies and the search for an HIV-1 vaccine: the end of the beginning. *Nature Reviews Immunology* 13, 693-701.
- Kwong, P.D., Wyatt, R., Robinson, J., Sweet, R.W., Sodroski, J., and Hendrickson, W.A. (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature* 393, 648-659.
- Lemmon, M.A., Flanagan, J.M., Hunt, J.F., Adair, B.D., Bormann, B.J., Dempsey, C.E., and Engelman, D.M. (1992). Glycophorin A dimerization is driven by specific interactions between transmembrane alpha-helices. *J Biol Chem* 267, 7683-7689.
- Leonard, C.K., Spellman, M.W., Riddle, L., Harris, R.J., Thomas, J.N., and Gregory, T.J. (1990). Assignment of intrachain disulfide bonds and characterization of potential glycosylation sites of the type 1 recombinant human immunodeficiency virus envelope glycoprotein (gp120) expressed in Chinese hamster ovary cells.
- Li, H., and Papadopoulos, V. (1998). Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. *Endocrinology* 139, 4991-4997.
- Li, Y., and Tamm, L.K. (2007). Structure and Plasticity of the Human Immunodeficiency Virus gp41 Fusion Domain in Lipid Micelles and Bilayers. *Biophys J* 93, 876-885.
- Liao, Z., Cimasky, L.M., Hampton, R., Nguyen, D.H., and Hildreth, J.E. (2001). Lipid rafts and HIV pathogenesis: host membrane cholesterol is required for infection by HIV type 1. *AIDS Res Hum Retroviruses* 17, 1009-1019.
- Lilie, H., Schwarz, E., and Rudolph, R. (1998). Advances in refolding of proteins produced in *E. coli*. *Current Opinion in Biotechnology* 9, 497-501.
- Liu, J., Deng, Y., Dey, A.K., Moore, J.P., and Lu, M. (2009). Structure of the HIV-1 gp41 Membrane-Proximal Ectodomain Region in a Putative Prefusion Conformation.
- Locatelli, S., and Peeters, M. (2012). Cross-species transmission of simian retroviruses: how and why they could lead to the emergence of new diseases in the human population. In *Aids (England)*, pp. 659-673.
- Long, Y., Meng, F., Kondo, N., Iwamoto, A., and Matsuda, Z. (2011). Conserved arginine residue in the membrane-spanning domain of HIV-1 gp41 is required for efficient membrane fusion. *Protein Cell* 2, 369-376.
- Lorenzi, P., Opravil, M., Hirschel, B., Chave, J.P., Furrer, H.J., Sax, H., Perneger, T.V., Perrin, L., Kaiser, L., and Yerly, S. (1999). Impact of drug resistance mutations on virologic response to salvage therapy. *Swiss HIV Cohort Study. AIDS* 13, F17-21.
- Lorizate, M., Huarte, N., Sáez-Ciri3n, A., and Nieva, J.L. (2008). Interfacial pre-transmembrane domains in viral proteins promoting membrane fusion and fission. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1778, 1624-1639.
- Makrides, S.C. (1996). Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol Rev* 60, 512-538.
- Manzo, G., Carboni, M., Rinaldi, A.C., Casu, M., and Scorciapino, M.A. (2013). Characterization of sodium dodecylsulphate and dodecylphosphocholine mixed micelles through NMR and dynamic light scattering. *Magn Reson Chem* 51, 176-183.

- Marley, J., Lu, M., and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. *J Biomol NMR* 20, 71-75.
- Marsan, M.P., Muller, I., Ramos, C., Rodriguez, F., Dufourc, E.J., Czaplicki, J., and Milon, A. (1999). Cholesterol orientation and dynamics in dimyristoylphosphatidylcholine bilayers: a solid state deuterium NMR analysis. *Biophys J* 76, 351-359.
- Merk, a., and Subramaniam, S. (2013). HIV-1 envelope glycoprotein structure. 23, 268–276.
- Miller, M.D., Geleziunas, R., Bianchi, E., Lennard, S., Hrin, R., Zhang, H., Lu, M., An, Z., Ingallinella, P., Finotto, M., *et al.* (2005). A human monoclonal antibody neutralizes diverse HIV-1 isolates by binding a critical gp41 epitope.
- Mobley, P.W., Waring, A.J., Sherman, M.A., and Gordon, L.M. (1999). Membrane interactions of the synthetic N-terminal peptide of HIV-1 gp41 and its structural analogs. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1418, 1-18.
- Montefiori, D.C., Karnasuta, C., Huang, Y., Ahmed, H., Gilbert, P., de Souza, M.S., McLinden, R., Tovanabutra, S., Laurence-Chenine, A., Sanders-Buell, E., *et al.* (2012). Magnitude and Breadth of the Neutralizing Antibody Response in the RV144 and Vax003 HIV-1 Vaccine Efficacy Trials. *J Infect Dis* 206, 431-441.
- Montero, M., Houten, N.E.v., Wang, X., and Scott, J.K. (2008). The Membrane-Proximal External Region of the Human Immunodeficiency Virus Type 1 Envelope: Dominant Site of Antibody Neutralization and Target for Vaccine Design.
- Muñoz-Barroso, I., Salzwedel, K., Hunter, E., and Blumenthal, R. (1999). Role of the Membrane-Proximal Domain in the Initial Stages of Human Immunodeficiency Virus Type 1 Envelope Glycoprotein-Mediated Membrane Fusion.
- Nabel, G.J. (2013). Designing Tomorrow's Vaccines. *N Engl J Med* 368, 551-560.
- NIAID (2013). Structure of HIV.
- Nicely, N.I., Dennison, S.M., Spicer, L., Scarce, R.M., Kelsoe, G., Ueda, Y., Chen, H., Liao, H.-X., Alam, S.M., and Haynes, B.F. (2010). Crystal structure of a non-neutralizing antibody to the HIV-1 gp41 membrane-proximal external region. *Nature Structural & Molecular Biology* 17, 1492-1494.
- Niezborala, C. (2008). Une nouvelle technique de mesure du dichroïsme circulaire : premiers pas vers l'étude du repliement de polypeptides modèles (Palaiseau, Ecole polytechnique).
- Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J.R., Wyatt, R., and Kwong, P.D. (2004a). Structure and Mechanistic Analysis of the Anti-Human Immunodeficiency Virus Type 1 Antibody 2F5 in Complex with Its gp41 Epitope.
- Ofek, G., Tang, M., Sambor, A., Katinger, H., Mascola, J.R., Wyatt, R., and Kwong, P.D. (2004b). Structure and Mechanistic Analysis of the Anti-Human Immunodeficiency Virus Type 1 Antibody 2F5 in Complex with Its gp41 Epitope. *J Virol* 78, 10724-10737.
- Pancera, M., Majeed, S., Ban, Y.E., Chen, L., Huang, C.C., Kong, L., Kwon, Y.D., Stuckey, J., Zhou, T., Robinson, J.E., *et al.* (2010). Structure of HIV-1 gp120 with gp41-interactive region reveals layered envelope architecture and basis of conformational mobility. *Proc Natl Acad Sci U S A* 107, 1166-1171.
- Pantophlet, R., and Burton, D.R. (2006). GP120: Target for Neutralizing HIV-1 Antibodies.
- Pines, A., Gibby, M.G., and Waugh, J.S. (1973). Proton-enhanced NMR of dilute spins in solids. *J Chem Phys.*

- Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258, 598-599.
- Rance, M., and Byrd, R.A. (1983). Obtaining high-fidelity spin-12 powder spectra in anisotropic media: Phase-cycled Hahn echo spectroscopy. *Journal of Magnetic Resonance* (1969) 52, 221-240.
- Saitô, H., Ando, I., and Naito, A. (2006). *Solid State NMR Spectroscopy for Biopolymers - Principles and Applications*.
- Salzwedel, K., West, J.T., and Hunter, E. (1999). A Conserved Tryptophan-Rich Motif in the Membrane-Proximal Region of the Human Immunodeficiency Virus Type 1 gp41 Ectodomain Is Important for Env-Mediated Fusion and Virus Infectivity.
- Samuelson, J. (2011). *Bacterial Systems*. In (Wiley-VCH Verlag GmbH & Co. KGaA), pp. 11-35.
- Sattentau, Q.J., and Weiss, R.A. (1988). The CD4 antigen: Physiological ligand and HIV receptor. *Cell* 52, 631-633.
- Schibli, D.J., Montelaro, R.C., and Vogel, H.J. (2001). The membrane-proximal tryptophan-rich region of the HIV glycoprotein, gp41, forms a well-defined helix in dodecylphosphocholine micelles. *Biochemistry* 40, 9570-9578.
- Schmick, S.D., and Weliky, D.P. (2010). Major Antiparallel and Minor Parallel  $\beta$  Sheet Populations Detected in the Membrane-Associated Human Immunodeficiency Virus Fusion Peptide.
- Senes, A., Engel, D.E., and DeGrado, W.F. (2004). Folding of helical membrane proteins: the role of polar, GxxxG-like and proline motifs. *Current Opinion in Structural Biology* 14, 465-479.
- Shen, H.-H., Lithgow, T., and Martin, L. (2013). Reconstitution of Membrane Proteins into Model Membranes: Seeking Better Ways to Retain Protein Activities. *International Journal of Molecular Sciences* 14, 1589-1607.
- Silvius, J.R. (2003). Role of cholesterol in lipid raft formation: lessons from lipid model systems. *Biochim Biophys Acta* 1610, 174-183.
- Song, Y., Hustedt, E.J., Brandon, S., and Sanders, C.R. (2013). Competition between homodimerization and cholesterol binding to the C99 domain of the amyloid precursor protein. *Biochemistry* 52, 5051-5064.
- Sparr, E., Ash, W.L., Nazarov, P.V., Rijkers, D.T., Hemminga, M.A., Tieleman, D.P., and Killian, J.A. (2005). Self-association of transmembrane alpha-helices in model membranes: importance of helix orientation and role of hydrophobic mismatch. *J Biol Chem* 280, 39324-39331.
- Sreerama, N., and Woody, R.W. (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* 287, 252-260.
- Studier, F.W., Rosenberg, A.H., Dunn, J.J., and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* 185, 60-89.
- Sun, Z.-Y.J., Oh, K.J., Kim, M., Yu, J., Brusica, V., Song, L., Qiao, Z., Wang, J.-h., Wagner, G., and Reinherz, E.L. (2008). HIV-1 Broadly Neutralizing Antibody Extracts Its Epitope from a Kinked gp41 Ectodomain Region on the Viral Membrane. *Immunity* 28, 52-63.
- Sundquist, W.I., and Kräusslich, H.-G. (2012). *HIV-1 Assembly, Budding, and Maturation*.
- Sørensen, H.P., and Mortensen, K.K. (2005). Advanced genetic strategies for recombinant protein expression in *Escherichia coli*. *Journal of Biotechnology* 115, 113-128.

Tan, K., Liu, J.-h., Wang, J.-h., Shen, S., and Lu, M. (1997). Atomic structure of a thermostable subdomain of HIV-1 gp41.

Tatko, C.D., Nanda, V., Lear, J.D., and Degrado, W.F. (2006). Polar networks control oligomeric assembly in membranes. *J Am Chem Soc* *128*, 4170-4171.

Teixeiraa, C., Gomes, J.R.B., Gomesc, P., and Maurela, F. (2011). Viral surface glycoproteins, gp120 and gp41, as potential drug targets against HIV-1: Brief overview one quarter of a century past the approval of zidovudine, the first anti-retroviral drug. *46*, 979-992.

Tiburu, E.K., Dave, P.C., and Lorigan, G.A. (2004). Solid-state <sup>2</sup>H NMR studies of the effects of cholesterol on the acyl chain dynamics of magnetically aligned phospholipid bilayers. *Magn Reson Chem* *42*, 132-138.

Tossonian, H.K., Raffa, J.D., Grebely, J., Viljoen, M., Mead, A., Khara, M., McLean, M., Krishnamurthy, A., DeVlaming, S., and Conway, B. (2009). Primary drug resistance in antiretroviral-naïve injection drug users. *Int J Infect Dis* *13*, 577-583.

Trauger, S.A., Webb, W., and Siuzdak, G. (2002). Peptide and protein analysis with mass spectrometry. *Journal of Spectroscopy* *16*, 15.

UNAIDS (2011). How to get to zero: faster, smarter, better (<http://www.unaids.org>).

van Meer, G., and de Kroon, A.I. (2011). Lipid map of the mammalian cell. *J Cell Sci* *124*, 5-8.

Viard, M., Parolini, I., Sargiacomo, M., Fecchi, K., Ramoni, C., Ablan, S., Ruscetti, F.W., Wang, J.M., and Blumenthal, R. (2002). Role of cholesterol in human immunodeficiency virus type 1 envelope protein-mediated fusion with host cells. *J Virol* *76*, 11584-11595.

Vidovic, V., Prongidi-Fix, L., Bechinger, B., and Werten, S. (2009). Production and isotope labeling of antimicrobial peptides in *Escherichia coli* by means of a novel fusion partner that enables high-yield insoluble expression and fast purification. *Journal of Peptide Science* *15*, 278-284.

Vincent, N., Genin, C., and Malvoisin, E. (2002). Identification of a conserved domain of the HIV-1 transmembrane protein gp41 which interacts with cholesterol groups. *Biochim Biophys Acta* *1567*, 157-164.

Volkov, V., and Bonn, M. (2013). Structural Properties of gp41 Fusion Peptide at a Model Membrane Interface.

Walker, B.D., and Burton, D.R. (2008). Toward an AIDS Vaccine.

Wallace, B.A. (2009). Protein characterisation by synchrotron radiation circular dichroism spectroscopy. *Q Rev Biophys* *42*, 317-370.

Wang, Y., Jing, L., and Xu, K. (2002). A unique approach for high level expression and production of a recombinant cobra neurotoxin in *Escherichia coli*. *J Biotechnol* *94*, 235-244.

Weissenhorn, W., Dessen, A., Harrison, S.C., Skehel, J.J., and Wiley, D.C. (1997). Atomic structure of the ectodomain from HIV-1 gp41. *Nature* *387*, 426-430.

White, T.A., Laboratory of Cell Biology, C.f.C.R., National Cancer Institute, NIH, Bethesda, Maryland, United States of America, Bartesaghi, A., Laboratory of Cell Biology, C.f.C.R., National Cancer Institute, NIH, Bethesda, Maryland, United States of America, Borgnia, M.J., Laboratory of Cell Biology, C.f.C.R., National Cancer Institute, NIH, Bethesda, Maryland, United States of America, Meyerson, J.R., Laboratory of Cell Biology, C.f.C.R., National Cancer Institute, NIH, Bethesda, Maryland, United States of America, Cruz, M.J.V.d.l., Laboratory of Cell Biology, C.f.C.R., National Cancer Institute, NIH, Bethesda, Maryland, United States of America, *et al.* (2010). Molecular Architectures

of Trimeric SIV and HIV-1 Envelope Glycoproteins on Intact Viruses: Strain-Dependent Variation in Quaternary Structure. *PLOS Pathogens* 6.

Whitmore, L., and Wallace, B.A. (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* 32, W668-673.

Whitmore, L., and Wallace, B.A. (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 89, 392-400.

# Annexe 1

## Gradient de purification des polypeptides

Temps (min)	% éluant B (Isopropanol 90%, H <sub>2</sub> O, 0,1% TFA 0,1%)
0	10
5	10
13	32,66
30	48,24
49	75,16
56	95
61	95
66	10
71	10

# Annexe 2

## Composition du milieu minimum M9

Composition du milieu utilisé pour produire fpTM marqué uniformément en <sup>15</sup> N et <sup>13</sup> C pour un litre de culture : Solutions stocks	Volume utilisé (mL)
<b>5 X M9</b>	200
<b>50 g/L <sup>15</sup>NH<sub>4</sub>Cl</b>	10
<b>20 % <sup>13</sup>C glucose</b>	10
<b>1M MgSO<sub>4</sub></b>	2
<b>1M CaCl<sub>2</sub></b>	0,1
<b>1g/L FeSO<sub>4</sub></b>	0,1
<b>10 g/L thiamine</b>	1
<b>10 g/L biotine</b>	1
<b>10 mg/mL kanamycine</b>	5
<b>préculture</b>	10
<b>eau</b>	Qsp 1000

Composition du 5XM9	Masse (g)
<b>Na<sub>2</sub>HPO<sub>4</sub>, 2H<sub>2</sub>O</b>	<b>42,5</b>
<b>KH<sub>2</sub>PO<sub>4</sub></b>	<b>15</b>
<b>NaCl</b>	<b>2,5</b>

**Production des polypeptides issus des glycoprotéines d'enveloppe du VIH-1 pour des études biophysique et structurale par RMN et DC.**

## Résumé

Quelques régions stables ont été découvertes sur les gp d'env du VIH-1 contre lesquelles des patients produisent des anticorps neutralisants. Les épitopes les plus prometteurs se trouvent dans la MPER et sont probablement exposés durant la fusion. Alors que les peptides isolés à partir de cette région ne sont pas parvenus à induire une réaction immunogène neutralisante, des études antérieures suggèrent que la membrane lipidique joue un rôle dans la structuration des antigènes et dans la réponse immunogénique.

C'est pourquoi nous étudions la structure de ces épitopes. Cela nécessite leur surexpression, leur purification et leur reconstitution dans des liposomes. Une étude de CD montre qu'ils pourraient changer de conformation, cela sera confirmé par RMN. En outre, leur immunogénicité sera vérifiée par vaccination des souris. En plus, nous trouvons que le cholestérol peut modifier l'orientation des peptides englobant le motif CRAC de la gp41.

Mot clés : VIH-1, gp120, gp41, surexpression en système bactérien, vaccination, RMN, CD.

## Résumé en anglais

A few stable regions have been discovered on the HIV-1 env gp against which some patients produce neutralizing antibodies. The most promising ones are located in the MPER and are probably exposed transiently during the fusion. Whereas the peptides isolated from this region failed to induce immunogenic response, previous studies suggest the lipid membrane plays a role in antigens structure and in the immunogenic response.

That is why we investigate the structure of these épitopes in membrane models. This requires the production of these épitopes by bacterial overexpression, their purification and their reconstitution in liposomes. A CD study shows that they could undergo a conformational change; this will be confirmed by NMR. Also their immunogenicity will be checked by mice immunization. In addition, we find that cholesterol could change the orientation of peptides encompassing a gp41 CRAC motif.

Key words: HIV-1, gp120, gp41, fusion peptide, CRAC, bacterial overexpression, vaccination, CD, RMN.