



**HAL**  
open science

# Etudes fonctionnelles et structurales des complexes Hélicase-Primase du virus Epstein-Barr

Eric Thierry

► **To cite this version:**

Eric Thierry. Etudes fonctionnelles et structurales des complexes Hélicase-Primase du virus Epstein-Barr. Sciences agricoles. Université de Grenoble, 2013. Français. NNT : 2013GRENV002 . tel-00957973

**HAL Id: tel-00957973**

**<https://theses.hal.science/tel-00957973>**

Submitted on 11 Mar 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## THÈSE

Pour obtenir le grade de

## DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : **Biologie Structurale et Nanobiologie**

Arrêté ministériel : 7 août 2006

Présentée par

**Eric THIERRY**

Thèse dirigée par le **Pr Wim BURMEISTER**

préparée au sein de l'**Unité d'interaction Virus – Cellule Hôte**  
dans l'**École Doctorale Chimie et Sciences du Vivant**

# Etudes fonctionnelles et structurales du complexe Hélicase – Primase du virus Epstein – Barr.

Thèse soutenue publiquement le **23 avril 2013**,  
devant le jury composé de :

**Dr Carlo PETOSA**

Institut de Biologie Structurale (Président et Examineur)

**PU-PH Sophie ALAIN**

Faculté de Médecine de Limoges (Rapporteur)

**Dr Laurent TERRADOT**

Institut de Biologie et de Chimie des Protéines (Rapporteur)

**Pr Wim BURMEISTER**

Unité d'Interaction Virus – Cellule Hôte (Examineur)



UNIVERSITÉ DE GRENOBLE



# **Préface :**

## **Ma perception de la recherche :**

Ce métier est un combat perpétuel face à un ennemi qui peut rester longtemps sans visage, obligeant à user de toutes les techniques pour le débusquer. Une fois que son visage commence à apparaître, c'est une satisfaction indescriptible. Ce visage encore flou permet de guider notre réflexion grâce aux résultats obtenus et permet de révéler au grand jour le visage de cet ennemi qui s'avèrera sûrement être au final un ami !!!



# Remerciements :

Je tiens tout d'abord à remercier le Pr Rob Ruigrok pour m'avoir permis d'effectuer mon stage de Master 2 et ma thèse au sein de l'institut UVHCI dont il est le directeur. J'ai pu ainsi bénéficier d'une qualité de recherche aussi bien pour les outils mis à ma disposition que pour la qualité du travail scientifique abordé dans ce laboratoire. J'ai essayé d'utiliser au maximum les moyens mis à ma disposition pour pouvoir mener à bien mon projet de thèse.

Je tiens aussi à remercier tous les membres de mon jury, merci aux deux rapporteurs de mon manuscrit, le Pr Sophie Alain et le Dr Laurent Terradot d'avoir accepté d'évaluer mes travaux de thèse. Ainsi qu'au président de mon Jury, le Dr Carlo Petosa.

Je n'ai pas pu atteindre les objectifs fixés en début de thèse, mais l'avancement de ma thèse est en partie dû au soutien des membres de mon équipe. Je tiens donc à remercier le Pr Wim Burmeister pour avoir accepté de superviser ma thèse. Je n'ai pas pu exploiter tes qualités de cristallographe au cours de ma thèse, les cristaux n'étant pas au rendez-vous !!! Je tiens à remercier chaleureusement le Dr Marlyse Buisson. Il n'est pas facile de trouver les bons termes afin d'exprimer toute la gratitude que j'ai pour toi. Je sais que si je me suis accroché et que j'ai pu aller aussi loin avec mon sujet de thèse c'est en grande partie grâce à ton soutien aussi bien scientifique que moral. Ta nature de voir toujours le côté positif des résultats qui ne sont pas forcément attendus, donne l'énergie pour poursuivre les expériences. Je n'oublie pas non plus le Dr Nicolas Tarbouriech qui lors des longs « group meeting » m'a fait partager sa vision critique et scientifique très précieuse. Il y a aussi Céleste, enfin Dr Céleste devrai-je dire !! Même si tu n'es plus au labo maintenant, on a fait un sacré bout de chemin ensemble !!!

Je n'oublie pas non plus les autres personnes qui m'ont aidé lors des étapes importantes de ma thèse. Je pense tout particulièrement à Frédéric Garzoni et le Dr Yan Nie qui m'ont conseillé et guidé tout au long des étapes de clonage MultiBac et d'expression en cellules d'insectes. Je tiens aussi à remercier sincèrement le Dr Julien Perard pour ces idées et son investissement pour m'épauler tout au long de ma thèse.

J'ai encore beaucoup de personnes à remercier et notamment tous les thésards, stagiaires, postdoc et autres permanents qui font de la vie à l'UVHCI un vrai bonheur. Je sais que je vais sûrement oublier des personnes mais je me lance quand même !! Je commence par les « anciens » qui ont quitté l'établissement, Euripédès, Lucy, Alexandre D, Julien L, Pierre-Alain, Ambroise, Cédric, Audrey et David. Et tous ceux encore présent, Charlotte, Ivan, Catherine, Alexandre M, Christopher, Jaka, Nicolas M, Aurélien, Emilie, Paulin, Nick, Filip, Léandro, Hélène ... Pour ceux que j'aurai oublié, désolé et un grand merci pour me faire pardonner !!!

Les collègues et les amis c'est important, mais la famille aussi !!! Et oui vous avez aussi le droit à votre tour d'honneur. Tout d'abord merci à tous mes grands-parents, ils me montrent la voie à suivre et si on me demande de signer pour avoir une vie aussi belle que la vôtre je dis OUI. Merci à mes parents qui sont toujours présents à mes côtés que ce soit dans les moments de bonheurs comme dans les moments de doutes. Vous êtes toujours là pour me montrer le chemin et me donner

envie de le parcourir. Pendant mes quatre années de thèse, la famille s'est agrandie, merci Elise de m'avoir fait connaître les rôles de tonton et de parrain avec tes deux magnifiques petits bouts, Nathan et Lalie. Mais je ne t'oublie pas non plus Christophe tu es le mari idéal pour ma sœur !!

Il y a une dernière personne que je souhaite remercier plus que quiconque. Même si tu es arrivée dans ma vie vers la fin de ma thèse, en l'espace d'une année tu m'as fait redécouvrir le vrai sens du mot « Aimer » et que la vie nous réserve toujours les plus belles surprises quand on s'y attend le moins !!! Merci Laura pour le bonheur que tu m'apportes de jours en jours.

Je ne sais pas trop comment finir ces remerciements, mais une chose est sûre ces quatre années de thèse passées à vos côtés vont rester gravées à jamais dans ma mémoire et mon cœur.

Merci.

ERIC

# **TABLE DES MATIERES**

<b>PARTIE I : INTRODUCTION</b>	<b>1</b>
<b>CHAPITRE 1 : INTRODUCTION GENERALE</b>	<b>2</b>
<b>I. Le Virus Epstein-Barr</b>	<b>3</b>
1. La découverte du virus Epstein-Barr	3
2. EBV est un membre de la famille des herpèsvirus	4
3. La structure d'EBV	6
<b>II. Le cycle viral d'EBV</b>	<b>6</b>
1. La transmission virale	6
1.1. Attachement sur les cellules cibles et internalisation	7
1.2. Tropisme cellulaire	9
1.3. Infection latente, maintien du génome	11
1.4. Infection lytique, un processus séquentiel et ordonné	13
2. Les pathologies associées à EBV	15
2.1. La mononucléose infectieuse	15
2.2. Les lymphomes des cellules B	16
2.2.1. Le lymphome de Burkitt	16
2.2.2. Le lymphome de Hodgkin	16
2.3. Les carcinomes des cellules épithéliales	17
2.3.1. Le carcinome indifférencié du rhinopharynx	17
2.3.2. Le carcinome gastrique	17
2.4. Les lymphoproliférations des patients immunodéprimés	17
2.5. L'implication du cycle lytique dans les maladies associées à EBV	19
<b>CHAPITRE 2 : LE COMPLEXE DE REPLICATION DES HERPESVIRUS</b>	<b>20</b>
<b>I. La réplication de l'ADN génomique et les protéines impliquées</b>	<b>21</b>
<b>II. Description structurale des protéines du complexe de réplication d'EBV</b>	<b>24</b>
1. Revue Virologie	24
2. Le complexe de réplication des herpèsvirus est fortement conservé	39
3. Complément d'informations sur les protéines du complexe Hélicase-Primase d'EBV	43
3.1. L'hélicase (BBLF4)	43
3.2. La primase (BSLF1)	45
3.3. La sous-unité accessoire (BBLF2/3)	45
<b>III. Le rôle charnière du complexe H-P</b>	<b>48</b>
1. Localisation cellulaire des protéines du complexe H-P	48
2. L'hélicase, des motifs primordiaux mais peu connus	50
3. La primase une protéine en relation étroite avec l'hélicase	52
4. Le recrutement du complexe dans les foyers réplicatifs et l'importance de sa stœchiométrie	55
5. L'implication des protéines BBLF4, BSLF1 et BBLF2/3 au sein du complexe de réplication du virus Epstein-Barr	57
<b>IV. Hélicases des herpèsvirus et Dda des bactériophages, un passé commun ?</b>	<b>58</b>
1. Les hélicases du bactériophage T4	58
2. Caractéristique de l'hélicase Dda et son implication dans la réplication des bactériophages	59

2.1.	Implication de la protéine Dda dans la réplication des bactériophages	59
2.2.	Caractéristique de la Dda du bactériophage T4	59
3.	Informations complémentaires apportées par la structure de la Dda du bactériophage T4	60
4.	Relation entre la Dda et les hélicases des herpèsvirus	62
<b>V.</b>	<b>Les objectifs du travail de thèse</b>	<b>62</b>
	<b>PARTIE II : MATERIEL &amp; METHODES</b>	<b>64</b>
	<b>CHAPITRE 1 : ETUDES PRELIMINAIRES</b>	<b>65</b>
<b>I.</b>	<b>Prédiction de désordre et caractéristiques des protéines</b>	<b>66</b>
1.	Prédiction d'ordre et de désordre	66
2.	Caractérisation bio-informatique complémentaire	67
	<b>CHAPITRE 2 : EXPRESSION DES PROTEINES RECOMBINANTES</b>	<b>68</b>
<b>I.</b>	<b>Expression et production de protéines recombinantes</b>	<b>69</b>
1.	Les prérequis	69
1.1.	Les cellules d'insectes	69
1.2.	La culture cellulaire	69
1.3.	Les milieux de cultures	69
2.	Production de protéines recombinantes en cellules adhérentes	70
2.1.	Principes du clonage par le système Bac-to-Bac® (Invitrogen)	70
2.2.	Constructions utilisées en cellules adhérentes	71
2.2.1.	Clonage de la protéine BBLF2/3	72
2.3.	Obtention des bacmides	75
2.4.	Transfection et mesure du titre viral	76
3.	Test d'expression en « Time course »	78
3.1.	Productions des protéines recombinantes en cellules adhérentes	78
4.	Clonage MultiBac et production de protéines en cellules en suspension	79
4.1.	Constructions utilisées pour le clonage MultiBac	79
4.2.	Transfection et amplification virale	83
4.3.	Production des protéines recombinantes en cellules en suspension :	83
	<b>CHAPITRE 3 : EXPERIENCES BIOCHIMIQUES</b>	<b>85</b>
<b>I.</b>	<b>Tampons et conditions de purification</b>	<b>86</b>
1.	Recherche de tampon de solubilisation en plaque 96 puits	86
2.	Lyse cellulaire	87
3.	Purification par affinité sur colonne de nickel	88
4.	Chromatographie d'exclusion de taille	90
5.	Optimisation du tampon par ThermoFluor	91
6.	Purification en conditions dénaturantes	92
<b>II.</b>	<b>Identification des protéines produites par surexpression en cellules d'insectes</b>	<b>94</b>
1.	Marquage par anticorps spécifique après Western Blot	94
2.	Identification par séquençage N-terminal	95
3.	Identification par spectrométrie de masse	95
3.1.	Digestion tryptique des protéines d'intérêt « in-gel »	95
3.2.	Analyse par spectrométrie de masse	97

<b>CHAPITRE 4 : CARACTERISATIONS BIOPHYSIQUES</b>	<b>98</b>
<b>I. Etude de la polydispersité et du rayon hydrodynamique par DLS</b>	<b>99</b>
1. Principe	99
2. Tests expérimentaux	99
<b>II. SEC-MALLS-RI</b>	<b>100</b>
1. Principe	100
2. Tests expérimentaux	101
<b>III. Test de cristallisation</b>	<b>101</b>
1. Principe	101
2. Tests de cristallisation de BBLF2/3 et du complexe H-P	103
2.1. BBLF2/3	103
2.2. Complexe H-P	104
<b>IV. Analyse par microscopie électronique et affinement par la technique GraFix</b>	<b>104</b>
1. Principe de la technique GraFix	105
2. Réalisation du GraFix	105
<b>CHAPITRE 5 : CARACTERISATIONS ENZYMATIQUES</b>	<b>107</b>
<b>I. Caractérisations enzymatiques</b>	<b>108</b>
1. Mesure de l'activité ATPase	108
2. Mesure de l'activité hélicase	109
<b>PARTIE III : RESULTATS</b>	<b>112</b>
<b>CHAPITRE 1 : ETUDES DES PROTEINES BBLF4, BSLF1 ET BBLF2/3 PRODUITES PAR TROIS BACULOVIRUS</b>	<b>113</b>
<b>I. Etudes des protéines du complexe H-P produites individuellement</b>	<b>114</b>
1. Caractéristiques des protéines composant le complexe H-P d'EBV	114
2. Prédiction d'ordre et de désordre des protéines BSLF1, BBLF2/3 et BBLF4	115
3. Expression des protéines individuelles	116
3.1. Clonage des gènes dans les vecteurs baculovirus	116
3.2. Production individuelle des protéines BSLF1, BBLF2/3 et BBLF4 en cellules adhérentes	117
3.3. Identification des protéines surexprimées	119
3.3.1. Western Blot anti étiquette hexahistidine	119
3.3.2. Séquençage N-terminal	120
4. Purification par affinité sur colonne de nickel	120
4.1. BBLF2/3	120
4.2. BBLF4 et BSLF1	122
5. Purification des protéines BBLF4 et BSLF1 en conditions dénaturantes et renaturation	123
5.1. L'hélicase (BBLF4)	123
5.2. La primase (BSLF1)	124
6. Co-production par co-infection avec les baculovirus codant chacun pour une protéine	125
6.1. Co-production BBLF4 + BBLF2/3	125
6.2. Co-production de BBLF4 +BSLF1	126
6.3. Co-production de BBLF2/3 + BSLF1	126
7. Triple co-infection	127

<b>II. Etude biophysique de BBLF2/3</b>	<b>128</b>
1. Optimisation du tampon de purification	128
2. Caractérisation biophysique par chromatographie d'exclusion de taille, MALLS et DLS	130
2.1. Diffusion dynamique de lumière (DLS, Dynamic Light Scattering)	130
2.2. Chromatographie d'exclusion de taille et MALLS	131
<b>III. Tests de cristallisation de BBLF2/3</b>	<b>132</b>
<b>IV. Synthèse des résultats pour les protéines produites par des baculovirus porteurs d'un seul gène d'intérêt</b>	<b>134</b>

## **CHAPITRE II : ETUDE DU COMPLEXE H-P D'EBV PRODUIT PAR UN BACULOVIRUS PORTANT LES TROIS GENES D'INTERETS**

<b>I. Production du complexe par un seul baculovirus</b>	<b>136</b>
1. Expression des protéines par le système MultiBac	136
1.1. Le système MultiBac	136
1.2. La stratégie de clonage	136
1.3. Test d'expression et de purification	137
2. Purification du complexe H-P d'EBV	138
2.1. Construction 1 (BSLF1 + BBLF2/3 + BBLF4his)	139
2.2. Construction 2 (BSLF1 + BBLF4 + BBLF2/3his)	140
3. Optimisation des conditions de lyse	141
3.1. Test de tampons de lyse en plaque 96 puits	141
3.2. Optimisation des conditions de solubilisation de la construction 1 (BSLF1 + BBLF2/3 + BBLF4his)	142
3.2.1. Dot Blot du test de solubilisation en plaque 96 puits	142
3.2.2. Optimisation des tampons de purifications sur colonne de nickel	143
3.2.3. Affinement des conditions d'utilisation des détergents	146
3.3. Optimisation des conditions de solubilisation de la construction 2 (BSLF1 + BBLF4 + BBLF2/3his)	147
3.3.1. Dot Blot du test de solubilisation en plaque 96 puits	147
3.3.2. Optimisation des tampons de purification sur colonne de nickel	148
<b>II. Confirmation des protéines produites par la construction MultiBac 1</b>	<b>149</b>
1. Identification par séquençage N-terminal	149
2. Identification par spectrométrie de masse	151
3. Identification par Western Blot	153
<b>III. Analyse du complexe par chromatographie d'exclusion de taille</b>	<b>154</b>
1. Tampon optimisé pour la purification par exclusion de taille	154
2. Conditions de lyse cellulaire	157
<b>IV. Etudes structurales du complexe H-P</b>	<b>159</b>
1. Tests de cristallisation	159
2. Etudes du complexe par microscopie électronique	160
2.1. Etudes préliminaires	160
2.2. GraFix	163
<b>V. Etudes des activités enzymatiques du complexe</b>	<b>166</b>
1. Test ATPase	166
2. Tests hélicase	167

<b>PARTIE IV : DISCUSSION ET PERSPECTIVES</b>	<b>168</b>
I. Discussion	169
II. Perspectives	173
<b>PARTIE V : CONCLUSION</b>	<b>176</b>
Conclusion :	177
<b>BIBLIOGRAPHIE</b>	<b>178</b>

## Liste des figures

Titre	Page
<b>Figure 1</b> : Représentation de la « ceinture de lymphomes »	3
<b>Figure 2</b> : Observation du virus Epstein-Barr par microscopie électronique	4
<b>Figure 3</b> : Classification du virus Epstein-Barr au sein de la taxonomie des virus	5
<b>Figure 4</b> : Représentation schématique de la structure d'EBV et cliché de microscopie électronique d'EBV	6
<b>Figure 5</b> : Mécanismes d'attachement et de fusion au lymphocyte B	8
<b>Figure 6</b> : Mécanisme d'infection des cellules épithéliales par EBV	9
<b>Figure 7</b> : Le tropisme d'EBV est déterminé par la présence de la protéine gp42	10
<b>Figure 8</b> : Structure du domaine bZip de ZEBRA en complexe avec un ADN contenant le motif AP-1	14
<b>Figure 9</b> : Modèle du cycle de vie d'EBV, de la persistance et de son association tumorale	18
<b>Figure 10</b> : Représentation schématique de la phase d'initiation de la réplication génomique	21
<b>Figure 11</b> : Représentation schématique de la phase d'élongation de la réplication génomique	22
<b>Figure 12</b> : Schéma de la phase de terminaison de la réplication de l'ADN génomique	23
<b>Figure 13</b> : Protéines composant le complexe de réplication d'HSV-1 et d'EBV	40
<b>Figure 14</b> : Représentation schématique de l'initiation de la réplication lytique d'HSV-1 par OBP	41
<b>Figure 15</b> : Représentation schématique des étapes nécessaires à l'initiation de la réplication d'EBV	42
<b>Figure 16</b> : Alignement des hélicases des herpèsvirus humains	44
<b>Figure 17</b> : Alignement des primases des herpèsvirus humains	46
<b>Figure 18</b> : Alignement des sous-unités accessoires des herpèsvirus humains	47
<b>Figure 19</b> : Représentation schématique de la localisation cellulaire des protéines du complexe H-P d'EBV en l'absence du facteur de transcription ZEBRA	49
<b>Figure 20</b> : Représentation de la localisation cellulaire des protéines du complexe H-P transfectées simultanément avec la protéine ZEBRA	50
<b>Figure 21</b> : Représentation des mutations réalisées sur les 6 motifs hélicase d'UL5	50
<b>Figure 22</b> : Structure des doigts de zinc des primases du bactériophage T7 et de la bactérie <i>Geobacillus stearothermophilus</i>	53
<b>Figure 23</b> : Schéma illustrant les résidus mutés dans la séquence du doigt de zinc d'UL52	53
<b>Figure 24</b> : La composition du complexe de répliatif est dépendante de la concentration du complexe H-P	56
<b>Figure 25</b> : Schéma de l'initiation de la réplication du génome viral du bactériophage T4	59
<b>Figure 26</b> : Structure de la protéine Dda (K38A) en interaction avec un substrat ADNs <sub>b</sub> de 8 nucléotides	61
<b>Figure 27</b> : Modèle de déplacement des protéines SF1B	61



<b>Figure 28</b> : Représentation schématique des étapes requises pour l'expression d'un gène d'intérêt par le système d'expression Baculovirus Bac-to-Bac®	71
<b>Figure 29</b> : Représentation schématique des trois constructions utilisées en cellules adhérentes	72
<b>Figure 30</b> : Représentation du principe de la RT-PCR	73
<b>Figure 31</b> : Principe de la délétion par PCR	74
<b>Figure 32</b> : Représentation schématique des vecteurs donneurs pUCDM des constructions 1 et 2	81
<b>Figure 33</b> : Détails des étapes permettant la formation des plasmides recombinants construction 1 et 2	82
<b>Figure 34</b> : Détails de l'équipement utilisé pour la lyse mécanique avec le « potter »	88
<b>Figure 35</b> : Représentation graphique d'une mesure en Thermofluor	92
<b>Figure 36</b> : Représentation schématique du système SEC-MALLS-RI	100
<b>Figure 37</b> : Courbes obtenues sur le système SEC-MALLS-RI pour la protéine BSA	101
<b>Figure 38</b> : Principe de cristallisation par échange de vapeur	102
<b>Figure 39</b> : Diagramme de phase représentant l'évolution de la composition d'une goutte par échange de vapeur au cours du temps	103
<b>Figure 40</b> : Microscope électronique à transmission utilisé à la plateforme de l'IBS	104
<b>Figure 41</b> : Représentation schématique de la préparation des solutions GraFix	105
<b>Figure 42</b> : Principe du test « ADP-Glo™ Kinase Assay »	108
<b>Figure 43</b> : Détails des substrats utilisés pour mesurer l'activité hélicase du complexe H-P d'HSV1	109
<b>Figure 44</b> : Substrat de type fourche pour le test hélicase	110
<b>Figure 45</b> : Substrat avec brin flottant en 5', hybridé au plasmide M13mp18ss pour le test hélicase	110
<b>Figure 46</b> : Profil de désordre des trois protéines composant le complexe Hélicase-Primase d'EBV obtenu par le script D-score	116
<b>Figure 47</b> : Contrôle par PCR de la présence des gènes BBLF2/3, BSLF1 et BBLF4 dans les bacmides	117
<b>Figure 48</b> : Profil d'expression des trois protéines du complexe Hélicase-Primase	118
<b>Figure 49</b> : Western Blot anti étiquette polyhistidine	119
<b>Figure 50</b> : Profil d'élution de la protéine BBLF2/3 d'une colonne de nickel	121
<b>Figure 51</b> : Digestion de l'étiquette hexahistidine de BBLF2/3 par la protéase TEV	121
<b>Figure 52</b> : Profil d'élution de la protéine BBLF4 d'une colonne de nickel	122
<b>Figure 53</b> : Profil d'élution de la protéine BSLF1 d'une colonne de nickel	123
<b>Figure 54</b> : Profil d'élution de la protéine BBLF4 d'une colonne de nickel après lyse des corps d'inclusion et repliement de BBLF4 par dialyse	124
<b>Figure 55</b> : Profil d'élution d'une colonne de nickel de la co-infection BBLF2/3 et BBLF4	126
<b>Figure 56</b> : Analyse de la co-expression des protéines BSLF1 et BBLF2/3	126

<b>Figure 57</b> : Analyse de la triple production BSLF1, BBLF4 et BBLF2/3 à partir d'une MOI stricte ou cumulée	127
<b>Figure 58</b> : Valeurs des $T_m$ de BBLF2/3 dans 24 tampons	129
<b>Figure 59</b> : Profil d'élution de BBLF2/3 d'une colonne de nickel après clivage de l'étiquette hexahistidine	130
<b>Figure 60</b> : Histogrammes représentant le pourcentage d'intensité et du pourcentage de masse en fonction du rayon d'hydrodynamique ( $R_h$ ) de BBLF2/3	131
<b>Figure 61</b> : Profil d'élution de BBLF2/3 d'une colonne S200 sur un système Basse Pression et MALLS	132
<b>Figure 62</b> : Clichés des inspections à J+3, J+7, J+15, J+32 et J+67 de BBL2/3 à 1,5 et 5 mg/mL	133
<b>Figure 63</b> : Représentation schématique des plasmides recombinants des constructions 1 et 2 contenant les trois gènes du complexe H-P d'EBV.	137
<b>Figure 64</b> : Test d'expression des baculovirus des constructions 1 et 2	137
<b>Figure 65</b> : Profils de purification par affinité sur colonne de nickel des constructions 1 et 2	139
<b>Figure 66</b> : Dot Blot de la construction 1 : BSLF1 + BBLF2/3 + BBLF4his dans 96 conditions de lyse	142
<b>Figure 67</b> : Dot Blot de la construction 2 : BSLF1 + BBLF4 + BBLF2/3his dans 96 conditions de lyse	147
<b>Figure 68</b> : Profil d'élution de la construction 2 (BSLF1 + BBLF4 + BBLF2/3his) sur une colonne de nickel	148
<b>Figure 69</b> : Membrane PVDF colorée au bleu portant les bandes utilisées pour le séquençage N-terminal	149
<b>Figure 70</b> : Profil de digestion ménagée du complexe H-P d'EBV par la Trypsine.	150
<b>Figure 71</b> : Couverture des séquences protéiques des protéines BSLF1, BBLF2/3 et BBLF4 par les peptides identifiés expérimentalement	152
<b>Figure 72</b> : Détection de la protéine BSLF1 par un anticorps anti BSLF1 par Western Blot	153
<b>Figure 73</b> : Purification du complexe H-P par chromatographie d'exclusion de taille avec le tampon n°XII	156
<b>Figure 74</b> : Etude des techniques de lyse du complexe dans le tampon n°XII	158
<b>Figure 75</b> : Clichés de microscopie électronique du pic 2 de gel filtration après concentration	161
<b>Figure 76</b> : Clichés de microscopie électronique en transmission du complexe H-P en coloration négative	162
<b>Figure 77</b> : Etude par microscopie électronique à transmission du complexe H-P après GraFix	164
<b>Figure 78</b> : Détails de particules isolées pour la condition en présence d'ATP et Mg.	165
<b>Figure 79</b> : Réaction hélicase avec des concentrations croissantes de complexe	167

## Liste des tableaux

Titre	Page
<b>Tableau 1</b> : Profils d'expression des gènes latents d'EBV	12
<b>Tableau 2</b> : Tableau de synthèse des mutations réalisées sur les motifs hélicase	51
<b>Tableau 3</b> : Tableau de synthèse des effets des mutations du motif DxD	52
<b>Tableau 4</b> : Etude de l'affinité des complexe H-P sauvage et mutés d'HSV-1 en fonction du substrat ADN.	54
<b>Tableau 5</b> : Liste des 11 prédicteurs utilisés dans le script « D-score »	66
<b>Tableau 6</b> : Détails des 96 conditions de solubilité	86
<b>Tableau 7</b> : Composition des tampons utilisés pour la purification protéique	89
<b>Tableau 8</b> : Détails des 5 tampons composants le gradient de GraFix.	106
<b>Tableau 9</b> : Cartes d'identités des protéines du complexe H-P d'EBV	114
<b>Tableau 10</b> : Tableau récapitulatif des niveaux de surexpression et de solubilité	120
<b>Tableau 11</b> : Tableau détaillant les 96 conditions de tampons de lyse	141
<b>Tableau 12</b> : Détails des tampons utilisés pour la purification par affinité sur colonne de nickel	143
<b>Tableau 13</b> : Analyse des protéines de la construction 1 (BSLF1 + BBLF2/3 + BBLF4his) éluées d'une colonne de nickel en utilisant 9 tampons différents	144
<b>Tableau 14</b> : Détails des conditions utilisées pour le test ATPasique	166

## Liste des abréviations

<b>aa</b> : acide aminé	<b>FPLC</b> : <i>fast protein liquid chromatography</i> , chromatographie en milieu liquide
<b>ADN</b> : acide désoxyribonucléique	<b>GTA</b> : glutaraldéhyde
<b>ADNc</b> : ADN complémentaire	<b>hCMV</b> : cytomégalovirus humain
<b>ADNsb</b> : ADN simple brin	<b>HHV</b> : herpès virus humain
<b>ADP</b> : adénosine di-phosphate	<b>HL</b> : <i>Hodgkin's lymphoma</i> , lymphome de Hodgkin
<b>AppNHp</b> : adénylyl désoxyribonucléique	<b>HRS</b> : cellule de type Hodgkin/Reed-Sternberg
<b>ARN</b> : acide ribonucléique	<b>HSV</b> : virus herpès simplex
<b>ARNm</b> : ARN messenger	<b>IBS</b> : institut de biologie structurale
<b>ATP</b> : adénosine tri-phosphate	<b>K<sub>m</sub></b> : constante d'affinité
<b>ATPyS</b> : adenosine 5'-[γ-thio]triphosphate	<b>KSHV</b> : sarcome de Kaposi associé à l'herpèsvirus
<b>BART</b> : BamH1 A region's rightward transcript	<b>LB-agar</b> : Luria Broth-agar
<b>BL</b> : <i>Burkitt's lymphoma</i> , lymphome de Burkitt	<b>LCL</b> : lignée cellulaire lymphoblastoïde
<b>BSA</b> : <i>bovine albumine serum</i> , albumine de sérum bovin	<b>LMP</b> : <i>latent membrane protein</i> , protéine de latence membranaire
<b>bZip</b> : motif de type basic/leucine Zipper fixant l'ADN	<b>m/v</b> : masse/volume
<b>CG</b> : centre germinatif	<b>m/z</b> : masse/charge
<b>CMC</b> : concentration micellaire critique	<b>MALDI-TOF</b> : matrix-assisted laserDesorption/ionisation - time-of-flight
<b>CMH</b> : complexe majeur d'histocompatibilité	<b>MALLS</b> : <i>multiple angle laser light scattering</i> , diffusion de la lumière par un laser à angle multiple.
<b>complexe H-P</b> : complexe Hélicase-Primase	<b>MCS</b> : <i>multiple cloning site</i> , site de clonage multiple
<b>CSF-1</b> : colony-stimulating factor 1	<b>ME</b> : microscopie électronique
<b>Da/kDa</b> : dalton/kilo-dalton	<b>miARN</b> : micro ARN
<b>Dda</b> : <i>DNA dependant ATPase</i> , ATPase ADN dépendante	<b>MM</b> : marqueur de poids moléculaire
<b>DHB</b> : 2,5-dihydroxybenzoïque	<b>MNI</b> : mononucléose infectieuse
<b>DLS</b> : <i>dynamic light scattering</i> , diffusion dynamique de lumière	<b>MOI</b> : multiplicity of infection
<b>DPA</b> : <i>day after proliferation arrest</i> , jour suivant l'arrêt de la prolifération	<b>NK</b> : natural killer
<b>EBER</b> : <i>EBV encoded RNA</i> , ARN encode par EBV	<b>NLS</b> : <i>nuclear localization signal</i> , signal de localisation nucléaire
<b>EBNA</b> : <i>Epstein-Barr nuclear antigen</i> , antigène nucléaire Epstein-Barr	<b>NPC</b> : <i>nasopharyngeal carcinoma</i> , carcinome du rhinopharynx
<b>EBV</b> : virus Epstein-Barr	<b>nt</b> : nucléotide
<b>EDTA</b> : acide éthylène diamine tétraacétique	<b>NTP</b> : nucléotide tri-phosphate
<b>EEF</b> : <i>eucaryote expression facility</i> , plateforme d'expression eucaryote	<b>oriLyt</b> : origine de réplication lytique (EBV)
<b>EMBL</b> : <i>european molecular biology laboratory</i> , laboratoire européen de biologie moléculaire	<b>oriP</b> : origine réplication latente (EBV)

**oriS** : origine de réplication lytique (HSV-1)  
**P10** : promoteur du polypeptide 10 kDa  
**pb/kpb** : paire de base/kilo paire de base  
**PBS** : *phosphate buffer saline*, tampon phosphate salin  
**PCR** : *polymerization chain reaction*, réaction de polymérisation en chaîne  
**pfu** : *plaque forming unit*  
**PMA** : phorbol-12-myristate-13-acétate  
**PoIH** : polyhédrine  
**PVDF** : polyfluorure de vinylidène  
**R/E** : tampon avec arginine/glutamate  
**RI** : *refractory index*, index de refraction  
**rpm** : rotation par minute  
**RT-PCR** : Transcription inverse par PCR  
**SDS** : sodium dodécyl sulfate  
**SDS-PAGE** : *sodium dodecyl sulfate polyacrylamide gel electrophoresis*, électrophorèse en condition dénaturante sur gel polyacrylamide

**SEC** : *size exclusion chromatography*, chromatographie d'exclusion de taille  
**SF1B** : super famille 1B  
**SIDA** : syndrome de l'immunodéficience acquise  
**SLPT** : syndromes de lymphoprolifération post-transplantation  
**TAE** : tampon tris, acétate, EDTA  
**TCEP** : Tris [2-carboxyethyl] phosphine  
**TEV** : tobacco etch virus  
**TFA** : acide trifluoroacétique  
**T<sub>m</sub>** : température de fusion  
**Tn7/Tn7 att** : transposon/transposon attachement  
**VIH** : virus de l'immunodéficience humain  
**vIL-10** : viral interleukine 10  
**VZV** : virus varicella zoster  
**YFP** : *yellow fluorescent protein*, protéine de fluorescence jaune  
**ZRE** : *ZEBRA responsive elements*, éléments de réponses à ZEBRA

#### Code à une et trois lettres pour les 20 acides aminés :

G - Glycine (Gly)	P - Proline (Pro)
A - Alanine (Ala)	V - Valine (Val)
L - Leucine (Leu)	I - Isoleucine (Ile)
M - Methionine (Met)	C - Cysteine (Cys)
F - Phenylalanine (Phe)	Y - Tyrosine (Tyr)
W - Tryptophan (Trp)	H - Histidine (His)
K - Lysine (Lys)	R - Arginine (Arg)
Q - Glutamine (Gln)	N - Asparagine (Asn)
E - Glutamic Acid (Glu)	D - Aspartic Acid (Asp)
S - Serine (Ser)	T - Threonine (Thr)

# Partie I : Introduction

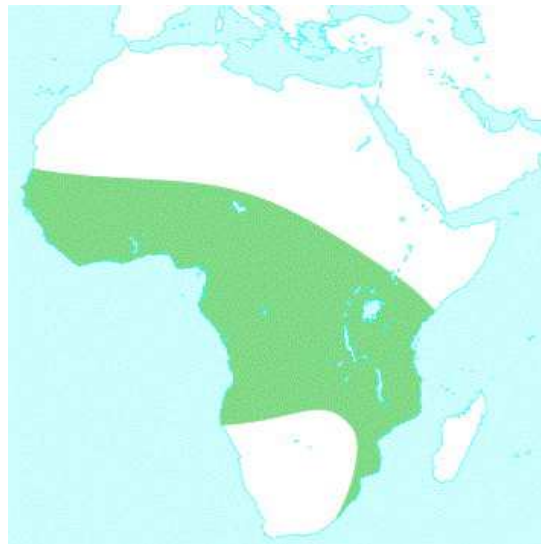
# CHAPITRE 1 :

## Introduction générale

# I. Le Virus Epstein-Barr

## 1. La découverte du virus Epstein-Barr

Le virus Epstein-Barr (EBV) fut découvert au cours des années 1960 à partir de lymphomes affectant les enfants de l'Afrique équatoriale. Pour mieux comprendre la découverte de ce virus, il faut remonter au début des années 1950 et les études menées en Ouganda par un médecin anglais nommé Denis Burkitt. Il décrivit des enfants ayant principalement des tumeurs du maxillaire supérieur mais aussi au sein d'autres organes (reins, surrénales, foie, thyroïde, cœur, estomac et ovaires). Ces premières observations le poussèrent à entreprendre une étude épidémiologique pour décrire la répartition et la prévalence de ces symptômes. Il put mettre en évidence que ces symptômes s'étaient de la côte Ouest à la côte Est africaine dans la zone équatoriale et décrivit cette répartition comme la « ceinture des lymphomes » (Figure 1).

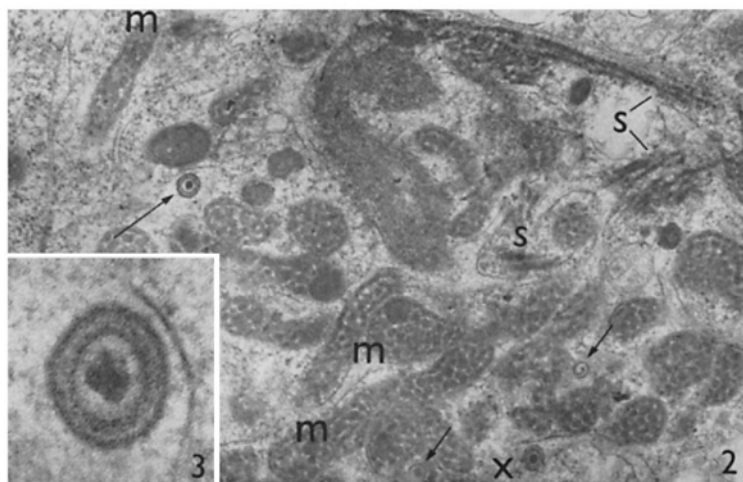


**Figure 1 : Représentation de la « ceinture de lymphomes ».** Le lymphome de Burkitt est endémique des régions ayant une température moyenne supérieure à 15,5°C et des précipitations annuelles supérieures à 50 mL/mm<sup>2</sup>. Cette ceinture s'étend du 10<sup>ème</sup> degré Nord au 10<sup>ème</sup> degré Sud de l'équateur. Source : The Lancet

Cette distribution géographique, correspond à celle du paludisme, ce qui suggéra à Burkitt que la transmission des lymphomes de Burkitt (BL) pouvait provenir d'un agent infectieux ayant pour vecteur un arthropode. Lorsque Burkitt exposa ses résultats au « Middlesex Hospital » de Londres en 1961, il attisa la curiosité de Michael Anthony Epstein. Tout comme Burkitt, Epstein semblait convaincu que ces lymphomes étaient d'origine infectieuse. Burkitt accepta de transmettre à Epstein des biopsies de BL pour rechercher la présence d'un agent infectieux au sein des cellules tumorales. Deux ans durant, Anthony Epstein, Yvonne Barr et Bert Achong cherchèrent à mettre en évidence la présence de particules virales au sein des lymphomes de Burkitt. C'est grâce à un échantillon



présentant des cellules tumorales en suspension qu'ils purent observer pour la première fois des particules virales contenues dans les cellules [1]. L'étude de ces particules permit d'identifier un nouveau membre de la famille des herpèsvirus qui fut nommé virus Epstein-Barr (EBV) (Figure 2). Cette découverte décrit le premier virus associé aux cancers chez l'être humain. Depuis sa découverte, EBV a été décrit comme étant l'agent causal de la mononucléose infectieuse (MNI) et associé à d'autres formes de cancers en plus de celle décrite par Burkitt. EBV est associé à des lymphomes des cellules B (lymphome de Burkitt et lymphome Hodgkin (HL)) et des carcinomes de cellules épithéliales (carcinome indifférencié du rhinopharynx (NPC) et carcinomes gastriques). Il est aussi associé à des lymphoproliférations chez les patients immunodéprimés (transplantés ou sidéens). L'importance d'EBV dans ces différentes maladies est décrite en fin de l'introduction générale.



**Figure 2 : Observation du virus Epstein-Barr par microscopie électronique** d'une coupe de lignée lymphoblastoïde. **2** : La flèche longue représente une particule mature formée de deux membranes refermant un nucléoïde. Les flèches courtes représentent des particules immatures n'ayant qu'une seule membrane. La lettre « m » désigne les mitochondries (x 42 000). **Encadré 3** représente une particule virale mature ayant un diamètre de 115 nm (x 213 500). Source : Epstein & al, 1965 [2].

## 2. EBV est un membre de la famille des herpèsvirus

EBV fait partie d'une très grande famille composée de plus de 120 virus connus, ayant un spectre d'hôtes relativement large, dont des mammifères, des oiseaux, des amphibiens, des reptiles et des poissons [3]. Les virus de type herpès forment l'ordre des *Herpesvirales*. Les herpèsvirus humains sont classés dans la famille des *Herpesviridae* et répartis en trois sous familles, *Alphaherpesvirinae*, *Betaherpesvirinae* et *Gammaherpesvirinae*. Ces trois sous familles sont composées de 8 herpèsvirus humains. Ils possèdent la capacité d'infecter leur hôte de façon persistante et de rester à l'état latent dans des réservoirs de cellules spécifiques. Sporadiquement les virus ont la capacité de sortir de la

latence et d'initier un cycle lytique permettant la production de nouveaux virions infectieux. EBV appartient au genre *Lymphocryptovirus* de la sous famille *Gammaherpesvirinae* (Figure 3). Le sarcome de Kaposi associé à l'herpèsvirus (KSHV) est aussi un *Gammaherpesvirinae* affectant l'homme, mais il est associé au genre *Rhadinovirus*.

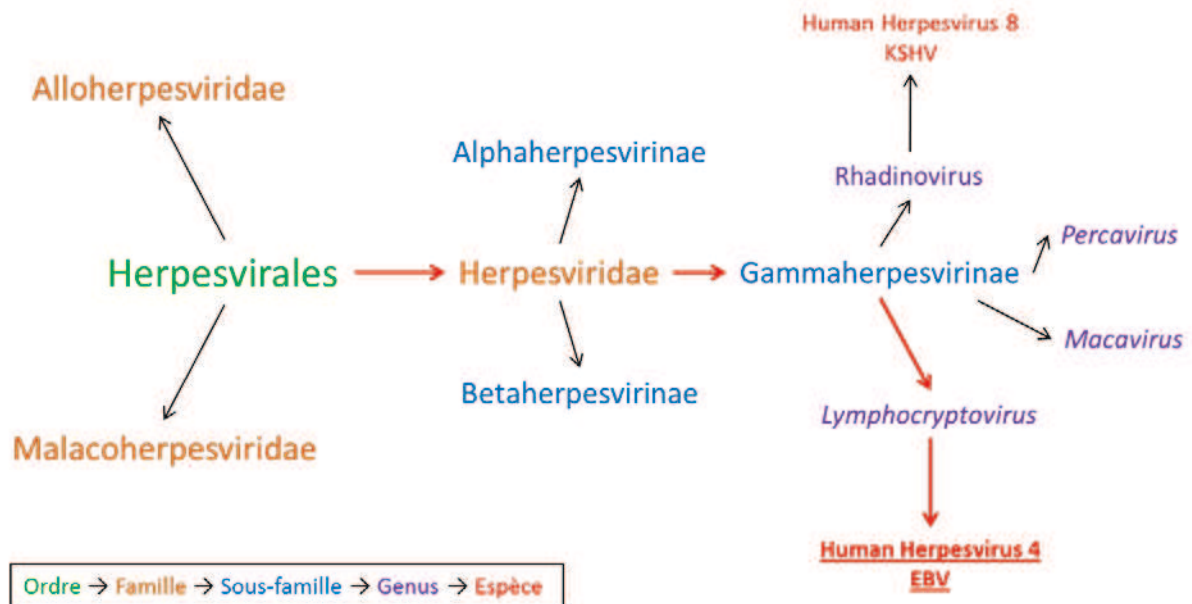


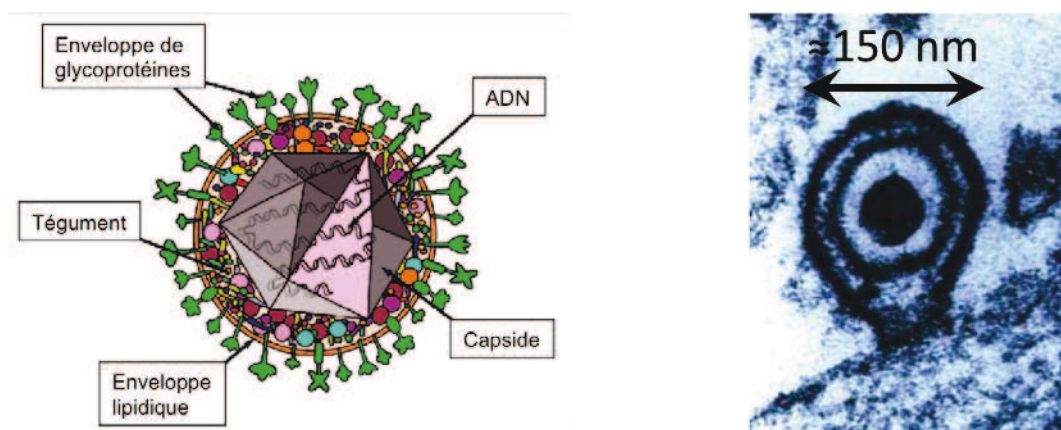
Figure 3 : Classification du virus Epstein-Barr au sein de la taxonomie des virus. Adapté de : <http://www.ictvonline.org>.

- **Les  $\alpha$ -herpèsvirus** infectent expérimentalement une large variété de cellules en culture et une gamme large d'animaux. Les  $\alpha$ -herpèsvirus se propagent rapidement en culture cellulaire et la forte cytotoxicité provoque la lyse des cellules infectées. Dans leurs hôtes naturels, l'infection latente est généralement établie au sein des ganglions sensoriels et l'infection lytique a majoritairement lieu dans les cellules épidermiques. Les virus de cette famille infectant l'Homme sont *Herpès simplex* (HSV) et *Varicella zoster* (VZV) [3].
- **Les  $\beta$ -herpèsvirus** ont un spectre d'hôtes beaucoup plus restreint que la famille  $\alpha$ . En culture cellulaire, ils possèdent des cycles de progression et de réplication plus lents que les  $\alpha$ -herpèsvirus. En culture de cellules, l'infection conduit à une cytomégalie. Dans leurs hôtes naturels, les  $\beta$ -herpèsvirus persistent de manière latente dans les glandes sécrétrices et les reins. Le représentant principal de cette famille infectant l'Homme est le *cytomégalovirus* (hCMV). Les herpèsvirus humain type 6 (HHV-6) et de type 7 (HHV-7) sont aussi des membres de la sous-famille  $\beta$  [3].

- Les  **$\gamma$ -herpèsvirus** possèdent le spectre d'hôte le plus restreint, expérimentalement ils infectent seulement les membres de la famille ou de l'ordre auquel appartient l'hôte naturel. Ils sont spécifiques des lymphocytes B et T où l'infection est généralement latente. Ces virus se répliquent dans les cellules lymphoblastoïdes et certains peuvent infecter de façon lytique les cellules épithéliales et les fibroblastes. Les deux virus de cette famille affectant l'Homme sont *EBV* et *KSHV*, EBV appartient au genre *Lymphocryptovirus* et KSHV au genre *Rhadinovirus* [3] (Figure 3).

### 3. La structure d'EBV

Le virion EBV est composé de 5 parties distinctes. Une **capside** icosaédrique de 150 nm environ, contenant le **génom viral** de 172 kb (lignée B95-8) codant pour environ 85 protéines. Un **tégument** composé de protéines virales est situé entre la capside et la **bicouche lipidique**, dans laquelle sont enchâssées des **glycoprotéines** (Figure 4).



**Figure 4 : Représentation schématique de la structure d'EBV** (réalisation Dr Géoui Thibault (2006)) et cliché de microscopie électronique d'EBV. Adapté de Young & al (2004) [4].

## II. Le cycle viral d'EBV

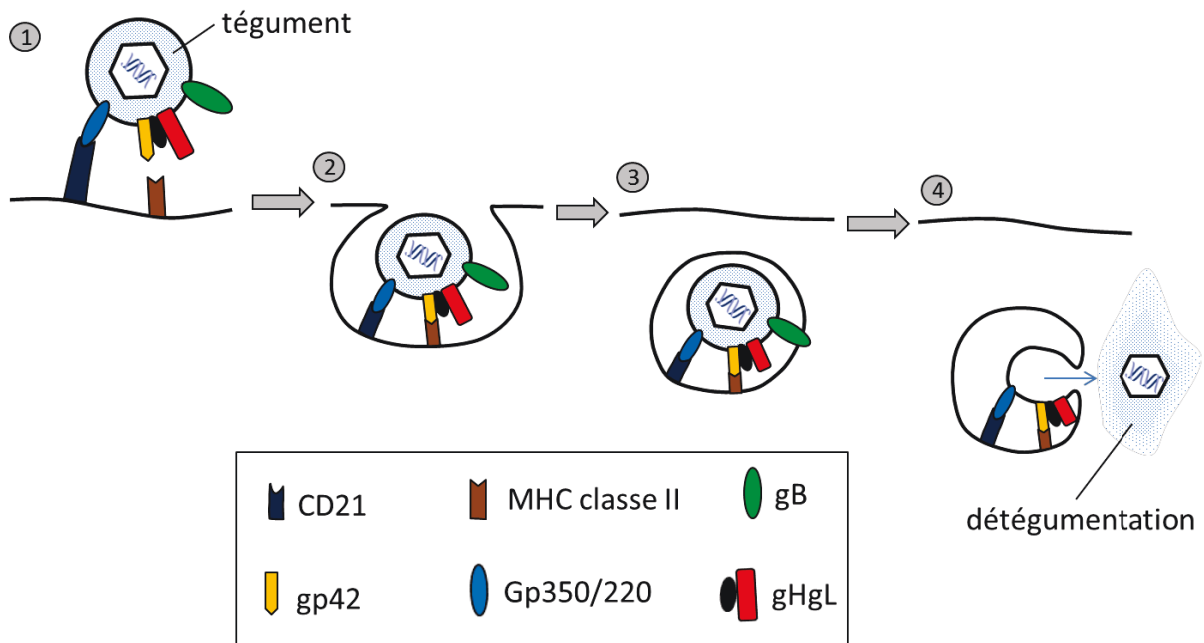
### 1. La transmission virale

*In vivo* l'infection par EBV est transmise le plus souvent par la salive, permettant une diffusion du virus dans la sphère oropharyngée [5]. D'autres modes de transmission sont possibles, par transfusion sanguine [6], lors des greffes d'organes [7] et des transferts entre la mère et l'enfant pendant la grossesse ou lors de l'allaitement [8, 9].

## 1.1. Attachement sur les cellules cibles et internalisation

La primo-infection d'EBV se situe essentiellement dans l'oropharynx, où il infecte préférentiellement les lymphocytes B [10]. Il peut aussi infecter les cellules épithéliales présentes dans la zone oropharyngée [11]. Les maladies auxquelles EBV est associé reflètent un tropisme prédominant d'EBV pour deux types cellulaires : les lymphocytes B et les cellules épithéliales. Les facteurs influençant le tropisme cellulaire d'EBV sont abordés après la description des processus d'attachement et d'entrée dans les cellules B et épithéliales. Quel que soit le type cellulaire infecté par EBV, le processus se décompose en deux étapes distinctes : l'attachement à la cellule cible puis la fusion membranaire permettant l'entrée du virus. La fusion membranaire n'est pas la même en fonction du type cellulaire ciblé par le virus. Miller & al (1992) [12] ont montré que l'entrée d'EBV dans les cellules B était majoritairement permise par endocytose puis fusion avec la membrane endosomale, alors que pour les cellules épithéliales la fusion a essentiellement lieu par l'interaction directe entre les membranes virales et cellulaires. Ces deux mécanismes sont indépendants du pH cellulaire pour initier la fusion membranaire tout comme les virus Sendai, herpès simplex et de l'immunodéficience humaine (VIH) [13].

**Infection des lymphocytes B (Figure 5):** L'étape d'attachement est permise par l'interaction de la glycoprotéine de l'enveloppe virale gp350/220 avec le récepteur cellulaire CD21 des lymphocytes B [14] **(1)**. Ce couple glycoprotéine/récepteur permet la fixation du virus sur la cellule hôte et le rapprochement des protéines virales de fusion de la membrane de la cellule cible. Le virus ne fusionne pas directement sur la membrane plasmique mais est endocyté avant le processus de fusion. Ce processus utilise un triplex composé des glycoprotéines virales gH-gL et gp42. Il est formé par l'interaction de l'extrémité N-terminale de la glycoprotéine gp42 sur le dimère gH-gL. La protéine gp42 interagit avec le complexe majeur d'histocompatibilité (CMH) de classe II présent à la surface des lymphocytes B **(2)**. Cette interaction induit un changement conformationnel de gp42 et/ou gH-gL provoquant l'interaction de gB, la glycoprotéine de fusion d'EBV, avec la membrane de la vésicule [15, 16] **(3)**. Cette dernière interaction permet la fusion des membranes virales et de l'endosome, libérant la capsidie dans le cytoplasme (Figure 5) **(4)**. La décapsidation permet la libération du génome viral linéaire et son importation dans le noyau. Les mécanismes de dissolution de la capsidie et d'importation du génome sont encore peu connus.

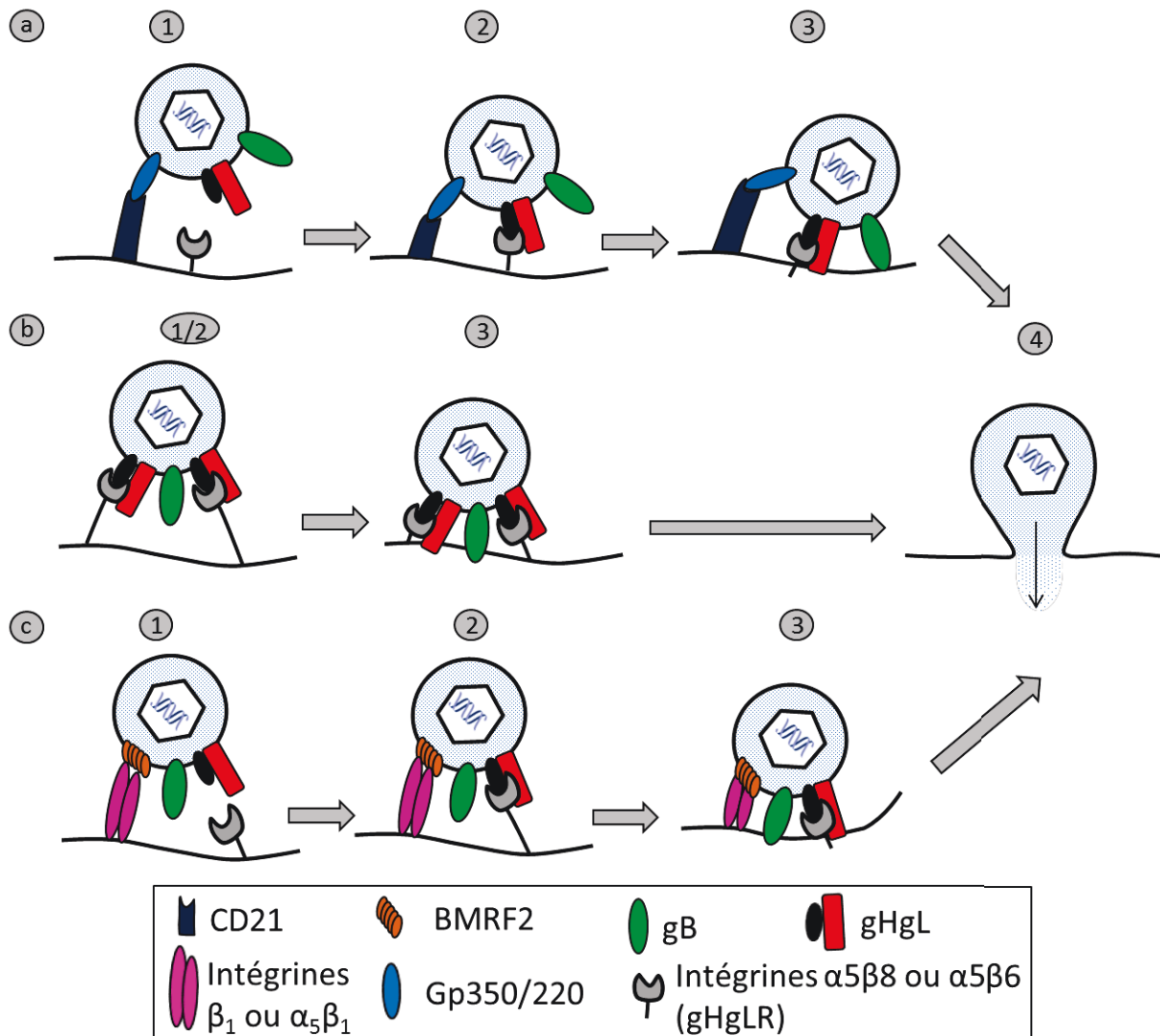


**Figure 5 : Mécanismes d'attachement et de fusion au lymphocyte B.** Les détails des étapes sont donnés dans le texte. Adapté de Hutt-Fletcher (2007) [17].

**Infection des cellules épithéliales (Figure 6):** La compréhension du mécanisme de pénétration dans les cellules épithéliales a fortement progressé dans les dix dernières années.

- Etape d'attachement : L'attachement d'EBV aux cellules épithéliales peut être réalisé par trois mécanismes différents dépendants des récepteurs présents à la surface de la cellule. Les cellules épithéliales exprimant le récepteur CD 21 permettent l'attachement par l'intermédiaire de gp350/220 (a), un processus identique aux lymphocytes B. Pour les cellules n'ayant pas de récepteur CD21 à leur surface, le virus utilise différentes intégrines pour l'attachement. Le dimère gH-gL peut permettre l'attachement direct aux intégrines  $\alpha 5\beta 6$  et  $\alpha 5\beta 8$  (précédemment annoté gH-gLR) présentes à la surface des cellules épithéliales (b). La glycoprotéine de surface BMRF2 permet l'interaction avec les intégrines  $\beta 1$  ou  $\alpha 5\beta 1$  [18] et l'attachement du virus à la cellule épithéliale (c).
- Etape de fusion : Les étapes permettant la fusion du virus à la cellule épithéliale suivent le même processus quel que soit le mode d'attachement utilisé. La gp42 n'est pas utile pour la pénétration dans les cellules épithéliales, sa présence provoque l'inhibition de la fixation et de la fusion [19, 20]. L'interaction entre gH-gL et le récepteur gH-gLR induit un changement de conformation du dimère gH-gL, permettant l'interaction de gB avec la membrane plasmique et d'initier la fusion (2

et 3). EBV est capable de se propager entre cellules épithéliales adjacentes en traversant les membranes latérales au niveau des jonctions serrées [18].

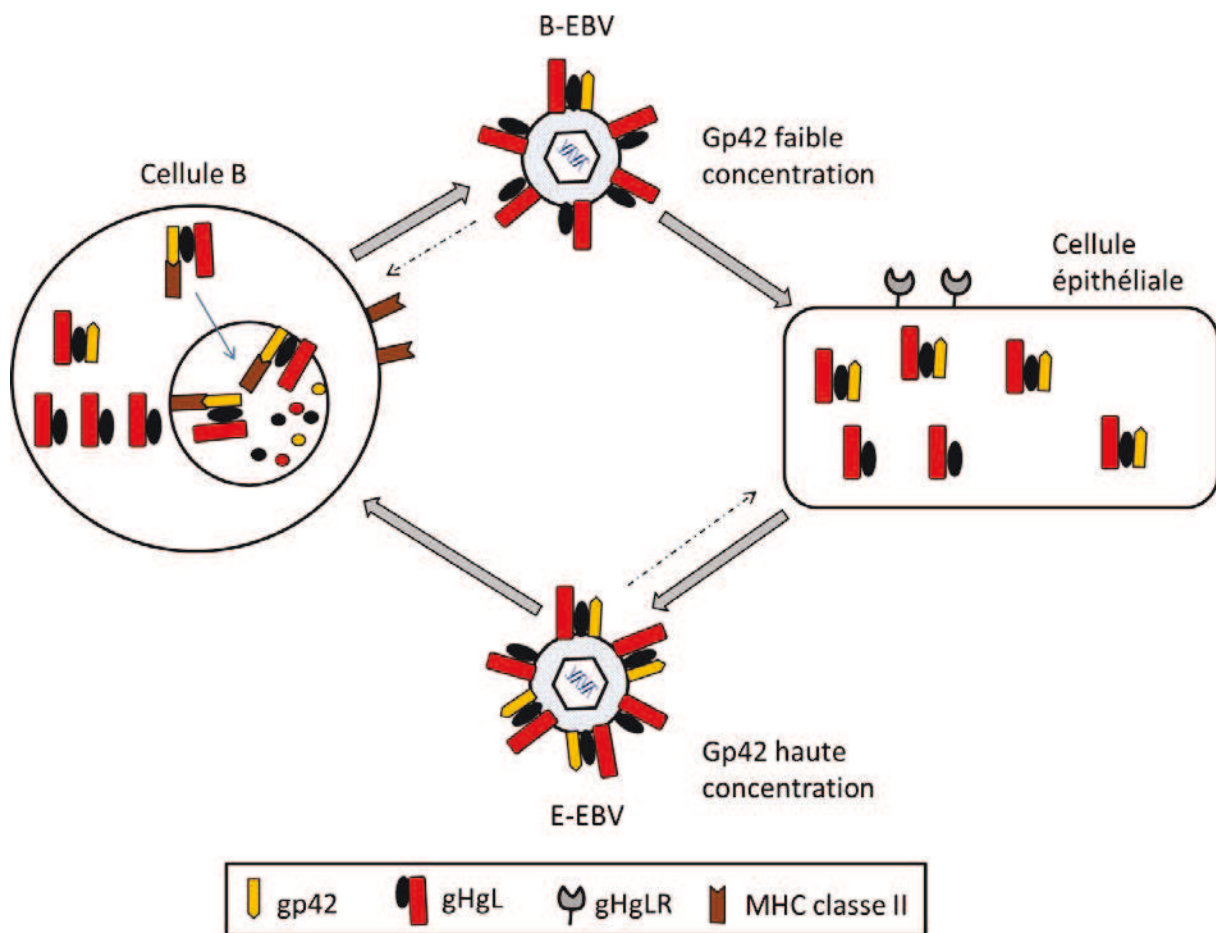


**Figure 6 : Mécanisme d'infection des cellules épithéliales par EBV.** L'attachement du virus aux cellules épithéliales dépend des récepteurs cellulaires de surface. Les détails des mécanismes sont donnés dans le texte. Adapté de Hutt-Fletcher (2007) [17].

## 1.2. Tropisme cellulaire

Les maladies associées à EBV ont permis de décrire les types cellulaires pouvant être infectés par EBV. Il infecte majoritairement deux types cellulaires, les lymphocytes B et les cellules épithéliales, mais dans certaines circonstances il est capable d'infecter d'autres types cellulaires tels que les cellules T, les cellules « natural killer » (NK), les cellules musculaires lisses et potentiellement les monocytes [17]. Les lymphocytes B servent de réservoir au virus. Chez les sujets sains c'est le site de la latence du virus où le génome viral est présent dans le noyau sous forme épisomale. Cette

forme épisomale ne persiste pas au sein d'une culture de cellules épithéliales, ce type cellulaire n'est donc pas considéré comme un site de la latence du virus, mais comme le site d'amplification virale [21]. Le tropisme pour les cellules B et épithéliales dépend essentiellement de la présence de la protéine gp42 à la surface du virus. La fusion d'EBV avec les lymphocytes B requiert la présence du triplex gp42/gH-gL à la surface du virus, alors que l'entrée dans les cellules épithéliales nécessite seulement la présence du duplex gH-gL. Ainsi gp42 est considérée comme la protéine aiguillant le tropisme cellulaire d'EBV [19]. Il a aussi été montré que les virions synthétisés par les cellules B contiennent une plus faible quantité de gp42 (B-EBV) dans leurs enveloppes que les virions produits par les cellules épithéliales (E-EBV). Ainsi les virions synthétisés par les cellules B sont plus aptes à infecter les cellules épithéliales [22], corroborant le rôle de ce type de cellules comme site d'amplification virale (Figure 7).



**Figure 7 : Le tropisme d'EBV est déterminé par la présence de la protéine gp42.** L'infection des cellules épithéliales nécessite un virus exprimant peu de protéines gp42 à sa surface (B-EBV). L'infection des cellules épithéliales permet la production de nouveaux virions présentant une forte densité de protéine gp42 à sa surface (E-EBV). Ces virions sont adaptés pour l'infection des cellules B. L'infection de ces cellules permet la génération de virions avec une faible densité de gp42 à sa surface. Ainsi le tropisme cellulaire d'EBV dépend de la densité de gp42 à sa surface. Adapté de Hutt-Fletcher (2007) [17].

### 1.3. Infection latente, maintien du génome

La latence d'EBV est établie dans les lymphocytes B où le génome persiste sous forme épisomale et n'exprime qu'un nombre très limité de gènes [23]. Pendant le cycle cellulaire, le génome est maintenu grâce à la réplication par une polymérase cellulaire à partir de l'origine de réplication oriP. Pendant la phase de latence jusqu'à 9 protéines peuvent être exprimées ainsi que trois types d'ARNs :

- **Les protéines nucléaires EBNA** (Epstein-Barr Nuclear Antigen) sont au nombre de six, EBNA-1, -2, -3A, -3B, -3C et -LP.
- **Les protéines de latence membranaire LMP** (Latent Membrane Protein) sont au nombre de trois, LMP-1, -2A, -2B.
- **Les ARN nucléaires non polyadénylés EBER** (EBV Encoded RNA) sont au nombre de deux, EBER-1 et EBER-2.
- **Les microARN** (miARN) sont issus de deux régions principales du génome, la première est proche du gène BHRF1 (BamHI fragment H Rightward Transcript), la deuxième est localisée dans la région BART (BamH1 A region's Rightward Transcripts). Les miARN sont généralement nommés BARTs.

Les combinaisons d'expression des protéines et ARNs latents ont permis de décrire 4 types de latence (0, I, II et III) retrouvés dans différentes pathologies *in vivo* et observées *in vitro* [24] (Tableau 1). *In vitro*, EBV a la capacité d'immortaliser les lymphocytes B et d'établir des lignées lymphoblastoïdes (LCL). Les LCL possèdent une latence de type III où l'ensemble des gènes latents sont exprimés. Dans les maladies associées à EBV l'expression des gènes de latence est variable mais plus restreinte. La latence de type I est caractérisée par l'expression d'EBNA-1 et des EBERs, l'expression des gènes EBNA-1, LMP-1, LMP-2A et EBERs caractérise la latence de type II. La latence de type 0, ou latence « vraie », n'exprime que les EBERs et elle est retrouvée au sein des lymphocytes B mémoires (pour revue Young and Rickinson, 2004 [4]).



**Tableau 1 : Profils d'expression des gènes latents d'EBV** correspondant aux programmes de latence et maladies associées. Adapté de Kutok & al (2006) [25].

Programme de latence	Gènes exprimés	maladies associées
0	EBERs	
I	EBERs, BARTs, EBNA-1	Lymphome de Burkitt
II	EBERs, BARTs, EBNA-1, LMP-2A, LMP-1	Lymphome de Hodgkin, Carcinome du rhinopharynx
III	EBERs, BARTs, EBNA-1, LMP-2A, LMP-1, EBNA-2, EBNA-3s, EBNA-LP	Lymphomes associés au VIH, Lymphoproliférations post-transplantation

La protéine EBNA-1 est absolument indispensable pour la réplication du génome viral latent lors du cycle cellulaire. Elle se fixe à oriP et permet le recrutement des protéines de la machinerie de réplication cellulaire [26, 27]. La présence d'EBNA-1 permet le maintien du génome viral sous forme épisomale [28]. La séquence d'EBNA-1 contient une région répétée Gly-Ala prévenant sa dégradation par le protéasome [29]. La protéine EBNA-2 est indispensable pour l'immortalisation des lymphocytes B *in vitro* [30]. Elle a une activité de transactivation de certains gènes viraux (LMP-1 et LMP-2A) [31] et cellulaires. Les protéines EBNA3-A, -B et -C sont des régulateurs de la transcription. La protéine EBNA3-C est essentielle pour la transformation des lymphocytes B *in vitro* et régule les facteurs cellulaires contrôlant la prolifération des lymphocytes B [32]. En présence d'EBNA-2, EBNA-3C favorise l'expression de LMP-1 [33]. La protéine EBNA-LP augmente l'activité de transactivation d'EBNA-2 pour LMP-1 [34].

LMP-1 est la protéine oncogène principale d'EBV, sa délétion abolit l'immortalisation des lymphocytes B *in vitro* [35]. En plus de sa capacité immortalisante, LMP-1 fonctionne comme un récepteur au TNF- $\alpha$  (TNFR) constitutivement actif. Ce mimétisme permet une cascade de signalisation permettant l'expression de cytokines et de protéines anti-apoptotiques favorisant la survie et la croissance cellulaire [36]. Les protéines LMP2 ne sont pas indispensables pour la transformation des lymphocytes B, mais interviennent dans la survie et la prolifération cellulaire. LMP-2A inhibe la voie de signalisation des récepteurs des lymphocytes B, stimulant leur prolifération en l'absence de signaux externes [37].

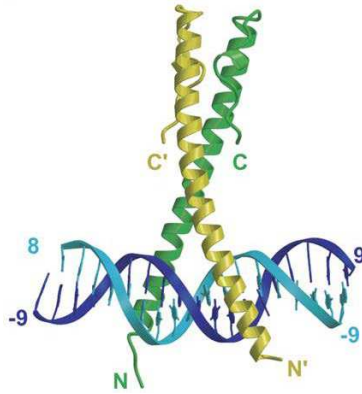
Les ARNs viraux EBERs et BARTs sont exprimés dans tous les types de latence. La séquence des EBERs est très conservée, mais ils ne jouent pas un rôle essentiel dans l'infection et la réplication virale [38]. Les EBERs sont impliqués dans la pathogénèse des maladies associées à EBV et dans la

persistance virale [39]. Les BARTs ne participent pas non plus à la transformation virale, mais permettent la persistance du virus à l'état épisomal dans les cellules infectées [40].

#### 1.4. Infection lytique, un processus séquentiel et ordonné

Dans les lymphocytes B, EBV peut se réactiver et initier sa réplication virale. Cette réplication nécessite l'activation successive des gènes « très précoces », « précoces » et « tardifs ». Le déclencheur de la réactivation *in vivo* n'est pas connu, mais elle peut être induite assez facilement *in vitro*. L'utilisation de composés chimiques comme le PMA (phorbol-12-myristate-13-acétate) ou l'acide butyrique, ainsi que le « cross-linking » des immunoglobulines de surface induisent le cycle lytique. Mais il est aussi possible d'utiliser des cytokines cellulaires (TGF- $\beta$ ) pour initier l'activation des gènes « très précoces », qui sont le point de départ de la réplication lytique [41, 42].

**Les protéines « très précoces » :** La réplication lytique est initiée par l'expression des gènes BZLF1 et BRLF1 codant respectivement pour les facteurs de transcription ZEBRA (aussi nommé Z, Zta ou EB1) et Rta (ou R). La combinaison des protéines ZEBRA et Rta induit l'expression de l'ensemble des protéines précoces permettant la réplication du virus. La structure cristallographique du domaine d'interaction à l'ADN de ZEBRA [43] a révélé que ce facteur de transcription est dimérique et de type bZip (basic/leucine Zipper) (Figure 8, structure obtenue à l'EMBL). ZEBRA se fixe aux motifs ZRE (ZEBRA Responsive Elements) de l'ADN et interagit de préférence avec les promoteurs méthylés des gènes BRRF1 et BRLF1 [44]. Le résidu sérine 186 du domaine basique de ZEBRA est primordial pour l'interaction avec les promoteurs méthylés [45]. Ainsi dans les premières étapes de la réactivation virale ZEBRA fixe préférentiellement le promoteur de Rta. L'activation de BRLF1 induit l'expression de Rta permettant l'activation de certains promoteurs précoces par une interaction directe. La protéine Rta active aussi le promoteur de ZEBRA indirectement grâce à des facteurs de transcription cellulaire (c-jun et ATF-2) [46]. Les facteurs de transcription « très précoces », ZEBRA et Rta sont essentiels dans la réactivation virale et la transcription des gènes « précoces ».



**Figure 8 : Structure du domaine bZip de ZEBRA en complexe avec un ADN contenant le motif AP-1.** Petosa & al (2006) [43]

**Les protéines « précoces » :** Les gènes « précoces » codent pour les protéines impliquées dans la réplication de l'ADN, l'inhibition de l'apoptose, l'évasion du système immunitaire et la régulation des gènes « tardifs ».

La protéine précoce BHRF1 est homologue à la protéine cellulaire anti-apoptotique BCL-2, et prévient l'apoptose de la cellule infectée. La protéine BARF1 semble favoriser l'évasion du système immunitaire, il a été montré *in vitro* que cette protéine est sécrétée à partir de cellules B transportant EBV et dont le cycle lytique a été induit [47]. BARF1 est hexamérique et peut interagir fortement avec la cytokine CSF-1 (colony-stimulating factor 1) [48]. La séquestration de CSF-1 par BARF1 *in vitro*, empêche la prolifération et l'activation de monocytes et de macrophages favorisant l'évasion du système immunitaire [47]. La protéine vIL-10 (viral interleukine 10) est aussi sécrétée et permet la répression de la réponse cytotoxique des cellules T [49]

Le complexe de réplication est composé de six protéines, l'ADN polymérase composé d'une sous unité catalytique (BALF5) et d'un facteur de processivité (BMRF1), la protéine d'adhésion à l'ADN simple brin (BALF2) et le complexe entre la primase (BSLF1), l'hélicase (BBLF4) et la sous-unité accessoire (BBLF2/3) [50, 51]. Ces protéines interagissent avec oriLyt (origine de réplication lytique) et permettent la réplication du génome viral par un mécanisme de « rolling-circle » ou cercle roulant [52].

D'autres protéines ne font pas directement partie du complexe de réplication mais sont importantes pour la réplication virale. Elles interviennent dans le métabolisme et la régulation des désoxynucléotides permettant la production de substrats nucléotidiques dans les cellules non répliquatives. Les connaissances actuelles et les rôles du complexe de réplication et des protéines associées seront développés dans les prochains chapitres.

**Les protéines « tardives » :** Les gènes « tardifs » sont exprimés après la réplication du génome viral et codent pour des protéines structurales impliquées dans la formation du virion. La capsid

virale d'EBV, d'HSV et de KSHV s'assemble de façon autonome par des interactions protéines-protéines. La capsid e d'EBV est composée des protéines suivantes : la protéine majeure de la capsid e (BcLF1), le triplex 1 (BORF1), le triplex 2 (BDLF1), la petite protéine de la capsid e (BFRF3), la protéine échafaudage (BdRF1) et la protéase (BVRF2) permettant l'obtention d'une capsid e mature [53]. La structure de la capsid e d'EBV a été déterminée récemment par microscopie électronique [54]. Les glycoprotéines permettant l'attachement et la pénétration dans la cellule hôte (gp350/220, gH-gL, gp42 et gB) sont exprimées par des gènes « tardifs ».

Les protéines de la latence ou de la réplication lytique d'EBV reflètent une implication profonde au sein des processus cellulaires. L'infection de l'hôte par EBV et sa survie face au système immunitaire de l'Homme sont basées sur des interactions complexes entre les protéines de l'hôte et les protéines virales.

## **2. Les pathologies associées à EBV**

Le virus Epstein-Barr est ubiquitaire et infecte plus de 95% de la population mondiale. La primo-infection est généralement asymptomatique quand elle a lieu pendant l'enfance, mais lorsqu'elle survient à l'adolescence ou à l'âge adulte elle peut être responsable de la mononucléose infectieuse (MNI). EBV n'est pas seulement l'agent causal de la MNI, mais il est aussi associé à plusieurs types de cancers. Il est estimé qu'EBV serait associé à 1% des tumeurs au niveau mondiale, touchant un grand nombre de types cellulaires différents.

### **2.1. La mononucléose infectieuse**

Les symptômes de la MNI sont de la fièvre, des adénopathies, une pharyngite et une fatigue importante. Les cellules épithéliales oropharyngées semblent jouer un rôle important dans les premières étapes de la réplication virale d'une primo infection à EBV. Les virus peuvent ensuite infecter directement les cellules B sous-épithéliales [55]. Au pic clinique de la MNI, entre 0,5 et 2% des cellules B périphériques peuvent contenir le virus [56] et elles présentent un programme de latence de type III contribuant à la multiplication des cellules infectées [57]. L'invasion du système immunitaire par des lymphocytes B infectés par EBV stimule une forte réponse des cellules T CD8+. Les symptômes de la MNI sont majoritairement dus à la prolifération des cellules T, plutôt qu'à la prolifération des cellules B infectées par EBV [58]. Il n'existe pas de traitement spécifique de la MNI, seuls les symptômes peuvent être traités.

## **2.2. Les lymphomes des cellules B**

### **2.2.1. Le lymphome de Burkitt**

Il existe trois variants cliniques de lymphome de Burkitt (BL) comprenant les formes, endémiques, sporadiques et celles liées à l'immunodéficience. La forme BL endémique est présente en Afrique équatoriale et Nouvelle Guinée-Papouasie, touchant principalement les enfants. EBV est retrouvé dans 98 % des cellules tumorales de la forme endémique. La forme BL sporadique est retrouvée dans les pays occidentaux (Etats-Unis et Europe) mais de manière beaucoup plus rare. Dans cette forme EBV n'est détectable que dans 15 à 20 % des cas et dans 30 à 40 % si le sujet est atteint du SIDA (Syndrome de l'immunodéficience acquise) [35]. La pathogénèse du lymphome de Burkitt, aussi bien endémique que sporadique, est caractérisée par la présence de translocations chromosomiques qui juxtaposent le proto-oncogène c-myc avec les gènes codant pour les chaînes lourdes (IgH) ou légères (IgL) des immunoglobulines. La translocation t(8;14) est la plus répandue (80% des cas), elle place le gène c-myc (chromosome 8) au voisinage des éléments activateurs des IgH (chromosome 14) [59]. Deux autres translocations t(2;8) et t(8;22) impliquant les IgL  $\kappa$  (chromosome 2) et les IgL  $\lambda$  (chromosome 22) ont une prévalence de 5 et 15% respectivement [60]. Ces translocations provoquent la surexpression de c-myc favorisant la prolifération des cellules tumorales.

Le lien entre les cellules B infectées par EBV et l'apparition des tumeurs malignes n'est pas encore connu, mais la détection d'EBNA-1 et des EBERs (latence de type I) dans les cellules malignes suggèrent l'importance d'EBV pour le développement de ces lymphomes [61]. Les BL peuvent être traités par chimiothérapie avec un taux de guérison de 90% pour les enfants [62] et 70% pour les adultes.

### **2.2.2. Le lymphome de Hodgkin**

La maladie de Hodgkin est due à une prolifération des lymphocytes B, elle est associée à EBV dans 40 % des cas. La structure des ganglions lymphatiques est perturbée et présente des cellules malignes de type Hodgkin/Reed-Sternberg (HRS) représentant 1 % de la masse tumorale. Les facteurs EBNA-1, LMP-1 et LMP-2 sont détectés dans les cellules HRS caractérisant une latence de type II [63]. EBV est détecté dans 30 à 50% des cellules HRS dans les pays occidentaux et dans 90% des cas dans les pays en développement [64]. Cette maladie se déclenche le plus souvent chez le jeune adulte et présente un second pic d'incidence autour de 45 ans. Les lymphomes de Hodgkin sont traités par

chimio et radiothérapie ; le taux de guérison est de 65 à 90% en fonction du stade d'avancement de la maladie [65].

## **2.3. Les carcinomes des cellules épithéliales**

### **2.3.1. Le carcinome indifférencié du rhinopharynx**

Le carcinome indifférencié du rhinopharynx (NPC) est un cancer des cellules épithéliales avec une répartition forte en Asie du Sud-Est et dans le bassin méditerranéen [66]. Le NPC est une tumeur essentiellement présente chez les adultes avec un pic d'incidence entre 40 et 60 ans. Le déclenchement de la maladie semble lié à des facteurs environnementaux et génétiques [67]. L'ADN viral d'EBV est présent dans 100% des cellules malignes du NPC [68]. Les NPC présentent un type de latence proche du type II : les EBERs 1 et 2 ainsi que la protéine EBNA-1 sont détectés dans 100% des cas [68]. On détecte aussi dans ces carcinomes les protéines latentes LMP-2 et les transcrits BARTs. La protéine LMP-1 est détectée dans 20% des cas [69]. La protéine ZEBRA, associée au cycle lytique du virus, est occasionnellement détectée dans les NPC [70]. Mais l'absence d'autres protéines virales à l'exception de BARF1 et BGLF5 ou de génomes viraux linéaires dans les cellules malignes indique que la tumeur n'est pas permissive à la réplication virale [69]. La grande sensibilité des NPC aux radiations fait de la radiothérapie le traitement principal de ces tumeurs, avec un taux de survie de 84 à 90 % en fonction de l'avancement de la maladie [71].

### **2.3.2. Le carcinome gastrique**

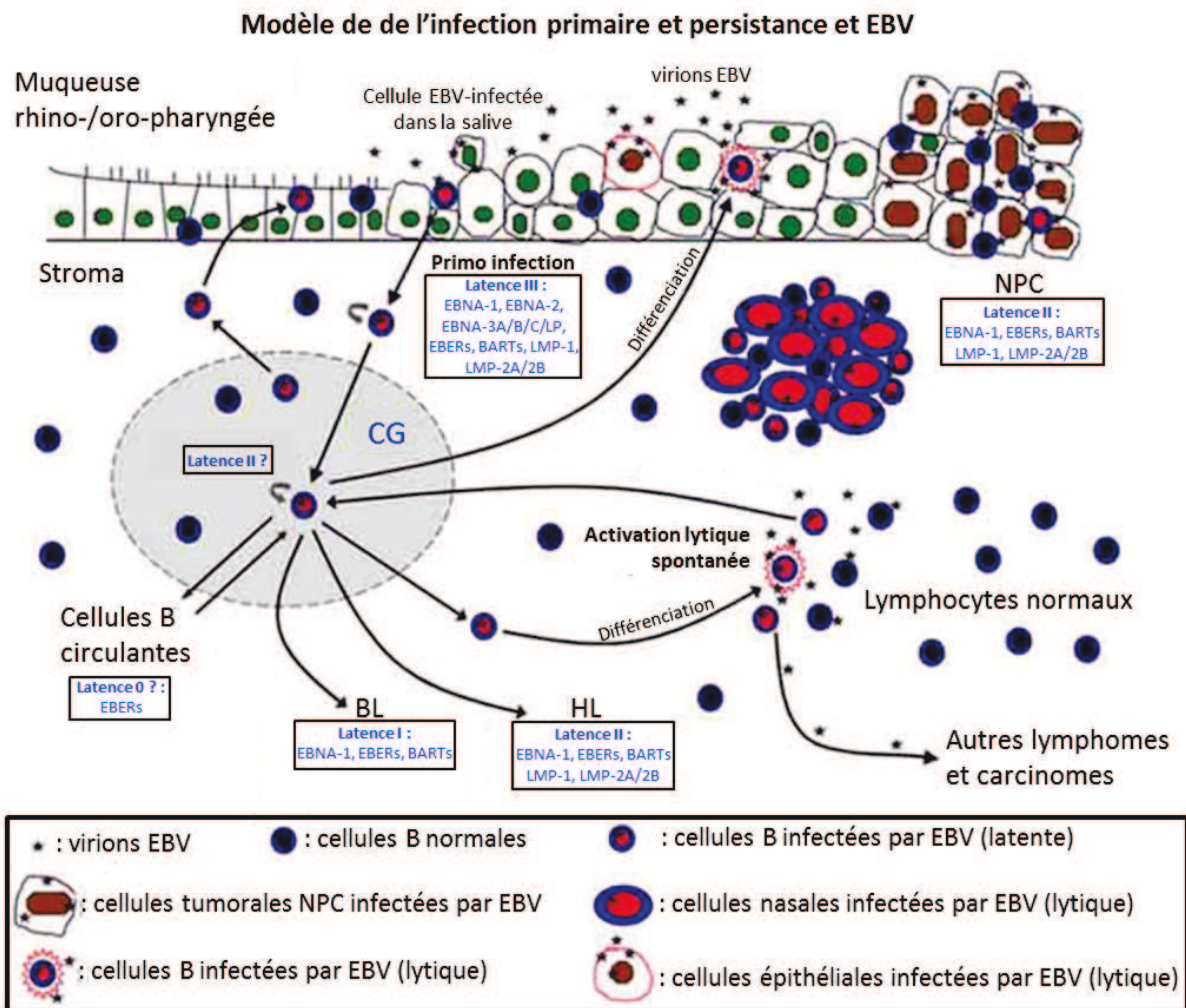
Environ 10% des carcinomes gastriques mondiaux impliquent une prolifération monoclonale de cellules tumorales infectées par EBV [72]. Le carcinome gastrique associé à EBV n'est pas endémique dans une région du globe, il représente 16 à 18% des cas de cancer gastrique aux Etats-Unis et en Allemagne et seulement 4,3% en Chine [73]. L'ARN EBER-1 est détecté dans 100% des cas de carcinomes associés à EBV. Les carcinomes expriment aussi les protéines latentes EBNA-1, LMP-2A et des BARTs. Actuellement la chirurgie est le seul moyen thérapeutique pour les carcinomes gastriques associés à EBV.

## **2.4. Les lymphoproliférations des patients immunodéprimés**

Les Syndromes de Lymphoprolifération Post-Transplantation (SLPT) surviennent chez 5 % des patients greffés. EBV est détecté dans 90% des SLPT avec en majorité une latence de type III, les latences de type I et II peuvent aussi être détectées [4, 74]. Ces lymphomes sont favorisés par les traitements immunosupresseurs nécessaires pour éviter le rejet du greffon. Le traitement réduit la

présence des lymphocytes T qui ne sont plus capables de contrôler la lymphoprolifération des cellules B infectées par EBV.

Les personnes atteintes du SIDA ou porteuses du VIH peuvent présenter des lymphomes associés à EBV dans 60 à 90% des cas. EBV induit une protection contre l'apoptose et favorise la survie cellulaire. Ces particularités propres à EBV en concert avec la suppression du système immunitaire favorisent le développement de lymphomes [75].



**Figure 9 : Modèle du cycle de vie d'EBV, de la persistance et de son association tumorale.** La primo-infection s'établit dans les cellules épithéliales ou dans les lymphocytes B présents dans le tissu épithélial. Le virus présent dans la salive permet la transmission d'hôte à hôte. Les cellules B infectées présentes un profil de latence III et prolifèrent. Après le passage dans le centre germinatif (CG), ils génèrent des lymphocytes B mémoire circulants où le virus est latent avec un profil de type 0. Des cellules B présentant des latences de type I sont présentes dans les lymphomes de Burkitt (BL) ou de type II dans les lymphomes de Hodgkin (HL). La présence de latence de type II dans les cellules épithéliales est caractéristique des carcinomes du rhinopharynx (NPC). Des activations spontanées de la réplication lytique permettent la propagation horizontale et la transmission hôte à hôte.

## 2.5. L'implication du cycle lytique dans les maladies associées à EBV

Les cancers associés à EBV décrits précédemment montrent l'importance des gènes de latence dans le développement ou l'apparition des tumeurs. Il est déjà connu que la réplication lytique d'EBV puisse induire la leucoplasie chevelue de la langue. Cette maladie est essentiellement présente chez les personnes atteintes du SIDA et est due à une réplication lytique massive du virus dans les cellules épithéliales de la langue [76]. Mais il semblerait que les protéines du cycle lytique d'EBV puissent jouer un rôle dans la présence des tumeurs associées à EBV. Une étude récente de Ma & al (2011) [77] utilisant un nouveau modèle de souris humanisées a permis d'étudier plus en détails l'apparition de lymphomes B chez ces animaux. Les souris ont été infectées par un virus EBV contrôle et un virus EBV ayant une délétion du gène codant pour ZEBRA (virus Z-KO). Les deux types de souris présentent des lymphocytes B avec l'établissement d'une infection latente à long-terme, indiquant que la réplication lytique n'est pas indispensable pour la transmission cellule à cellule. Certaines souris infectées par le virus contrôle développent des lymphomes B (6 souris sur 11 infectées) et les souris infectées par le virus Z-KO présentent moins de lymphome B (2 souris sur 14 infectées). Ce résultat suggère qu'une ou plusieurs protéines lytiques favorisent le développement de lymphome B dans ce modèle de souris. Les protéines ZEBRA (protéine « très précoce ») et BRMF1 (protéine « précoce ») sont détectées dans les souris infectées par le virus contrôle, mais pas la protéine « tardive » gp350/220. Ce résultat indique que les premières phases de la réplication lytique sont importantes pour la formation des tumeurs. D'autres arguments plaident en faveur de l'importance du cycle lytique dans les maladies lymphoprolifératives associées à EBV. Dans certains lymphomes ou cancers les patients atteints développent une réponse anticorps importante vis-à-vis des antigènes du cycle lytique suggérant une expression importante de ces antigènes.

Ces résultats récents impliquant un rôle des protéines lytiques d'EBV dans la formation de lymphomes B révèlent toute l'importance de l'étude des protéines lytiques d'EBV et le déroulement de sa réplication virale. Une connaissance accrue des protéines lytiques d'EBV servirait de point de départ pour la confection de nouveaux antiviraux.



## CHAPITRE 2 :

# Le complexe de réplication des herpèsvirus

# I. La réplication de l'ADN génomique et les protéines impliquées

La réplication du génome est un processus biologique qui se déroule en trois grandes étapes.

- **L'initiation** : Ce processus débute par l'adhésion d'un complexe initiateur à l'origine de réplication (1). Cette interaction permet la déstabilisation de la double hélice et de former une région non appariée. L'hélicase va créer une fourche de réplication rendant possible la réplication des deux brins parents. Les parties simples brins de la fourche sont stabilisées par une protéine de liaison à l'ADN simple brin (2). Une fois la fourche répliquative formée, la primase se lie à la fourche de réplication et permet la synthèse d'amorces d'ARN qui seront utilisées par la polymérase pour initier la réplication (3).

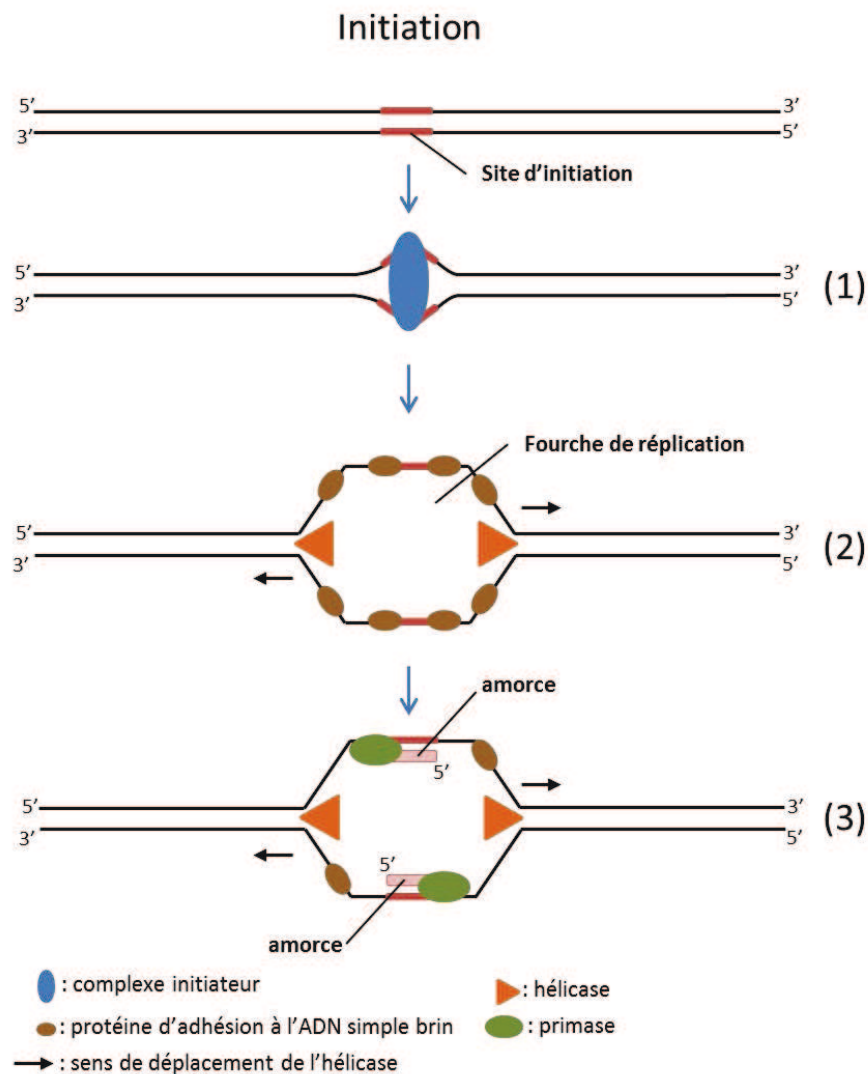


Figure 10 : Représentation schématique de la phase d'initiation de la réplication génomique.

- **L'élongation** : La polymérase réplique l'ADN de façon unidirectionnelle dans le sens 5'-3'. Cette particularité de la polymérase induit un comportement différent selon le brin qu'elle réplique. Sur le brin précoce la polymérase peut répliquer l'ADN de manière continue car l'ouverture de la fourche s'effectue dans le sens 5'-3'(5). Sur le brin tardif la polymérase réplique toujours l'ADN dans le sens 5'-3' mais de façon discontinue car la fourche s'ouvre dans le sens opposé 3'-5'. Sur le brin tardif, la polymérase initie la réplication sur les amorces d'ARN synthétisés par la primase (6) et donne lieu à la formation de fragments d'Okazaki [78].

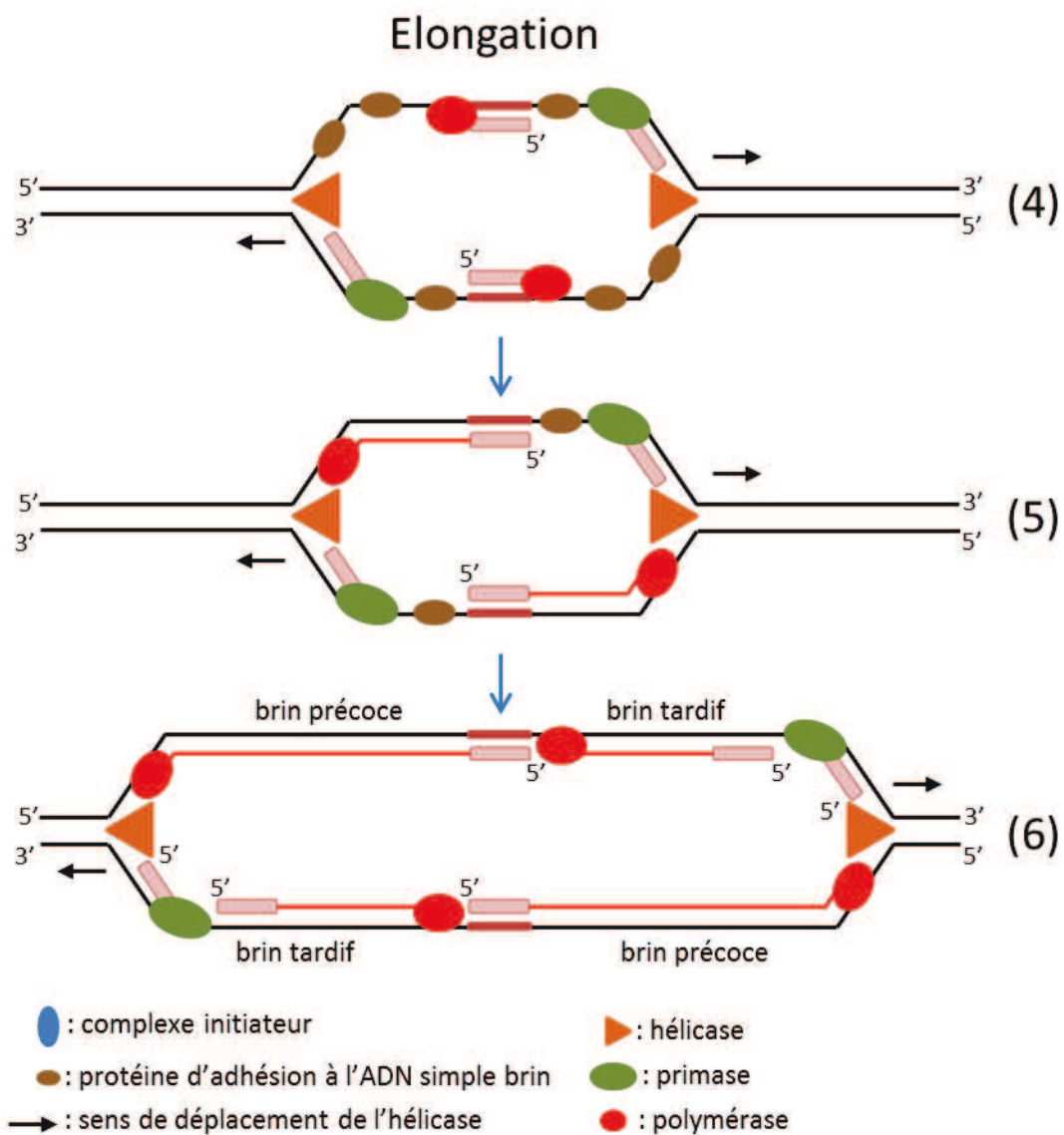
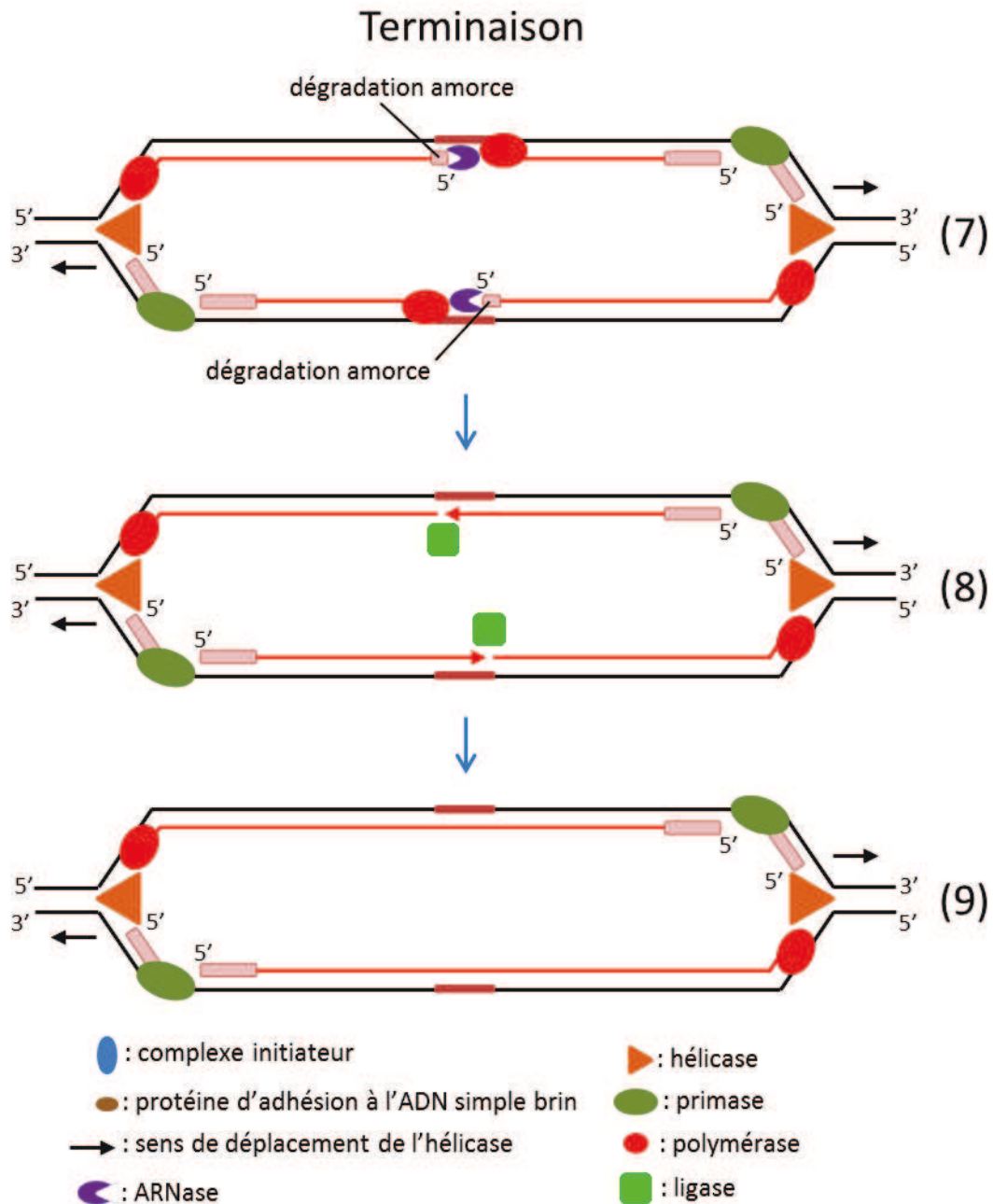


Figure 11 : Représentation schématique de la phase d'élongation de la réplication génomique.

- **La terminaison** : Elle est réalisée par l'action successive de l'ARNase H permettant la dégradation des amorces d'ARN (7). Le remplissage des brèches est réalisé par une polymérase (8) et une ligase permet la jonction des brins d'ADN limitrophes (9).



**Figure 12 : Schéma de la phase de terminaison de la réplication de l'ADN génomique**

Les principes de la réplication exposés précédemment sont généraux et sont présents pour la réplication des herpèsvirus. En fonction de l'état latent ou lytique d'EBV, deux processus de réplication différents sont utilisés. En phase de latence le génome viral est répliqué par le complexe de réplication cellulaire à partir de l'origine de réplication oriP d'EBV. En phase lytique, l'initiation de

la réplication a lieu au niveau de l'origine de réplication oriLyt. Elle serait permise par la présence du facteur de transcription ZEBRA et d'un transcrit d'ARN. La réplication lytique est réalisée par les protéines répliquatives virales selon un mécanisme dit de cercle roulant ou « rolling-circle ». Les protéines répliquatives d'EBV et son mécanisme de réplication sont détaillés dans le paragraphe suivant.

## II. Description structurale des protéines du complexe de réplication d'EBV

### 1. Revue Virologie

Cette section du manuscrit reprend la revue que j'ai écrite en collaboration avec les Pr Patrice Morand et Wim Burmeister. Elle a été publiée dans la revue virologie dans un numéro spécial intitulé : « **Originalité et diversité des mécanismes de réplication chez les virus : cibles de choix pour la conception rationnelle d'antiviraux** ».

La revue détaille les connaissances structurales actuelles des protéines du complexe de réplication d'EBV et des protéines virales importantes pour la réplication mais n'interagissant pas directement avec le complexe répliquatif. Elle présente les protéines dont la structure est connue et ayant un intérêt thérapeutique. La revue décrit aussi les protéines dont la détermination de la structure apporterait une aide considérable pour l'élaboration de nouvelles molécules antivirales.

Référence de la revue : Thierry E, Morand P, Burmeister WP. Structure des enzymes de réplication du virus Epstein-Barr. *Virologie* 2012; 16(4) : 185-98 doi:10.1684/vir.2012.0454



# Structure des enzymes de réplication du virus Epstein-Barr

Eric Thierry<sup>1</sup>  
Patrice Morand<sup>1,2</sup>  
Wim P. Burmeister<sup>1</sup>

<sup>1</sup> Unit of Virus Host Cell Interactions (UVHCI),  
UJF–Grenoble 1/EMBL/CNRS  
UMI 3265,  
6, rue Jules Horowitz,  
BP 181, 38042 Grenoble cedex 9,  
France  
<wpb@embl.fr>

<sup>2</sup> CHU de Grenoble,  
BP 217, 38043 Grenoble cedex 9,  
France

**Résumé.** Le virus Epstein-Barr (EBV) est un pathogène humain ubiquitaire qui établit une infection persistante à vie dans la vaste majorité de la population. Pendant la latence virale, la machine cellulaire assure la réplication de l'épisome viral. En revanche, EBV comme les autres herpesvirus, code de nombreuses enzymes nécessaires pour la réplication lytique de son ADN, ce qui permet la réplication dans des cellules au repos. Récemment, plusieurs structures tridimensionnelles d'enzymes d'EBV ont été résolues, mais aussi pour d'autres herpesvirus. Ainsi on dispose d'informations structurales pour la plupart des enzymes répliquatives. Le mécanisme de réplication lytique d'EBV et la structure des protéines impliquées sont revus par rapport aux possibilités de conception de médicaments et par rapport à l'évolution des herpesvirus. Les structures des protéines impliquées dans la réplication lytique révèlent la proximité évolutive des herpesvirus et des bactériophages à queue, elles montrent aussi la complexité des protéines des herpesvirus vis-à-vis des homologues d'autres organismes. Une analyse phylogénique dans le cadre de cette revue montre pour la première fois la proximité des hélicases répliquatives des herpesvirus avec les hélicases d'initiation de réplication Dda des caudovirales.

**Mots clés :** herpesvirus, cycle lytique, structure de protéines, virologie structurale, réplication, structure cristallographique, virus à ADN, Dda, caudovirales, hélicase, SF1, SF1B

**Abstract.** Epstein-Barr virus (EBV) is a ubiquitous human pathogen which establishes life-long persistent infection in the large majority of the human population. During viral latency, the cellular machinery takes care of the replication of the viral episome. But EBV, as well as herpesviruses in general, codes for numerous enzymes required for lytic DNA replication which allow viral replication in resting cells. Recently, several tridimensional structures of these enzymes became available for EBV as well as for other herpesviruses so that structural information now exists for most of them. The replication process and the structures of the proteins involved in replication are reviewed in the light of potential drug development and of herpesvirus evolution. The structures of the proteins involved in lytic replication show the relationship between herpesviruses and tailed bacteriophages, furthermore they show that EBV proteins tend to be more complex than their counterparts in other organisms. In this review, we could show the phylogenetic position of the herpesvirus helicase close to the Dda helicases involved in initiation of replication of the caudovirales.

**Key words:** herpesvirus, lytic cycle, protein structure, structural virology, replication, crystal structure, DNA virus, helicase, Dda, caudovirales, SF1, SF1B

Tirés à part : W. P. Burmeister



**L**e virus Epstein-Barr (EBV) appartient à la famille des herpèsvirus, qui se divise en trois sous-familles  $\alpha$ ,  $\beta$  et  $\gamma$ . EBV et l'herpèsvirus associé au sarcome de Kaposi (KSHV ou HHV8) sont les deux  $\gamma$ -herpèsvirus humains. EBV possède un génome à ADN double brin d'environ 184 kb codant pour environ 85 protéines. Le génome est contenu à l'intérieur d'une capsidie icosaédrique d'une centaine de nanomètres de diamètre. Cette capsidie est entourée par un ensemble de protéines formant le tégument, lui-même entouré d'une membrane lipidique dans laquelle sont enchâssées des glycoprotéines.

La primo-infection et la persistance d'EBV concernent plus de 95 % de la population mondiale et restent le plus souvent asymptomatiques ou sans gravité. Néanmoins, EBV peut être associé à des pathologies graves qui justifient comme souligné récemment par le National Institut of Health américain (NIH) [1], une recherche sur les moyens de prévenir ou de lutter contre les formes sévères de l'infection. Ainsi, la mononucléose infectieuse aiguë (MNI), généralement bénigne, peut aboutir à des complications aiguës ou chroniques invalidantes [2]. De même, les propriétés immortalisantes d'EBV sont le plus souvent contrôlées par le système immunitaire mais 1 % des cancers humains (lymphome de Burkitt endémique, lymphomes chez les immunodéprimés, carcinome indifférencié du rhinopharynx [NPC] etc.) sont liés à un mauvais contrôle de la persistance virale dans les lymphocytes B ou certains tissus épithéliaux de l'oropharynx [3]. Comme pour les autres herpèsvirus, l'histoire naturelle de l'infection et la physiopathologie des maladies associées à EBV sont le résultat d'interactions complexes entre nos systèmes de défenses innés et acquis et les différentes formes de l'infection virale : infection latente et infection lytique. L'infection latente se caractérise par l'expression de quelques gènes viraux avec production de protéines ou d'ARN non codants mais par défaut sans production de particules virales. L'infection lytique au niveau des plasmocytes ou des cellules épithéliales de l'oropharynx se traduit par l'expression d'un nombre beaucoup plus grand de gènes aboutissant, lorsque le cycle est complet, à la production de particules virales infectieuses et la lyse des cellules infectées. Certaines de ces protéines exprimées très précocement au début du cycle lytique sont importantes pour l'établissement de la latence et pourraient être impliquées, avec les protéines de latence, dans certains des cancers associés à l'EBV [4].

Comme pour les autres herpèsvirus humains, les protéines du cycle lytique impliquées dans le complexe de réplication propre au virus apparaissent comme des cibles potentielles pour des antiviraux spécifiques d'EBV. Les protéines directement associées à la réplication du génome viral sont (tableau 1) : la sous-unité catalytique de la polymérase (BALF5) liée à son facteur de processivité (BMRF1),

la protéine d'adhésion à l'ADN simple brin (BALF2), et le complexe entre l'hélicase (BBLF4), la primase (BSLF1) et la sous-unité accessoire (BBLF2/3) [5]. D'autres enzymes jouent un rôle important lors de la réplication : la nucléase (BGLF5), l'enzyme de réparation uracile-ADN-glycosylase (BKRF3). Les enzymes thymidine kinase (BXLF1), ribonucléotide réductase composée de BaRF1 et de BORF2 et d'UTPase (BLLF3) facilitent la réplication du virus au niveau des cellules non-prolifératives.

Seuls les aspects structuraux de ces protéines permettant de mieux comprendre l'évolution d'EBV et les possibilités d'inhibition de la réplication virale feront l'objet de cette revue. Les autres protéines du cycle lytique impliquées dans l'assemblage des capsides et l'encapsulation de l'ADN, dans la production du tégument ainsi que les glycoprotéines de la membrane virale, bien que constituant des cibles antivirales potentielles, font objet d'autres revues [6, 7] et ne seront pas abordées ici.

## Mécanisme de réplication lytique du virus Epstein-Barr

Le mécanisme de réplication des herpèsvirus est différent, si le virus est en phase lytique ou en phase de latence. Lorsque le virus est en phase de latence, un ADN double brin circulaire appelé épisome est présent au niveau du noyau et se réplique une fois par cycle cellulaire [8]. En présence de la protéine virale EBNA1, l'épisome d'EBV est répliqué par le complexe de réplication cellulaire de façon bidirectionnelle à partir de l'origine de réplication oriP. En phase lytique, c'est un tout autre mécanisme qui est mis en place et le génome d'EBV est amplifié 100 à 1 000 fois [9]. La réplication se déroule entièrement dans le noyau de la cellule. Même s'il a surtout été étudié chez le virus *herpes simplex* (HSV), le mécanisme de réplication lytique est sûrement similaire au sein de la famille des herpèsvirus où les protéines formant le complexe de réplication sont relativement conservées.

L'épisome d'EBV possède deux copies de l'origine de réplication lytique OriLyt (figure 1) [1], mais une seule est suffisante pour la réplication. Par ailleurs, la souche de laboratoire principale B95-8 n'en contient qu'une seule [9]. L'initiation de la réplication dépend de la protéine BZLF1 d'EBV (aussi appelée ZEBRA, Zta ou EB1) [10] et utilise aussi des transcrits d'ARN des gènes *BHLF1*, *LF3* ou *BHRF1*. L'initiation nécessite la formation de duplexes ADN-ARN et le déplacement d'un brin d'ADN dans une structure appelée R-loop [11]. L'ADN simple brin présent au niveau de cette structure permettrait le recrutement de la protéine de liaison à l'ADN simple brin (BALF2), du



**Tableau 1. Homologues des protéines de réplication chez les différentes familles des herpesvirus. Si les structures tridimensionnelles ont été résolues, le nom de l'entrée dans la Protein data bank (PDB) est indiqué et souligné en gras. Les tailles des protéines EBV sont également données.**

	$\gamma$ -herpèsvirus		$\alpha$ -herpèsvirus		$\beta$ -herpèsvirus
Polymérase <sup>a</sup>	<i>EBV</i> <i>BALF5</i> 1015 aa	<i>KSHV</i> <i>UL30</i>	<i>HSV</i> <i>UL30</i> <u><b>2GV9</b></u>	<i>VZV</i> <i>Orf28</i>	<i>hCMV</i> <i>UL54</i>
Facteur de processivité <sup>a</sup>	<i>BMRF1</i> <u><b>2Z0L</b></u> 404 aa	<i>PF-8</i> <u><b>3HSL</b></u>	<i>UL42</i> <u><b>1DML</b></u>	<i>Orf16</i>	<i>UL44</i> <u><b>1T6L</b></u>
Protéine d'adhésion à l'ADN simple brin	<i>BALF2</i> 1128 aa	<i>ORF6</i>	<i>UL29</i> <u><b>1URJ</b></u>	<i>Orf29</i>	<i>UL57</i>
Hélicase <sup>b</sup>	<i>BBLF4</i> 810 aa	<i>ORF44</i>	<i>UL5</i>	<i>Orf55</i>	<i>UL105</i>
Primase <sup>b</sup>	<i>BSLF1</i> 875 aa	<i>ORF56</i>	<i>UL52</i>	<i>Orf6</i>	<i>UL70</i>
Protéine accessoire <sup>b</sup>	<i>BBLF2/3</i> 710 aa	<i>ORF40</i>	<i>UL8</i>	<i>Orf52</i>	<i>UL102</i>
Uracile-ADN glycosylase	<i>BKRF3</i> <u><b>2J8X</b></u> 255 aa	<i>ORF46</i>	<i>UL2</i> <u><b>1UDI</b></u>	<i>Orf59</i>	<i>UL114</i>
Exonucléase	<i>BGLF5</i> <u><b>2W45</b></u> 470 aa	<i>ORF37,</i> <i>SOX,</i> <u><b>3FHD</b></u>	<i>UL12</i>	<i>Orf48</i>	<i>UL98</i>
dUTPase	<i>BLLF3</i> <u><b>2BSY</b></u> 278 aa	<i>ORF54</i>	<i>UL50</i>	<i>Orf8</i>	
Thymidine kinase	<i>BXLF1</i> 607 aa	<i>ORF21</i>	<i>UL23</i> <u><b>1E2J</b></u>	<i>Orf36</i> <u><b>1OSN</b></u>	
Ribonucléotide réductase R1 <sup>c</sup>	<i>BORF2</i> 826 aa	<i>ORF61</i>	<i>UL39</i>	<i>Orf19</i>	<i>UL45</i>
Ribonucléotide réductase R2 <sup>c</sup>	<i>BaRF1</i> <u><b>4A5E</b></u> 302 aa	<i>ORF60</i>	<i>UL40</i>	<i>Orf18</i>	

<sup>a,b,c</sup>Protéines appartenant à un même complexe.

complexe hélicase-primase (*figure 1*) [2] et finalement de la polymérase (*BALF5*) avec son cofacteur (*BMRF1*, *figure 1*) [2].

Au début, la réplication est similaire à une réplication  $\theta$  avec une bulle réplivative (*figure 1*) [3,4] qui devient une réplication de type « rolling circle » après le premier cycle de réplication (*figure 1*) [6,7]. L'hélicase-primase permet la séparation des deux brins devant le complexe polymérase. Elle dépose aussi périodiquement des amorces d'ARN sur le brin complémentaire pour permettre la synthèse du brin retardé. La protéine d'adhésion à l'ADN simple brin (*BALF2*) interagit avec les régions monocaténaïres présentes temporairement au niveau du brin complémentaire. Finalement, l'ARNase H cellulaire dégrade les amorces d'ARN ; c'est probablement la polymérase virale qui remplit les brèches et une ligase cellulaire achève la synthèse

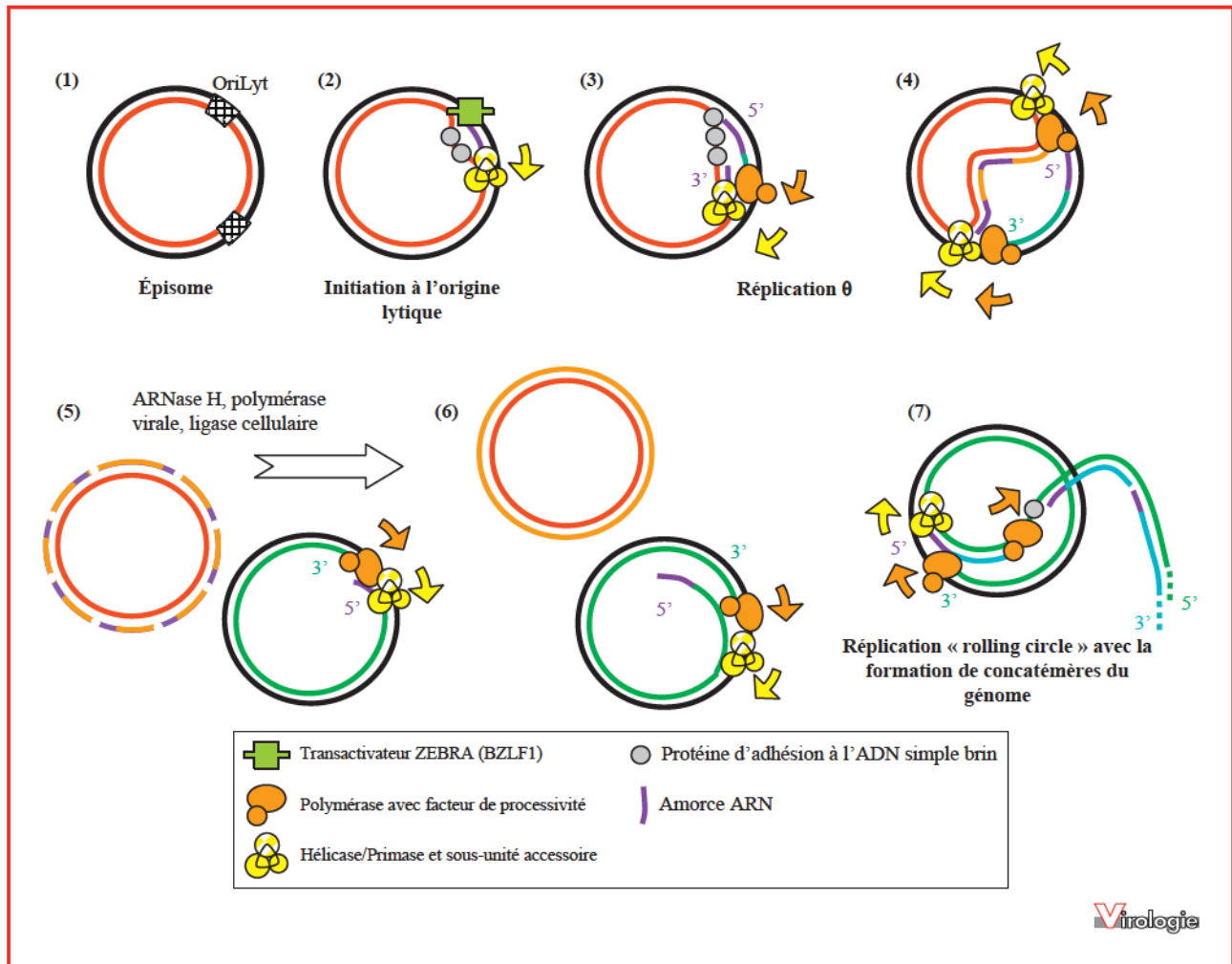
de l'ADN double brin (*figure 1*) [4,5]. Dans ce processus, il y a une formation d'intermédiaires branchés à partir d'événements de recombinaison ou d'invasion de brin. L'exonucléase (*BGLF5*) intervient afin de restaurer un ADN double brin [12]. Ce mécanisme permet de produire de longs concatémères de génomes qui sont par la suite clivés en unités par le complexe terminase au moment de l'encapsidation.

## Les protéines du complexe de réplication du virus Epstein-Barr

### La polymérase (*BALF5*)

En absence d'une structure tridimensionnelle de la polymérase d'EBV, on dispose de la structure de la polymérase

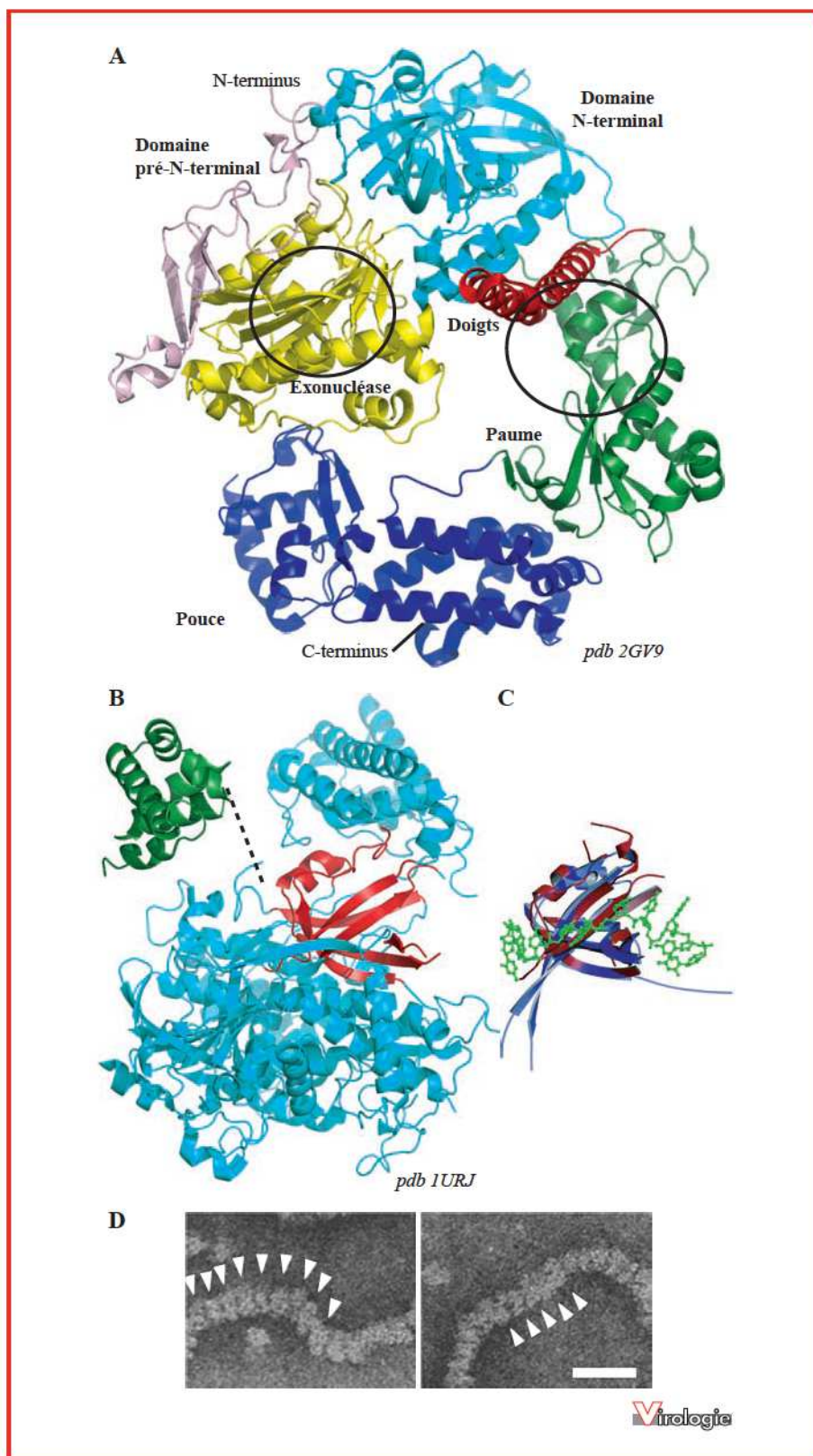




**Figure 1.** Mécanisme de réplication lytique du virus EBV avec initiation, réplication  $\theta$ , et la transition vers la réplication en cercle roulant. Les brins en rouge et noir représentent l'épisome parental, les amorces ARN sont colorées en violet et l'ADN nouvellement synthétisé en orange, vert et bleu. Les flèches colorées indiquent le sens de progression des complexes de l'hélicase-primase et de la polymérase sur l'ADN. (1) Représentation schématique de l'épisome. (2) Initiation de la réplication au niveau d'OryLyt avec le recrutement du transactivateur ZEBRA, d'un ARNm, de la protéine d'adhésion à l'ADN simple brin et du complexe hélicase-primase. (3) Début de la réplication  $\theta$  avec ouverture de la fourche de réplication et début de la synthèse du brin avancé. (4) Progression de la réplication  $\theta$  avec synthèse du brin avancé et du brin retardé. (5) Produits à la fin de la réplication  $\theta$  : l'épisome répliqué, représenté encore avec les brèches entre les fragments Okazaki du brin retardé ainsi que le brin avancé en transition vers une réplication de type cercle roulant. (6) Finalisation de la réplication de l'épisome par des enzymes cellulaires ARNase H, ligase et la polymérase virale. On retrouve la structure d'ADN double brin circulaire ainsi que le brin avancé qui continue d'être répliqué par le mécanisme du cercle roulant. Un brin retardé complémentaire est synthétisé. (7) Génération de concatémères du génome viral par la réplication de type cercle roulant.

de HSV (figure 2A) [13]. L'identité de séquences entre les deux polymérases est de 35 % et elles s'alignent quasiment sur toute la longueur à l'exception de quelques insertions pour EBV. Ces polymérases appartiennent à la famille des polymérases  $\alpha$ . Elles comportent un domaine polymérase proprement dit et un domaine 3'-exonucléase permettant la vérification du brin nouvellement synthétisé (« proof-reading »). L'interaction avec l'ADN suivant les deux modes d'action – polymérase et 3'-exonucléase – a

été largement étudiée pour la polymérase POL du bactériophage RB69 [14, 15] qui ressemble à celle des herpesvirus. La polymérase de HSV est composée de six domaines, un domaine pré-aminoterminal, un domaine aminoterminal, le domaine exonucléase et le domaine polymérase avec paume, doigts et pouce [13]. Cette organisation est partagée plus largement au sein de la famille des polymérases  $\alpha$ . Les extensions de séquence spécifique aux herpesvirus sont surtout localisées au niveau du domaine



**Figure 2.**

Virologie, Vol 16, n° 4, juillet-août 2012



pré-aminoterminal ou dans des boucles éloignées des sites actifs. Le C-terminus de la polymérase se lie au facteur de processivité viral. Cette interaction est décrite pour HSV [16], EBV [17], mais aussi pour le bactériophage RB69 [14]. Sans la présence de ce facteur, ces polymérases sont peu processives et peu de nucléotides (quelques dizaines pour EBV [17]) sont incorporés lors de la rencontre entre la polymérase et son substrat d'ADN simple brin associé à une amorce. En revanche, en présence du facteur de processivité la synthèse peut incorporer des dizaines de milliers de bases (plus de 7 200 pour EBV [17]).

Récemment, la polymérase de HSV a été impliquée dans la réparation des sites abasiques générés par l'excision d'une base par l'enzyme de réparation uracile-ADN glycosylase (UDG) avec une activité APase (activité endonucléase au niveau des sites abasiques), une activité polymérase réparatrice et le retrait du 5'-désoxyribose phosphate par une activité lyase [18]. Cette activité n'a pas encore été confirmée pour d'autres herpesvirus.

Le domaine polymérase est la cible des antiviraux anti-herpétiques de la famille des analogues nucléosidiques qui agissent comme des terminateurs de chaîne. Ces analogues sont incorporés dans l'ADN et ne peuvent pas être retirés d'une façon efficace par l'activité 3'-exonucléase de la polymérase. Des souches HSV ou du cytomégalovirus humain (hCMV) peuvent devenir résistantes à ces analogues nucléosidiques par des mutations localisées dans le site actif polymérase ou au niveau du site actif exonucléase (*figure 2A*) jouant ainsi sur le degré d'incorporation ou sur la facilitation de l'hydrolyse [19]. La polymérase de ces deux virus a fait également l'objet du développement d'inhibiteurs non-nucléosidiques [13, 20].

#### Le facteur de processivité (BMRF1)

Cette protéine, appelée également EA-D chez EBV (antigène précoce diffus) ou *sliding clamp* est considérée comme sous-unité de la polymérase dont l'extrémité C-terminale se lie à BALF5 [16] et augmente la processivité et l'activité 3'-exonucléase de la polymérase. BMRF1 est exprimée en large excès par rapport à la polymérase lors du cycle lytique. La forme libre de BMRF1 est essentiellement associée à l'ADN viral nouvellement synthétisé [21]. Elle joue aussi un rôle dans la transcription des gènes d'EBV,

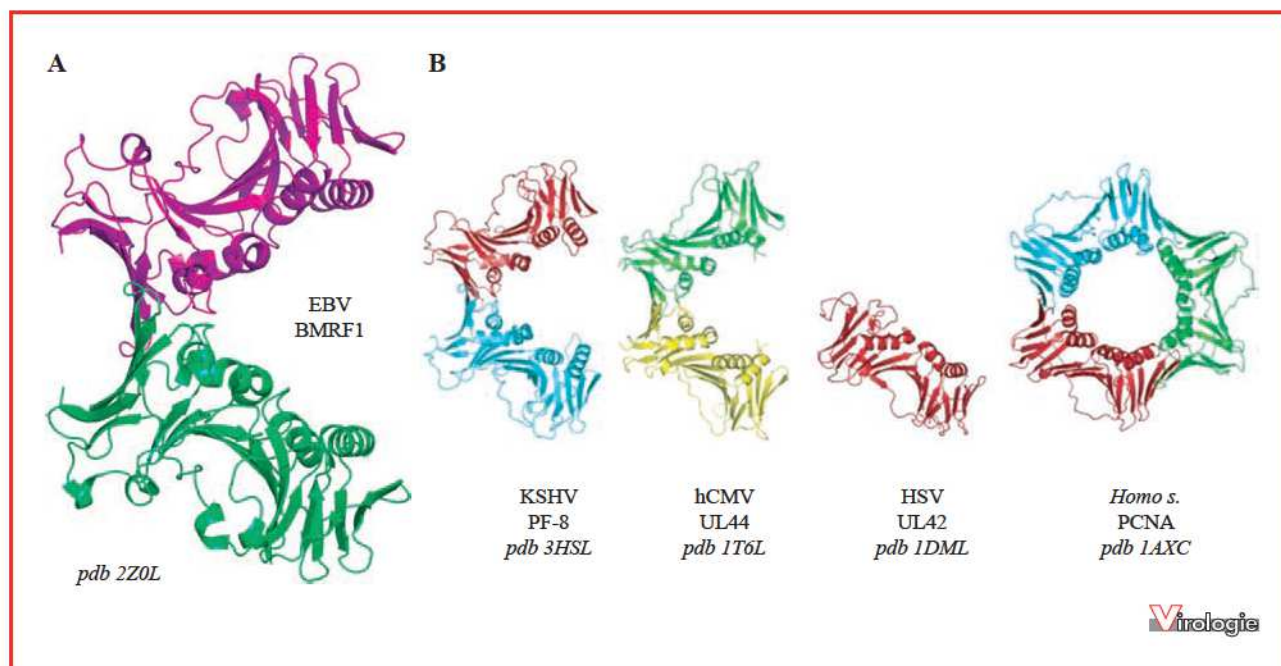
surtout au niveau de BHLF1 [22]. La structure cristallographique de nombreux facteurs de processivité a été déterminée (HSV [16], hCMV [23], KSHV [24] et EBV [25]). Les formes cristallisées (*figure 3*) sont tronquées d'environ 90 résidus au niveau du C-terminus mais ces résidus ne comportent aucun élément de structure secondaire prédite. Les facteurs de processivité présentent une face portant une forte charge positive [24] qui interagit étroitement avec l'ADN double brin. Malgré cette forte affinité, une grande mobilité des facteurs de processivité sur l'ADN double brin est assurée [26]. Alors que la protéine de HSV agit sous forme de monomère en solution [27], les protéines d'EBV, de KSHV et de hCMV agissent sous forme de dimères (*figure 3B*). Ces dimères sont observés au niveau des structures cristallographiques. Les structures tridimensionnelles des facteurs de processivité ont révélé une grande similitude par rapport au facteur de processivité eucaryote PCNA [28] (*figure 3B*) et aussi vis-à-vis de la protéine correspondante gp45 du bactériophage RB69 [14], malgré un niveau d'identité de séquence extrêmement faible (*figure 3A*). Contrairement aux protéines des herpesvirus, PCNA [28] et la protéine RB69 forment des anneaux trimériques qui entourent complètement l'ADN, d'où le nom de *sliding clamp*, et qui nécessitent une protéine de chargement dédiée (*clamp loader*), tandis que les dimères des herpesvirus conservent un arrangement similaire des sous-unités mais forment des anneaux incomplets (*figure 3B*).

#### La protéine d'adhésion à l'ADN simple brin (BALF2)

Lors de la synthèse d'ADN, les protéines d'adhésion à l'ADN monocaténaire maintiennent l'ADN à l'état simple brin et empêchent la formation d'éventuelles structures secondaires lors des processus de réplication, de réparation et de recombinaison. BALF2 n'a pas d'influence sur la processivité de la polymérase mais elle permet, en supprimant les structures secondaires de l'ADN, une élongation complète du substrat monocaténaire par le complexe polymérase BALF5-BMRF1 [29]. Cette protéine semble être monomérique en solution [29]. Mumtsido *et al.* [30] ont exploré plus en détails la structure de BALF2 en utilisant une version BALF2- $\Delta$ C tronquée coté C-terminus réduisant l'auto-agrégation. La microscopie électronique a permis de caractériser l'interaction entre cette protéine

**Figure 2.** Structures des homologues HSV des protéines de réplication EBV avec une structure inconnue. **A)** Structure de la polymérase UL30 de HSV avec les différents domaines de la structure. Le cercle gauche indique la position du site exonucléase, le cercle droit celle du site de l'activité polymérase. **B)** Structure de la protéine d'adhésion à l'ADN simple brin ICP8 de HSV [31]. Domaine C-terminal (vert) et domaine N-terminal (cyan et rouge). La partie correspondante à la protéine d'adhésion à l'ADN simple brin HsmtSSB est colorée en rouge. **C)** Superposition de HsmtSSB (bleu) [32] avec le motif d'adhésion à l'ADN simple brin d'ICP8 (en rouge) montré dans la même orientation qu'au niveau du panneau B. L'ADN modélisé est montré en vert (reproduit avec autorisation de [30]). **D)** Interaction de BALF2 avec un ADN monocaténaire circulaire en microscopie électronique. La barre indique 20 nm, les triangles indiquent des molécules individuelles de BALF2- $\Delta$ 60 alignées sur l'ADN monocaténaire (reproduit avec autorisation de [30]).





**Figure 3.** Structure des protéines de réplication d'EBV. **A)** Un dimère du facteur de processivité BMRF1. **B)** Comparé aux autres facteurs de processivité des herpesvirus et du PCNA eucaryote (reproduit avec autorisation de [24]).

et l'ADN monocaténaire (figure 2D). BALF2 recouvre environ 15 nucléotides par monomère. Elle a une identité de 25 % avec la protéine correspondante de HSV, ICP8. La structure cristallographique d'ICP8 a été déterminée (figure 2B) après délétion de 60 résidus coté C-terminal, permettant une interaction coopérative avec l'ADN simple brin [31]. La structure montre un domaine N-terminal très large avec un repliement unique et un petit domaine C-terminal hélicale connecté par une séquence flexible. Le domaine N-terminal comporte des éléments de structure qui ressemblent fortement à la protéine d'adhésion à l'ADN simple brin mitochondriale humaine HsmtSSB [32] (figure 2B et C). En revanche, cette dernière n'est composée que de 109 résidus comparés aux 1 128 résidus de BALF2.

#### Le complexe hélicase-primase

Ces protéines sont nécessaires à l'ouverture de la fourche de réplication par l'activité hélicase et à la synthèse du brin d'ADN retardé par l'activité primase. Les produits des cadres de lecture BBLF4 (hélicase), BSLF1 (primase) et BBLF2/3 (sous-unité accessoire) interagissent pour former le complexe avec un ratio 1:1:1 [33]. Aucune information structurale n'est disponible pour ce complexe et les travaux sur les activités fonctionnelles du complexe ont été menés avec les protéines équivalentes de HSV. Pour ce virus, le sous-complexe hélicase-primase (UL5-UL52) est

suffisant pour afficher les activités hélicase et primase alors que séparément les protéines ne présentent aucune activité [34]. La présence de la sous-unité accessoire (UL8) augmente l'efficacité des activités hélicase et primase du complexe binaire et permet sa localisation au niveau du noyau cellulaire. En présence d'UL42 (sous-unité catalytique de la polymérase) et d'ICP8 (protéine d'adhésion à l'ADN simple brin), le complexe hélicase-primase de HSV produit des amorces de 8 à 9 nucléotides en moyenne. Ce complexe, crucial pour la réplication virale, a été ciblé par des nouvelles molécules antiherpétiques en cours de développement [35, 36].

#### L'hélicase du virus Epstein-Barr (BBLF4)

L'activité hélicase a été attribuée au produit du gène *BBLF4* par homologie de séquence avec la protéine UL5 de HSV (34 % d'identité). Elle possède six motifs conservés [37] dont les deux premiers correspondant à des motifs Walker A et B permettent clairement son identification au sein de la famille des hélicases. Les autres motifs ont suggéré un positionnement au sein des hélicases de la super-famille 1 (SF1) mais la position exacte est restée longtemps obscure. La SF1 contient surtout des hélicases qui se déplacent dans le sens 3' vers 5' sur un ADN simple brin (directionnalité 3'-5') alors que les hélicases des herpesvirus ont une directionnalité 5'-3' comme cela a été montré pour HSV

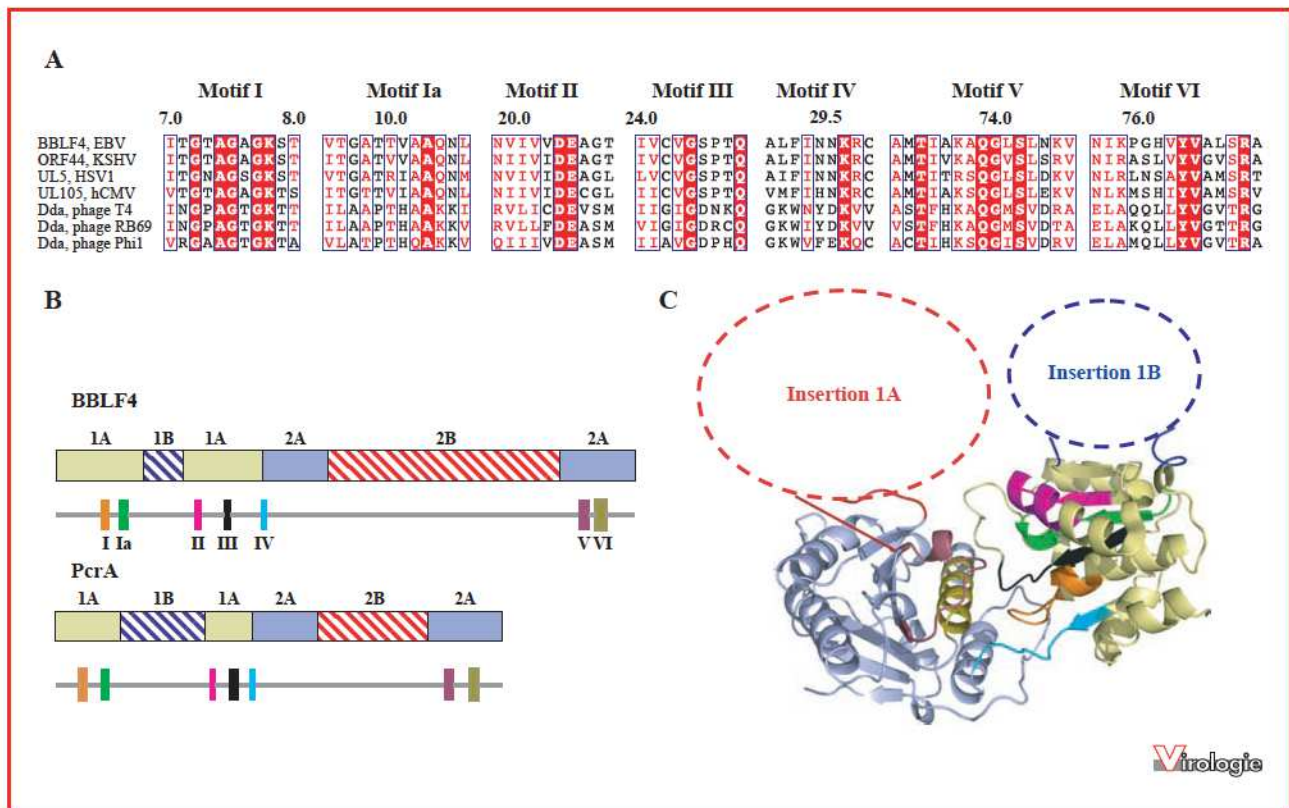


[38]. Au sein de la SF1, il existe un groupe d'hélicases avec une directionnalité 5'-3', les hélicases Dda des bactériophages [39, 40]. Effectivement, les motifs Walker A et B ainsi que le motif III de BBLF4 ressemblent plus au groupe Dda qu'aux autres groupes d'hélicases. En utilisant le logiciel récent CLUSTAL-omega [41], l'alignement des séquences des hélicases d'herpèsvirus avec des hélicases Dda permet de confirmer cette parenté au niveau des motifs V et VI au niveau du C-terminus avec certitude, et de manière moins certaine au niveau des motifs IA et IV (figure 4A). Ainsi les hélicases répliquatives des herpèsvirus se trouvent dans le même groupe que les hélicases Dda. Les hélicases Dda sont des hélicases monomériques qui, chez les bactériophages à queue (ordre des caudovirales, avec les exemples des bactériophages T4, RB69 et phi), permettent l'initiation de la réplication en séparant les deux brins de l'ADN afin d'ouvrir la fourche de réplication. Il n'y a pas de structure tridimensionnelle d'hélicase Dda connue, mais la répartition des motifs conservés sur la séquence de BBLF4 suggère une architecture similaire aux autres hélicases de la SF1 comme RecD ou PcrA [42] (figure 4B et C) [40]. Elles comporteraient ainsi deux domaines similaires

à RecA liant l'ADN monocaténaire, avec le site de liaison d'ATP situé entre les deux. Un domaine spécifique est inséré dans chaque domaine de RecA (figure 4C). En tout état de cause, cette analyse phylogénique permet d'affirmer que ce n'est pas l'hélicase répliquative hexamérique de la SF4 des bactériophages [43] qui est similaire à celle des herpèsvirus mais Dda qui a donc changé de fonction durant l'évolution.

#### La primase du virus Epstein-Barr (BSLF1)

Le gène *BSLF1* code pour une protéine qui partage 23 % d'identité de séquence avec la protéine UL52 d'HSV, qui a été caractérisée. Les primases des herpèsvirus possèdent deux motifs conservés, le motif DxD (acide aspartique-x-acide aspartique) et le doigt de zinc [33, 44] composé de trois cystéines et une histidine. Le motif DxD a un rôle majeur pour l'activité primase d'UL52 [45] tandis que le motif doigt de zinc, très conservé, a un effet aussi bien sur l'activité primase qu'hélicase : l'UL52 mutée au niveau du doigt de zinc est incapable de produire des amorces ARN et le complexe perd son activité hélicase.



**Figure 4.** L'hélicase BBLF4. **A)** Alignement des motifs des hélicases des herpèsvirus avec ceux des hélicases Dda des bactériophages. **B)** Emplacement des motifs au niveau de la séquence de BBLF4 comparée à celle de PcrA. **C)** Modèle structural de BBLF4 basé sur PcrA (pdb 1PJR). Motifs et domaines sont colorés comme dans (B).



### La sous-unité accessoire du virus Epstein-Barr (BBLF2/3)

Le gène codant pour la sous-unité accessoire d'EBV est issu d'un ARN épissé comportant les cadres de lecture BBLF2 et BBLF3, identifié initialement de façon erronée comme cadres de lecture indépendants. Contrairement à l'hélicase et la primase, il y a très peu d'homologie de séquence entre BBLF2/3 d'EBV et les protéines correspondantes d'autres herpesvirus (identité de 24 % pour KSHV sur une partie de la séquence). Par rapport à la protéine correspondante UL8 d'HSV, cette conservation est limitée à une séquence similaire de 55 acides aminés sur les 710 résidus de BBLF2/3 [5].

## Les enzymes accessoires

### L'exonucléase ou nucléase alcaline (BGLF5)

Ne faisant pas directement partie du complexe de réplication, la nucléase alcaline est néanmoins essentielle à la réplication. Son absence entraîne une baisse de l'infectivité d'un facteur 100 à 1 000 [46]. Chez HSV, elle intervient dans la réparation d'intermédiaires de réplication branchés [12]. *In vitro*, elle présente principalement une activité exonucléase 5-3' sur des substrats d'ADN simple ou double brin et une faible activité endonucléase.

Une action sur la synthèse protéique de la cellule hôte a été identifiée plus récemment pour cette enzyme chez les  $\gamma$ -herpesvirus humains [47] où une fraction de l'enzyme est localisée au niveau du cytosol [47] alors que l'enzyme agit sur l'ADN double brin au niveau du noyau. *In vitro*, BGLF5 montre une activité exonucléase dirigée contre les ARN monocaténares et les ARN messagers [48] seulement en présence de  $Mn^{2+}$ . Cette activité est commune à l'exonucléase de KSHV, où la dégradation d'ARN est observée même en présence de  $Mg^{2+}$  [49].

Sa structure a été résolue par cristallographie dans notre laboratoire [48] (figure 5A). Elle révèle l'appartenance à la famille des nucléases DExK, dans laquelle on trouve aussi des enzymes de restriction (figure 5C). En plus du domaine catalytique, que l'on retrouve dans les endonucléases de restriction, BGLF5 possède un site de fixation de nucléotides 5' similaire à l'exonucléase du bactériophage  $\lambda$  (figure 5B) [50], qu'on retrouve aussi au niveau de la séquence de l'exonucléase des baculovirus. Contrairement à l'exonucléase trimérique du bactériophage  $\lambda$ , qui forme une structure toroïdale, l'enzyme d'EBV semble agir comme monomère, comme l'exonucléase SOX de KSHV dont la structure a également été déterminée [49]. Ces exonucléases ne semblent pas avoir d'équivalent eucaryote mais sont présentes seulement au sein des virus et des bactériophages à ADN double brin ce qui en fait

des cibles thérapeutiques potentielles. La structure de l'exonucléase de KSHV en complexe avec un ADN double brin montre l'interaction avec son substrat [51] (figure 4D) sous forme de monomère et confirme le modèle proposé pour EBV [48]. La structure des complexes entre l'ADN et l'exonucléase  $\lambda$  précise le mécanisme catalytique et confirme le site de liaison du nucléotide positionné en 5' du site de coupure [52]. Ces travaux montrent également que le « pont » qui traverse le site actif (figure 5A) porte des résidus nécessaires à la séparation des deux brins de l'ADN invalidant les modèles où ce « pont » jouerait un rôle au niveau de la processivité en piégeant l'enzyme sur le brin non-hydrolysé [48]. En revanche, les bases structurales de l'interaction avec l'ARNm ne sont toujours pas établies.

### L'uracile-ADN-glycosylase (BKRF3)

La structure de l'enzyme d'EBV a été déterminée en complexe avec Uracil-ADN-Glycosylase Inhibitor (UGI) du bactériophage PBS1 dans notre laboratoire [53] (figure 5E). Les UGI sont des enzymes de réparation de l'ADN ubiquitaires. Elles enlèvent les bases d'uracile présentes dans l'ADN, provenant soit d'une déamidation spontanée de la cytosine, soit de l'incorporation de dUTP dans l'ADN nouvellement synthétisé. Les uraciles-ADN-glycosylases (UNG) parcourent l'ADN double brin en perturbant sa structure ce qui expulse momentanément une par une les bases de la double hélice permettant l'hydrolyse de la liaison N-glycosylique de la base s'il s'agit d'une base uracile (figure 5F). Après hydrolyse, l'UNG semble rester bloquée au niveau du site abasique, probablement pour favoriser le recrutement des enzymes de réparation. Plusieurs motifs de séquence sont impliqués dans cette fonction (figure 5E) : une boucle élargissant le sillon majeur (bleu) qui porte aussi un résidu leucine « bloque-porte » qui prend la place de la base expulsée (figure 5F), ainsi que des résidus qui compriment le squelette phosphatique. L'uracile interagit d'une façon spécifique avec le site catalytique (figure 5F) où sa liaison N-glycosylique est hydrolysée. Les UNG des herpesvirus sont aussi similaires entre elles (entre 40 et 50 % d'identité) que par rapport aux UNG d'autres organismes. La seule différence de l'UNG d'EBV est la boucle élargissant le sillon majeur qui est plus longue pour EBV. Cette conservation du site actif rend toute conception de médicament difficile. La situation de l'UNG est singulière par rapport à la plupart des protéines des herpesvirus qui sont toutes le résultat d'une évolution très divergente et qui n'ont que peu de similitude au niveau de la séquence avec des protéines d'autres royaumes du vivant. Cette conservation se traduit par l'inhibition commune par l'UGI du bactériophage PBS1 (figure 5E) [53, 54]. Les seules UNG non-inhibées sont celles des  $\beta$ -herpesvirus



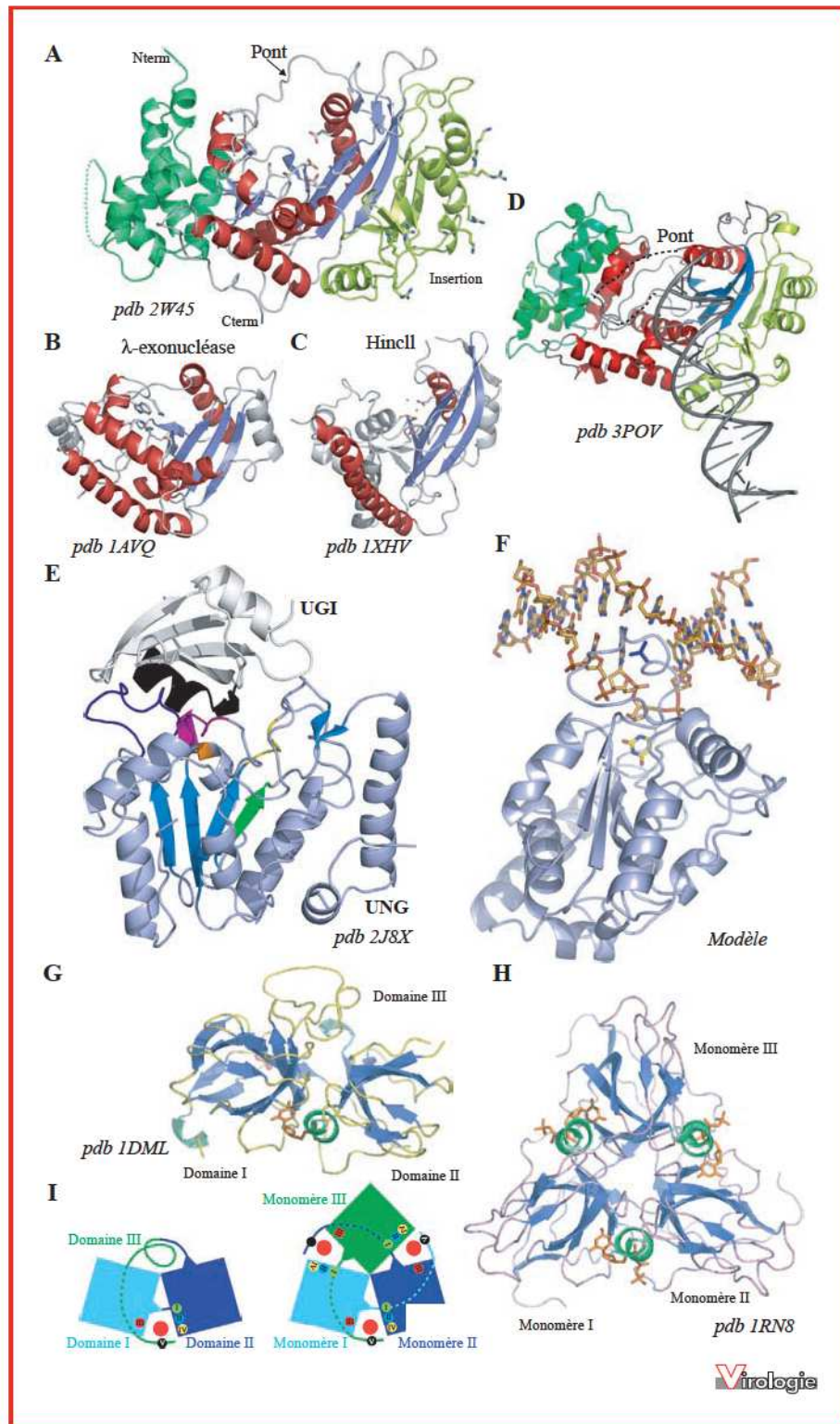


Figure 5.



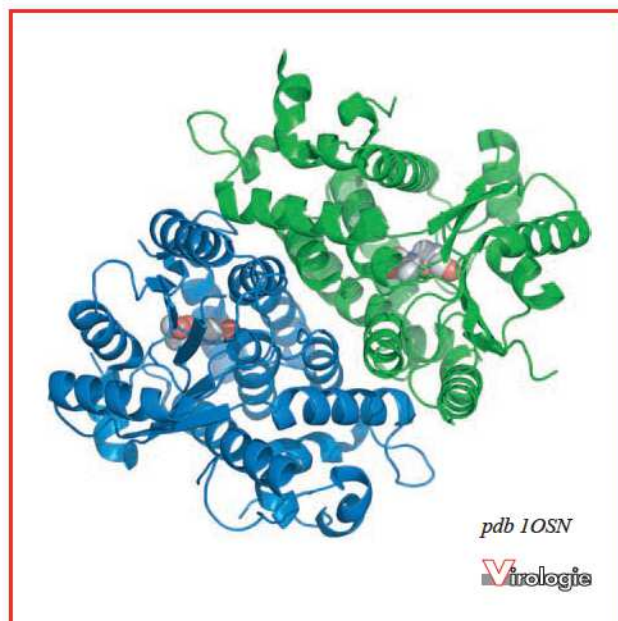
et des poxvirus où le résidu de leucine « bloque-porte » est remplacé par une arginine ou une lysine abolissant la reconnaissance par UGI.

Au niveau de HSV, une interaction physique directe entre UNG (UL2) et la polymérase (UL30) a été décrite [55], ce qui suggère une action coordonnée des deux enzymes. Par ailleurs, la polymérase UL30 intervient à plusieurs niveaux pour la réparation des sites abasiques [18] : l'activité AP endonucléase d'UL30 coupe l'ADN côté 5' du site abasique généré par l'UNG. Ensuite l'activité réparatrice de la polymérase incorpore une thymidine côté 3' chevauchant le site abasique qui est ensuite retiré grâce à l'activité AP-lyase de la même polymérase. Finalement, la ligase humaine III $\alpha$ -XRCC1 referme le squelette phosphatique.

#### dUTPase (BLLF3)

Deux enzymes virales servent à combattre la mésincorporation d'uridine (qui peut s'apparier de la même façon que la thymidine) au niveau du génome viral lors de la répllication lytique : l'UNG qui enlève les bases de l'uracile intégrées au niveau de l'ADN et la dUTPase qui réduit le pool cellulaire de dUTP. Cette dernière hydrolyse le dUTP en dUMP et pyrophosphate. La dUTPase est absente chez les  $\beta$ -herpèsvirus [56].

Contrairement aux autres dUTPases trimériques apparentées, celle d'EBV est monomérique. La détermination de cette structure par notre laboratoire [57] (*figure 5G*) a montré que le site catalytique est extrêmement bien conservé par rapport à d'autres organismes mais que l'organisation en domaines et sous-unités est extrêmement originale : l'enzyme trimérique classique, comportant trois sites actifs (*figure 5H*), a été transformée en enzyme monomérique avec un seul site actif où une sous-unité a été dédoublée et la troisième est devenue à peine identifiable (*figure 5G et I*). Le site actif est composé d'un site de reconnaissance spécifique de l'uridine et d'une reconnaissance du desoxyribose. L'hydrolyse du dUTP est induite par le repliement du motif V de l'enzyme au niveau de l'extrémité C-terminale flexible [58].



**Figure 6.** Structure d'une enzyme du métabolisme des nucléotides, la thymidine kinase du virus VZV [60] homologue à celle du virus Epstein-Barr. Cette enzyme est dimérique comme indiqué par les couleurs différentes des deux sous-unités. Les molécules de thymidine liées aux sites actifs sont montrées en modèle boule.

#### Thymidine kinase (BXLF1)

Cette structure est toujours inconnue pour EBV. La thymidine kinase a pu être cristallisée chez HSV [59] et le virus de la varicelle (VZV) [60] (*figure 6*), où la séquence est nettement plus courte. Le rôle de l'extension N-terminale de BXLF1 reste inconnu. La structure des deux thymidine kinases des  $\alpha$ -herpèsvirus est relativement similaire, mais l'architecture du site actif de ces deux protéines n'est pas tout à fait la même ce qui explique des différences de comportement vis-à-vis des analogues nucléosidiques de type ganciclovir et aciclovir. Chez EBV, ces analogues nucléosidiques ne sont pas phosphorylés par la thymidine kinase [61] mais par la sérine/thréonine kinase BGLF4 [62].

**Figure 5.** Structure des protéines d'EBV accessoires à la répllication. **A)** Exonucléase BGLF5. Les domaines spécifiques à BGLF5 sont montrés en vert, des résidus du site actif (centre) et du signal de localisation nucléaire sont montrés en mode bâton. **B)** Comparaison avec l'exonucléase du bactériophage  $\lambda$ . **C)** Comparaison avec une endonucléase de restriction. Les éléments de structure secondaire conservés entre ces trois structures sont montrés en bleu et rouge. **D)** Structure du complexe de l'exonucléase de KSHV avec l'ADN double brin. On voit la grande similitude de ces deux exonucléases des  $\gamma$ -herpèsvirus. Couleurs comme dans (A). **E)** Uracile-ADN-glycosylase (bleu clair) en complexe avec la protéine UGI (gris et noir). Les motifs de l'UNG sont colorés (la boucle activant l'eau catalytique, jaune ; la boucle 4-proline, rose ; le brin reconnaissance d'uracile, vert ; la boucle GS, orange et la boucle leucine, bleu). **F)** Modèle de l'interaction de l'UNG d'EBV avec l'ADN basé sur la comparaison de l'enzyme de virus Epstein-Barr (bleu clair) avec la structure de l'UNG humain en complexe avec ADN (orange) [67]. Une molécule d'uracile (jaune) est modélisée au niveau du site actif. La leucine « bloque-porte » est colorée en bleu. **G)** La dUTPase monomérique d'EBV en complexe avec dUMP (orange) comparé à la **H)** dUTPase trimérique d'*E. coli*. avec un analogue de dUTP (orange). **I)** Arrangement schématique des sous-unités pour la dUTPase monomérique d'EBV et une dUTPase trimérique. Les points rouges entourés des numéros de motifs marquent les sites actifs fonctionnels. Illustration reproduite avec autorisation de [57].



### Ribonucléotide réductase (BORF2 et BaRF1)

Elle convertit les ribonucléotides en deoxyribonucléotides utilisés pour la synthèse d'ADN. Elle est nécessaire lors de la réplication lytique dans des cellules au repos ayant un réservoir de deoxyribonucléotides faible comme les cellules épithéliales différenciées dans le cas d'EBV. Les  $\alpha$ -herpèsvirus et les  $\gamma$ -herpèsvirus codent pour deux sous-unités d'une ribonucléotide réductase de la classe Ia, R1 et R2 (cadres de lecture BORF2 et BaRF1) organisés en hétérotétramère  $\alpha_2\beta_2$  tandis que les  $\beta$ -herpèsvirus n'encodent qu'une sous-unité R1 inactive. La sous-unité R2 génère un radical de tyrosine à partir de l'oxygène moléculaire à l'aide de son site de liaison de fer. Pour la catalyse, ce radical est transféré sur la sous-unité R1 qui réduit le ribonucléotide. L'analyse de la séquence montre la conservation des résidus essentiels et suggère que l'enzyme d'EBV soit fonctionnelle. Très récemment, la cristallisation de la petite sous-unité R2 de KSHV a été décrite [63], et au moment de la correction des épreuves, sa structure a été rendue disponible par la pdb.

Les sous-unités portent souvent des fonctions anti-apoptiques ou d'inhibition de l'immunité innée [64].

### Conclusion

À ce jour, les structures tridimensionnelles de 12 protéines du cycle lytique de l'EBV ont été résolues (dont six par notre laboratoire) ; 5 concernent les enzymes du cycle répliatif. Ces structures apportent des éléments intéressants concernant l'évolution d'EBV et la possibilité d'inhiber la réplication virale.

La détermination de ces structures a fait apparaître des similitudes avec des protéines d'espèces différentes que les alignements de séquences ne pouvaient déceler. C'est le cas pour le facteur de processivité BMRF1 par rapport aux protéines de bactériophages et pour l'exonucléase BGLF5 par rapport à celle du bactériophage  $\lambda$ . Les structures des protéines de réplication suggèrent que les herpèsvirus proviennent des caudovirales comme par exemple les bactériophages T4, T7 et  $\lambda$ . On retrouve la même relation pour l'hélicase basée pour l'instant uniquement sur l'alignement de séquences, à défaut de données structurales.

De manière très intrigante, les protéines de réplication d'EBV et des autres herpèsvirus ont souvent une taille relativement grande par rapport aux protéines correspondantes des autres virus ou des bactériophages.

La dUTPase monomérique avec ses deux domaines est encodée par une séquence d'ADN deux fois plus longue comparée aux dUTPases trimériques d'autres organismes, sans que la fonction gagnée soit connue. L'exonucléase

des herpèsvirus est également deux fois plus grande que l'exonucléase trimérique du bactériophage  $\lambda$  sans qu'une fonction connue ne soit portée par les domaines supplémentaires, mis à part le signal de localisation nucléaire. À noter aussi, la très grande taille de la protéine liant l'ADN simple brin (1 128 résidus), comparé par exemple aux  $\sim 200$  résidus des protéines bactériennes et les 269 résidus du virus de la vaccine. Les facteurs de processivité ont également des extensions C-terminales d'environ 90 résidus par rapport aux protéines eucaryotes PCNA ou du phage RB69. Un dernier exemple est la thymidine kinase BSLF1 qui avec 607 acides aminés dépasse largement la taille minimale de 350 acides aminés pour d'autres herpèsvirus et de 230 résidus pour d'autres organismes. Les raisons de cette augmentation de complexité ne sont pas connues.

La connaissance structurale peut certainement contribuer au développement de nouveaux médicaments contre les herpèsvirus. Certaines cibles comme l'UNG et la dUTPase sont probablement peu intéressantes avec des sites actifs trop hydrophiles ou trop similaires à ceux des enzymes eucaryotes. D'autres protéines dont la structure n'est pas encore résolue comme l'hélicase-primase d'EBV, à l'instar de l'hélicase-primase de HSV dont on connaît déjà des inhibiteurs, pourraient être plus intéressantes [36].

En 2012, tous les antiviraux commercialisés contre les infections au virus *herpes simplex*, varicelle-zona ou cytomégalovirus (aciclovir et son dérivé valaciclovir, ou famciclovir, cidofovir, ganciclovir et foscarnet) ciblent efficacement l'ADN polymérase des herpèsvirus. Ces antiviraux inhibent également l'ADN polymérase d'EBV et sont efficace sur la réplication du virus *in vitro* et *in vivo* (notamment dans l'oropharynx) [65]. Cependant, leur efficacité clinique dans les maladies associées à EBV, y compris la MNI, est faible ou nulle [66]. L'hypothèse principale pour cette inefficacité est que la réplication virale d'EBV n'est pas responsable directement ni des symptômes de la MNI, liés à la réaction immunitaire, ni de l'émergence des cancers-EBV induits qui seraient favorisés uniquement par l'expression des gènes viraux de latence. Cependant, il est possible que l'inhibition des protéines très précoces ou précoces du cycle lytique (avant la synthèse de l'ADN viral) soit plus importante que le blocage de la réplication virale car ces protéines semblent impliquées dans les lymphomes chez l'immunodéprimé associé à l'EBV et les carcinomes indifférenciés du rhinopharynx. La détermination de la structure de ces protéines très précoces ou précoces du cycle lytique et leur comparaison avec leur équivalent cellulaire peut donc apporter des informations précieuses pour définir les meilleures cibles antivirales potentielles.

**Conflits d'intérêts :** aucun.

## Références

1. Cohen JI, Fauci AS, Varmus H, Nabel GJ. Epstein-Barr virus: an important vaccine target for cancer prevention. *Sci Transl Med* 2011 ; 3 : 107fs7.
2. Odumade OA, Hogquist KA, Balfour Jr. HH. Progress and problems in understanding and managing primary Epstein-Barr virus infections. *Clin Microbiol Rev* 2011 ; 24 : 193-209.
3. Thompson MP, Kurzrock R. Epstein-Barr virus and cancer. *Clin Cancer Res* 2004 ; 10 : 803-21.
4. Ma SD, Hegde S, Young KH, *et al.* A new model of Epstein-Barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas. *J Virol* 2011 ; 85 : 165-77.
5. Fixman ED, Hayward GS, Hayward SD. Trans-acting requirements for replication of Epstein-Barr virus ori-Lyt. *J Virol* 1992 ; 66 : 5030-9.
6. Müller G, El-Guindy A, Countryman J, Ye J, Gradoville L. Lytic cycle switches of oncogenic human gammaherpesviruses. *Adv Cancer Res* 2007 ; 97 : 81-109.
7. Gershburg E, Pagano JS. Epstein-Barr virus infections: prospects for treatment. *J Antimicrob Chemother* 2005 ; 56 : 277-81.
8. Yates JL, Guan N. Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. *J Virol* 1991 ; 65 : 483-8.
9. Hammerschmidt W, Sugden B. Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus. *Cell* 1988 ; 55 : 427-33.
10. Fixman ED, Hayward GS, Hayward SD. Replication of Epstein-Barr virus oriLyt: lack of a dedicated virally encoded origin-binding protein and dependence on Zta in cotransfection assays. *J Virol* 1995 ; 69 : 2998-3006.
11. Rennekamp AJ, Lieberman PM. Initiation of Epstein-Barr virus lytic replication requires transcription and the formation of a stable RNA-DNA hybrid molecule at OriLyt. *J Virol* 2011 ; 85 : 2837-50.
12. Goldstein JN, Weller SK. *In vitro* processing of herpes simplex virus type 1 DNA replication intermediates by the viral alkaline nuclease, UL12. *J Virol* 1998 ; 72 : 8772-81.
13. Liu S, Knafels JD, Chang JS, *et al.* Crystal structure of the herpes simplex virus 1 DNA polymerase. *J Biol Chem* 2006 ; 281 : 18193-200.
14. Shamoo Y, Steitz TA. Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell* 1999 ; 99 : 155-66.
15. Wang M, Xia S, Blaha G, Steitz TA, Konigsberg WH, Wang J. Insights into base selectivity from the 1.8 Å resolution structure of an RB69 DNA polymerase ternary complex. *Biochemistry* 2010 ; 50 : 581-90.
16. Zuccola HJ, Filman DJ, Coen DM, Hogle JM. The crystal structure of an unusual processivity factor, herpes simplex virus UL42, bound to the C terminus of its cognate polymerase. *Mol Cell* 2000 ; 5 : 267-78.
17. Tsurumi T, Daikoku T, Kurachi R, Nishiyama Y. Functional interaction between Epstein-Barr virus DNA polymerase catalytic subunit and its accessory subunit *in vitro*. *J Virol* 1993 ; 67 : 7648-53.
18. Bogani F, Chua CN, Boehmer PE. Reconstitution of uracil DNA glycosylase-initiated base excision repair in herpes simplex virus-1. *J Biol Chem* 2009 ; 284 : 16784-90.
19. Chou S, Lurain NS, Thompson KD, Miner RC, Drew WL. Viral DNA polymerase mutations associated with drug resistance in human cytomegalovirus. *J Infect Dis* 2003 ; 188 : 32-9.
20. Hartline CB, Harden EA, Williams-Aziz SL, Kushner NL, Bri-deau RJ, Kern ER. Inhibition of herpesvirus replication by a series of 4-oxo-dihydroquinolines with viral polymerase activity. *Antiviral Res* 2005 ; 65 : 97-105.
21. Daikoku T, Kudoh A, Fujita M, *et al.* Architecture of replication compartments formed during Epstein-Barr virus lytic replication. *J Virol* 2005 ; 79 : 3409-18.
22. Zhang Q, Hong Y, Dorsky D, *et al.* Functional and physical interactions between the Epstein-Barr virus (EBV) proteins BZLF1 and BMRF1: Effects on EBV transcription and lytic replication. *J Virol* 1996 ; 70 : 5131-42.
23. Appleton BA, Loregian A, Filman DJ, Coen DM, Hogle JM. The cytomegalovirus DNA polymerase subunit UL44 forms a C clamp-shaped dimer. *Mol Cell* 2004 ; 15 : 233-44.
24. Baltz JL, Filman DJ, Ciustea M, *et al.* The crystal structure of PF-8, the DNA polymerase accessory subunit from Kaposi's sarcoma-associated herpesvirus. *J Virol* 2009 ; 83 : 12215-28.
25. Murayama K, Nakayama S, Kato-Murayama M, *et al.* Crystal structure of Epstein-Barr virus DNA polymerase processivity factor BMRF1. *J Biol Chem* 2009 ; 284 : 35896-905.
26. Randell JC, Coen DM. Linear diffusion on DNA despite high-affinity binding by a DNA polymerase processivity factor. *Mol Cell* 2001 ; 8 : 911-20.
27. Randell JC, Coen DM. The herpes simplex virus processivity factor, UL42, binds DNA as a monomer. *J Mol Biol* 2004 ; 335 : 409-13.
28. Gulbis JM, Kelman Z, Hurwitz J, O'Donnell M, Kuriyan J. Structure of the C-terminal region of p21 (WAF1/CIP1) complexed with human PCNA. *Cell* 1996 ; 87 : 297-306.
29. Tsurumi T, Kobayashi A, Tamai K, *et al.* Epstein-Barr virus single-stranded DNA-binding protein: purification, characterization, and action on DNA synthesis by the viral DNA polymerase. *Virology* 1996 ; 222 : 352-64.
30. Mumtsidu E, Makhov AM, Konarev PV, Svergun DI, Griffith JD, Tucker PA. Structural features of the single-stranded DNA-binding protein of Epstein-Barr virus. *J Struct Biol* 2008 ; 161 : 172-87.
31. Mapelli M, Panjekar S, Tucker PA. The crystal structure of the herpes simplex virus 1 ssDNA-binding protein suggests the structural basis for flexible, cooperative single-stranded DNA binding. *J Biol Chem* 2005 ; 280 : 2990-7.
32. Yang C, Curth U, Urbanke C, Kang C. Crystal structure of human mitochondrial single-stranded DNA binding protein at 2.4 Å resolution. *Nat Struct Biol* 1997 ; 4 : 153-7.
33. Yokoyama N, Fujii K, Hirata M, *et al.* Assembly of the Epstein-Barr virus BBLF4, BSLF1 and BBLF2/3 proteins and their interactive properties. *J Gen Virol* 1999 ; 80 : 2879-87.
34. Dodson MS, Lehman IR. Association of DNA helicase and primase activities with a subassembly of the herpes simplex virus 1 helicase-primase composed of the UL5 and UL52 gene products. *Proc Natl Acad Sci U S A* 1991 ; 88 : 1105-9.
35. Andrei G, Snoeck R. Emerging drugs for varicella-zoster virus infections. *Expert Opin Emerg Drugs* 2011 ; 16 : 507-35.
36. Field HJ, Biswas S. Antiviral drug resistance and helicase-primase inhibitors of herpes simplex virus. *Drug Resist Updat* 2011 ; 14 : 45-51.
37. Zhu LA, Weller SK. The six conserved helicase motifs of the UL5 gene product, a component of the herpes simplex virus type 1 helicase-primase, are essential for its function. *J Virol* 1992 ; 66 : 469-79.
38. Chen Y, Bai P, Mackay S, *et al.* Herpes simplex virus type 1 helicase-primase: DNA binding and consequent protein oligomerization and primase activation. *J Virol* 2011 ; 85 : 968-78.
39. Rajagopal V, Patel SS. Viral Helicases. In : Cameron CE, Götte M, Raney KD, eds. *Viral Genome Replication*. New York : Springer, 2009, p. 429-66.
40. Singleton MR, Dillingham MS, Wigley DB. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem* 2007 ; 76 : 23-50.
41. Sievers F, Wilm A, Dineen D, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011 ; 7 : 539.
42. Subramanya HS, Bird LE, Brannigan JA, Wigley DB. Crystal structure of a DEXX box DNA helicase. *Nature* 1996 ; 384 : 379-83.
43. Mueser TC, Hinerman JM, Devos JM, Boyer RA, Williams KJ. Structural analysis of bacteriophage T4 DNA replication: a review in the

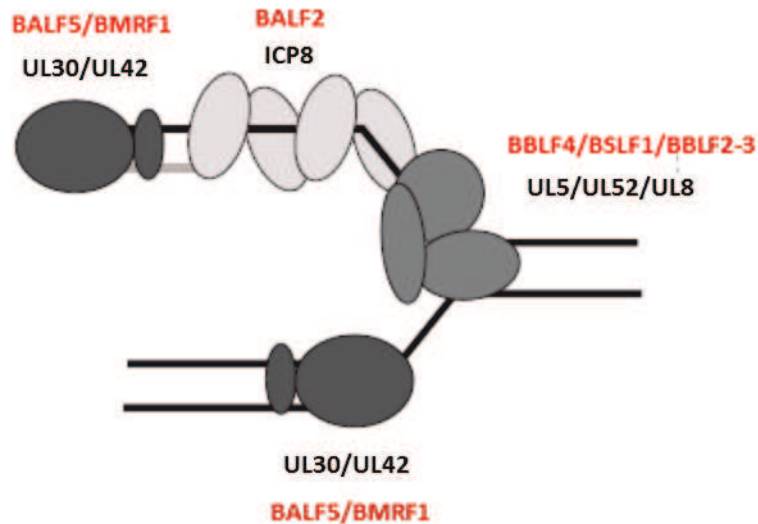


- Virology Journal series on bacteriophage T4 and its relatives. *Virology* 2010 ; 7 : 359.
44. Chen Y, Carrington-Lawrence SD, Bai P, Weller SK. Mutations in the putative zinc-binding motif of UL52 demonstrate a complex interdependence between the UL5 and UL52 subunits of the human herpes simplex virus type 1 helicase/primase complex. *J Virol* 2005 ; 79 : 9088-96.
45. Dracheva S, Koonin EV, Crute JJ. Identification of the primase active site of the herpes simplex virus type 1 helicase-primase. *J Biol Chem* 1995 ; 270 : 14148-53.
46. Goldstein JN, Weller SK. The Exonuclease activity of HSV-1 UL12 is required for *in vivo* function. *Virology* 1998 ; 244 : 442-57.
47. Rowe M, Glaunsinger B, van Leeuwen D, *et al.* Host shutoff during productive Epstein-Barr virus infection is mediated by BGLF5 and may contribute to immune evasion. *Proc Natl Acad Sci U S A* 2007 ; 104 : 3366-71.
48. Buisson M, Geoui T, Flot D, *et al.* A Bridge Crosses the Active-Site Canyon of the Epstein-Barr Virus Nuclease with DNase and RNase Activities. *J Mol Biol* 2009 ; 391 : 717-28.
49. Dahlroth SL, Gurmu D, Haas J, Erlandsen H, Nordlund P. Crystal structure of the shutoff and Exonuclease protein from the oncogenic Kaposi's sarcoma-associated herpesvirus. *FEBS J* 2009 ; 276 : 6636-45.
50. Kovall R, Matthews BW. Toroidal structure of lambda-Exonuclease. *Science* 1997 ; 277 : 1824-7.
51. Bagneris C, Briggs LC, Savva R, Ebrahimi B, Barrett TE. Crystal structure of a KSHV-SOX-DNA complex: insights into the molecular mechanisms underlying DNase activity and host shutoff. *Nucleic Acids Res* 2011 ; 39 : 5744-56.
52. Zhang J, McCabe KA, Bell CE. Crystal structures of {lambda} Exonuclease in complex with DNA suggest an electrostatic ratchet mechanism for processivity. *Proc Natl Acad Sci U S A* 2011 ; 108 : 11872-7.
53. Geoui T, Buisson M, Tarbouriech N, Burmeister WP. New insights on the role of the gamma-herpesvirus uracil-DNA glycosylase leucine loop revealed by the structure of the Epstein-Barr virus enzyme in complex with an inhibitor protein. *J Mol Biol* 2007 ; 366 : 117-31.
54. Savva R, Pearl LH. Nucleotide mimicry in the crystal structure of the uracil-DNA glycosylase-uracil glycosylase inhibitor protein complex. *Nat Struct Biol* 1995 ; 2 : 752-7.
55. Bogani F, Corredeira I, Fernandez V, *et al.* Association between the herpes simplex virus-1 DNA polymerase and uracil DNA glycosylase. *J Biol Chem* 2011 ; 285 : 27664-72.
56. Davison AJ, Stow ND. New genes from old: redeployment of dUTPase by herpesviruses. *J Virol* 2005 ; 79 : 12880-92.
57. Tarbouriech N, Buisson M, Seigneurin JM, Cusack S, Burmeister WP. The monomeric dUTPase from Epstein-Barr virus mimics trimeric dUTPases. *Structure* 2005 ; 13 : 1299-310.
58. Freeman L, Buisson M, Tarbouriech N, Van der Heyden A, Labbe P, Burmeister WP. The flexible motif V of Epstein-Barr virus deoxyuridine 5'-triphosphate pyrophosphatase is essential for catalysis. *J Biol Chem* 2009 ; 284 : 25280-9.
59. Wild K, Grafmuller R, Wagner E, Schulz GE. Structure, catalysis and supramolecular assembly of adenylate kinase from maize. *Eur J Biochem* 1997 ; 250 : 326-31.
60. Bird LE, Ren J, Wright A, *et al.* Crystal structure of varicella zoster virus thymidine kinase. *J Biol Chem* 2003 ; 278 : 24680-7.
61. Gustafson EA, Chillemi AC, Sage DR, Fingerroth JD. The Epstein-Barr virus thymidine kinase does not phosphorylate ganciclovir or acyclovir and demonstrates a narrow substrate specificity compared to the herpes simplex virus type 1 thymidine kinase. *Antimicrob Agents Chemother* 1998 ; 42 : 2923-31.
62. Meng Q, Hagemeyer SR, Fingerroth JD, Gershburg E, Pagano JS, Kenney SC. The Epstein-Barr virus (EBV)-encoded protein kinase, EBV-PK, but not the thymidine kinase (EBV-TK), is required for ganciclovir and acyclovir inhibition of lytic viral production. *J Virol* 2010 ; 84 : 4534-42.
63. Gurmu D, Dahlroth SL, Haas J, Nordlund P, Erlandsen H. Expression, purification, crystallization and preliminary X-ray analysis of ORF60, the small subunit (R2) of ribonucleotide reductase from Kaposi's sarcoma-associated herpesvirus (KSHV). *Acta Crystallogr Sect F Struct Biol Cryst Commun* 2010 ; 66 : 734-7.
64. Lembo D, Brune W. Tinkering with a viral ribonucleotide reductase. *Trends Biochem Sci* 2009 ; 34 : 25-32.
65. Balfour Jr. HH, Hokanson KM, Schacherer RM, *et al.* A virologic pilot study of valacyclovir in infectious mononucleosis. *J Clin Virol* 2007 ; 39 : 16-21.
66. Pagano JS, Sixbey JW, Lin JC. Acyclovir and Epstein-Barr virus infection. *J Antimicrob Chemother* 1983 ; 12 Suppl B : 113-21.
67. Parikh SS, Mol CD, Slupphaug G, Bharati S, Krokan HE, Tainer JA. Base excision repair initiation revealed by crystal structures and binding kinetics of human uracil-DNA glycosylase with DNA. *EMBO J* 1998 ; 17 : 5214-26.

## **2. Le complexe de réplication des herpèsvirus est fortement conservé**

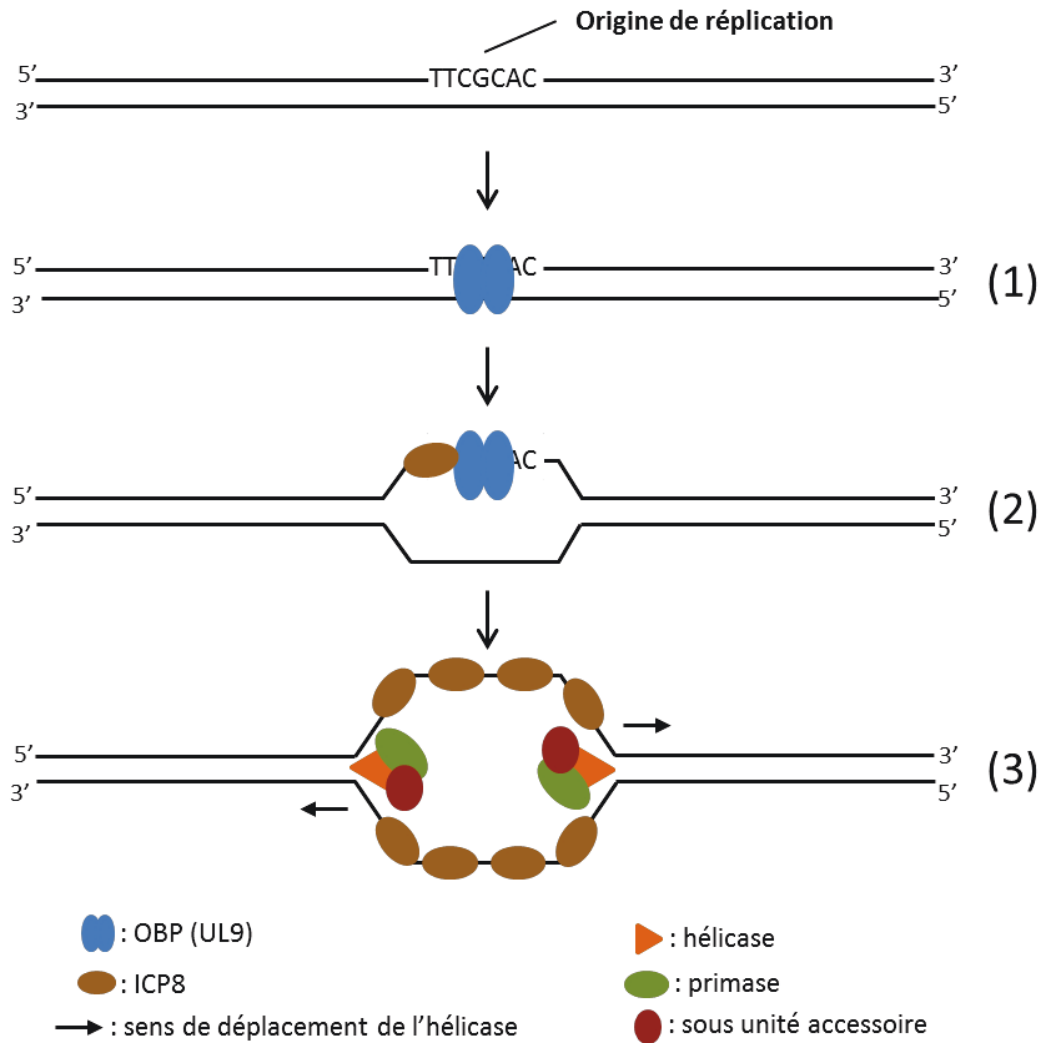
La revue précédente permet une description globale des protéines composant le complexe de réplication des herpèsvirus et des protéines accessoires associées à la réplication virale. La description structurale de ces protéines est une donnée très importante pour mieux comprendre les interactions régissant la formation du complexe réplicatif et permettant d'envisager la recherche d'antiviraux spécifiques. Cependant l'approche structurale est insuffisante, et doit être complétée par les études des activités et des interactions des protéines répliquatives permettant de décrire le fonctionnement de ce processus hautement contrôlé. Les connaissances sur les protéines répliquatives sont différentes en fonction des sous-familles  $\alpha$ ,  $\beta$  et  $\gamma$  des herpèsvirus, mais la forte homologie des protéines composant ces différentes sous familles permet d'étendre les résultats obtenus pour une famille à l'ensemble de la superfamille des herpèsvirus humains.

Tous les herpèsvirus expriment 6 gènes totalement indispensables à la réplication virale et codant pour : une polymérase et son facteur de processivité, une protéine d'adhésion à l'ADN simple brin (ADNsb) et trois protéines composant le complexe Hélicase-Primase (HSV-1 [79], EBV [80], hCMV [81] et KSHV [82]). Ces 6 gènes forment la base commune à la réplication des herpèsvirus, si bien qu'il est possible d'effectuer des expériences de complémentation de réplication. Elles ont été menées avec les protéines de réplication d'HSV-1 et l'origine de réplication oriLyt d'EBV. La substitution de la polymérase (UL30/UL42) ou du complexe hélicase-primase (UL5/UL52/UL8) ne permet pas l'observation d'une activité répliquative. En revanche l'utilisation de l'intégralité du complexe réplicatif d'HSV-1 en présence des facteurs de transcription d'EBV (ZEBRA, Rta et Mta), il y a mise en route du mécanisme de réplication. Il existe donc une équivalence fonctionnelle entre les complexes répliquatifs d'HSV et d'EBV, mais les protéines individuelles ne sont pas directement interchangeables [51].



**Figure 13 : Protéines composant le complexe de réplication d'HSV-1 et d'EBV.** En noir, les protéines appartenant à HSV-1, en rouge celles d'EBV : la polymérase est composée de deux sous-unités, une catalytique (UL30, BALF5) et une processive (UL42, BMRF1). La protéine d'adhésion à l'ADN simple brin (ICP8, BALF2). Les protéines du complexe hélicase-primase, l'hélicase (UL5, BBLF4), la primase (UL52, BSLF1) et la sous-unité accessoire (UL8, BBLF2/3)

Le processus réplcatif de chaque herpèsvirus possède ses propres spécificités. Le complexe réplcatif des herpèsvirus nécessite la présence d'un septième gène absolument indispensable à la réplication virale. La septième protéine d'HSV-1 ne trouve aucun homologue au sein des autres herpèsvirus humains. Le gène UL9 code pour la protéine d'adhésion à l'origine (OBP) qui permet l'initiation de la réplication de façon active (Figure 14). Elle forme un dimère qui se fixe sur la séquence consensus TTCGCAC [83] de l'origine de réplication oriS (1). L'interaction de cette protéine avec ICP8 (protéine de liaison à l'ADN simple brin, UL29) induit l'activité hélicase d'OBP conduisant à l'ouverture de l'origine de réplication en présence d'ATP (2). ICP8 permet le maintien de l'ADN sous forme simple brin [84] pour le recrutement des autres protéines de réplication (3). L'initiation utilise un mécanisme actif par une protéine d'adhésion à l'origine, d'autres virus présentent le même mécanisme mais il semble spécifique à la sous-famille des  $\alpha$ -herpèsvirus.



**Figure 14 : Représentation schématique de l'initiation de la réplication lytique d'HSV-1 par OBP.**

Pour les autres herpèsvirus humains, le septième gène code pour un facteur de transcription jouant un rôle primordial dans l'initiation de la réplication (ZEBRA pour EBV [51], K-Rta pour KSHV [85] et IE1 pour hCMV [86]). EBV possède deux origines oriLyt à partir desquelles la réplication lytique du génome viral peut être initiée. Une seule des origines lytiques est suffisante pour permettre la réplication du génome viral [87], d'ailleurs la souche de laboratoire principale B95-8 en possède qu'une seule [88]. Des études récentes menées par Rennekamp & al (2011) [89] ont révélées que la présence d'un transcrite d'ARN est nécessaire pour l'initiation de la réplication avec oriLyt. En fonction de l'oriLyt utilisé pour la réplication du génome viral, le transcrite d'ARN est exprimé soit par le gène BHLF1 soit par le gène BHRF1. Les travaux de Rennekamp & al (2011) [46] portent principalement sur le transcrite de BHLF1 car il est abondamment détecté pendant la réplication lytique.

La première étape du processus d'initiation est l'interaction de ZEBRA avec l'origine de réplication oriLyt (1). La présence de ZEBRA induit l'interaction du transcrite d'ARN, exprimé par BHLF1 ou BHRF1, avec oriLyt [89]. L'hybride ADN-ARN permet l'ouverture d'oriLyt et le recrutement de BALF2 stabilisant la forme ouverte de l'origine (2). La formation de ce complexe initial permet le recrutement du complexe H-P (BSLF1 + BBLF2/3 + BBLF4) à l'origine de réplication grâce à l'interaction entre ZEBRA et le complexe H-P [90]. La protéine d'adhésion à l'ADN simple brin interagit alors avec le pré-complexe ZEBRA-BBLF4-BSLF1-BBLF2/3 [90] (3). L'ensemble de ces protéines permet la formation d'une fourche de réplication et de synthétiser les premières amorces ARN. La dernière interaction finalisant le complexe réplcatif est réalisé par la fixation de la polymérase sur l'ensemble des protéines précédentes [91] (4). Le complexe réplcatif est alors complet et peut effectuer la réplication du génome viral (Figure 15). La présence d'un hybride ARN-ADN a aussi été décrit pour les herpesvirus KSHV [92] et hCMV [93].

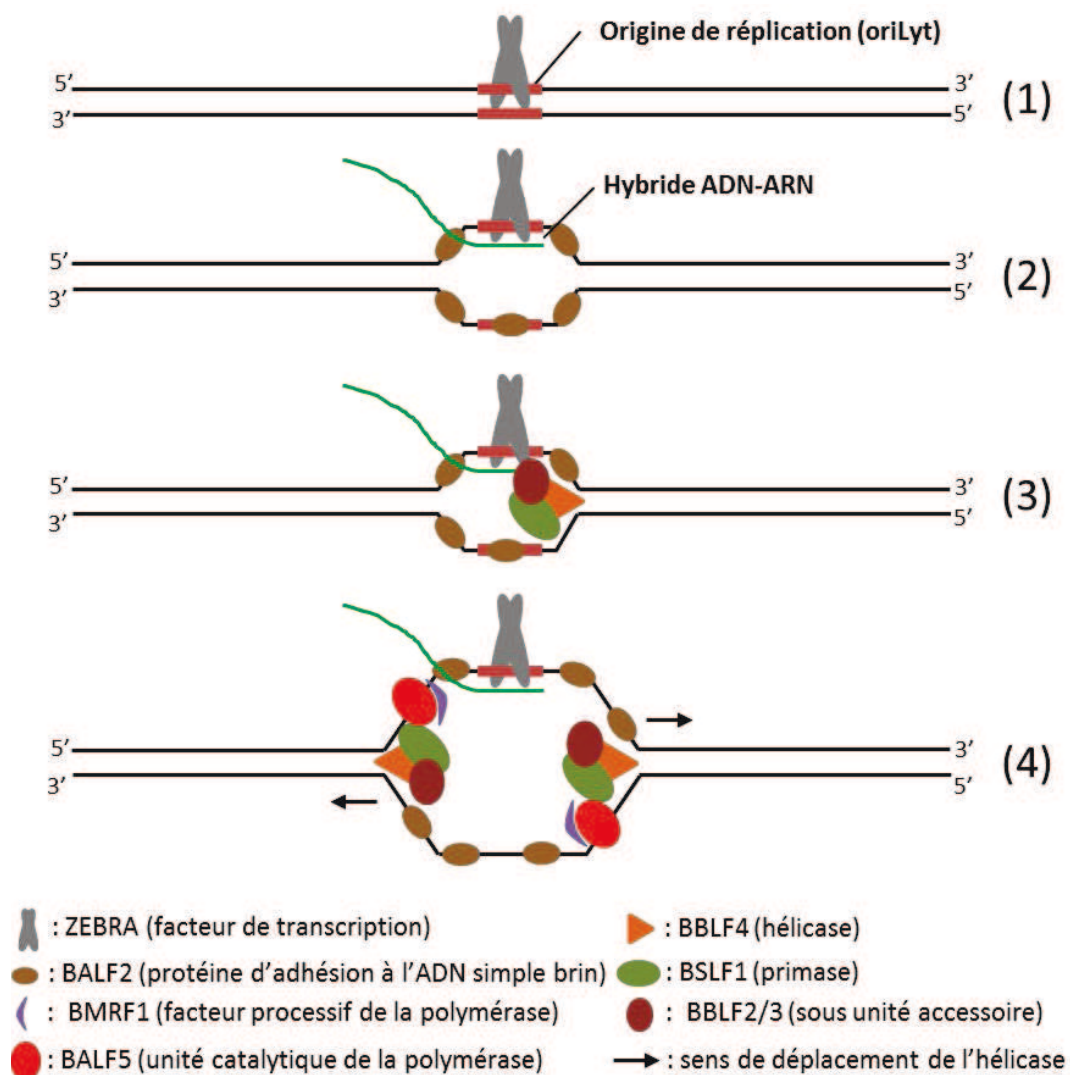


Figure 15 : Représentation schématique des étapes nécessaires à l'initiation de la réplication d'EBV.

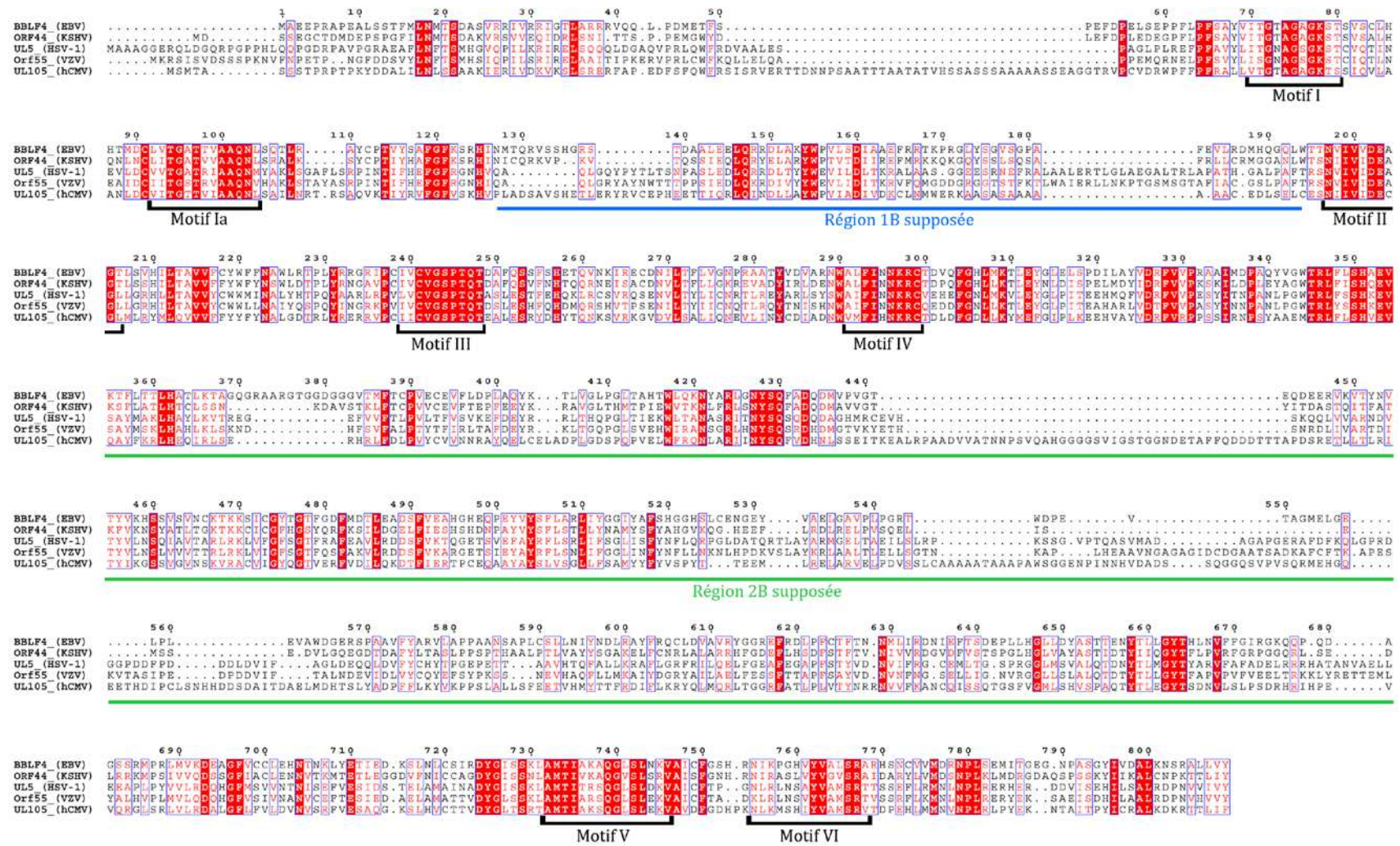
Ces différents exemples montrent que les informations obtenues sur le complexe de réplication d'un herpèsvirus peuvent être étendues aux autres membres de la famille mais qu'il faut rester prudent car le processus d'initiation varie.

### **3. Complément d'informations sur les protéines du complexe Hélicase-Primase d'EBV**

#### **3.1. L'hélicase (BBLF4)**

Le gène BBLF4 (2430 pb) code pour une protéine ayant une masse théorique de 89 kDa (809 aa). La fonction de cette protéine a été attribuée par homologie de séquence à partir de l'hélicase UL5 d'HSV-1. L'alignement de l'hélicase d'EBV avec celles des herpèsvirus KSHV, HSV-1, VZV et hCMV révèle que cette protéine est relativement bien conservée au sein des *Herpesviridae* humains. Les hélicases des herpèsvirus ont été classées dans la super famille 1B des hélicases (SF1B), on retrouve les 7 motifs caractéristiques de cette famille sur l'alignement. Les régions les moins conservées de l'alignement correspondent sûrement aux insertions 1B et 2B, qui sont connues pour varier fortement entre les protéines de la SF1B. Cet alignement permet aussi de montrer que les extrémités N-terminales des différentes hélicases sont variables, alors que les extrémités C-terminales sont relativement bien conservées.





**Figure 16 : Alignement des hélicases des herpesvirus humains.** Les 7 motifs hélicase sont facilement observés sur cet alignement. Les zones avec une faible identité de séquence correspondant aux insertions 1B et 2B. Ces deux régions sont peu conservées et sont probablement à l'origine des spécificités de chaque hélicase. Alignement réalisé avec les hélicases d'EBV, de KSHV, d'HSV-1, de VZV et de hCMV.

### **3.2. La primase (BSLF1)**

La primase est encodée par le gène BSLF1 (2625 pb), elle possède une masse théorique de 98 kDa (874 aa). Tout comme l'hélicase, la fonction primase est attribuée à BSLF1 par homologie avec la primase UL52 d'HSV-1. Les primases des *Herpesviridae* humains sont moins conservées que les hélicases. L'alignement complet entre les primases des herpèsvirus humains est possible, il existe des régions avec une conservation importante. Seules deux de ces régions sont décrites pour participer à l'activité primase : le motif DxD et le doigt de zinc situé en C-terminal (Figure 17).

### **3.3. La sous-unité accessoire (BBLF2/3)**

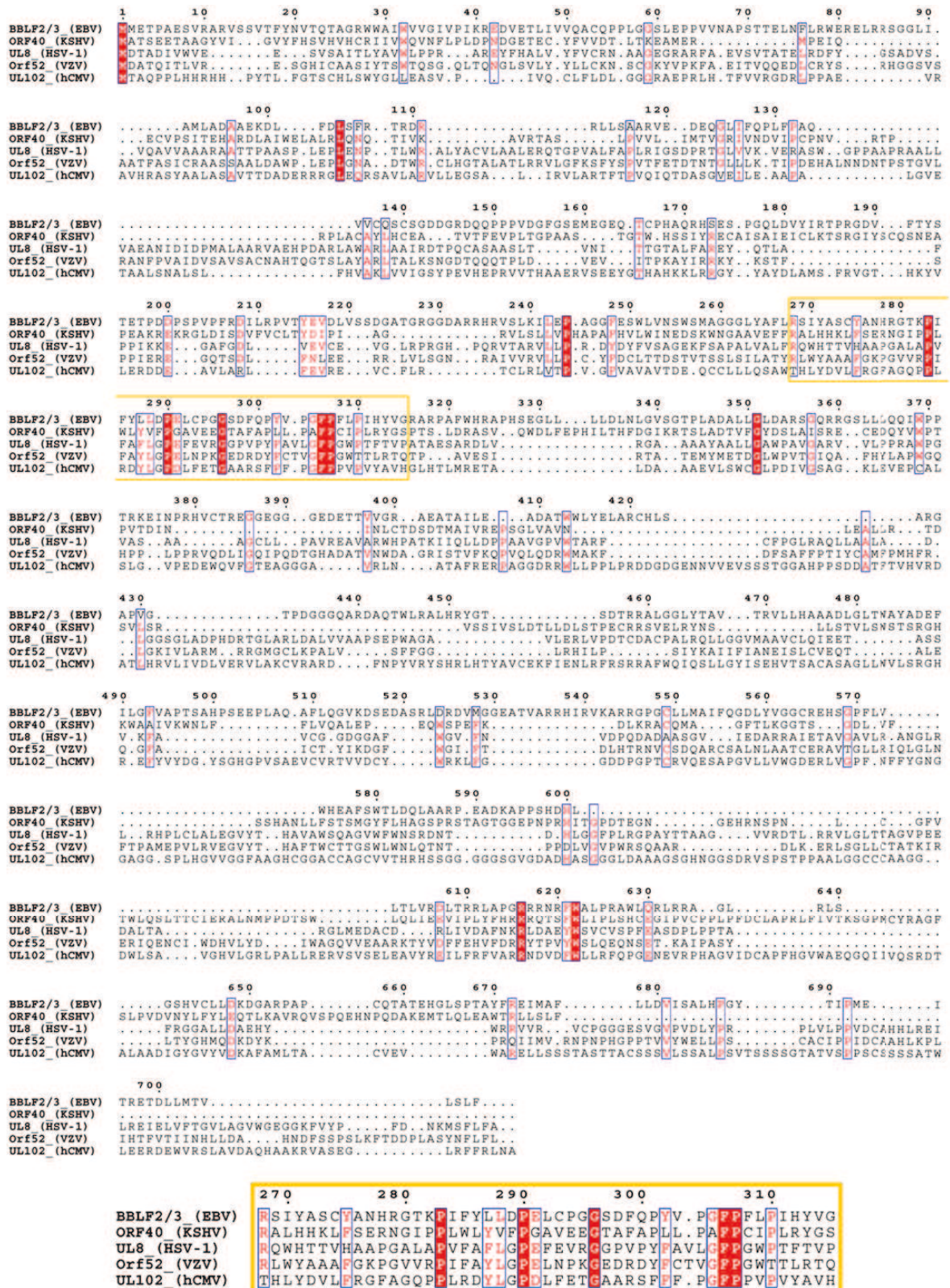
La sous-unité accessoire d'EBV est issue d'un épissage d'un ARN provenant des cadres de lectures BBLF2 et BBLF3. La protéine BBLF2/3 possède une taille théorique de 78 kDa (709 aa). L'alignement des sous-unités des *Herpesviridae* humains révèle une faible conservation de ces protéines. Les extrémités N et C-terminales des sous-unités accessoires sont que peu conservées. Cet alignement semble révéler une région conservée (268-315), mais aucune étude n'a été réalisée pour estimer son importance au sein du complexe H-P des herpèsvirus (Figure 18). Lorsque BBLF2/3 est alignée deux à deux avec ses homologues, une certaine conservation est observée, mais elle est beaucoup plus faible que pour la primase et l'hélicase. Les sous-unités accessoires des complexes H-P des herpèsvirus semblent avoir fortement divergées. Malgré les divergences cette protéine fait toujours partie du complexe H-P. Les singularités portées par cette protéine rendent son étude attractive car elle pourrait apporter des informations importantes sur son implication dans le complexe H-P et au niveau de la réplication du génome viral des herpèsvirus.





Figure 17 : Alignement des primases des herpesvirus humains. Seules deux régions sont connues pour participer à l'activité hélicase : les motifs DxD et doigt de zinc. Alignement réalisé avec les primases d'EBV, de KSHV, d'HSV-1, de VZV et de hCMV.





### III. Le rôle charnière du complexe H-P

Le complexe H-P des herpesvirus est composé de trois protéines, une hélicase, une primase et une sous-unité accessoire. Les protéines forment un complexe ayant une stœchiométrie de type 1/1/1 où chaque protéine interagit avec ses deux partenaires, cette spécificité a été retrouvée au sein des trois sous-groupes  $\alpha$ ,  $\beta$  et  $\gamma$  des herpesvirus (EBV [94], hCMV [95] et HSV [96]). La présence des protéines du complexe H-P au niveau de l'origine de réplication permet le lancement du processus réplcatif. Les activités portées par ce complexe sont indispensables à la réplication, l'hélicase permet l'ouverture de la fourche réplcative, la primase dépose des amorces d'ARN permettant la formation des fragments d'Okazaki sur le brin tardif et la sous-unité accessoire optimise les fonctions des deux protéines précédentes et permet le recrutement d'autre facteur de la réplication [97, 98]. Néanmoins le rôle de la sous-unité accessoire d'EBV (BBF2/3) reste encore vague. La description des différents rôles du complexe H-P exposés dans les paragraphes suivants rassemble les données obtenues à partir de plusieurs herpesvirus.

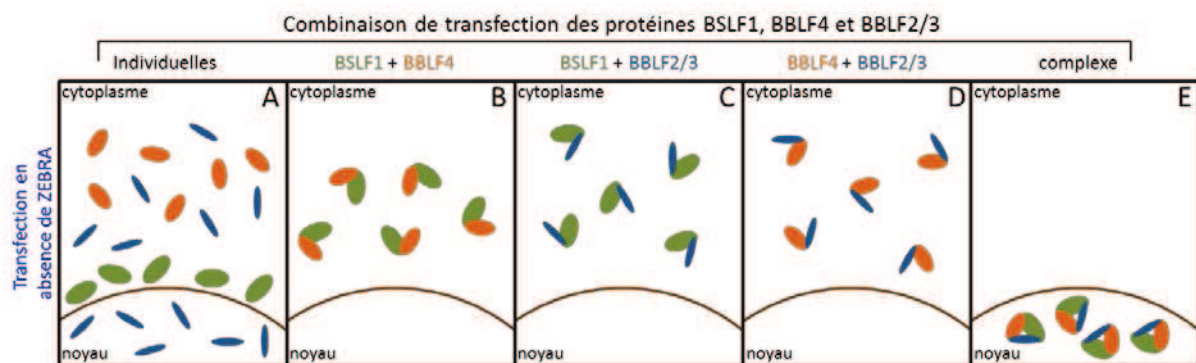
#### 1. Localisation cellulaire des protéines du complexe H-P

La réplication virale se déroule au sein de sites réplcatifs dans le noyau (CMV [99], HSV [100], KSHV [101]). Les protéines virales étant synthétisées au sein du cytoplasme, il faut qu'elles soient importées dans le noyau pour répliquer le génome viral.

Les études réalisées par Gao & al (1998) [90] ont permis de décrire la localisation des protéines du complexe H-P d'EBV dans des cellules Véro transfectées *in vitro*. Lorsque les protéines H-P sont exprimées individuellement dans la cellule, elles possèdent des localisations spécifiques. BBLF4 est exclusivement cytoplasmique, BSLF1 est périnucléaire et BBLF2/3 est répartie de façon uniforme entre le cytoplasme et le noyau [90] (Figure 19, encadré A). La même observation a été faite pour les protéines homologues du complexe H-P d'HSV-1. Les protéines UL5 (hélicase), UL52 (primase) et UL8 (sous-unité accessoire) ont une localisation cytoplasmique lorsqu'elles sont exprimées individuellement [98].

La localisation cellulaire des protéines est en partie modifiée lorsque les complexes binaires et le complexe ternaire sont formés au sein de la cellule. Les différents complexes binaires (BSLF1 + BBLF2/3, BBLF4 + BSLF1 et BBLF2/3 + BBLF4) ont tous une localisation exclusivement cytoplasmique (Figure 19, encadré B, C et D). La localisation partiellement nucléaire de BBLF2/3 est probablement supprimée par les protéines partenaires exclusivement cytoplasmiques. Le complexe entier (BSLF1 + BBLF2/3 + BBLF4) possède une localisation exclusivement nucléaire (Figure 19, encadré E). La

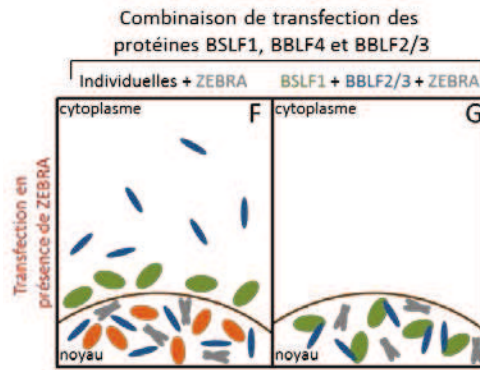
présence du complexe H-P d'EBV au niveau du noyau nécessite impérativement la présence des trois protéines, BSLF1, BBLF4 et BBLF2/3 [90]. Le complexe H-P d'HSV-1 est aussi nucléaire lorsque ses trois protéines sont exprimées simultanément, cette localisation est essentiellement due à la présence de la protéine accessoire (UL8), homologue de BBLF2/3 [98]. Que ce soit pour le complexe H-P d'EBV ou d'HSV-1, aucun signal NLS (Signal de Localisation Nucléaire) canonique n'a été identifié clairement pour une des trois protéines. Deux conclusions sont donc possibles : soit le signal NLS est formé lors de l'interaction des trois protéines à partir de résidus appartenant aux trois sous-unités du complexe, soit l'importation dans le noyau se fait par l'intermédiaire d'une interaction directe du complexe avec une importine.



**Figure 19 : Représentation schématique de la localisation cellulaire des protéines du complexe H-P d'EBV en l'absence du facteur de transcription ZEBRA. A :** transfection individuelle de chaque protéine. **B :** co-transfection des protéines BLSF1 et BBLF4. **C :** co-transfection des protéines BSLF1 et BBLF2/3. **D :** co-transfection des protéines BBLF4 et BBLF2/3. **E :** co-transfection triple, formation de l'holo-complexe. D'après l'étude de Gao & al (1998) [90].

Gao & al (1998) [90] ont obtenu des données supplémentaires lorsque ZEBRA est co-exprimée avec les protéines du complexe H-P. Transfectée seule la protéine ZEBRA est localisée exclusivement dans le noyau. La présence de ce facteur de transcription semble influencer la localisation cellulaire des protéines du complexe H-P. Lors de la co-expression de ZEBRA et BBLF4, l'hélicase est relocalisée au sein du noyau cellulaire (Figure 20, encadré F). La localisation des protéines monomériques BBLF2/3 et BSLF1 n'est pas modifiée par la présence de ZEBRA, mais le sous-complexe BBLF2/3 + BSLF1 devient nucléaire en présence de ZEBRA (Figure 20, encadré G). Lorsque le domaine de transactivation N-terminal de ZEBRA est délété, la relocalisation nucléaire de BBLF4 monomérique et du sous-complexe BSLF1 + BBLF2/3 n'est plus observée, indiquant que l'interaction entre ZEBRA et les protéines du complexe H-P est permise par ce domaine [90]. Le complexe entier est localisé dans le noyau en absence de ZEBRA et la présence de ZEBRA permettrait de placer le complexe H-P sur oriLyt.

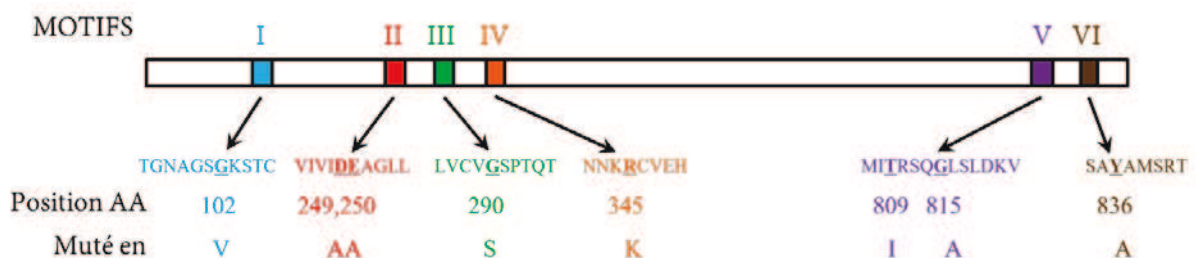




**Figure 20 :** Représentation de la localisation cellulaire des protéines du complexe H-P transfectées simultanément avec la protéine ZEBRA. F : transfection des protéines individuelles + ZEBRA. G : co-transfection BSLF1 + BBLF2/3 et ZEBRA.

## 2. L'hélicase, des motifs primordiaux mais peu connus

Les connaissances sur les activités hélicases et primases des herpèsvirus proviennent principalement des études réalisées sur le complexe H-P d'HSV-1 et en particulier sur l'hélicase (UL5) d'HSV-1. L'activité hélicase est seulement observée lors de la présence du sous complexe UL52-5, aucune activité hélicase ou primase n'est mesurée lorsque les protéines sont produites individuellement [102, 103]. Les tests hélicase sont généralement réalisés à une température comprise entre 30 et 37°C en présence d'ATP, de  $MgCl_2$  et de glycérol [104-108]. L'hélicase contient 6 motifs particuliers la reliant à la superfamille 1 (SF1) des hélicases [109] et possède une activité d'ouverture de l'ADN dans le sens 5'-3' [110]. Les motifs I (Walker A) et II (Walker B) sont très conservés et interviennent dans la stabilisation du NTP par une interaction directe du motif I et l'ion  $Mg^{2+}$  présent au niveau du motif II [111-113]. Pour évaluer l'importance des 6 motifs hélicase, Graves-Woodward & al (1997) [109] ont réalisé des mutations sur les résidus conservés au sein des herpèsvirus humains (Figure 21).



**Figure 21 :** Représentation des mutations réalisées sur les 6 motifs hélicase d'UL5. Adapté de Graves-Woodward & al (1997) [109].

Ces mutations abolissent la réplication lors d'expériences de complémentations répliquatives [110]. Elles réduisent fortement l'activité ATPase dépendante de l'ADN simple brin et abolissent complètement l'activité hélicase portée par la protéine. Le sous-complexe UL52-5 possède toujours une activité primase, indiquant que les mutations ne modifient pas fortement la structure d'UL5 et permettent de conserver l'interaction avec UL2. Tous les mutants sont encore capables d'interagir avec des substrats ADN de type fourche. La présence des protéines, sous-unité accessoire H-P (UL8), polymérase/facteur de processivité (UL30-42) et de la protéine d'adhésion à l'ADN simple brin (ICP8) en plus des protéines UL52-5 ne permet pas l'observation d'une réplication *in vitro*. Ainsi l'absence d'activité hélicase est sûrement due à l'incapacité de coupler les activités ATPase et hélicase présentes sur le complexe [109]. L'hélicase, non mutée, est capable d'ouvrir des duplex d'ADN supérieur à 30 pb seulement si le complexe H-P entier (UL52-5-8) est formé [97]. La cinétique d'ouverture d'un duplex d'ADN varie si le substrat est linéaire (24 pb/s) ou circulaire (60 pb/s), ces vitesses sont atteintes seulement en présence du complexe entier et de la protéine ICP8. La cinétique d'ouverture est réduite à 20 pb/s lorsque la polymérase d'HSV1 est présente, mais pas en présence de la polymérase T7 (65 pb/s). Pour HSV-1 c'est donc la polymérase qui limite la vitesse d'ouverture d'un duplex d'ADN [106].

**Tableau 2 : Tableau de synthèse des mutations réalisées sur les motifs hélicase** et leurs effets sur les activités ATPase, primase et hélicase.

Motifs	I	II	III	IV	V	VI
Mutations	G102V	D249A, E250A	G290S	R345K	T809I, G815A	Y836A
Activité ATPase dépendante de l'ADN simple brin	Faible					
Activité hélicase	nulle					
Activité primase	normale					

L'activité hélicase est mesurable seulement avec le sous complexe UL52-5, mais les motifs essentiels pour cette activité semblent être seulement portés par UL5. La présence de la sous-unité accessoire (UL8) permet de stimuler l'activité hélicase du complexe et favorise la réplication de longs brins d'ADN recouverts par la protéine d'adhésion à l'ADN simple brin (ICP8).



### 3. La primase une protéine en relation étroite avec l'hélicase

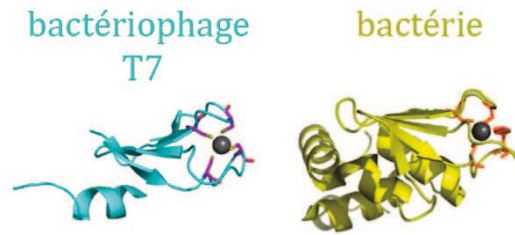
Comme pour l'hélicase, les données sur les activités et les motifs primases proviennent de l'étude de la primase (UL52) d'HSV-1. Les primases des herpèsvirus possèdent un motif particulier appelé DxD, il est aussi présent dans des primases procaryotes et eucaryotes [114-116]. Le rôle de ce motif pour l'activité primase a été évalué par les mutations des deux acides aspartique en alanine (D628A et D630A) et la mutation en glycine d'une asparagine située 4 résidus avant ce motif (N624G). Les trois mutations ne perturbent pas l'activité hélicase du sous complexe, indiquant qu'elles ne modifient pas fortement la structure d'UL52. L'interaction entre UL5 et UL52 est préservée. La présence de l'activité hélicase portée par les complexes UL52-5 mutés suggère que l'activité ATPase est toujours effective. Par contre l'activité primase et le couplage de la synthèse d'ADN sont fortement réduits par les mutations D628A et D630A. Le mutant N624G permet toujours la production d'amorces ARN de 12 nucléotides équivalent à l'action de l'holoenzyme et utilisables par la polymérase pour la synthèse d'ADN [117]. Les deux acides aspartiques se situent dans une zone hydrophobe favorisant la stabilisation du NTP par la présence de l'ion divalent  $Mg^{2+}$ . Ces données montrent que les activités hélicase et primase sont portées respectivement par les protéines UL5 et UL52.

**Tableau 3 : Tableau de synthèse des effets des mutations du motif DxD** sur les activités hélicase et primase du complexe UL52-5.

Motifs	DxD	pré-DxD
Mutations	D628A, D630A	N624G
Activité hélicase	normale	normale
Activité primase	nulle	normale

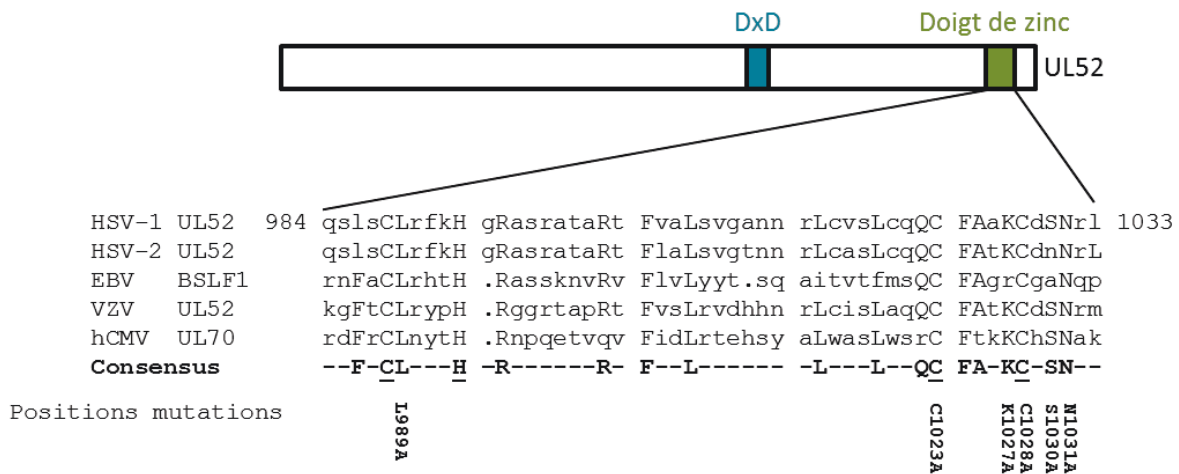
La protéine UL52 possède un motif doigt de zinc présumé au niveau de sa partie C-terminale. Ce motif est fortement conservé au sein des herpèsvirus et pour d'autres primases comme celles des bactéries, des bactériophages, des levures ou des mammifères (études chez la souris) [116, 118]. Les motifs doigt de zinc ont déjà été impliqués dans d'autres processus comme la reconnaissance de séquences spécifiques d'ADN des facteurs de transcription [119, 120]. La structure du doigt de zinc des primases est relativement similaire entre organismes (Figure 22). Les tests primases d'UL52 sont généralement réalisés à une température comprise entre 30 et 37°C en présence de  $MgCl_2$ , de NTP et à un pH compris entre 7,5 et 8,5 [105-108].

## Structure doigt de zinc



**Figure 22 : Structure des doigts de zinc des primases du bactériophage T7 et de la bactérie *Geobacillus stearothermophilus*.** La structure globale des doigts de zinc à proximité de l'ion  $Zn^{2+}$  est similaire. Les atomes de zinc, représentés par des sphères grises, sont coordonnés par les chaînes latérales de cystéines et histidines localisées dans deux boucles flanquées par des brins  $\beta$ . Code PDB (Protein Data Bank) bactériophage T7 : 1NUI, *G. stearothermophilus* : 1D0Q. Adapté de Lee & al (2012) [121].

Différents mutants ont été créés pour estimer la contribution du doigt de zinc de la primase sur les activités du complexe H-P. Les mutations des cystéines C1023 et C1028 en alanines (C1023A,C1028A) abolissent la formation du doigt de zinc, les trois autres mutations (L989A, K1027A et la double mutation S1030A,N1031A) impliquent des résidus conservés au sein des herpesvirus ne contribuant pas directement au maintien de la structure en doigt de zinc (Figure 23). Les tests menés avec les différents mutants ont été réalisés en présence des protéines hélicase (UL5) et sous-unité accessoire (UL8).



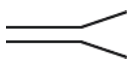
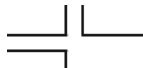
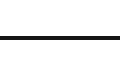

**Figure 23 : Schéma illustrant les résidus mutés dans la séquence du doigt de zinc d'UL52.** Les séquences contenant le doigt de zinc et les résidus mutés sont montrés. Adapté de Chen & al (2005) [107].

Les mutants L989A et K1027A présentent une activité primase tout juste détectable, le double mutant S1030A,N1031A retient une activité primase mais elle est très affaiblie, le dernier double mutant C1023A,C1028A ne présente aucune activité primase [107]. Les mutants C1023A,C1028A et L989A ne permettent pas la réplication virale lors d'expériences de

complémentation transitoires, alors que les mutants K1027A et S1030A,N1031A affichent respectivement 7 et 20% de complémentation. Le doigt de zinc est donc indispensable pour l'activité primase et par conséquent nécessaire pour l'activité de réplication.

Le complexe hélicase primase sauvage interagit avec des affinités variables avec l'ADN en fonction de sa structure. Le complexe possède une affinité décroissante pour les substrats suivants : substrat type fourche > substrat jonction triple > substrat ADN simple brin (Tableau 4). De plus le complexe requiert absolument une région simple brin pour avoir une interaction avec l'ADN, aucune interaction n'est observée avec des substrats ayant des bouts francs [107]. Les mutants K1027A et S1030A,N1031A conservent respectivement 5 à 10 % et 15 à 30 % de l'affinité du complexe sauvage en fonction des substrats d'ADN utilisés [107]. Les deux autres mutants (L989A et C1023A,C1028A) n'interagissent plus avec l'ADN ou de façon extrêmement faible. La capacité de liaison à l'ADN corrèle les profils de complémentation transitoire.

**Tableau 4 : Etude de l'affinité des complexe H-P sauvage et mutés d'HSV-1 en fonction du substrat ADN.** Mesure effectuée par gel retard avec 4 types de substrats : fourche, jonction triple, simple brin et bouts francs. Adapté de Chen & al (2005) [107].

	 Fourche	 Jonction triple	 ADN simple brin	 Bouts francs	
L989A	0,8	1,2	0,1	0	% de complexe observé
K1027A	2	3	0,5	0	
S1030A,N1031A	11,9	7,9	1,4	0	
C1023A,C1028A	0	0	0	0	
Sauvage	44,1	38,5	9,5	0	

Contrairement au motif DxD les mutations sur le doigt de zinc influent sur l'activité ATPase portée par le complexe. L'activité ATPase dépendante de l'ADN simple brin du complexe sauvage permet d'hydrolyser 19 nmol d'ATP pour 1 pmol de complexe en 30 min, cette valeur chute drastiquement en absence d'ADN (0,7 nmol d'ATP hydrolysé en 30 min par 1 pmol de complexe) [107]. Les mutations engendrent une baisse de l'activité ATPase dépendante de l'ADN simple brin. Les mutants K1027A et S1030A,N1031A possèdent 15 % de l'activité du complexe sauvage et pour les deux autres mutants (L989A et C1023A,C1028A) moins de 5%. L'affinité pour les NTPs des différents mutants n'est pas altérée, les  $K_m$  sont équivalents à ceux du complexe sauvage [122]. Leurs activités ATPase ADN indépendante sont similaires à celle du complexe sauvage. Il y a ainsi une corrélation

entre la diminution de l'affinité pour l'ADN et de l'activité ATPase dépendante de l'ADN simple brin [107].

Les comportements de ces mutants par rapport à l'activité hélicase du complexe sont très variés. Le mutant C1023A,C1028A ne présente aucune activité hélicase aussi bien en absence ou en présence de la protéine ICP8. Les mutants L989A et K1027A possèdent tous deux 15% (sans ICP8) et respectivement 30 et 70% de l'activité hélicase sauvage en présence d'ICP8. Un résultat surprenant provient du mutant S1030A,N1031A qui possède 50% (sans ICP8) et 120% (avec ICP8) de l'activité hélicase du complexe sauvage. Ce mutant fixe faiblement un substrat ADN de type fourche (30% de l'activité sauvage) et présente une très faible capacité à hydrolyser l'ATP (10 à 20 % de l'activité sauvage). Malgré cette réduction d'activité le mutant est capable d'ouvrir des duplexes d'ADN avec une efficacité supérieure à celle du complexe sauvage. Une des raisons expliquant ce phénomène serait la faible affinité de ce mutant pour un substrat d'ADN de type fourche conduisant à un « turnover » plus rapide induisant une augmentation de la vitesse d'ouverture [107]. Ce résultat révèle qu'il est délicat de lier l'activité hélicase à la seule fonction d'une ATPase.

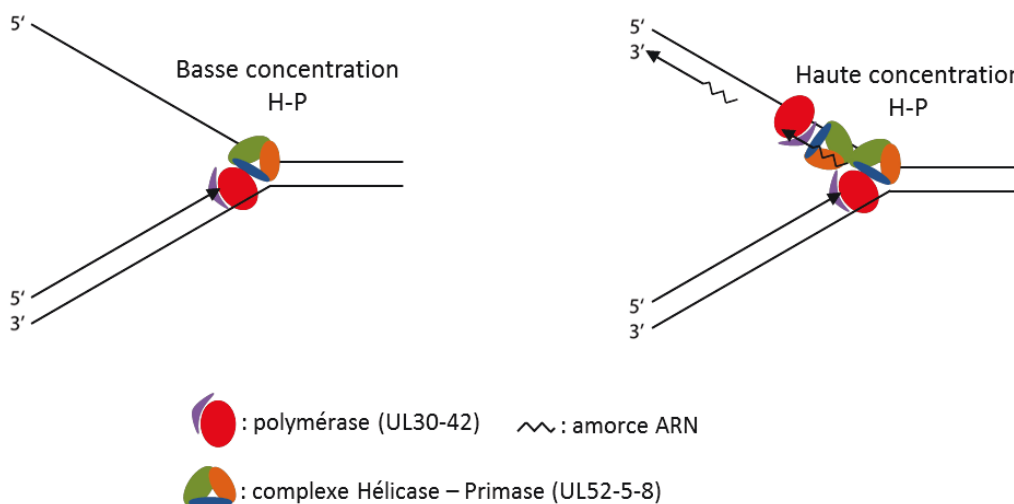
Le doigt de zinc du complexe H-P est essentiel pour l'activité primase du complexe, mais il influe aussi sur les activités ATPase et hélicase. Même si les protéines UL52 et UL5 semblent porter respectivement les motifs primase et hélicase, ces activités semblent couplées. Cette dépendance des deux protéines pour les activités hélicase et primase peut provenir d'un partage des sites de liaison à l'ADN ou bien les sites actifs dépendent directement du bon repliement du domaine doigt de zinc.

#### **4. Le recrutement du complexe dans les foyers réplicatifs et l'importance de sa stœchiométrie**

Lorsque que la protéine d'adhésion à l'origine (UL9 ou OBP) et la protéine de l'adhésion à l'ADN simple brin (ICP8) ont induit la déstabilisation de l'origine de réplication, le complexe H-P est recruté pour ouvrir le duplex d'ADN et synthétiser de courtes amorces d'ARN pour l'initiation de la réplication. Cette première étape est spécifique d'HSV-1, les autres herpèsvirus humains ne codent pas de protéine homologue à UL9, mais ils semblent utiliser un transcrit d'ARN pour déstabiliser l'origine de réplication [89, 92, 93]. Les étapes post-initiation sont similaires pour les herpèsvirus humains. Un complexe H-P recombinant est capable d'ouvrir *in vitro* un duplex d'ADN s'il possède une région simple brin faisant au moins six nucléotides [123]. L'activité primase peut être détectée dans les expériences utilisant des ADN simple brin, mais elle n'est pas toujours observée en présence de substrat de type fourche. Une explication à ce résultat a été récemment apportée par des tests de

retards sur gel et de résonance plasmonique de surface. Ils ont révélé qu'en présence d'une fourche d'ADN, des complexes d'ordre supérieur peuvent se former entre les protéines H-P et le substrat de type fourche [123]. De plus, l'activité primase du complexe H-P présente un effet coopératif en fonction de la concentration en protéines, mais pas pour les activités ATPase et hélicase [123]. Ces nouvelles données indiquent que la forme fonctionnelle du complexe H-P contient au moins deux copies de complexe H-P.

Le recrutement de la polymérase à la fourche de réplication nécessite la présence d'une activité primase. Le changement de conformation au sein du complexe H-P ou la présence d'amorces ARN sont indispensables pour le recrutement de la polymérase et de son facteur de processivité [122]. Le recrutement de la polymérase utilise probablement une interaction avec la sous-unité accessoire du complexe H-P [124]. Une fois la polymérase liée au complexe H-P il est possible d'initier la réplication des brins précoces et tardifs. Les essais de reconstitution de la réplication d'HSV *in vitro* ont révélé une synthèse efficace du brin précoce, mais qui reste très faible pour le brin tardif [109, 125]. Le couplage de la synthèse des brins précoces et tardifs est possible avec des petits substrats d'ADN circulaire en présence des protéines H-P, de la protéine d'adhésion à l'ADN simple brin (ICP8) et du complexe polymérase (UL30-42). Ce couplage est seulement observé avec une concentration élevée en complexe H-P [126]. La présence de deux complexes H-P permet de résoudre le problème des sens de déplacement opposés des activités hélicase et primase portées par le complexe H-P (Figure 24).



**Figure 24 : La composition du complexe de répliatif est dépendante de la concentration du complexe H-P.** A basse concentration en protéines H-P seule la synthèse du brin précoce est possible. A forte concentration en protéines H-P, il y a la formation de complexe d'ordre supérieur entre la fourche et le complexe H-P (au moins deux complexes UL52-5-8) permettant la synthèse des brins précoce et tardif. Adapté de Chen & al (2011) [123].



La cinétique de formation du complexe répliatif décrite précédemment se déroule au niveau des foyers de réplication au sein du noyau de la cellule infectée. Dans un premier temps, un complexe pré-répliatif composé de la protéine d'adhésion à l'ADN simple brin (ICP8), des protéines H-P et de la protéine d'adhésion à l'origine (UL9 ou OBP) est retrouvé au niveau des foyers répliatifs. Le recrutement de la polymérase (UL30-42) nécessite la présence d'une activité primase dans les foyers [122]. La présence des protéines d'adhésion à l'ADN simple brin (ICP8), H-P et polymérase forme un complexe répliatif fonctionnel permettant la réplication du génome viral au niveau des foyers répliatifs.

## **5. L'implication des protéines BBLF4, BSLF1 et BBLF2/3 au sein du complexe de réplication du virus Epstein-Barr**

Les études réalisées sur le complexe H-P d'EBV permettent de décrire le déroulement de la réplication et la formation du complexe répliatif sur oriLyt.

Le facteur de transcription ZEBRA interagit directement sur les protéines BBLF2/3 et BBLF4 composant le complexe H-P. BBL2/3 interagit sur les résidus 11 à 25 du domaine de transactivation de ZEBRA situé en N-terminal [90]. L'interaction de BBLF4 avec ZEBRA est située sur deux zones N-terminales comprenant les résidus 22 à 26 et 74 à 86 de ZEBRA [127]. Cette interaction n'est pas que structurale, mais elle conditionne aussi le comportement de ZEBRA. En présence de BBLF4, l'activité de transactivation de ZEBRA est favorisée et permet l'activation du promoteur de BHLF1. De plus lorsque BBLF4 est absente, ZEBRA n'est pas localisée dans les compartiments répliatifs formés au sein du noyau. Ainsi l'interaction de ZEBRA avec le complexe H-P ne permet pas seulement son recrutement au niveau de l'origine de réplication, mais influe aussi directement sur l'activité de ZEBRA [127].

La polymérase (BALF5) interagirait directement avec chacune des protéines formant le complexe H-P [91]. Pour former l'holoenzyme, BALF5 interagit avec son facteur de processivité (BMRF1), qui n'interagit pas directement avec le complexe H-P, mais par l'intermédiaire de BALF5 [91]. L'absence de structure tridimensionnelle pour le complexe H-P d'EBV ne permet pas encore d'estimer si l'interaction avec BBLF4 est transitoire ou constante lors de la progression sur l'ADN.

Plus récemment il a été montré que la formation du complexe de réplication permet un rétro contrôle positif en début de réplication en augmentant l'activation de la transcription de Rta par ZEBRA [128]. L'augmentation de la transcription de Rta conduit à un renforcement de la reconnaissance d'oriLyt et de la réplication virale. Ce rétro contrôle positif est permis lorsque les trois protéines, BSLF1, BALF2 et BRMF1 sont présentes au niveau d'oriLyt par l'intermédiaire de ZEBRA

[128]. La présence simultanée de ces trois protéines indique que le complexe de réplication est complet et prêt à répliquer le génome viral.

Les protéines formant le complexe de réplication permettent aussi le recrutement de protéines cellulaires. La protéine BBLF2/3 semble recruter la protéine ZBRK1 et son co-represseur KAP-1 au niveau du complexe de réplication d'EBV [129]. Ces deux protéines permettraient l'interaction avec la protéine BRCA1 qui possède une activité de réparation sur les cassures double brin. BBLF2/3 semble contribuer à la réplication lytique d'EBV aussi par le recrutement de protéines cellulaires en relation avec la réplication.

Le processus permettant la réplication du génome viral d'EBV en phase lytique implique un grand nombre de protéines virales. L'interaction séquentielle des protéines permet la formation d'un complexe de réplication actif, mais permet aussi un contrôle et une régulation de la réplication. Le mécanisme de réplication des herpèsvirus humains se révèlent être plus compliqué que la plupart des virus.

## **IV. Hélicases des herpèsvirus et Dda des bactériophages, un passé commun ?**

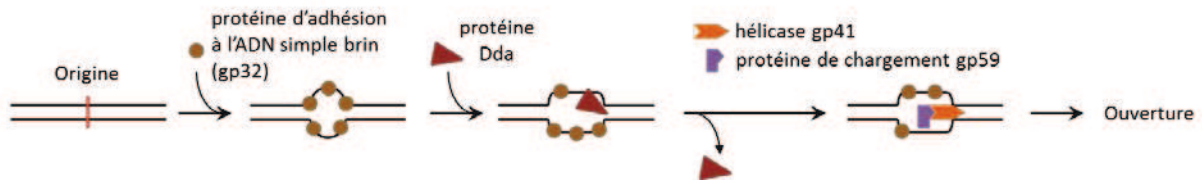
Lors de l'écriture de la revue nous avons cherché à relier les hélicases des herpèsvirus et les différentes familles d'hélicases. Les hélicases des herpèsvirus étant monomériques et possédant une directionnalité 5'-3', elles sont proches des hélicases de la super famille 1B (SF1B). Pour affiner ce positionnement, nous avons aligné les hélicases des herpèsvirus avec d'autres membres de cette famille, ce qui nous a permis de rapprocher ces hélicases des hélicases Dda (ATPase ADN dépendante) des bactériophages. Les hélicases de type Dda sont essentiellement retrouvées dans le genre « T4-like viruses » (ordre : *Caudovirales*, famille : *Myoviridae*, sous famille : *Tenenvirinae*). Le bactériophage T4 étant le représentant principal de ce genre, je vais développer les connaissances actuelles de la Dda de ce bactériophage.

### **1. Les hélicases du bactériophage T4**

Le bactériophage T4 encode deux hélicases nécessaires pour la réplication de son génome, la Dda et la gp41 (gène protéine 41) possèdent 20 % d'identité de séquence. La présence de Dda au niveau de la fourche de réplication n'augmente pas le taux d'ouverture exercé par gp41 ; ainsi les deux hélicases ne possèdent pas d'action synergique au niveau de la fourche. L'action de l'hélicase

Dda est distributive et elle n'interagit pas avec la primase (gp61), contrairement à la gp41 qui est processive et interagit avec gp61 pour former le primosome [130]. Ainsi Dda semble être utilisée pour l'initiation de la réplication, alors que gp41 permet la réplication complète du génome.

Phase d'initiation de la réplication :



**Figure 25 : Schéma de l'initiation de la réplication du génome viral du bactériophage T4.** En début d'infection par le phage T4, l'origine est ouverte par les protéines de liaison à l'origine. Cette fourche pré-répllicative est maintenue par l'interaction avec la protéine de liaison à l'ADN simple brin (gp32). La protéine Dda permet d'amorcer la formation d'une fourche de réplication. L'hélicase (gp41) permet l'extension de cette fourche répllicative. La protéine gp49 permet le chargement de l'hélicase répllicative gp41 sur le brin non apparié menant à la formation d'une fourche de réplication opérationnelle. Adapté de Gauss & al (1994) [131].

## 2. Caractéristique de l'hélicase Dda et son implication dans la réplication des bactériophages

### 2.1. Implication de la protéine Dda dans la réplication des bactériophages

La protéine Dda interagit fortement et spécifiquement avec la protéine d'adhésion à l'ADNsb du phage T4 nommée gp32 (gène protéine 32) [130, 132]. L'activité de synthèse de déplacement de brin effectuée par Dda est seulement observée en présence de gp32 [133]. La présence de Dda stimule la synthèse du brin avancé *in vitro*, mais sa délétion n'abolit pas la réplication du phage T4 *in vivo* [134]. Dda est importante pour l'initiation de la réplication, mettant en jeu l'origine répllicative, mais elle n'est pas indispensable pour les étapes tardives de la réplication [135]. L'ouverture de l'origine induite par les protéines d'adhésion à l'origine est maintenue par la présence de la protéine d'adhésion à l'ADN simple brin (gp32). L'hélicase Dda cible les portions d'ADN recouvertes par la protéine gp32 permettant la formation d'un complexe pré-répliatif. Ce complexe permet l'ouverture de l'ADN favorisant une région simple brin plus importante, favorisant le chargement de l'hélicase répllicative (gp41) par sa protéine de chargement (gp59).

### 2.2. Caractéristique de la Dda du bactériophage T4

La protéine Dda du bactériophage T4 est composée de 439 acides aminés et semble majoritairement monomérique en solution. Cette protéine a la capacité de se déplacer sur des ADN simple brin avec une directionnalité 5'-3' [136] et d'ouvrir des duplex d'ADN dans le même sens celui

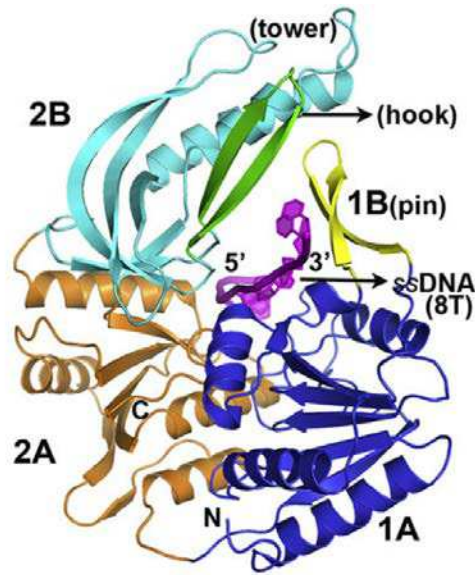
d'ouverture de la fourche [130, 132]. Pour observer cette capacité d'ouverture d'un ADN double brin, le substrat doit obligatoirement avoir un brin flottant en 5' d'au moins six nucléotides [137, 138]. La fonction de moteur moléculaire est observée en présence de Dda monomérique, bien qu'il ait été montré que la présence de plusieurs molécules améliore sa fonction de moteur moléculaire *in vitro* [137].

La Dda du bactériophage T4 est une hélicase peu processive. Sa faible processivité est due à un taux de dissociation rapide entre l'ADN et la protéine Dda [139]. L'analyse de la cinétique de Dda suggère que cette protéine effectue une pause entre deux évènements rapides de séparation de brin. Cette cinétique présentant des pauses ne limite pas le couplage du déplacement sur l'ADN ( $233 \pm 27$  nt/s) et de l'ouverture de duplex ( $258 \pm 17$  pb/s) qui est proche de la perfection *in vitro* [140]. Une autre capacité est portée par la protéine Dda : elle est capable de déplacer des protéines présentes sur le duplex ADN en cours d'ouverture [138].

La protéine Dda est importante pour l'initiation de la réplication des bactériophages T4, mais ne permet pas à elle seule la réplication du génome viral.

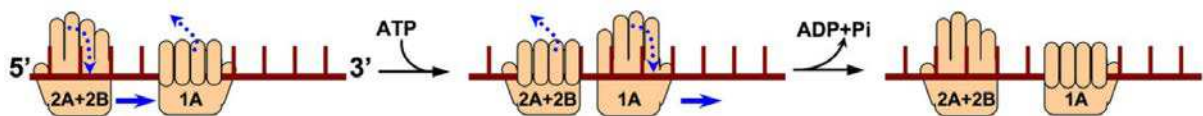
### **3. Informations complémentaires apportées par la structure de la Dda du bactériophage T4**

He & al [140] ont récemment dévoilé la structure de la protéine Dda en interaction avec un ADN simple brin de 8 nucléotides (structure parue après la publication de ma revue). Elle possède les six motifs caractéristiques des hélicases (I à VI) et a une structure globale similaire à RecD2 faisant aussi partie de la famille SF1B. Cette structure est composée des 4 domaines, 1A, 2A, 1B et 2B. Les domaines 1B et 2B sont des insertions présentes respectivement dans les domaines 1A et 2A. La structure de Dda révèle trois motifs particuliers, « pin » (1B), « hook » (2B) et « tower » (2B) (Figure 26). Les motifs « tower » et « pin » interagissent à leur extrémité distale. Ces deux motifs sont très bien conservés au sein des phages du genre « T4-like ». Le motif « hook » est plus important que celui de RecD2 (autre protéine SF1B) et le motif « tower » n'est pas retrouvé dans RecD2. L'interaction avec l'ADN simple brin est fortement similaire entre les structures de Dda et de RecD2. Le motif « pin » est absolument indispensable pour l'ouverture d'un duplex d'ADN. L'interaction observée entre les motifs « tower » et « pin » permet le couplage des activités de liaison à l'ATP et d'hydrolyse avec l'ouverture de l'ADN.



**Figure 26 : Structure de la protéine Dda (K38A) en interaction avec un substrat ADNsb de 8 nucléotides.** La structure comprend 4 domaines : 1A (bleu), 2A (marron), 1B (jaune) et 2B (turquoise et vert). L'ADNsb est coloré en violet. Source He & al (2012) [140]

Les activités reliées aux motifs « pin », « hook » et « tower » sont en accord avec le modèle de déplacement des protéines SF1B proposé par Saikrishnan & al (2009) [141] (Figure 27). Pour Dda le « pin » permet la séparation du duplex et contribue l'entrée du brin dans le sens 5'-3'. Les domaines 2A et 2B se déplacent en une seule unité, ce qui maintient l'interaction « pin »/« tower » permettant le couplage de la liaison à l'ATP et son hydrolyse avec l'ouverture du double brin d'ADN.



**Figure 27 : Modèle de déplacement des protéines SF1B.** Le déplacement est permis par l'hydrolyse de l'ATP. Avant l'hydrolyse de l'ATP, les domaines 2A+2B se déplacent le long de l'ADN pour se rapprocher du domaine 1A. La fixation de l'ATP permet de fixer les domaines 2A+2B à l'ADN et de libérer le domaine 1A qui se déplace le long de l'ADN. L'hydrolyse de l'ATP permet la fixation de l'ADN par le domaine 1A et la libération des domaines 2A+2B. Un nouveau cycle de déplacement peut débuter. Les mains ouvertes représentent les domaines avec une faible interaction sur l'ADN et peuvent se déplacer selon le sens indiqué par les flèches pleines. Les mains fermées représentent une interaction forte du domaine avec l'ADN. Les flèches en pointillées représentent la conversion entre les interactions forte et faible des domaines pour l'ADN. Source Saikrishnan & al (2009) [141].



## 4. Relation entre la Dda et les hélicases des herpèsvirus

Les spécificités de la protéine Dda exposées dans le paragraphe précédent permettent de faire un parallèle au niveau fonctionnel avec la protéine UL9 d'HSV-1. L'interaction d'UL9 avec la protéine d'adhésion à l'ADN simple brin (ICP8) permet de stimuler les activités hélicase et ATPase ADN dépendantes portée par UL9. De plus le complexe UL9-ICP8 a été impliqué dans l'ouverture de l'origine de réplication d'HSV-1 [142, 143]. Mais la protéine UL9 est spécifique d'herpès simplex et aucun homologue de cette protéine n'est retrouvé au sein des autres herpèsvirus humains.

La protéine Dda peut aussi être rapprochée de l'hélicase répliquative des herpèsvirus. Ces deux types d'hélicase sont monomériques et affichent un déplacement et une ouverture de l'ADN avec une directionnalité 5'-3'. Elles possèdent les six motifs caractéristiques des hélicases et sont attribuées à la superfamille 1B. Dans le paragraphe II.1 de ce chapitre (p.38) il est montré que les alignements de séquences des motifs V et VI entre les hélicases répliquatives des herpèsvirus et les Dda des bactériophages permet d'établir que ces deux types de protéines font parties du même groupe d'hélicase. Ainsi il semble possible que l'hélicase Dda permettant l'initiation de la réplication du génome virale pour le phage T4 soit devenue l'hélicase répliquative des herpèsvirus.

## V. Les objectifs du travail de thèse

L'équipe au sein de laquelle j'ai effectué ma thèse possède comme thématique principale l'étude des protéines du virus Epstein-Barr. Les protéines étudiées sont surtout celles participant directement ou étant associées à la réplication. L'équipe a fortement contribué à la détermination des structures des protéines répliquatives d'EBV (détails dans le paragraphe II.1 p.38-52). Pour continuer à approfondir la connaissance des protéines intervenant dans la réplication d'EBV l'équipe s'est intéressée aux protéines formant le complexe Hélicase-Primase d'EBV. Ce complexe est composé des protéines : hélicase (BBLF4), primase (BSLF1) et sous-unité accessoire (BBLF2/3). Aucune information structurale n'est disponible pour le complexe ou les protéines individuelles d'EBV ni pour les autres herpèsvirus. L'obtention de la structure du complexe permettrait une avancée majeure de la compréhension du complexe H-P des herpèsvirus et permettra potentiellement d'affiner les connaissances sur le mécanisme de réplication des herpèsvirus. Ce complexe est essentiel pour la génération de nouveaux virions infectieux lors de la phase de réplication lytique du virus. La connaissance sur la structure des protéines et les interactions permettant la formation du complexe H-P peuvent favoriser la découverte de nouvelles molécules possédant une activité antivirale ou d'optimiser les molécules disponibles actuellement. Au début

des années 2000, des composés dérivés de thiazolylphényle se sont révélés être des inhibiteurs du complexe H-P d'HSV-1, HSV-2 et VZV. Des études pré-cliniques récentes ont révélées que le composé ASP2151 (anemavir) possède une plus grande efficacité que le valaciclovir (analogue nucléosidique) pour traiter des lésions cutanées orales [108]. La connaissance structurale d'un complexe H-P d'herpèsvirus humain favoriserait une optimisation rationnelle de ces composés. Le travail qui sera réalisé sur le complexe H-P d'EBV permettra peut-être d'évaluer l'efficacité de cet inhibiteur vis-à-vis de ce complexe.

Au début de ma thèse l'essentiel des informations disponibles pour le complexe H-P des herpèsvirus provenaient principalement des études menées sur herpès simplex. Les activités hélicase et primase étaient relativement bien décrites pour les protéines d'HSV-1. Il y a moins d'informations sur la protéine accessoire, mais elle est importante pour la stimulation des activités hélicase et primase. Le manque d'informations structurales pour les complexes H-P des herpèsvirus humains et l'absence totale de connaissances sur les activités de ces protéines chez EBV, m'ont poussé à entreprendre l'étude du complexe H-P d'EBV pendant ma thèse.

Des tests de clonages en bactéries (*E. coli*) avaient déjà été réalisés au sein de notre équipe pour les protéines BSLF1 et BBLF4, mais aucune de ces deux protéines ne s'exprimait. Les compétences de l'équipe pour la production de protéines recombinantes en cellules d'insectes [144, 145] et les travaux menés par Yokoyama & al (1999) [94] sur l'expression du complexe H-P en cellules d'insectes ont été le point de départ de ma thèse. Pour pouvoir étudier ces trois protéines j'ai entrepris leur clonage dans trois baculovirus indépendants. Cette approche fut réalisée dans l'optique de pouvoir d'abord étudier les protéines individuellement et de pouvoir aussi les produire en complexe (binaire ou ternaire) par co-infection des cellules d'insectes dans l'optique d'études structurales et fonctionnelles des protéines seules et en complexe.

Partie II :  
Matériel & Méthodes

# CHAPITRE 1 :

## Etudes préliminaires

# I. Prédiction de désordre et caractéristiques des protéines

## 1. Prédiction d'ordre et de désordre

La prédiction d'ordre et de désordre permet d'estimer si une protéine possède des motifs structurés ou des boucles flexibles. Un grand nombre de scripts évaluant l'ordre et le désordre d'une protéine sont disponibles sur internet. Les résultats d'un seul prédicteur ne peuvent être suffisants pour décrire le désordre d'une protéine. Chaque prédicteur utilise des scripts propres pour évaluer le désordre, basés sur différents paramètres (reconnaissance de motif « hydrophobes », recherche de structures en hélice $\alpha$ , en feuillet $\beta$  ou en boucles dites « chaudes »...). Ils décrivent l'état ordonné ou désordonné de chaque acide aminé. Pour une prédiction fiable, il faut prendre en compte les résultats de plusieurs prédicteurs. Le script « D-score », développé au sein de notre laboratoire par le Dr Ambroise Desfosses, permet de normaliser puis combiner les résultats de 11 prédicteurs de désordre et d'obtenir ainsi une vue globale de l'état de la protéine d'intérêt. Les prédicteurs de structures utilisés sont listés dans le Tableau 5.

**Tableau 5 : Liste des 11 prédicteurs utilisés dans le script « D-score »** pour estimer le désordre des protéines du complexe H-P d'EBV. L'accès aux serveurs se fait depuis internet, le script « D-score » est utilisé en local. std : utilisation des paramètres par défaut du script.

Prédicteurs	Paramètres	Liens	publication
DisEMBL <sup>TM</sup>	std	<a href="http://dis.embl.de/">http://dis.embl.de/</a>	[146]
DISPOPRED2	std	<a href="http://bioinf.cs.ucl.ac.uk/disopred/">http://bioinf.cs.ucl.ac.uk/disopred/</a>	[147]
FoldIndex <sup>®</sup>	std	<a href="http://bip.weizmann.ac.il/fldbin/findex/">http://bip.weizmann.ac.il/fldbin/findex/</a>	[148]
GlobPlot 2.3	std	<a href="http://globplot.embl.de/">http://globplot.embl.de/</a>	[149]
IUPred	short disorder long disorder	<a href="http://iupred.enzim.hu/">http://iupred.enzim.hu/</a>	[150]
PONDR1	std	N'existe plus	
POODLE	S, S(High B factor), L, W, I	<a href="http://mbs.cbrc.jp/poodle/poodle.html">http://mbs.cbrc.jp/poodle/poodle.html</a>	[151-153]
PrDOS	std	<a href="http://prdos.hgc.jp/cgi-bin/top.cgi">http://prdos.hgc.jp/cgi-bin/top.cgi</a>	[154]
RONN	std	<a href="http://www.strubi.ox.ac.uk/RONN">http://www.strubi.ox.ac.uk/RONN</a>	[155]
SPRITZ	short disorder long disorder	<a href="http://distill.ucd.ie/spritz/">http://distill.ucd.ie/spritz/</a>	[156]
DisProt	VL2, VL3, VL3H, VSL2B	<a href="http://www.ist.temple.edu/disprot/Predictors.html">http://www.ist.temple.edu/disprot/Predictors.html</a>	[157-159]

Le script « D-score » calcule pour chaque acide aminé la moyenne des résultats donnés par les 11 prédicteurs. En fonction du prédicteur utilisé, les résultats peuvent être formatés de deux manières différentes. Soit le prédicteur donne deux états possibles, l'acide aminé est ordonné



(valeur 1) ou l'acide aminé est désordonné (valeur 0). Certains prédicteurs donnent des valeurs intermédiaires entre 0 et 1, dans ce cas les résultats sont normalisés à un seuil fixé à 0,5. Les acides aminés ayant une valeur supérieure à 0,5 sont considérés comme ordonnés et leur valeur devient 1. Dans le cas contraire pour une valeur inférieure à 0,5, l'acide aminé est considéré comme désordonné et sa valeur devient 0. Cette normalisation des résultats est faite par le script « D-score », puis il calcule la valeur moyenne de l'ensemble des prédicteurs pour chaque acide aminé. Les résultats du « D-score » sont donnés sur un graphique permettant de visualiser la probabilité d'ordre de chaque acide aminé de la séquence protéique.

## **2. Caractérisation bio-informatique complémentaire**

Pour affiner la connaissance de chaque protéine, leurs séquences ont été analysées sur le site PredictProtein (<http://www.predictprotein.org/>)[160]. Ce site regroupe un grand nombre de script permettant de décrire diverses caractéristiques d'une protéine à partir de sa séquence (structures secondaires, accès au solvant, région globulaire, région « coiled-coil », ponts disulfures, interaction protéine-protéine ou protéine-ADN ...). Pour compléter les informations de ce site, nous avons utilisé le site ProtParam (<http://web.expasy.org/protparam/>) pour analyser la composition de la séquence protéique [161] et le site cNLS Mapper (<http://nls-mapper.iab.keio.ac.jp/>) pour évaluer la présence d'un signal NLS (Signal de Localisation Nucléaire) sur les protéines du complexe [162, 163].

## CHAPITRE 2 :

# Expression des protéines recombinantes

# I. Expression et production de protéines recombinantes

## 1. Les prérequis

### 1.1. Les cellules d'insectes

Deux types de lignées cellulaires ont été utilisés au cours des expériences de clonage, d'expression et de production protéique :

**Les cellules SF21** : C'est une lignée continue obtenue à partir d'ovaire de chenille (*Spodoptera frugiperda*). Ces cellules sont rondes et de petites tailles, elles ont été utilisées pour les approches par cellules adhérentes et en suspension. Les SF21 sont optimisées pour la production virale et sont donc parfaitement adaptées pour les étapes de transfection et d'amplification du titre viral.

**Les cellules HighFive (Hi5)** : C'est une lignée cellulaire issue d'ovaire de chenille (*Trichoplusia ni*). Les Hi5 sont rondes et de petites tailles en suspension, alors qu'elles adoptent une forme de type fibroblaste en culture adhérente. Elles sont très bien adaptées pour la production protéique. Leur phénotype fibroblaste en culture adhérente change au cours de l'infection, elles deviennent rondes et certaines sont en suspension.

### 1.2. La culture cellulaire

- Cellules adhérentes :
  - Entretien de cellules en flasques de 75 cm<sup>2</sup> (BD Falcon)
  - Production protéique en flasques de 300 cm<sup>2</sup> (BD Falcon)
  - Transfection et « plaque assay » en plaque six puits (BD Falcon)
- Cellules en suspension :
  - Entretien des cellules dans des erlenmeyers de 250 mL (Dutscher)
  - Production protéique dans des erlenmeyers de 2 L (Dutscher)
  - La transfection est réalisée en plaque 6 puits (BD Falcon)
  - L'amplification virale est réalisée dans des erlenmeyers de 500 mL (Dutscher)

### 1.3. Les milieux de cultures

**TC100 (Gibco)** : Le milieu TC100 est utilisé pour la culture des cellules adhérentes. En fonction des expériences le milieu est utilisé non supplémenté ou supplémenté. La supplémentation est réalisée par l'ajout de 10% sérum de veau fœtal (PAN Biotech GmbH) et l'ajout des deux antibiotiques pénicilline/streptomycine (Gibco) à 50 µg/mL. Dans la suite du « matériels et

méthodes » la référence à du milieu TC100 supplémenté correspond à la présence du sérum et des antibiotiques.

**HyClone SFM4Insect™ (Thermo Scientific)** : Ce milieu est utilisé pour la culture des cellules en suspension. Le milieu est utilisé non supplémenté et sans antibiotiques.

## 2. Production de protéines recombinantes en cellules adhérentes

### 2.1. Principes du clonage par le système Bac-to-Bac® (Invitrogen)

La génération d'un baculovirus contenant un gène d'intérêt comporte 4 grandes étapes

- **Etape 1** : Le gène d'intérêt est cloné dans un vecteur donneur.
- **Etape 2** : Le vecteur donneur est introduit par transformation dans des bactéries compétentes contenant le génome du baculovirus (bacmide). La transposition est permise par les transposons Tn7 présents sur le plasmide donneur, de part et d'autre du gène d'intérêt, et la présence de Tn7 *att* (Tn7 attachment) sur le bacmide. Le plasmide helper présent dans la bactérie exprime les protéines requises pour la transposition. La transposition du gène dans le bacmide inactive la production de la protéine LacZ $\alpha$  et permet de sélectionner les bacmides recombinant par un criblage blanc/bleu. Le bacmide recombinant est purifié en prévision de sa transfection.
- **Etape 3** : Le bacmide contenant le gène d'intérêt est transfecté dans les cellules d'insectes grâce à un lipofectant. Les cellules d'insectes produisent la première génération de baculovirus recombinant.
- **Etape 4** : La première génération de virus est amplifiée pour une production à grande échelle. La production de la protéine recombinante est établie par infection des cellules d'insectes par le baculovirus.

Le protocole Bac-to-Bac® a été suivi pour la génération des trois baculovirus exprimant chacun une protéine du complexe H-P d'EBV. Les adaptations du protocole que j'ai effectué pour le clonage des protéines seront détaillées dans les sections suivantes.

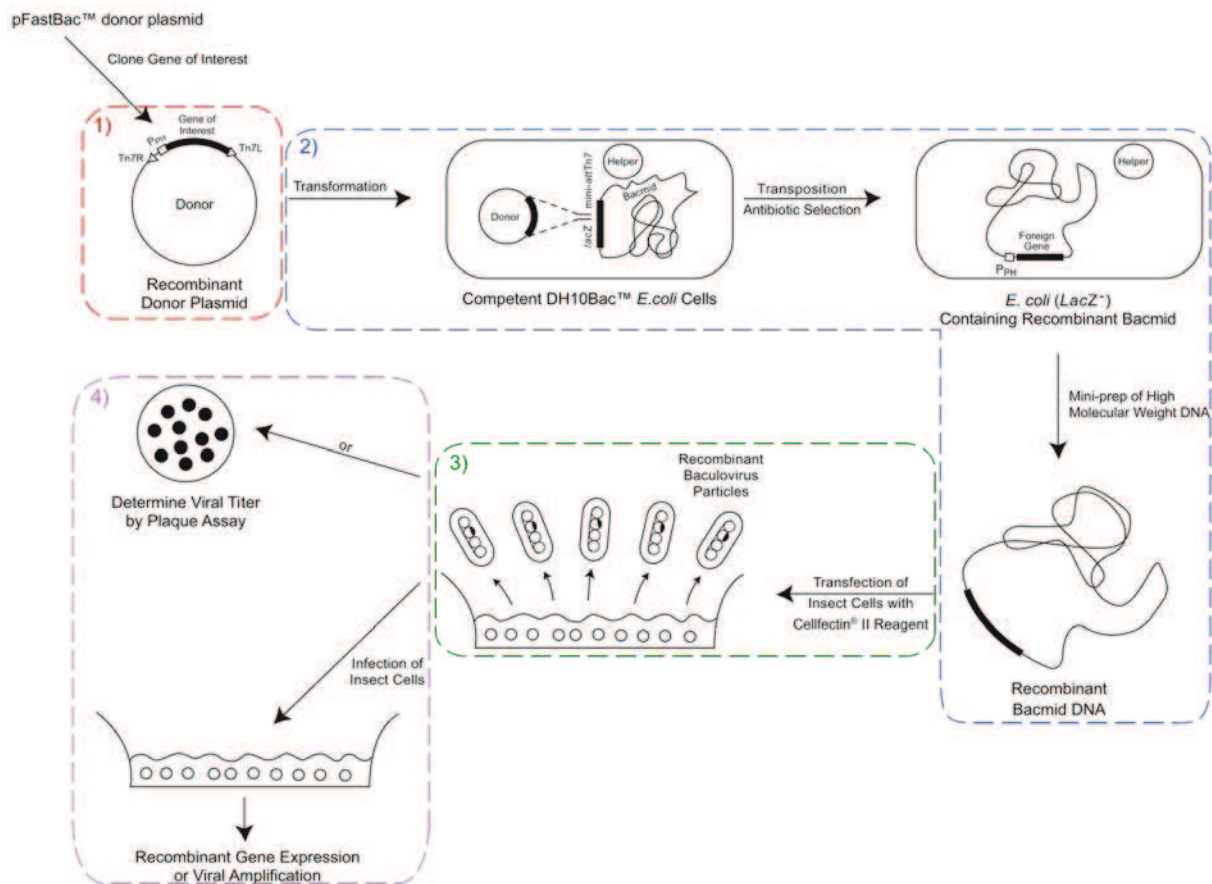
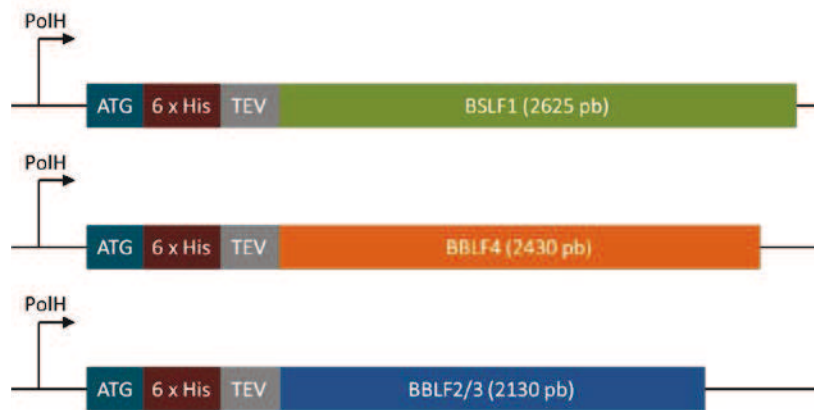


Figure 28 : Représentation schématique des étapes requises pour l'expression d'un gène d'intérêt par le système d'expression Baculovirus Bac-to-Bac®. Adapté de Bac-to-Bac® manuel version F.

## 2.2. Constructions utilisées en cellules adhérentes

La culture en cellules d'insectes adhérentes a permis de générer un baculovirus recombinant pour chaque protéine du complexe H-P d'EBV. Les gènes entiers de chaque protéine ont été isolés par PCR à partir de la lignée cellulaire B95-8. C'est une lignée de référence totalement séquencée [164]. Les étapes ayant permis d'obtenir le gène codant pour la protéine BBLF2/3 sont détaillées à la fin de ce paragraphe. Les gènes d'intérêts sont ensuite clonés dans le plasmide donneur pFastBac™HTB (Invitrogen). L'insertion des gènes dans ce vecteur permet de produire des protéines possédant une étiquette hexahistidine N-terminale clivable par la protéase TEV (*Tobacco Etch Virus*). Les gènes sont sous le contrôle du promoteur tardif PolH (polyhédrine) permettant la production de protéine en grande quantité.



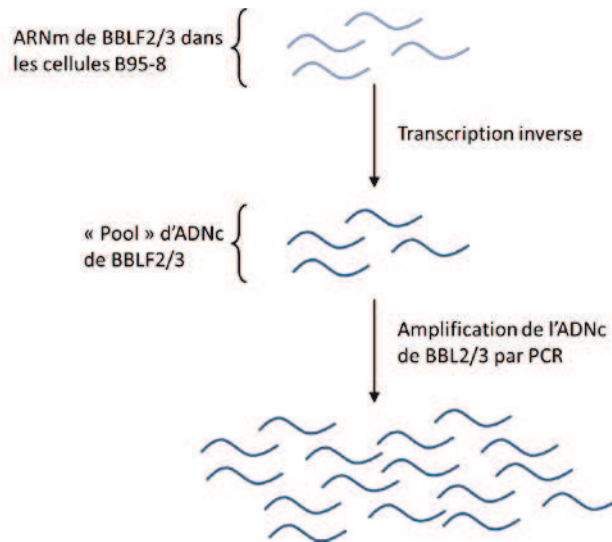


**Figure 29 : Représentation schématique des trois constructions utilisées en cellules adhérentes.** Chaque gène du complexe H-P est inséré dans le plasmide pFastBac™HTB, ils possèdent tous une étiquette hexahistidine et un site de clivage en 5' du gène codant pour la protéine d'intérêt. ATG : méthionine initiant la transcription. 6 x His : étiquette hexahistidine. TEV : site de clivage pour la protéase TEV.

### 2.2.1. Clonage de la protéine BBLF2/3

La protéine BBLF2/3 provient de l'épissage des cadres de lecture ouverts BBLF2 et BBLF3 après l'élimination d'un intron de 129pb. Il n'est donc pas possible d'amplifier par PCR le gène codant directement pour la protéine à partir de la lignée cellulaire B95-8. Deux approches différentes ont été utilisées pour le clonage de ce gène.

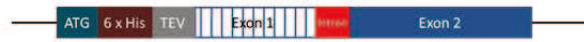
- **Isolation de l'ARNm et amplification de l'ADNc :** La RT-PCR (Transcription Inverse par Réaction de Polymérisation en Chaîne) permet d'effectuer une transcription inverse de l'ARNm codant pour la protéine BBLF2/3 et d'obtenir son ADNc. L'ADNc est ensuite amplifié par PCR (Fig13). Le kit « Titan One tube RT-PCR » (Roche Applied Science) a été utilisé pour réaliser la RT-PCR. Il permet d'amplifier en une seule étape l'ADNc de BBLF2/3 en couplant la RT-PCR et la PCR. La lignée B95-8 a été utilisée comme support de départ pour la RT-PCR. Dans cette lignée, 2 à 5% des cellules sont en phase de réplication et expriment l'ARNm de BBLF2/3. Un traitement au PMA et/ou à l'acide butyrique a été appliqué aux cellules B95-8 permettant l'entrée en réplication lytique d'un plus grand nombre de cellules. Et favorisant donc la présence de l'ARNm de BBLF2/3. Pour l'amplification de l'ADNc utilisation d'amorces spécifiques de BBLF2/3



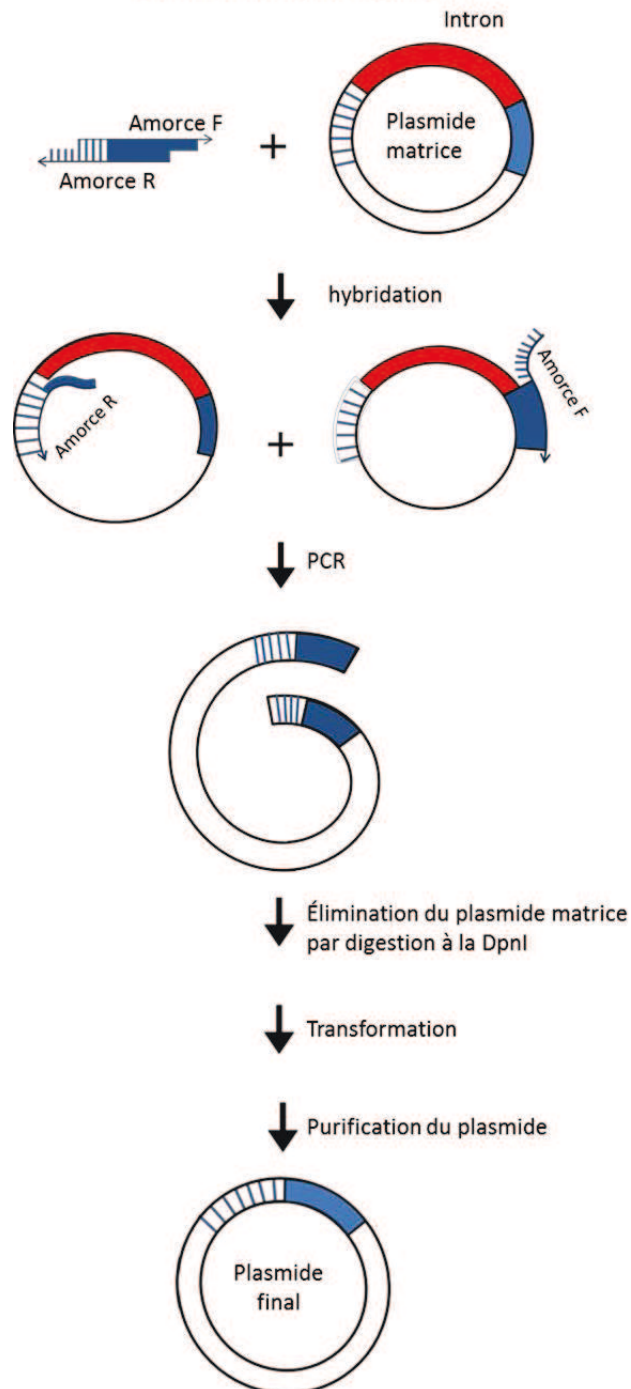
**Figure 30 : Représentation du principe de la RT-PCR.** L'ARNm de BBLF2/3 est utilisé comme matrice pendant la phase de transcription inverse, générant l'ADNc de BBLF2/3. Cet ADNc sert de matrice pour la phase d'amplification par PCR.

- **Délétion de l'intron par PCR :** Le gène BBLF2/3 entier (2 exons + 1 intron) a été amplifié par PCR à partir de la lignée B95-8. Le gène entier est cloné dans le vecteur pFatsBac™HTB. Le plasmide recombinant néo formé sert de matrice pour la délétion par PCR. Cette PCR utilise des amorces ayant leur extrémité 5' complémentaire et se fixant chacune de part et d'autres de l'intron. Pendant la PCR l'ensemble du plasmide, excepté l'intron, est amplifié. Après la PCR l'échantillon est traité par l'enzyme de restriction DpnI pour éliminer le plasmide matrice. Les produits de PCR sont ensuite transformés puis purifiés (Figure 31). La délétion est visualisée sur un gel d'agarose 1%, TAE 0,5% après coloration dans un bain de BET pendant 15min.

### Détails du gène présent sur le plasmide matrice



### Etapes de la délétion par PCR



**Figure 31 : Principe de la délétion par PCR.** Utilisation d'amorce fixant les extrémités des exons 1 et 2, permettant l'amplification complète du plasmide matrice sans l'intron.

### 2.3. Obtention des bacmides

Les étapes nécessaires pour l'obtention des bacmides ont été réalisées en suivant le protocole Bac-to-Bac® d'Invitrogen.

- **Ligation** : Chaque gène d'intérêt a été ligué dans un plasmide pFastBac™HTB
- **Transformation** : Chaque produit de ligation est ensuite transformé dans des bactéries Top10 (Invitrogen) par choc thermique. Les bactéries sont étalées sur des boîtes Luria Broth-agar (LB-agar) contenant l'antibiotique ampicilline (100 µg/mL).
- **Purification des plasmides portant les gènes d'intérêts** : Après pré-culture des bactéries Top10 en présence de l'antibiotique ampicilline (100 µg/mL), Les plasmides sont purifiés à partir des pré-cultures grâce au kit miniprep (Qiagen).
- **Séquençage des clones positifs** : Une PCR contrôle (avec les amorces spécifiques de chaque gène) est réalisée pour identifier les clones positifs avant envoi à l'entreprise Eurofins MWG Operon (Allemagne). Le séquençage est réalisé à partir d'amorces spécifiques au plasmide pFastBac™.

Après la vérification par séquençage que chaque gène est bien orienté et en phase avec l'étiquette de purification N-terminale, l'insertion des gènes d'intérêts dans les bacmides est possible.

- **Transformation des pFastBac™HTB positifs** : Les trois pFastBas™HTB contenant chacun un gène d'intérêt sont transformés dans trois souches *E. coli* DH10Bac™(Invitrogen). La transposition des gènes d'intérêt est permise par les mini-transposon Tn7 présent sur le plasmide et le bacmide.
- **Contrôle de la transposition des gènes d'intérêt** : Le criblage est permis par croissance des bactéries sur des boîtes LB-agar contenant du Bluo-gal et trois antibiotiques (50 µg/ml kanamycine, 7 µg/ml gentamicine, 10 µg/ml tetracycline). La transposition du gène dans le bacmide interrompt l'expression de la protéine LacZ $\alpha$ , les colonies sont alors blanches. Si aucune transposition n'a eu lieu, la protéine LacZ $\alpha$  dégrade le Bluo-gal et la colonie sera bleue.
- **Purification des bacmides** : Pour être sûr de l'absence de faux-positifs, les colonies utilisées pour les pré-cultures (50 µg/ml kanamycine, 7 µg/ml gentamicine, 10 µg/ml tetracycline) sont ré-étalées sur boîte blanc/bleu. Seules les pré-cultures ayant une colonie blanche après ré-étalement sont purifiées. La purification du bacmide suit le protocole décrit dans le manuel Bac-to-Bac®.

- **Vérification de la présence des gènes d'intérêt dans les bacmides :** La vérification est réalisée par PCR en utilisant les amorces pUC/M13 (Invitrogen) s'hybridant sur le bacmide. Les tailles observées sur gel d'agarose après PCR correspondent à la taille des gènes d'intérêt + 2300 pb. Cette taille de 2300 pb correspond à la séquence du plasmide pFastBac™HTB située entre les deux Tn7 et transposée dans le bacmide en même temps que le gène d'intérêt.

Les bacmides positifs après analyse par PCR sont utilisés pour la transfection en cellules d'insectes.

## 2.4. Transfection et mesure du titre viral

Les transfusions et les mesures des titres viraux sont réalisées avec des cellules d'insectes SF21.

- **Transfection :** La transfection est réalisée en plaque 6 puits. Les cellules SF21, entre 80 et 100 % de confluence dans une boîte 75 cm<sup>2</sup>, sont décollées par 20 mL de TC100 supplémenté à l'aide d'une pipette Boy. On dépose ensuite 2 mL de cellules SF21 dans chaque puits de la plaque. La plaque est laissée 1 h à température ambiante pour permettre aux cellules d'adhérer au fond des puits. Pour chaque bacmide trois ratios de transfection sont réalisés :
  - **Ratio 1 :** 5 µl bacmide + 6 µl Cellfectine®II à 1 mg/mL (Invitrogen)
  - **Ratio 2 :** 5 µl bacmide + 10 µl Cellfectine®II
  - **Ratio 3 :** 10 µl bacmide + 6 µl Cellfectine®II

Pour chaque ratio les volumes de bacmides et de Cellfectine®II indiqués sont chacun repris par 100 µL de TC100 non supplémentés puis mélangés afin d'avoir un volume de 200 µL par ratio. Le mélange est laissé 30 à 45 min à température ambiante. La suite de la transfection suit le protocole Bac-to-Bac®. Le mélange de transfection est déposé sur les cellules, l'incubation est réalisée entre **3 et 5 h à 27°C**. Le milieu de transfection est retiré et chaque puits est complété par 2 mL de TC100 supplémenté. La plaque est conservée pendant **6 j à 27°C**.

Les surnageants de chaque ratio sont prélevés, ils constituent les stocks baculovirus P1 de chaque construction.

- **Amplification virale :** Les volumes stocks de P1 sont faibles et ont un titre viral souvent faible, de l'ordre de  $1.10^7$  plaque forming unit (pfu)/mL. L'amplification est réalisée avec des cellules SF21 ayant une confluence de 80% dans une flasque de 175 cm<sup>2</sup>. Pour connaître le volume de virus à utiliser, la formule suivante est utilisée :

$$\text{Inoculum (mL)} = \frac{(\text{MOI (pfu/Cell)} \times \text{nombre de cellules})}{\text{Titre (pfu/mL)}} \quad (1)$$

La MOI ou *multiplicity of infection* représente le ratio entre le nombre d'agents infectieux (e.g. le baculovirus) vis à vis de la cible (e.g. la cellule d'insecte) et permet ainsi de reproduire la même infection virale à chaque production protéique. Pour l'amplification virale la MOI est fixée à 0,05 et le titre n'étant pas encore connu la valeur de  $2.10^7$  pfu/mL est fixée arbitrairement pour obtenir la valeur d'inoculum. Le volume d'inoculum obtenu par la formule (1) est complété à 3 mL par du TC100 supplémenté pour recouvrir uniformément la surface cellulaire de la boîte 175 cm<sup>2</sup>. L'incubation de l'inoculum avec les cellules est effectuée 1h à 27°C. Après l'heure d'incubation 25 mL de TC100 supplémenté sont ajoutés par flasques 175 cm<sup>2</sup> et l'infection est poursuivie à 27°C pendant 7 j. En fin d'infection les nouveaux stocks de baculovirus, nommés P2, sont récupérés par centrifugation. Les stocks P2 sont conservés à 4°C.

- **Mesure du titre viral par « plaque assay »** : La mesure du titre viral est effectuée sur les lots de baculovirus P2. Pour évaluer le titre viral un « plaque assay » est réalisé. Les cellules SF21, entre 80 et 100 % de confluence dans une boîte 75 cm<sup>2</sup>, sont décollées par 20 mL de TC100 supplémenté à l'aide d'une pipette Boy. On dépose ensuite 2 mL de cellules SF21 dans chaque puits d'une plaque 6 puits. Une dilution en cascade du virus est réalisée de  $10^{-1}$  à  $10^{-9}$ , seules les dilutions de  $10^{-5}$  à  $10^{-9}$  sont utilisées pour le « plaque assay ». Pour les 5 dilutions, 200 µL d'inoculum sont déposés sur les cellules SF21 et incubé 1 h à température ambiante. Après l'heure d'incubation, les puits sont recouverts par de l'agarose 2% (Gibco). Après polymérisation de l'agarose (15 min), 1 mL de TC100 supplémenté est ajouté dans chaque puits, la plaque est stockée 4 j à 27°C. Après les 4 j d'infection, 1 mL de rouge neutre (0,1 mg/mL) est ajouté dans chaque puits. Le rouge neutre permet de visualiser les plages de lyses. Il colore les cellules vivantes en rouge, mais pas les cellules mortes. Les plages de lyses restent blanches alors que le reste du puits possède une teinte rouge. Le calcul du titre viral est obtenu grâce à l'équation suivante :

$$\text{Titre (pfu/mL)} = \text{nbre de plaques} \times \text{facteur de dilution} \times \frac{1}{\text{mL inoculum par puits}} \quad (2)$$



La connaissance du titre viral permet la reproductibilité des productions protéiques. Les « plaques assay » sont toujours réalisés en duplicata.

### **3. Test d'expression en « Time course »**

Les conditions optimales d'infection sont déterminées par deux paramètres : la durée de l'infection et la MOI. Pour estimer ces paramètres, un test d'expression en « Time course » a été réalisé à 24, 48, 72 et 96h, et pour les MOI 1, 5, 10, 15 et 20. Les différentes conditions d'infections sont réalisées sur des cellules Hi5 dans les puits d'une plaque 6 puits. Sur la plaque, 5 puits sont dédiés aux 5 MOI et le dernier sert de contrôle cellulaire non infecté. Chaque puits est infecté par le volume d'inoculum obtenu à partir des valeurs de MOI et du titre de chaque virus selon l'équation (2). L'inoculum est complété par du milieu TC100 supplémenté pour un volume final de 0,8 mL. A la fin de chaque temps, le même protocole est appliqué pour récupérer les cellules. Elles sont décollées par aspiration-refoulement à l'aide d'un pipetteMan, puis lysées par deux cycles successifs de choc thermique : 3 min azote liquide et 10 min à 37°C. Les cellules sont ensuite centrifugées 30 min à 13400 rpm (centrifugeuse de paillasse MiniSpin Eppendorf). Le surnageant correspond la fraction soluble et le culot la fraction brute. Les échantillons sont conservés à -80°C jusqu'à la fin de l'expérience de « Time course ». Les résultats sont observés sur un gel SDS-PAGE 8% coloré au bleu de coomassie (15 min).

#### **3.1. Productions des protéines recombinantes en cellules adhérentes**

L'infection est réalisée sur des flasques de 300 cm<sup>2</sup> avec des cellules Hi5 à 80 % de confluence. Les cellules sont infectées à une MOI 10, le volume d'inoculum à utiliser par flasques 300 cm<sup>2</sup> et déduit à partir de l'équation (2). L'inoculum est complété à 4mL par du TC100 supplémenté, est laissé 1h sur le tapis cellulaire pour que les virus s'adsorbent aux cellules. Après l'heure d'adsorption, 40 mL de milieu TC100 sont ajoutés dans chaque flasque. Pendant toute la durée de l'infection les flasques sont stockées à 27°C. Les cellules sont récoltées mécaniquement à l'aide d'un grattoir de cellules 72h post-infection. Elles sont centrifugées à température ambiante pendant 30 min à 2300 rpm (centrifugeuse de paillasse 5840 Eppendorf). Le culot est resuspendu dans du PBS 1x (tampon phosphate salin : selon le protocole de « Molecular Cloning: A Laboratory Manual »[165]) pour laver les cellules. Elles sont centrifugées une seconde fois, le surnageant est retiré et les culots humides sont congelés à -80°C.

En utilisant des baculovirus codant pour chaque protéine du complexe H-P d'EBV, il est possible de réaliser diverses co-infections permettant la formation des complexes binaires ou du complexe ternaire *in vivo*. Le protocole pour les co-infections est identique à l'infection simple, la

seule divergence repose sur la MOI à utiliser. Pour estimer si la charge virale influe sur la production protéique nous avons mis en place deux types de MOI. Une co-infection par MOI stricte consiste à utiliser chaque virus à sa MOI standard déterminée pendant les tests de « Time course ». Une co-infection par MOI cumulée, consiste à utiliser une MOI plus faible pour chaque virus afin que la somme des MOI corresponde à la MOI optimale pour l'infection des cellules d'insectes. Par exemple, nous utilisons une MOI de 10 pour les infections simple, ainsi pour la co-infection par MOI cumulée nous estimons que la MOI optimale est de 10. Pour une infection double nous utilisons chacun des virus à une MOI de 5 pour que la somme de ces MOI soit égale à la MOI optimale.

## **4. Clonage MultiBac et production de protéines en cellules en suspension**

La technique de clonage MultiBac a été développée par la plateforme EEF (Eucaryote Expression Facility) de l'EMBL, sous la direction du Dr BERGER Imre [166]. Cette technique consiste à créer un baculovirus exprimant plusieurs gènes d'intérêt simultanément. Ce baculovirus a pour but de faciliter la formation du complexe, qui peut être guidée par la présence des trois protéines et/ou la présence de protéines chaperonnes. Ces deux critères sont très importants pour la formation du complexe entier et soluble.

Le complexe H-P étant composé de trois protéines plusieurs constructions sont réalisables. Le but principal de la thèse étant l'étude du complexe entier, nous avons décidé d'ajouter l'étiquette de purification sur une seule des trois protéines. L'utilisation d'une seule protéine du complexe portant l'étiquette de purification permet d'observer dès la phase de purification si les protéines sont capables de former le complexe *in vivo*. Si le complexe est formé, on observera les trois protéines dans les fractions d'élution de la colonne d'affinité. Dans le cas contraire seule la protéine portant l'étiquette de purification sera détectable lors de l'élution.

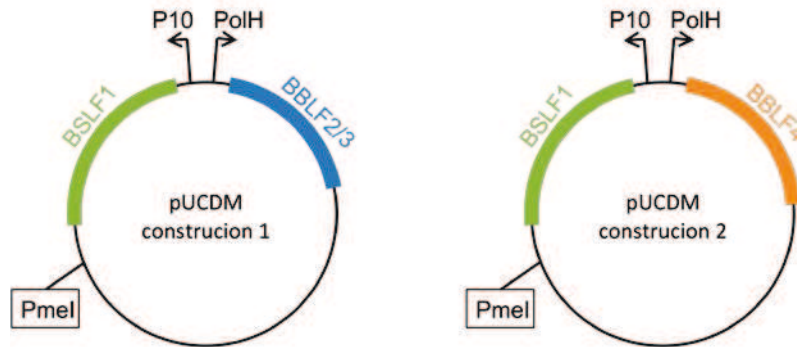
### **4.1. Constructions utilisées pour le clonage MultiBac**

Deux constructions différentes ont été réalisées avec les gènes BSLF1, BBLF2/3 et BBLF4. Le choix de la première construction est basé sur l'analogie des résultats obtenus par Schreiner & al (2009) [167] sur le complexe H-P d'HSV-1. Ils ont montré que le complexe H-P d'HSV-1 est facilement purifiable par affinité sur colonne de nickel lorsque l'étiquette hexahistidine est portée par l'hélicase. Pour la première construction réalisée, c'est la protéine BBLF4 qui porte l'étiquette hexahistidine en N-terminal (construction 1 = BSLF1 + BBLF2/3 + BBLF4his). Le choix de la seconde construction est basé sur les résultats que nous avons obtenus lors de la production individuelle de la sous unité

accessoire. Cette protéine se purifiant très facilement par affinité sur colonne de nickel nous avons décidé de lui faire porter l'étiquette hexahistidine (construction 2 = BSLF1 + BBLF4 + BBLF2/3his).

Le clonage Multibac est différent de celui réalisé en cellules adhérentes avec le protocole Bac-to Bac® (Invitrogen). Pour le MultiBac le clonage se déroule en deux temps. Tout d'abord les deux gènes codant pour les protéines non étiquetées sont insérés dans un plasmide donneur, nommé pUCDM. Pour les protéines portant l'étiquette de purification, nous réutilisons les plasmides pFastBac™HTB générés lors du clonage avec le protocole Bac-to-Bac®. La seconde étape consiste à fusionner le plasmide recombinant pUCDM avec le plasmide recombinant pFastBac™HTB, pour former qu'un seul et grand plasmide qui sera utilisé pour la transformation dans des bactéries DH10EMBacY contenant le bacmide du baculovirus. Les étapes de clonage sont identiques pour les deux constructions, seul celui de la construction 1 est détaillé dans les paragraphes suivants.

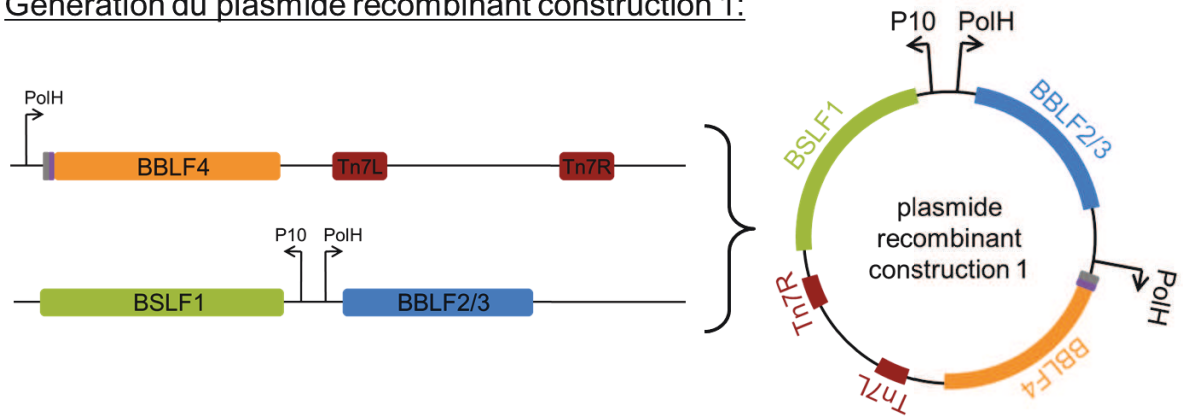
- **Génération du plasmide donneur pUCDM :** Le plasmide pUCDM contient deux sites de clonage multiple ou MCS. Dans un premier temps nous avons amplifié par PCR le gène BBLF2/3. Le gène BBLF2/3 et le pUCDM sont digérés par les enzymes de restrictions RsrII et XbaI. Après digestion le gène et le plasmide sont ligués selon le protocole du kit de ligation « Rapid DNA Ligation »(Thermo Scientific). Les produits de ligation sont transformés par choc thermique dans les bactéries 8B. Ces bactéries permettent de reconnaître l'origine de réplication R6Kgamma portée par le plasmide donneur pUCDM. Les bactéries sont alors étalées sur boîtes LB-agar (chloramphénicol 34 µg/mL). Après pré-culture des bactéries 8B dans du milieu LB (chloramphénicol 34 µg/mL), le plasmide est purifié à partir des pré-cultures grâce au kit miniprep (Qiagen). La présence du gène dans le plasmide est vérifiée par double digestion avec les enzymes RsrII et XbaI. Une fois le BBLF2/3 inséré dans le MCS2, le gène BSLF1 peut être inséré dans le MCS1 du plasmide pUCDM. Le clonage est identique à celui réalisé pour le gène BBLF2/3 dans le MCS2 sauf que le couple d'enzyme de restriction utilisé est XhoI/NsiI. Les résultats intermédiaires des deux constructions sont représentés en Figure 32.



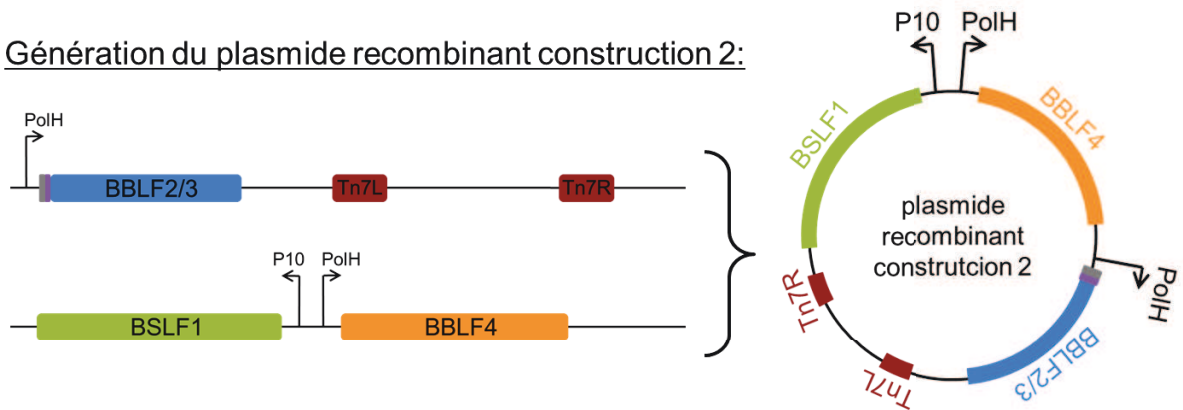
**Figure 32 : Représentation schématique des vecteurs donneurs pUCDM des constructions 1 et 2.** PmeI site de restriction enzymatique. P10 et PolH promoteurs tardifs.

- **Génération du plasmide hybride pUCDM/pFastBac™HTB :** Le plasmide recombinant est généré par la fusion des plasmides pUCDM (BSLF1 + BBLF2/3) et pFastBac™HTB (BBLF4). Les plasmides pUCDM (BSLF1 + BBLF2/3) et pFastBac™HTB (BBLF4) sont respectivement digérés par les enzymes de restriction PmeI et SnaBI, générant des bouts francs sur chacun des plasmides. Les deux plasmides sont ligués en un seul plasmide recombinant à l'aide du kit « Rapid DNA Ligation Kit » (Thermo Scientific). Les produits de ligation sont transformés par choc thermique dans les bactéries 8B. Les bactéries sont alors étalées sur boîtes LB-agar (chloramphénicol 34 µg/mL et ampicilline 100 µg/mL). Après pré-culture des bactéries 8B dans du milieu LB (chloramphénicol 34 µg/mL et ampicilline 100µg/mL), le plasmide est purifié à partir des pré-cultures grâce au kit miniprep (Qiagen). La vérification du plasmide recombinant par hybridation des plasmides pUCDM (BSLF1 + BBLF2/3) et pFastBac™HTB (BBLF4) est réalisée par digestion avec l'enzyme XbaI. Le profil de digestion est différent suivant si on a la présence du plasmide recombinant ou des plasmides intermédiaires pUCDM et pFastBac™HTB. Les détails des plasmides recombinants des constructions 1 et 2 sont représentés sur la Figure 33.

### Génération du plasmide recombinant construction 1:



### Génération du plasmide recombinant construction 2:



**Figure 33 : Détails des étapes permettant la formation des plasmides recombinants construction 1 et 2.** Les plasmides pUCDM et pFastBac™HTB sont linéarisés, respectivement, par les enzymes de restriction PmeI et SnaBI. Les plasmides digérés sont fusionnés pour former les plasmides recombinants construction 1 et 2 finaux. Rectangle violet : site de clivage à la protéase TEV, Rectangle gris : étiquette hexahistidine N-terminale.

- **Transposition du plasmide recombinant dans le bacmide :** Pour la transposition du plasmide recombinant ; une souche de bactéries DH10 modifiée est utilisée. La souche de bactéries DH10EMBacY contient le bacmide du génome baculovirus (AcMNPV), auquel a été ajouté un gène rapporteur codant pour la YFP (protéine fluorescente jaune) sous le contrôle du promoteur PolH. Il y a aussi la présence du plasmide helper permettant la transposition à l'aide des mini-transposons Tn7. Les trois gènes situés sur le plasmide recombinant sont localisés entre les deux mini-transposons Tn7. La transformation du plasmide recombinant dans les bactéries DH10EMBacY est permise par choc thermique. Seule la partie du plasmide recombinant située entre les Tn7 est transposée sur le bacmide. L'isolation des bacmides recombinants et leur purification suivent les mêmes protocoles que ceux décrits dans le manuel Bac-to-Bac.

## 4.2. Transfection et amplification virale

La transfection est réalisée en plaque 6 puits avec des cellules SF21 cultivées en milieu HyClone SFM4Isect™. Chaque puits contient 3mL de cellules à une densité comprise entre 0,5 et  $1.10^6$  cellules/mL. Le calcul de la densité cellulaire est réalisé à l'aide d'une cellule de comptage Neubauer (Dutscher). Deux clones pour chaque construction sont transfectés en parallèle. Comme contrôles, un puits est chargé de cellules non infectées et un autre avec du milieu seul. La transfection est réalisée en suivant le protocole de l'agent transfectant FuGENE® (Promega). Le surnageant de chaque puits (3 mL) est récupéré entre 38 et 60h post-transfection. Il contient le baculovirus recombinant initial, aussi appelé Vo. Chaque puits est complété à 3 mL par du milieu frais pour poursuivre l'infection pendant 2 jours. Au bout des 2 jours les cellules sont récoltées et utilisées pour la mesure de fluorescence (YFP) et l'analyse par Western Blot avec un anticorps anti étiquette hexahistidine.

L'amplification virale est effectuée à partir du virus Vo. Une culture cellulaire de 25 mL à  $0,5.10^6$  cellules/mL est infectée par 3 mL de Vo. La densité cellulaire de la culture en suspension est maintenue à  $0,5.10^6$  cellules/mL jusqu'à arrêt de la division cellulaire. Le maintien s'effectue par comptage cellulaire toutes les 24 h. Généralement 24 h post-infection le nombre de cellules a doublé, la densité cellulaire est maintenue en augmentant le volume final à 50 mL. 48 h post-transfection, les cellules se divisent presque plus, elles sont maintenues à  $0,5.10^6$  cellules/mL en éliminant l'excès de cellules et en maintenant le volume final à 50 mL. L'arrêt de la division constitue le temps DPA (Day Post Arrest). A partir de la DPA,  $1.10^6$  cellules sont prélevées toutes les 24 h pour mesurer la fluorescence de la YFP. Les cellules sont culotées 48h après la DPA pour récolter le surnageant contenant le virus amplifié, V1. Les cellules sont resuspendues dans 25 mL de milieu frais permettant de poursuivre les mesures de YFP jusqu'à apparition d'un plateau. Le protocole utilisé pour les mesures de YFP est spécifique de l'appareil utilisé au laboratoire et de la protéine YFP produite par la plateforme EEF et n'a donc que très peu d'intérêt à être détaillé. Les cellules sont récupérées et la production protéique est analysée sur gel SDS-PAGE.

## 4.3. Production des protéines recombinantes en cellules en suspension :

La production protéique en cellules en suspension ne nécessite pas la connaissance du titre viral. Chaque production protéique est infectée par le même volume de V1. Une culture cellulaire de 200 mL à  $0,6.10^6$  cellules/mL (erlenmeyer de 2 L) est infectée par 200  $\mu$ L de virus V1. 24 h post-infection, la densité cellulaire est maintenue à  $0,6.10^6$  cellules/mL, en augmentant le volume de



culture à 400 mL. Généralement le temps 48 h correspond à la DPA, la division cellulaire est presque nulle, la densité cellulaire est surveillée afin qu'elle n'excède pas  $1,2 \cdot 10^6$  cellules/mL. A partir de la DPA,  $1 \cdot 10^6$  cellules/mL sont récoltées toutes les 24 h et la fluorescence de la YFP est mesurée jusqu'à obtention d'un plateau. Les cellules sont alors centrifugées 15 min à 2300 rpm (centrifugeuse de paillasse 5840 Eppendorf), le culot cellulaire est resuspendu dans du PBS 1x et centrifugé de nouveau. Le surnageant est retiré, le culot de cellule est congelé à  $-80^{\circ}\text{C}$ . Une production protéique de 400 mL de cellules permet d'obtenir un culot de 3 à 4 g de cellules.

## CHAPITRE 3 :

# Expériences biochimiques

# I. Tampons et conditions de purification

## 1. Recherche de tampon de solubilisation en plaque 96 puits

L'étude en plaque 96 puits permet de tester 96 conditions de tampons de solubilisation en une seule expérience. Six conditions de pH, quatre conditions en sel et huit conditions d'additifs peuvent être testées pour améliorer la solubilisation des protéines formant le complexe (Tableau 6). Cette approche a été utilisée pour le complexe produit en cellules en suspension. Elle a été menée en parallèle sur les constructions 1 (BSLF1 + BBLF2/3 + BBLF4his) et 2 (BSLF1 + BBLF4 + BBLF2/3his).

**Tableau 6 : Détails des 96 conditions de solubilité** testées pour améliorer la solubilité des protéines du complexe H-P.

	1	2	3	4	5	6	7	8	9	10	11	12
<b>A</b>	Acide citric pH 5,5 NaCl 150mM	Acide citric pH 5,5 NaCl 500mM	Acide citric pH 5,5 NaCl 2M	Acide citric pH 5,5 SulfAmo 500mM	Acide citric pH 5,5 NaCl 500mM 10% glycerol	Acide citric pH 5,5 NaCl 500mM 25mM Glucose	Acide citric pH 5,5 NaCl 500mM 0,1% tritonx100	Acide citric pH 5,5 NaCl 500mM 250mM NDSB	Acide citric pH 5,5 NaCl 500mM 10mM Glycine	Acide citric pH 5,5 NaCl 500mM 50mM Imidazole	Acide citric pH 5,5 NaCl 500mM 2mM MgCl2	Acide citric pH 5,5 NaCl 500mM 0,5M UREA
<b>B</b>	MES pH 6 NaCl 150mM	MES pH 6 NaCl 500mM	MES pH 6 NaCl 2M	MES pH 6 SulfAmo 500mM	MES pH 6 NaCl 500mM 10% glycerol	MES pH 6 NaCl 500mM 25mM Glucose	MES pH 6 NaCl 500mM 0,1% tritonx100	MES pH 6 NaCl 500mM 250mM NDSB	MES pH 6 NaCl 500mM 10mM Glycine	MES pH 6 NaCl 500mM 50mM Imidazole	MES pH 6 NaCl 500mM 2mM MgCl2	MES pH 6 NaCl 500mM 0,5M UREA
<b>C</b>	MES pH 6,5 NaCl 150mM	MES pH 6,5 NaCl 500mM	MES pH 6,5 NaCl 2M	MES pH 6,5 SulfAmo 500mM	MES pH 6,5 NaCl 500mM 10% glycerol	MES pH 6,5 NaCl 500mM 25mM Glucose	MES pH 6,5 NaCl 500mM 0,1% tritonx100	MES pH 6,5 NaCl 500mM 250mM NDSB	MES pH 6,5 NaCl 500mM 10mM Glycine	MES pH 6,5 NaCl 500mM 50mM Imidazole	MES pH 6,5 NaCl 500mM 2mM MgCl2	MES pH 6,5 NaCl 500mM 0,5M UREA
<b>D</b>	HEPES pH 7 NaCl 150mM	HEPES pH 7 NaCl 500mM	HEPES pH 7 NaCl 2M	HEPES pH 7 SulfAmo 500mM	HEPES pH 7 NaCl 500mM 10% glycerol	HEPES pH 7 NaCl 500mM 25mM Glucose	HEPES pH 7 NaCl 500mM 0,1% tritonx100	HEPES pH 7 NaCl 500mM 250mM NDSB	HEPES pH 7 NaCl 500mM 10mM Glycine	HEPES pH 7 NaCl 500mM 50mM Imidazole	HEPES pH 7 NaCl 500mM 2mM MgCl2	HEPES pH 7 NaCl 500mM 0,5M UREA
<b>E</b>	HEPES pH 7,5 NaCl 150mM	HEPES pH 7,5 NaCl 500mM	HEPES pH 7,5 NaCl 2M	HEPES pH 7,5 SulfAmo 500mM	HEPES pH 7,5 NaCl 500mM 10% glycerol	HEPES pH 7,5 NaCl 500mM 25mM Glucose	HEPES pH 7,5 NaCl 500mM 0,1% tritonx100	HEPES pH 7,5 NaCl 500mM 250mM NDSB	HEPES pH 7,5 NaCl 500mM 10mM Glycine	HEPES pH 7,5 NaCl 500mM 50mM Imidazole	HEPES pH 7,5 NaCl 500mM 2mM MgCl2	HEPES pH 7,5 NaCl 500mM 0,5M UREA
<b>F</b>	Tris pH 8 NaCl 150mM	Tris pH 8 NaCl 500mM	Tris pH 8 NaCl 2M	Tris pH 8 SulfAmo 500mM	Tris pH 8 NaCl 500mM 10% glycerol	Tris pH 8 NaCl 500mM 25mM Glucose	Tris pH 8 NaCl 500mM 0,1% tritonx100	Tris pH 8 NaCl 500mM 250mM NDSB	Tris pH 8 NaCl 500mM 10mM Glycine	Tris pH 8 NaCl 500mM 50mM Imidazole	Tris pH 8 NaCl 500mM 2mM MgCl2	Tris pH 8 NaCl 500mM 0,5M UREA
<b>G</b>	Tris pH 8,5 NaCl 150mM	Tris pH 8,5 NaCl 500mM	Tris pH 8,5 NaCl 2M	Tris pH 8,5 SulfAmo 500mM	Tris pH 8,5 NaCl 500mM 10% glycerol	Tris pH 8,5 NaCl 500mM 25mM Glucose	Tris pH 8,5 NaCl 500mM 0,1% tritonx100	Tris pH 8,5 NaCl 500mM 250mM NDSB	Tris pH 8,5 NaCl 500mM 10mM Glycine	Tris pH 8,5 NaCl 500mM 50mM Imidazole	Tris pH 8,5 NaCl 500mM 2mM MgCl2	Tris pH 8,5 NaCl 500mM 0,5M UREA
<b>H</b>	Tris pH 9 NaCl 150mM	Tris pH 9 NaCl 500mM	Tris pH 9 NaCl 2M	Tris pH 9 SulfAmo 500mM	Tris pH 9 NaCl 500mM 10% glycerol	Tris pH 9 NaCl 500mM 25mM Glucose	Tris pH 9 NaCl 500mM 0,1% tritonx100	Tris pH 9 NaCl 500mM 250mM NDSB	Tris pH 9 NaCl 500mM 10mM Glycine	Tris pH 9 NaCl 500mM 50mM Imidazole	Tris pH 9 NaCl 500mM 2mM MgCl2	Tris pH 9 NaCl 500mM 0,5M UREA

Il faut un minimum de 4 millions de cellules par puits d'une plaque PCR Multi (Dominique Dutcher). Une production protéique à partir de 200 mL de cellules à  $1.10^6$  cellules/mL, permet d'obtenir en fin d'infection un volume de 400 mL de cellules à  $1.10^6$  cellules/mL. Les cellules sont culottées à température ambiante pendant 15 min à 2300 rpm (centrifugeuse de paillasse 5840 Eppendorf) puis resuspendues dans 10 mL de PBS 1x. Elles sont aliquotées par fraction de 100 µL dans chaque puits. La plaque est centrifugée à température ambiante pendant 15 min à 2300 rpm

(centrifugeuse de paillasse 5840 Eppendorf), le PBS 1x est retiré et chaque puits est resuspendu par 140 µL de tampon de lyse respectif. La lyse cellulaire est effectuée par 4 cycles successifs de choc thermique (2 min Azote liquide, 5 min à 37°C). Après lyse, les cellules sont centrifugées à 4000 rpm pendant 1 h et à 10°C (centrifugeuse de paillasse 5840 Eppendorf). Pendant la centrifugation, on dépose 100 µL de résine de type G25 dans plaque 96 puits de purification munie de filtre (Chromabond® Multi 96, Macherey-Nagel). Après centrifugation, les surnageants sont prélevés et déposés sur la plaque contenant la résine G25. La résine G25 permet de retenir les gros agrégats et pouvoir juger plus facilement du pouvoir de solubilité de chaque tampon. Les surnageants sont récupérés en sortie de résine G25 dans une plaque 96 puits à fond conique (Greiner).

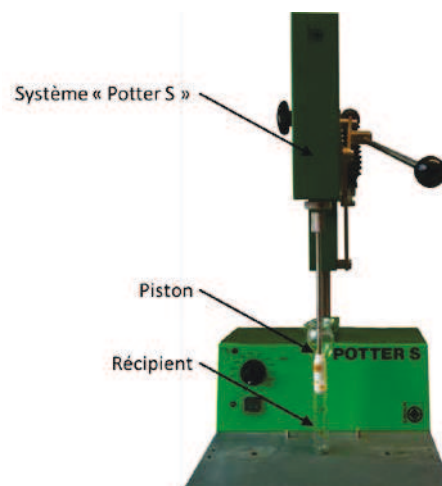
L'analyse des surnageants est effectuée par Dot Blot sur une membrane PVDF (Millipore). Dépôt de 3 µL de surnageant par conditions sur la membrane. Elle est ensuite saturée pendant 1 h avec une solution de PBS (1x)/0,1% Tween20/5% Lait écrémé en poudre. Après 1 h, elle est rapidement rincée par une solution de PBS (1x)/0,1% Tween20. Un anticorps monoclonal anti-polyhistidine couplé à la peroxydase (Sigma Aldrich) est dilué au 1/2500 dans du PBS/Tween/Lait et incubé sur la membrane pendant 1h. La membrane est lavée rapidement par du PBS/Tween puis à l'eau avant d'être révélée par chimioluminescence sur film photo (procédure identique à celle décrite pour le Western Blot).

## 2. Lyse cellulaire

Diverses techniques de lyse sont disponibles pour les cellules d'insectes. Le choix du type de lyse doit être adapté à l'organisme de surexpression pour éviter une dégradation prématurée des protéines d'intérêts. Des trois techniques décrites, ci-dessous, seules les approches par choc thermique et ultrason ont été utilisées pour la lyse des culots cellulaires produits en cellules adhérentes. Quelle que soit la technique de lyse utilisée, le culot cellulaire est repris par un volume de tampon de lyse équivalent à 10 fois son poids.

- **Lyse par choc thermique** : Le culot cellulaire est resuspendu dans un tampon de lyse supplémenté par le cocktail antiprotéase « Complete » (Roche). La solution protéique est répartie dans des tubes falcons de 50 mL. La lyse thermique est alors effectuée par 3 à 5 cycles successifs : Azote liquide/bain marie 37°C. Les temps de chaque cycle sont dépendant des volumes de culot à lyser. Pour ce type de lyse il faut bien veiller à ce que l'ensemble du culot soit congelé après l'étape dans l'azote liquide et que l'ensemble du culot est bien décongelé après passage dans le bain marie.

- **Lyse par ultrason (Sonicateur Labsonic® P, Sartorius)** : Le culot cellulaire est resuspendu dans du tampon de lyse supplémenté. La lyse par ultrason est réalisée dans un bécher métallique plongé dans la glace pour éviter tout échauffement de la solution protéique. La lyse est composée de 2 cycles successifs d'ultrason (2min30, 0,5 pulsation par seconde, 60% d'intensité). Le choix de la taille de la sonde et du bécher métallique dépend du volume de cellules à lyser, mais les paramètres de lyse par ultrason restent les mêmes.
- **Lyse mécanique, le « potter »** : Le culot cellulaire est resuspendu dans le tampon de lyse supplémenté. La solution protéique est lysée par fraction de 15 mL dans un potter (B.Braun). Chaque fraction de 15 mL est lysée par une succession d'aller-retour du piston. La partie « piston » du potter est fixée au système « Potter S » (B.Braun) permettant une rotation à 150 rpm et la partie « récipient » est plongée dans la glace (Figure 34).



**Figure 34** : Détails de l'équipement utilisé pour la lyse mécanique avec le « potter ». La vitesse de rotation du piston est maintenue à 150 rpm par le système potter S. Le récipient contient la fraction protéique à lyser.

### 3. Purification par affinité sur colonne de nickel

Le protocole de purification est similaire pour les productions protéiques en cellules adhérentes ou en suspension. Les paramètres communs aux deux modes de productions sont écrits en noir, ceux spécifiques aux cellules adhérentes en bleu et en rouge pour les cellules en suspension.

**Tableau 7 : Composition des tampons utilisés pour la purification protéique.** Bleu : tampons spécifiques de BBLF2/3 produite en cellules adhérentes. Rouge : tampons spécifiques du complexe produit en cellules en suspension.

matériel biologique	BBLF2/3	BSLF1 + BBLF2/3 + BBLF4his
Tampon de lyse et d'équilibration	20 mM Tris HCl pH 7,9 ; 500 mM NaCl ; 10 mM imidazole	50 mM Hépès pH 7,6 ; 500 mM NaCl ; 10 mM imidazole ; 0,01% Triton X-100 ; 10% glycérol
Tampon de lavage 1	20 mM Tris HCl pH 7,9 ; 500 mM NaCl ; 10 mM imidazole	50 mM Hépès pH 7,6 ; 500 mM NaCl ; 10 mM imidazole ; 0,01% Triton X-100 ; 10% glycérol
Tampon de lavage 2	20 mM Tris pH 7,9 ; 1 M NaCl ; 20 mM imidazole	20 mM Hépès pH 7,6 ; 1 M NaCl ; 15 mM imidazole ; 10% glycérol
Tampon d'élution	Tris 20 mM pH 7,9 ; 500 mM NaCl ; 500 mM imidazole	20 mM Hépès pH 7,6 ; 500 mM NaCl ; 500 mM imidazole ; 10% glycérol
Tampon de dessalage	Tris 20 mM pH 7,9 ; 250 mM NaCl ; 10% glycérol ; 15 mM imidazole	
Tampon de chromatographie d'exclusion	Tris 20 mM pH 7,9 ; 250 mM NaCl ; 10% glycérol	20 mM Hépès pH 7,6 ; 500 mM NaCl ; 50 mM R/E ; 10% glycérol

Le culot cellulaire est resuspendu par environ 10x son poids dans du tampon de lyse complété. En fonction des résultats obtenus pour BBLF2/3 ou le complexe, trois types de lyse ont été utilisés : choc thermique, sonication et potter.

Après lyse la solution protéique est centrifugée à 20000 rpm pendant 30 min à 4°C. Le surnageant est filtré sur 0,45 µm (Millipore) avant dépôt sur la colonne d'affinité. La colonne nickel, 1ml tassé, est formée d'une résine His Select (Sigma Aldrich). Le surnageant traverse la colonne par gravité permettant la fixation de l'étiquette hexahistidine sur les ions Ni<sup>2+</sup>.

Après fixation de la protéine sur la colonne, elle est lavée par 10x son volume, avec deux tampons successifs. Le premier lavage est effectué par le tampon de lyse et le deuxième par le tampon Lavage 2. Le second tampon possède une force ionique élevée, permettant l'élution des protéines contaminantes qui interagissent de manière non spécifique sur la colonne.

La protéine est éluee de la colonne par le tampon d'élution. Ce tampon contient une concentration élevée d'imidazole. Cette molécule possède une forte affinité pour le nickel et entre en compétition avec l'étiquette polyhistidine portée par la protéine et permet son élution. L'élution est effectuée par 10x le volume de la colonne. La protéine éluee est récupérée par fraction de 1 mL



ou en **une seule fraction de 10 mL**.

**Pour les productions par cellules en suspension**, le volume d'élution de 10 mL est réduit jusqu'à 800 µL environ par centrifugation sur des concentrateurs Amicon Ultra 4 10 k (Millipore). Un gel SDS-PAGE est réalisé avec les fractions brutes, solubles, de lavages et d'élution obtenues au cours de la purification par affinité. La fraction d'élution concentrée ( $\pm$  800µL) est injectée sur une colonne de chromatographie d'exclusion de taille SuperDex200 10/300 (GE-Healthcare). Les fractions d'intérêt de la chromatographie d'exclusion sont déposées sur gel SDS-PAGE 8 % et colorées au bleu de coomassie (15 min). Les fractions correspondant aux pics du chromatogramme sont concentrées par centrifugation sur Amicon, dosées en mesurant l'absorbance à 280 nm et conservées 4°C jusqu'à utilisation.

**Pour les productions en cellules adhérentes**, les fractions d'élution de la colonne d'affinité sont déposées sur gel SDS-PAGE. Après coloration au bleu de coomassie, les fractions contenant le plus de protéines sont rassemblées. On procède alors au changement de tampon (20 mM Tris HCl pH 7,5 ; 250 mM NaCl ; 10% glycérol) de la protéine sur une colonne de dessalage Econo-Pac<sup>®</sup> 10DG (Biorad). L'étiquette polyhistidine est clivée par la protéase TEV (10 µL de protéase (1 mg/mL) par mg de protéine) sur la nuit à température ambiante. Le lendemain, la protéine clivée est passée sur une seconde colonne de nickel, utilisation  $\pm$  250 µL de résine tassée en fonction de la quantité de protéine. Cette seconde colonne de nickel permet de retenir la protéase (présence d'une étiquette polyhistidine) sur la colonne. La protéine d'intérêt n'a plus d'étiquette, elle n'est donc pas retenue par la colonne. La protéine est ensuite concentrée sur Amicon en vue d'une injection sur chromatographie d'exclusion de taille SuperDex200 10/300 (GE-Healthcare). Les fractions d'intérêt de la chromatographie d'exclusion sont déposées sur gel SDS-PAGE 10 % et colorées au bleu de coomassie (15 min). Les fractions correspondant aux pics du chromatogramme sont concentrées par centrifugation sur Amicon, dosées en mesurant l'absorbance à 280 nm et conservées 4°C jusqu'à utilisation.

#### **4. Chromatographie d'exclusion de taille**

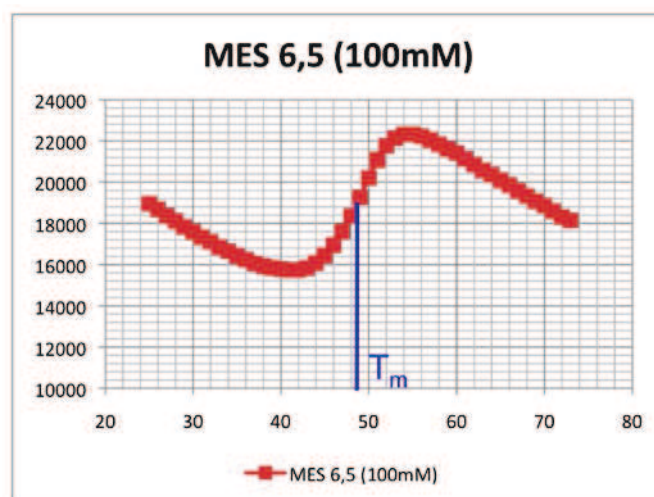
Pour isoler le complexe et augmenter sa pureté, une étape supplémentaire de purification est réalisée par chromatographie d'exclusion de taille. Les échantillons protéiques concentrés sont centrifugés 10 min à 13000 rpm (centrifugeuse de paillasse MiniSpin Eppendorf) puis injectés sur une colonne de chromatographie d'exclusion de taille SuperDex200 10/300 (GE-Healthcare). Une colonne de type S200 permet la séparation des protéines de grandes tailles, la taille théorique du complexe est de 260 kDa. La colonne est, au préalable, équilibrée avec le tampon chromatographie d'exclusion (Tableau 7), grâce au système Biologic DuoFlow (Bio-Rad). Ce système permet de mesurer

l'absorbance à 280 nm sur toute la durée de la chromatographie et de collecter la solution protéique par fractions de 500 µL. Les fractions correspondant aux pics d'absorbance du chromatogramme sont analysées sur gel SDS-PAGE 8 ou 10 %. Les fractions d'intérêts sont concentrées par centrifugation sur concentrateur Amicon (4000 rpm à 15°C). La fraction finale est dosée à 280 nm et conservée à 4°C. Pour une concentration de 1 mg/mL l'absorbance du complexe est de 1,2 et de 1,5 pour la protéine BBLF2/3.

## 5. Optimisation du tampon par Thermofluor

La structure d'une protéine *in vivo* est dépendante d'un grand nombre de paramètres, présence d'autres protéines, d'ADN, niveau d'encombrement ... Lors d'une purification il faut trouver les conditions *in vitro* permettant de bien replier la protéine malgré l'absence des paramètres *in vivo*. Le Thermofluor permet d'évaluer la stabilité de la protéine d'intérêt dans divers tampons. De plus cette technique a été adaptée pour effectuer des tests à haut débit, 48 tampons peuvent être testés simultanément. Le principe du Thermofluor repose sur les caractéristiques du fluorophore Sypro orange (Invitrogen), qui est sensible au milieu environnant. En milieu aqueux, il émet que très peu de fluorescence (phénomène de « quenching »), mais lorsqu'il est proche de zones hydrophobes sa fluorescence augmente. Ainsi lorsque que l'on met le Sypro orange dans une solution protéique il possède une fluorescence basale plutôt faible. Très peu de zones hydrophobes de la protéine sont accessibles. L'augmentation progressive de la température va provoquer la dénaturation de la protéine exposant son cœur hydrophobe. Le fluorophore se lie à ces régions qui sont maintenant accessibles, augmentant la fluorescence émise. Puis la fluorescence décroît globalement avec la température.

Les résultats sont présentés graphiquement par des courbes représentant l'intensité de fluorescence en fonction de la température (Figure 35). La température de fusion ( $T_m$ ) peut être estimée à partir de ces courbes, elle correspond à la température représentant 50% du signal maximal de fluorescence.



**Figure 35 : Représentation graphique d'une mesure en Thermofluor.** Le  $T_m$  est la valeur de température représentant 50% de la valeur de fluorescence maximale.

Les essais de Thermofluor ont été réalisés sur BBLF2/3 pour la recherche de tampons améliorant la stabilité de la protéine. Mais ils ont aussi été effectués en routine par la plateforme de cristallisation à haut débit (HTX Lab), pour évaluer la stabilité de BBLF2/3 et du complexe (BSLF1 + BBLF2/3 + BBLF4his) lors des tests de cristallisation. Pour la recherche de tampons stabilisant la protéine, le test est réalisé en plaque 96 puits (Stratagène) avec une gamme de 48 tampons dont le pH varie de 5,5 à 9.0. Les tampons sont testés avec deux concentrations en sel, 100 mM et 500 mM NaCl. Pour le test nous avons utilisé BBLF2/3 à 37,5  $\mu$ M dans 20 mM Tris HCl pH 7,5, 250 mM NaCl, avec ou sans 10 % glycérol. Les tampons testés sont à une concentration de 100 mM pour produire un effet tampon fort qui n'est pas modifié lors de l'ajout de la protéine. La mesure par Thermofluor a été réalisée par la plateforme Haut Débit de l'EMBL sur un appareil PCR en temps réel MX3005P Q-PCR (Stratagene). Les longueurs d'ondes d'excitation et d'émission sont respectivement de 470 nm et 570 nm.

## 6. Purification en conditions dénaturantes

Les problèmes de solubilité lors de la surexpression peuvent avoir plusieurs causes. Soit la protéine d'intérêt est habituellement produite en petite quantité et une surexpression provoque son agrégation. Soit la protéine a besoin d'un intermédiaire ou d'un partenaire pour se replier correctement. Si la protéine agrège, l'organisme qui l'exprime l'isole dans des corps d'inclusion pour qu'elle ne perturbe pas son fonctionnement. Les corps d'inclusion existent aussi bien chez les procaryotes (e.g. la bactérie) que chez les eucaryotes (e.g. la cellule d'insecte). La protéine isolée dans ces corps d'inclusions est partiellement ou pas repliée. Lors d'une purification traditionnelle les corps d'inclusion restent dans la fraction insoluble. Pour purifier la protéine contenue dans des corps

d'inclusion, il faut dans un premier temps isoler les corps d'inclusion puis utiliser des solutions dénaturantes pour obtenir la protéine dans la fraction soluble. Pour faciliter la démarche de purification en conditions dénaturantes nous avons utilisé le produit « protein refolding kit Novagen® » (Millipore).

**Solubilisation des protéines :** Les corps d'inclusion sont isolés à partir d'une production à grande échelle. Les culots cellulaires sont resuspendus par 1/10 du volume de culture dans du tampon 1X IB Wash Buffer (20 mM Tris-HCl pH 7,5 ; 10 mM EDTA ; 1% Triton X-100), puis lysés par ultrason. Trois cycles de centrifugation (10000 g, 10min) et de lavage dans du tampon 1X IB Wash Buffer (1/10 du volume de culture) permettent de collecter les corps d'inclusion. Pour solubiliser les protéines, le culot de corps d'inclusion est resuspendu dans « 1X IB Solubilization Buffer » supplémenté par le détergent N-lauroylsarcosine à 0,3%. Après 15 min à température ambiante, une centrifugation (10000 g, 10 min) permet de récupérer la fraction soluble.

**Dialyse et repliement de la protéine d'intérêt :** Pour replier correctement la protéine d'intérêt plusieurs bains de dialyse successifs sont réalisés. Le volume du bain de dialyse doit être 50 fois le volume de surnageant récupéré en fin de solubilisation. Le surnageant est déposé dans une membrane de dialyse avec une taille de pore de 3500 Da (Spectrum Labs).

- **Dialyse N°1 :** 20 mM Tris HCl pH 8,5 ; « 1X Dialysis Buffer » ; 0,1 mM DTT ; 150 mM NaCl pendant 3h à 4°C. Le DTT est un agent réducteur permettant d'éviter la formation de pont disulfure.
- **Dialyse N°2 :** Remplacement du bain par du tampon frais et poursuite de la dialyse pour 3h minimum. Elle peut aussi être réalisée sur la nuit.

Un affinement des conditions de dialyse a été réalisé, en ajoutant 50 mM Arginine/Glutamate (R/E) aux 2 dialyses décrites précédemment. La présence des acides aminés R/E dans un tampon permet d'améliorer la solubilité et la stabilité des protéines, ils permettent aussi de limiter l'agrégation pendant les étapes de concentration et d'augmenter la stabilité à long-terme de la protéine [168, 169]. Après dialyse la solution protéique est purifiée par affinité sur colonne de nickel.

## II. Identification des protéines produites par surexpression en cellules d'insectes

Pour affirmer que la protéine surexprimée correspond bien au gène cloné, différentes techniques d'identification sont utilisables. Cette étape d'identification est très importante pour les études menées sur le complexe, car elle permet de relier directement les bandes que l'on observe sur gel SDS-PAGE à leur protéine correspondante.

### 1. Marquage par anticorps spécifique après Western Blot

Cette technique permet la détection et l'identification de protéines spécifiques. Les protéines migrent au préalable sur un gel SDS-PAGE. Elles sont ensuite transférées du gel polyacrylamide à une membrane PVDF. Après transfert la membrane est saturée par une solution de PBS (1x)/0,1% Tween20/5% Lait écrémé en poudre pendant 1h. Cette étape permet de limiter les interactions non spécifiques de l'anticorps primaire sur la membrane PVDF. L'anticorps primaire est dilué dans du PBS/Tween/Lait et laissé au contact de la membrane pendant 1h.

- **Détection de l'étiquette Hexahistidine** : L'utilisation de l'anticorps « Anti-His Antibody » de souris (GE Healthcare) permet la détection de l'étiquette hexahistidine. Il est dilué au 1/2000 dans le tampon PBS (1x)/0,1% Tween20/5% Lait écrémé en poudre. Après l'heure d'incubation avec l'anticorps primaire, la membrane est lavée par le tampon PBS (1x)/0,1% Tween20 (3 x 5min). L'anticorps secondaire utilisé est un IgG anti souris couplé à la phosphatase alcaline (Promega). Cet anticorps est dilué au 1/4000 dans du PBS/Tween/Lait puis appliqué sur la membrane pendant 1 h. Après l'heure d'incubation la membrane est rincée par du PBS/Tween (3 x 5min) puis par du PBS (2 x 1min). L'application sur la membrane de 10 mL de réactif « Western Blue stabilized substrate for alkaline phosphatase » (Promega) (1 à 30 min) permet la production d'un composé brun par la phosphatase alcaline et ainsi de localiser les protéines possédant l'étiquette hexahistidine. Une fois la réaction achevée, la membrane est rincée abondamment à l'eau, séchée, puis scellée dans une pochette plastique pour la conservation.
- **Détection de BSLF1** : Utilisation d'un anticorps anti BSLF1 de lapin (donné par le Dr Tatsuya TSURUMI) pour la détection de BSLF1, il est dilué au 1/100 dans du PBS/Tween/Lait. Après une heure d'incubation avec l'anticorps primaire, la membrane est lavée par le tampon PBS /Tween (3 x 5min).L'anticorps secondaire utilisé est un anticorps anti-lapin couplé à la peroxidase (Jackson ImmunoResearch). Il est dilué au

1/2000 dans du PBS/Tween/Lait puis appliqué sur la membrane pendant 1h. Après l'heure d'incubation la membrane est rincée par du PBS/Tween (3 x 5min) puis par du PBS (2 x 1min). L'anticorps secondaire est révélé par chimioluminescence grâce au couplage avec la peroxydase. Le réactif de révélation est composé de deux solutions, 5 mL SuperSignal WestPico Luminol/Enhancer solution (Thermo Scientific) et 5 mL WestPico Stable peroxide solution (Thermo Scientific). La membrane est égouttée et scellée dans une pochette plastique. Un film photo CL-XPosure (Thermo Scientific) est exposé à la membrane (5 min jusqu'à une nuit), le film est ensuite révélé en chambre noire.

## **2. Identification par séquençage N-terminal**

L'identification des protéines par séquençage N-terminal a été réalisée par la plateforme de l'IBS dirigée par J.P Andrieu. Le séquençage N-terminal permet l'identification des premiers résidus d'une protéine coté N-terminal. La séquence obtenue et comparée à celle de la protéine d'intérêt ou à une base de données protéiques telle que SwissProt. La plateforme de séquençage N-terminal exige un protocole spécifique à appliquer aux échantillons.

Les protéines sont séparées sur un gel SDS-PAGE 8 ou 10 %, puis transférées sur une membrane PVDF. L'utilisation de membrane de nitrocellulose est à proscrire car elle est détruite par les réactifs utilisés lors du séquençage. La membrane PVDF est activée dans le méthanol (1min) puis rincée par le tampon de transfert (CAPS 1x pH 11, méthanol 10%) (3 x 6min). Le gel SDS-PAGE est saturé 10 min dans le tampon de transfert. Le transfert des protéines, du gel à la membrane, est réalisé à 50V pendant 1h30 dans du tampon de transfert. Après transfert, la membrane est rincée à l'eau puis colorée (15 min) par du bleu de Coomassie (0,1% bleu coomassie R250, 40% méthanol, 1% acide acétique). La membrane est décolorée par des bains successifs à 50% de méthanol et abondamment rincée à l'eau avant d'être stockée à l'abri de la lumière à 4°C. La membrane entière est transmise à la plateforme pour la réalisation du séquençage sur les bandes d'intérêt.

## **3. Identification par spectrométrie de masse**

### **3.1. Digestion tryptique des protéines d'intérêt « in-gel »**

L'identification est effectuée sur des bandes de gels digérées à la trypsine. La trypsine permet de cliver la protéine aux extrémités C-terminales des résidus lysine et arginine. Les peptides digérés sont analysés par spectrométrie de masse. L'ensemble des masses obtenues forment la carte



massique de la protéine. Elle est comparée aux cartes peptidiques théoriques calculées pour les protéines présentes dans les bases de données. La comparaison est permise par des moteurs de recherche d'empreinte peptidique (Mascot, <http://www.matrixscience.com/>). Cette expérience exige beaucoup de rigueur afin d'éviter la contamination par de la kératine (peau, cheveux) qui donne un fort signal en spectrométrie de masse et rend difficile l'interprétation des résultats. Il est donc conseillé de travailler sur une paillasse préalablement nettoyée à l'éthanol et de porter des gants et une blouse tout au long de l'expérience.

**Digestion tryptique « in-gel » :** Les protéines du complexe sont séparées sur gel SDS-PAGE 8 %, puis colorées au bleu de Coomassie. Le gel est abondamment rincé à l'eau sur la nuit et à plusieurs reprises avant l'expérience de digestion tryptique. Les trois bandes d'intérêt sont découpées, puis chaque morceau de gel est à nouveau découpé en petits carrés de 1 mm<sup>2</sup>. Les morceaux de gel sont lavés dans 50 µL de NH<sub>4</sub>HCO<sub>3</sub> (25mM) pendant 30 min sous agitation dans un bloc chauffant. La fraction soluble est éliminée, les fragments de gel sont lavés avec 50 µL de 50% NH<sub>4</sub>HCO<sub>3</sub> (25 mM) et 50% acétonitrile (30 min sous agitation). Le surnageant est éliminé, un dernier lavage est effectué par 50 µL de NH<sub>4</sub>HCO<sub>3</sub> (25 mM) pendant 30 min sous agitation. Ces étapes de lavages permettent de retirer le SDS et le bleu présent dans le gel. Ils perturbent l'activité protéolytique de la trypsine.

Les étapes de réduction et d'alkylation sont réalisées après le lavage des fragments de gel. Elles permettent de détruire les ponts disulfures par réduction des cystéines et empêcher leur reformation par l'alkylation des cystéines réduites. Cette procédure permet d'optimiser au maximum la digestion tryptique. Les fragments de gel sont réduits par ajout de 10 µL de TCEP (10 mM) dilué dans 90µL H<sub>2</sub>O (30 min à température ambiante). L'agent réducteur est retiré, les morceaux de gel sont alkylés par ajout de 100 µL d'iodoacétamide (20 mM) pendant 45 min et à l'abri de la lumière sans agitation. Les morceaux de gel sont lavés dans les mêmes conditions que décrites précédemment.

La digestion tryptique est effectuée par ajout de 10 µL de trypsine à 20 ng/µL (Sequencing Grade Modified Trypsin, Product code: V5113, Promega). Pour que la digestion soit efficace, les fragments de gel doivent être recouverts par du NH<sub>4</sub>HCO<sub>3</sub> (25mM). La digestion est effectuée sur la nuit à 37°C sous agitation. Pour stopper la réaction de digestion du TFA dilué est ajouté aux morceaux de gel. Le milieu est prélevé et transféré dans un nouveau tube. Les fragments de gel sont lavés par 50 µL de 5% acétonitrile/0.1% TFA pendant 15 min à température ambiante et sous agitation. Cette étape est répétée plusieurs fois, les liquides de lavages sont transférés dans le tube contenant la solution de digestion sur la nuit. La dernière étape de lavage est effectuée par ajout de 50µL de 50% acétonitrile/0.1% TFA (15 min sous agitation), le surnageant est ajouté au tube de

digestion à la fin du lavage. Le volume de digestion est réduit jusqu'à 10  $\mu$ L en utilisant un « speedvac ». L'échantillon est ensuite conservé à -80°C en attendant l'étude par MALDI-TOF.

### **3.2. Analyse par spectrométrie de masse**

L'analyse des peptides est effectuée par MALDI-TOF Autoflex (Bruker Daltonics) à la plateforme de spectrométrie de masse de l'IBS. Pour l'analyse, 1  $\mu$ L de la solution de digestion est mélangé avec 1  $\mu$ L de la matrice DHB (Fluka), le mix est déposé sur la plaque MALDI. Les peptides sont volatilisés et ionisés à l'aide d'un laser (MALDI) puis analysés par mesure de temps de vol (TOF). L'analyseur TOF trie les ions en fonction de leur  $m/z$ ,  $m$  étant la masse de l'ions et  $z$  la valeur absolue du nombre de charge. Pour chaque protéine nous obtenons une liste de peptides spécifiques. Ces listes sont comparées à une base de données composée des peptides théoriques calculées *in silico* à partir des séquences des trois protéines du complexe H-P d'EBV (base de données H-P). La comparaison est effectuée par le logiciel Mascot. Pour les trois listes de peptides expérimentaux soumises à la base de données H-P, le logiciel donne la protéine ayant le score le plus élevé. Nous avons étendu l'analyse aux cartes d'empreintes peptidiques des protéines présentes dans la base de données « SwissProt », avec le critère de taxonomie « virus ». Et pour le contrôle de la qualité de l'expérience cette même base de données a été analysée avec le critère taxonomie « humain », pour identifier les peptides provenant de la kératine.

## CHAPITRE 4 :

# Caractérisations biophysiques

# I. Etude de la polydispersité et du rayon hydrodynamique par DLS

## 1. Principe

La diffusion dynamique de lumière (DLS) permet d'étudier la polydispersité d'une solution protéique et de déterminer le rayon hydrodynamique des particules présentes dans la solution. C'est une technique spectrométrique non destructive. La source de lumière utilisée est un laser cohérent. Lorsque cette lumière traverse l'échantillon protéique, elle est diffusée par les particules présentes dans la cuvette d'analyse. L'intensité de diffusion est mesurée à un angle fixe (90° généralement). La fluctuation de l'intensité de diffusion est liée au mouvement brownien des particules et elle est reliée au coefficient de diffusion par une fonction d'autocorrélation. Une particule de petite taille aura un coefficient de diffusion élevé et donc une intensité de diffusion qui varie rapidement au cours du temps. Alors que les particules de grandes tailles possèdent un coefficient de diffusion faible et donc une intensité de diffusion qui varie peu dans le temps. Le lien entre la taille des particules et le coefficient de diffusion est permis par l'équation de Stokes-Einstein sur le mouvement Brownien des particules sphériques. Cette équation permet de déterminer le rayon hydrodynamique de la particule grâce à son coefficient de diffusion. En assimilant la particule mesurée à une sphère, il est possible de déterminer une masse moléculaire approximative. Les protéines sont rarement sphériques, mais plutôt globulaires ou allongées. La taille et la masse mesurées peuvent être différentes des valeurs réelles de la protéine. En plus des renseignements sur la taille et la masse apparentes de la solution protéique, cette technique permet d'estimer sa polydispersité. La DLS est très sensible à la présence de gros agrégats (moins de 0.01% en poids) et la présence de plusieurs espèces dans une solution protéique. Si la polydispersité de la solution est élevée ou s'il y a la présence d'une grande quantité d'agrégats, il est difficile voire impossible d'interpréter les résultats.

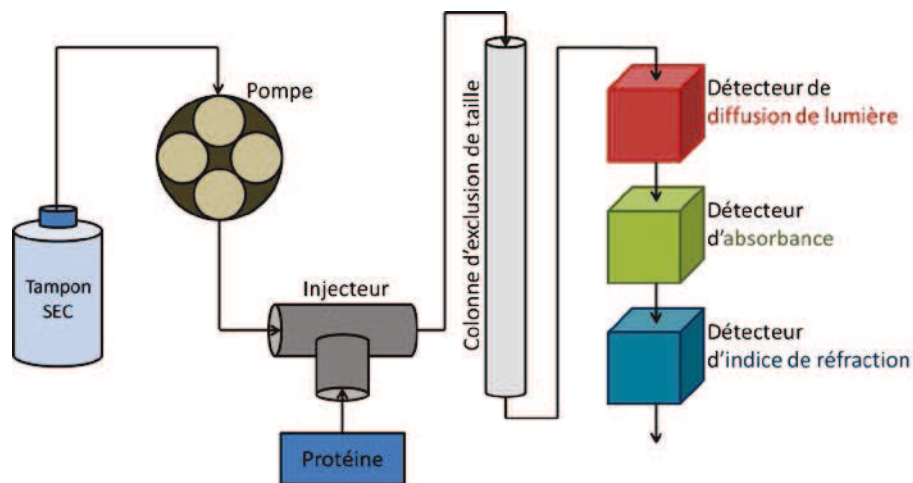
## 2. Tests expérimentaux

Les mesures ont été réalisées pour la protéine BBLF2/3 seule et le complexe (BSLF1 + BBLF2/3 + BBLF4his). Elles ont été réalisées sur un système ProteinSolutions DynaPro (Wyatt Technology) et analysé par le logiciel Dynamics pour BBLF2/3. Un système Zetasizer (Malvern) et le logiciel du même nom ont été utilisés pour étudier le complexe. Quelle que soit la solution protéique étudiée, nous avons déposé 45µL entre 0.3 et 1 mg/mL dans une cuve en quartz (Hellma).

## II. SEC-MALLS-RI

### 1. Principe

C'est une méthode permettant de mesurer le rayon hydrodynamique et la masse moléculaire d'une particule. Pour réaliser ces mesures le système combine plusieurs techniques. Une chromatographie d'exclusion de taille (SEC), la mesure de la diffusion de la lumière à différents angles (MALLS) et la réfractométrie (RI) (Figure 36). La chromatographie d'exclusion de taille et la DLS permettent de mesurer la masse moléculaire des protéines, mais la forme des protéines influe sur la valeur calculée. Le système SEC-MALLS-RI regroupe les avantages des deux techniques précédentes et permet de mesurer la masse moléculaire des protéines sans être influencé par leur forme.

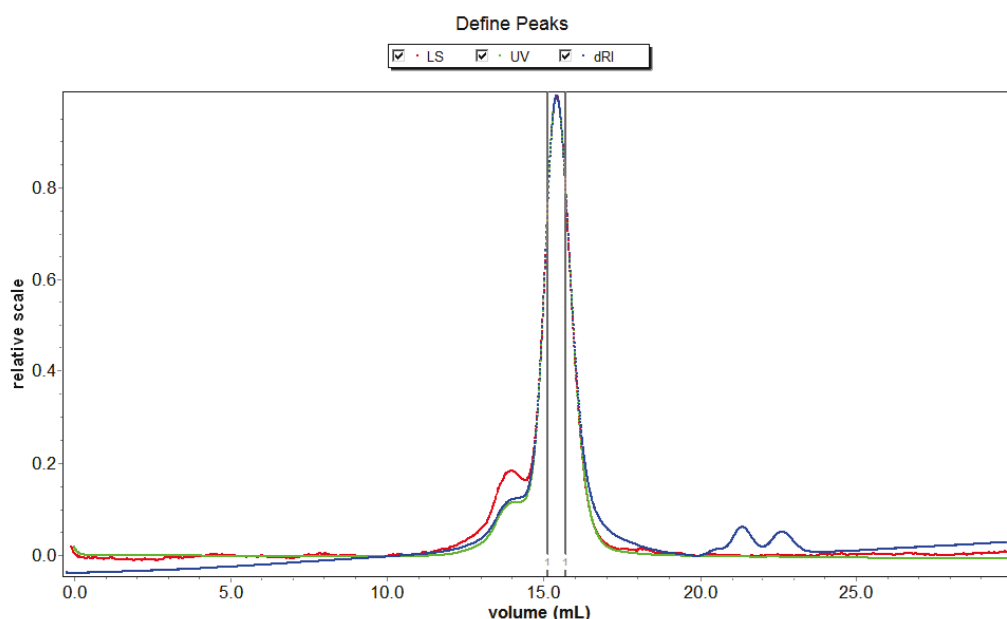


**Figure 36 : Représentation schématique du système SEC-MALLS-RI**, la protéine est poussée par le solvant à travers la colonne d'exclusion de taille puis les trois détecteurs (diffusion de lumière, absorbance et réfractomètre).

Le système FPLC couplé avec une pompe permet d'injecter l'échantillon et d'assurer sa progression grâce au flux de tampon. Les particules présentes dans l'échantillon sont séparées par une colonne d'exclusion de taille SuperDex200 10/300 (GE-Healthcare), puis elles passent à travers un détecteur DAWN-EOS (Wyatt Technology) muni d'un laser émettant à 690nm permettant la mesure de la diffusion de lumière à 18 angles. Enfin les particules traversent un réfractomètre différentiel RI2000b (Schambeck) mesurant l'indice de réfraction. Les données sont traitées par le logiciel ASTRA V (Wyatt Technology).

Le changement de la valeur d'indice de réfraction permet de mesurer la concentration de la protéine. L'intensité de la diffusion de la lumière et la concentration obtenue par le réfractomètre permettent de calculer la masse molaire de la protéine. Cette technique permet aussi d'estimer l'oligomérisation d'une protéine ou la stœchiométrie d'un complexe. La précision des mesures

dépend de la concentration de l'échantillon et de la qualité du tampon. Dans des conditions optimales l'erreur en masse est de 0,5% (Figure 37).



**Figure 37 : Courbes obtenues sur le système SEC-MALLS-RI pour la protéine BSA (Albumine de sérum bovin). Bleu :** courbe de détection de l'indice de réfraction. **Rouge :** courbe de détection de la diffusion de lumière. **Vert :** Courbe de détection de l'absorbance à 280 nm.

## 2. Tests expérimentaux

- **Etude de la protéine BBLF2/3 :** Le système SEC-MALLS-RI a permis l'étude la protéine BBLF2/3. Un échantillon de 50  $\mu$ L à 5 mg/mL dans un tampon Tris HCl pH 7,5 ; 250 mM NaCl ; 10% glycérol a été injecté sur le système avec un flux de tampon à 0,5 mL/min.
- **Etude du complexe entier :** Une injection du complexe H-P entier d'EBV a été réalisée sur le système SEC-MALLS-RI. Un échantillon de 50  $\mu$ L à 1,3 mg/mL dans le tampon XII (20 mM Hépès pH 7,6 ; 500 mM NaCl ; 50 mM R/E ; 5% glycérol) a été injecté sur le système avec un flux de tampon à 0,5 mL/min.

## III. Test de cristallisation

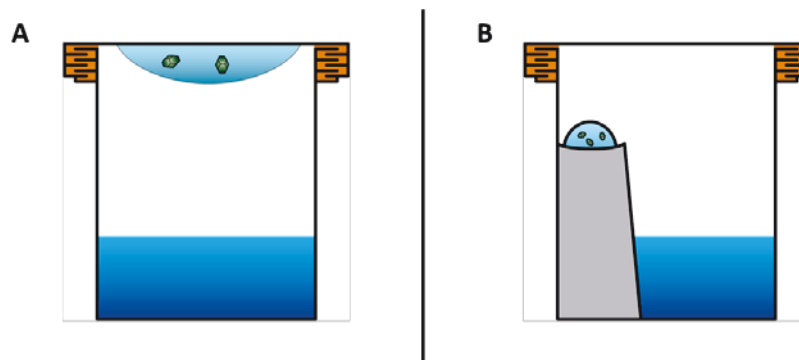
### 1. Principe

La cristallisation des protéines est un processus très empirique et dépend d'un grand nombre de paramètres. La pureté, le  $T_m$ , la concentration de la protéine, mais aussi la température, la présence de cofacteur ou de substrat et la composition du tampon peuvent influencer sur la cristallisation d'une protéine.



Depuis les débuts de la cristallographie des protéines, un grand nombre de techniques de cristallisation ont été développées. A l'heure actuelle, il n'existe aucune technique miracle permettant de cristalliser une protéine à coup sûr, la cristallisation reste très empirique. Il existe quatre techniques de cristallisation : en batch, par échange de vapeur, par dialyse et par diffusion libre. Seule la technique de cristallisation par échange de vapeur est détaillée ici, car c'est la technique la plus utilisée et celle que nous avons menée pour les tests de cristallisation de BBLF2/3 et du complexe (BLSF1 + BBLF2/3 + BBLF4his).

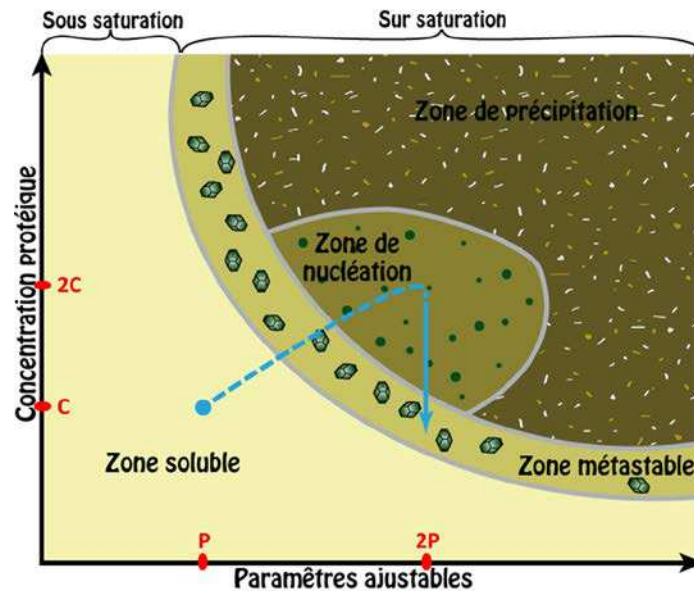
La technique de cristallisation par échange de vapeur consiste à effectuer des échanges gazeux entre un réservoir contenant la solution de cristallisation et une goutte formée par un mélange avec 50% de solution protéique et 50% de solution de réservoir en général. La solution de réservoir est composée de divers constituants pouvant favoriser la cristallisation de la protéine. Il existe deux variantes d'échange de vapeur, la goutte suspendues et la goutte assise (Figure 38).



**Figure 38 : Principe de cristallisation par échange de vapeur. A :** Méthode cristallisation par goutte suspendue. **B :** Méthode de cristallisation par goutte assise.

Quel que soit la variante utilisée le principe de cristallisation reste le même. La cristallisation suit les règles imposées par le diagramme de phase des protéines (Figure 39). Il est divisé en deux grandes zones, la première zone est dite « zone de sous saturation » et la seconde zone est appelée « zone de sur saturation ». En zone de sous saturation aucun cristal ne peut se former car la protéine est totalement soluble. La zone de sur saturation est plus complexe, elle se divise en trois sous parties. La zone de précipitation contient les protéines sous forme précipitées. La zone de nucléation permet d'initier la formation d'un cristal. La dernière zone, appelée zone métastable, permet la croissance cristalline. Pour une goutte suspendue ou assise, à l'instant  $T=0$  de l'expérience, elle se situe à un point du diagramme de phase décrit par sa concentration en protéine (C) et en solution de réservoir (P). L'échange de vapeur entre la goutte et la solution de réservoir va faire évoluer la composition de la goutte et donc modifier sa position au sein du diagramme de phase. Au cours du temps, la concentration en précipitant de la goutte va se rééquilibrer avec celle du réservoir (2P).

Dans certaines conditions la protéine peut débuter une nucléation conduisant à la croissance cristalline. Pendant la phase de croissance la concentration en protéines libres dans la goutte diminue (Figure 39). La technique par échange de vapeur est la plus utilisée actuellement et beaucoup de systèmes automatisés ont été développés pour cette technique.



**Figure 39 :** Diagramme de phase représentant l'évolution de la composition d'une goutte par échange de vapeur au cours du temps. C et 2C: concentration en protéine dans la goutte. P et 2P : concentration en précipitant dans la goutte.

## 2. Tests de cristallisation de BBLF2/3 et du complexe H-P

Pour les études de cristallisation de BBLF2/3 et du complexe (BSLF1 + BBLF2/3 + BBLF4his) nous avons utilisé la plateforme de cristallisation à haut débit du laboratoire. Le HTX Lab permet d'effectuer des tests de cristallisation par échange de vapeur avec la méthode de la goutte assise. Cette plateforme permet de tester 1728 conditions de cristallisation réparties en 19 « screens ». Les gouttes (100 nL + 100 nL) et les solutions de réservoir sont réparties sur une plaque 96 puits Crystal Quick plates (Greiner) par un robot de cristallisation Cartesian PixSys 4200 (Genomic Solutions). La plaque est entreposée dans un robot RoboDesign Minstrel III (Rigaku Msc) à 20°C. Ce robot prend des photos des gouttes sur une durée de 3 mois.

### 2.1. BBLF2/3

Pour l'étude de BBLF2/3, nous avons effectué trois séries de 576 conditions cristallisation (Crystal Screen I & II, Crystal Screen Lite & PEG/Ion, MembFac & Natrix, QuickScreen & Grid screens Ammonium Sulfate, Sodium Malonate - Sodium Formate, Grid screens PEG 6K, PEG/LiCl, MPD - Screen Mme, Index Screen, les screens utilisés proviennent de Hampton Research ou des screens

équivalent de Qiagen). Les tests de cristallisation ont été réalisés à deux concentrations, 1 et 5 mg/mL dans un tampon 20 mM Tris HCl pH7,5 ; 250 mM NaCl ; 10% glycérol et une condition à 1,5 mg/mL sans glycérol.

## 2.2. Complexe H-P

Pour l'étude du complexe (BSLF1 + BBLF2/3 + BBLF4his) 576 conditions de cristallisation ont été testées (les mêmes que pour BBLF2/3) pour une concentration de 0,6 mg/mL dans un tampon 20 mM Hépès pH7,6 ; 500 mM NaCl ; 50 mM R/E ; 5% glycérol. Une seconde concentration à 1,3 mg/mL (même condition de tampon) a été utilisée avec 96 conditions de cristallisation (The JCSG screen).

## IV. Analyse par microscopie électronique et affinement par la technique GraFix

Les analyses de microscopie électronique sont réalisées par coloration négative. Le sel d'atome lourd acétate d'uranyl est utilisé à 2% (m/v). Il permet d'enrober les protéines et de produire des images avec un fort contraste. Cette méthode de coloration renseigne seulement sur les informations structurales de l'enveloppe, elle est surtout utilisée pour le contrôle qualité des échantillons. Les échantillons protéiques sont appliqués sur la face propre d'une couche de carbone vaporisée sur une feuille de mica (interface carbone/mica) et colorés négativement par 2% (m/v) d'acétate d'uranyl (pH 4,5). Les échantillons sont visualisés sur un microscope électronique à transmission CM12 120kV LaB6 (Philips) équipé d'une caméra CCD 2,6k x 4k.



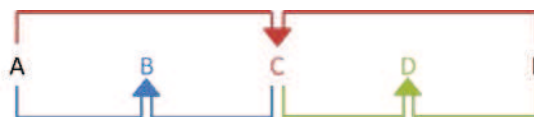
Figure 40 : Microscope électronique à transmission utilisé à la plateforme de l'IBS. CM12 120kV LaB6 (Philips).

## 1. Principe de la technique GraFix

Les résultats obtenus sur le complexe ont révélés une variabilité importante et une forte hétérogénéité entre les expériences. Pour améliorer la qualité du complexe nous avons utilisé la technique de GraFix [170]. Les complexes macromoléculaires sont fragiles, le changement de tampon, la procédure de coloration négative peuvent induire des changements de composition ou de conformation augmentant l'hétérogénéité de l'échantillon. Ces problèmes peuvent être limités en renforçant la stabilité du complexe par des agents fixant (glutaraldéhyde, GTA). L'utilisation directe des agents de fixation chimique sur le complexe purifié est déconseillé car elle conduit à la formation d'artéfacts (fixation inter-molécule et agrégation). La technique de GraFix combine la centrifugation et la fixation du complexe. Un gradient de glycérol et de glutaraldéhyde permet de séparer les différentes formes du complexe selon leur masse et de fixer l'état du complexe pour une masse donnée.

## 2. Réalisation du GraFix

Le gradient de glycérol utilisé pour le GraFix est composé de 5 tampons nommés de A à E (Tableau 8). Pour faciliter la création du gradient, tous les tampons sont réalisés par des mélanges volume à volume à partir des solutions A et E (Figure 41).



**Figure 41 : Représentation schématique de la préparation des solutions GraFix B, C et D à partir des solutions A et E.**

Le gradient est réalisé dans des tubes Ultra-Clear™ (Beckman) de 750  $\mu$ L. Il est réalisé à la main par ajouts successifs de 100  $\mu$ L de chaque tampon. Le tampon A est déposé en premier, suivi par le dépôt des tampons B, C, D et E. Chaque tampon est déposé au fond du tube et pousse les tampons précédents vers le haut. Les tampons B et D sont refroidis 30 min à 4°C, les solutions refroidies ont une viscosité plus importante et limite ainsi le mélange de la fraction avec celles adjacentes. Avant le dépôt de 100  $\mu$ L de solution protéique, le gradient est stabilisé 30 min à 4°C.

La centrifugation (40000 rpm, 4°C, 18 h) est réalisée dans une ultracentrifugeuse Optima XPN-80 (Beckman Coulter) munie d'un rotor SW55Ti avec adaptateur. Après les 18h de centrifugation, le gradient est fractionné en aliquotes de 50  $\mu$ L en partant du haut du tube. Chaque fraction est mélangée à 10  $\mu$ L de Tris-base (1 M) pour stopper la réaction de fixation du GTA. Les

fractions (30  $\mu$ L) sont analysées sur gel SDS-PAGE à 6%, puis colorées dans du bleu de Coomassie (15 min). Les 30  $\mu$ L restant sont transmis à la plateforme de microscopie électronique.

**Tableau 8 : Détails des 5 tampons composants le gradient de GraFix.**

Tampon	A	B	C	D	E
Composition	50 mM Hépès pH 7,5 ; 150 mM NaCl ; 5 mM BME ; 0,1 mM PMSF ; 10% glycérol	50 mM Hépès pH 7,5 ; 150 mM NaCl ; 5 mM BME ; 0,1 mM PMSF ; 15% glycérol ; 0,0125% GTA	50 mM Hépès pH 7,5 ; 150 mM NaCl ; 5 mM BME ; 0,1 mM PMSF ; 20% glycérol ; 0,05% GTA	50 mM Hépès pH 7,5 ; 150 mM NaCl ; 5 mM BME ; 0,1 mM PMSF ; 25% glycérol ; 0,0375% GTA	50 mM Hépès pH 7,5 ; 150 mM NaCl ; 5 mM BME ; 0,1 mM PMSF ; 30% glycérol ; 0,1% GTA

## CHAPITRE 5 :

# Caractérisations enzymatiques

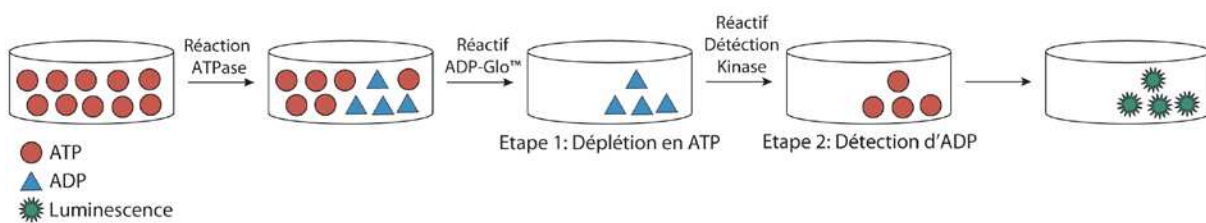


# I. Caractérisations enzymatiques

## 1. Mesure de l'activité ATPase

Les tests ATPase sur le complexe ont été réalisés à l'aide du kit « ADP-Glo™ Kinase Assay » (Promega). Le kit mesure la quantité d'ADP produit lors d'un test kinase ou ATPase. La réaction ATPase est réalisée avant le test « ADP-Glo™ Kinase Assay ». Le test complet comprend trois étapes :

- **Etape 1 :** L'ATP est dégradé par le complexe H-P d'EBV à 37°C pendant 30 min.
- **Etape 2 :** Le réactif « ADP-Glo™ Reagent » est ajouté à la réaction ATPase. Il permet de stopper la réaction ATPasique et de déphléter l'ATP restant en solution.
- **Etape 3 :** L'ajout du réactif « Kinase Detection Reagent » permet de simultanément convertir l'ADP en ATP et mesurer l'ATP nouvellement synthétisé par une réaction luciférase/luciférine (Figure 42).



**Figure 42 :** Principe du test « ADP-Glo™ Kinase Assay », détails des étapes permettant de mesurer le pourcentage de conversion d'ATP en ADP par la réaction ATPase (adapté du manuel technique « ADP-Glo™ Kinase Assay »).

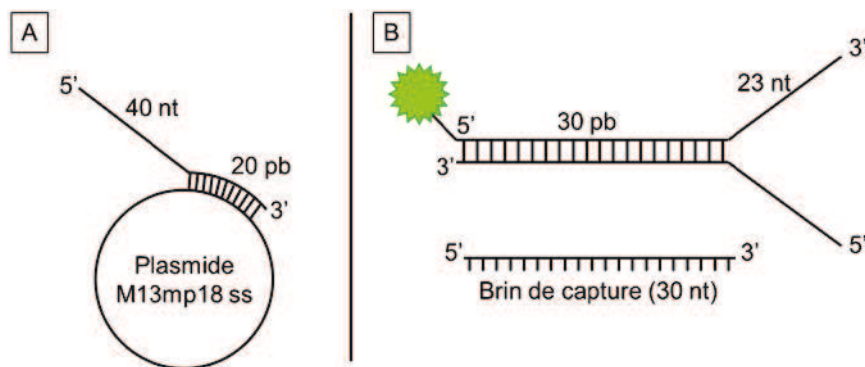
Le taux de conversion de l'ATP en ADP est calculé à partir d'une courbe de calibration réalisée pour des ratios d'ATP et d'ADP à 50  $\mu$ M final. La gamme de calibration suit les mêmes étapes du test « ADP-Glo™ Kinase Assay » que la réaction ATPase.

La réaction ATPase est réalisée dans un plaque 96 puits blanche à fond plat (Nunc) en présence de tampon ATPase 1x (20 mM Hépès pH 7,6 ; 1 mM DTT ; 0,1 mg/mL BSA ; 5 mM MgCl<sub>2</sub> ; 10% glycérol), 200 nM ou 1  $\mu$ M ADNSb (53 bp), 50  $\mu$ M ATP, 410 nM ou 1  $\mu$ M de complexe (BSLF1 + BBLF2/3 + BBLF4his), le volume réactionnel est de 25  $\mu$ L. La réaction est effectuée pendant 30 min à 37°C. La plaque doit revenir à température ambiante avant ajout de 25  $\mu$ L de réactif « ADP-Glo™ Reagent ». Le réactif est incubé 40 min à température ambiante. Après les 40 min d'incubation, on ajoute 50  $\mu$ L de réactif « Kinase Detection Reagent ». La valeur de luminescence est lue à 30, 45 et 60 min sur un lecteur de plaque Victor<sup>2</sup>V (Wallac).

L'équation de la droite de calibration permet d'estimer le pourcentage de conversion d'ATP en ADP du test ATPase.

## 2. Mesure de l'activité hélicase

Les tests d'activité hélicase ont été inspirés de ceux décrits dans la littérature pour le complexe H-P d'HSV1 (UL52 + UL8 + UL5). Les tests réalisés par Falkenberg & al (1998) [106] et Chono & al (2010) [108] ont servis de support pour la création des substrats ADN. Les deux publications reportent des substrats ayant l'extrémité 5' flottante et 3' appariée. Dans le test de Falkenberg & al [106] l'ADNsb (60 nt) marqué par au  $^{32}\text{P}$  est hybridé sur le plasmide simple brin M13mp18 (Biolab), formant une séquence double brin de 20 pb à l'extrémité 3' et une partie flottante de 40 nt à l'extrémité 5' (Figure 43, A). Le test de Chono & al (2010) [108] décrit un substrat ADN double brin de type fourche. Un brin fluorescent de 53 nt couplé en 5' avec l'Alexa Fluor<sup>®</sup> 488 (Invitrogen), est hybridé à un brin non marqué de 53 nt. La fourche comprend une zone d'hybridation de 30 pb et une fourche flottante de 23 nt. Le test comprend un brin capture qui s'hybride sur le brin non marqué après séparation de la fourche par l'hélicase, le brin marqué devient simple brin (Figure 43, B).



**Figure 43 : Détails des substrats utilisés pour mesurer l'activité hélicase du complexe H-P d'HSV1 (UL52 + UL8 + UL5). A : Falkenberg & al (1998) [106] ; B : Chono & al (2010) [108].**

Pour les deux tests le suivi de la réaction est réalisé sur des gels de polyacrylamide non dénaturant par autoradiographie (Falkenberg & al (1998) [106]) ou par mesure de fluorescence (Chono & al (2010) [108]).

Nous avons adapté les brins du test de Chono & al pour qu'ils puissent s'hybrider sur le plasmide M13mp18ss. L'objectif de cette adaptation est de pouvoir tester l'activité hélicase sur des substrats avec un brin flottant 5' ou une fourche, en utilisant une mesure de fluorescence de la Cyanine 5 pour analyser les résultats.

Séquence du brin portant la cyanine 5 en 5' :

5' -ACATGATAAGATACATGGATGAGTTTGGACGAGGATCCCCGGGTACCGAGCTCGAATTTCG-3'

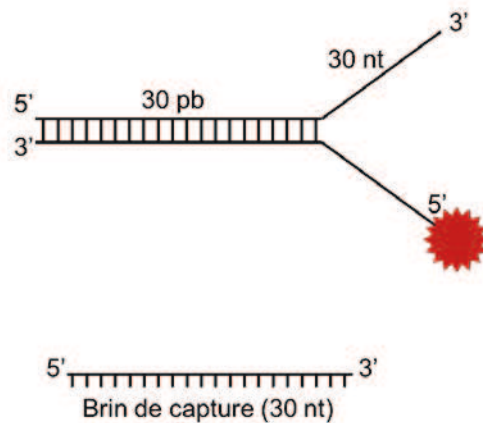
Séquence du brin non marqué :

3' - CGAATTCGAGCTCGGTACCCGGGGATCCTCCAGGTTTGAGACTTATTGGCGAATAGTACA -5'

Séquence du brin de capture :

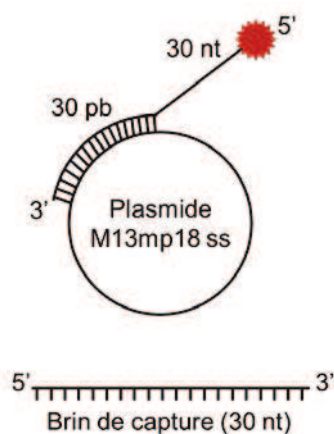
5' -GAGGATCCCCGGGTACCGAGCTCGAATTTCG-3'

La fourche est formée par un mélange volume à volume de brin non marqué (100 pmol/ $\mu$ L) et de brin marqué (100 pmol/ $\mu$ L). Le mélange est chauffé 5 min à 95°C, le retour à température ambiante est réalisé sur la nuit pour permettre l'hybridation (Figure 44).



**Figure 44 : Substrat de type fourche pour le test hélicase.** La cyanine 5 est présente sur le brin 5'. Lors de la séparation de la fourche par l'hélicase, le brin de capture peut se lier au brin non marqué, laissant le brin portant la cyanine 5 sous forme simple brin.

L'hybridation sur le plasmide M13mp18ss (tebu-bio) est réalisé de la même façon que pour la fourche. Un mélange entre le plasmide M13mp18ss (10 pmol) et le brin marqué cyanine 5 en 5' (10 pmol) est chauffé pendant 5 min à 95°C, le retour à température ambiante est réalisé sur la nuit.



**Figure 45 : Substrat avec brin flottant en 5', hybridé au plasmide M13mp18ss pour le test hélicase.** La cyanine 5 est présente sur le brin 5'. Lors de la séparation du brin marqué et du plasmide par l'hélicase, le brin de capture peut se lier au plasmide, laissant le brin portant la cyanine 5 sous forme simple brin.

La réaction hélicase a été réalisée dans un volume de 12,5  $\mu$ L contenant 20 mM Hépès pH 7,5 ; 250 mM NaCl ; 1 ou 3mM ATP ou 1 mM analogue ATP (AppNHp, ATP $\gamma$ S) ou 1 mM ADP ; 2.5 mM Mg<sup>2+</sup> ou Mn<sup>2+</sup> ; 10 nM ou 80 nM Duplex CY5 ou 1,1 pM Duplex M13 CY5 ; 8  $\mu$ M brin de capture (large excès). La réaction est effectuée à 37°C pendant 1 h. Les résultats sont analysés sur gels

polyacrylamides non dénaturants 20% (Duplex CY5) ou agarose 1% (Duplex M13 CY5). La fluorescence est mesurée sur un Typhoon<sup>TRIO</sup> (Amersham Biosciences) avec une longueur d'onde d'excitation de 550 nm et de 570 nm pour l'émission.

# Partie III : Résultats

## CHAPITRE 1 :

Etudes des protéines BBLF4,  
BSLF1 et BBLF2/3 produites  
par trois baculovirus

# I. Etudes des protéines du complexe H-P produites individuellement

En clonant chacun des gènes codant pour le complexe Hélicase-Primase (H-P) d'EBV, dans des baculovirus individuels, nous pouvons étudier chacune des protéines de manière indépendante. Cette approche permet aussi l'étude du complexe en réalisant des co-infections doubles ou triples des cellules d'insectes. Les résultats décrits dans la littérature révèlent que le système cellules d'insectes/baculovirus est adapté pour la production des trois protéines du complexe H-P d'EBV [90, 91, 94], mais l'objectif de ces études ne nécessitait ni la production de grande quantité de protéine ni le besoin d'une pureté élevée, ces deux paramètres sont prérequis pour une étude par cristallographie aux rayons X. Nous avons entrepris, dans un premier temps, d'effectuer une étude bio-informatique des trois protéines afin d'estimer si une étude par cristallographie était envisageable sur les protéines entières. Le degré d'ordre d'une protéine est un des facteurs principal pour sa cristallisation. Le désordre des protéines a été étudié à l'aide du script « D-score », les résultats sont détaillés dans le paragraphe suivant.

## 1. Caractéristiques des protéines composant le complexe H-P d'EBV

Les trois protéines composant le complexe H-P d'EBV sont, l'hélicase (BBLF4), la primase (BSLF1) et la sous unité accessoire (BBLF2/3). La séquence protéique des trois protéines permet d'obtenir un certain nombre de caractéristiques propres à chacune. Les informations obtenues sur les sites ProtParam tool et PredictProtein sont des prédictions et permettent d'établir une carte d'identité de chaque protéine servant de base aux futures expériences.

**Tableau 9 : Cartes d'identités des protéines du complexe H-P d'EBV.** Les informations sont obtenues à partir de l'analyse des séquences des protéines par les sites, ProtParam tool et PredictProtein.

	Hélicase	Primase	Sous-unité Accessoire
Nom gène/protéine	BBLF4	BSLF1	BBLF2/3
Taille du gène (pb)	2430	2625	2130
Nombre de résidus (aa)	809	874	709
Masse moléculaire (Da)	89853,3	98040,9	78225,7
pI	6,53	8,88	6,23
Coefficient d'extinction molaire	1,129	1,085	1,458
Prédiction de structure secondaire : Hélice, Feuillet, Boucle (%)	36 ; 22 ; 42	40 ; 18 ; 41	24 ; 22 ; 54



La recherche de signaux NLS (Signal de Localisation Nucléaire) a été réalisée à l'aide des logiciels NLS Mapper et PredictProtein. Aucune des protéines du complexe ne contient de peptide signal dans sa séquence correspondant à un signal NLS, BBLF2/3 obtient le score le plus élevé dans NLS Mapper, mais pas suffisant en déduire avec certitude un signal NLS. Le complexe H-P permet la réplication du génome virale au sein du noyau. Un signal NLS apparaît peut être lors de la formation du complexe, ou il est pris en charge par une protéine de type importine pour permettre son entrée dans le noyau.

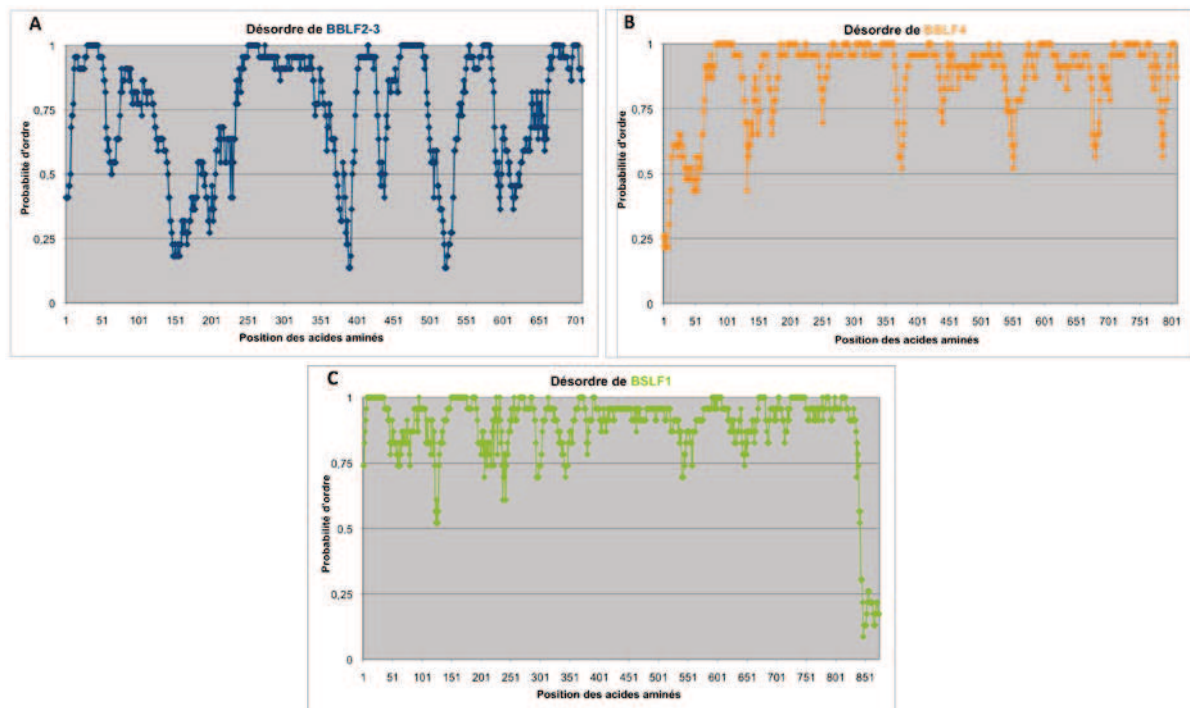
## **2. Prédiction d'ordre et de désordre des protéines BSLF1, BBLF2/3 et BBLF4**

Il est très important lors de l'étude d'une protéine ou d'un ensemble de protéines de réaliser une étude préliminaire par bio-informatique. Cette investigation est d'autant plus importante pour les études visant à déterminer la structure protéique par cristallographie aux rayons X. Cette technique biophysique nécessite des protéines stables et largement repliées. Il est donc indispensable de connaître la prédiction de désordre d'une protéine avant de commencer les phases de clonage et d'expression en vue des expériences de cristallographie.

Un très grand nombre de logiciels de prédiction d'ordre et de désordre sont disponibles sur internet. L'utilisation d'un seul de ces programmes pour décrire le désordre d'une protéine est une vue potentiellement erronée. En effet chacun de ces logiciels utilisent des scripts propres ne mettant pas en avant les mêmes propriétés des acides aminés et des polypeptides pour évaluer le désordre d'une protéine. Pour éviter de se baser sur un seul prédicteur un script, appelé « D-score », a été développé au sein du laboratoire par le Dr Ambroise Desfosses. Comme décrit dans la partie Matériels et Méthodes de ce manuscrit, le D-score permet de combiner les résultats de plusieurs prédicteurs pour obtenir un profil moyenne probabilité de désordre des protéines étudiées.

Le D-score a été appliqué aux trois protéines du complexe H-P d'EBV. A partir de la séquence en acides aminés de chacune des protéines nous avons pu établir leur prédiction d'ordre (Figure 46). La prédiction de structures indique peu de désordre pour les protéines BBLF4 (Figure 46 B) et BSLF1 (Figure 46 C). Ces protéines présentent chacune une zone désordonnée à une extrémité de leur séquence, en N-terminal pour BBLF4 et en C-terminal pour BSLF1. Les prédictions observées pour ces deux protéines sont intéressantes car elles montrent qu'elles ne possèdent pas ou très peu de boucles flexibles pouvant être gênantes pour la cristallisation. La prédiction de désordre de BBLF2/3 semble révéler une alternance de motifs ordonnés et de boucles désordonnées. On pourrait ainsi

décrire BBLF2/3 avec 5 zones ordonnées (1-140, 225-370, 394-506, 622-701) sûrement reliés entre eux par des boucles désordonnées.



**Figure 46 : Profil de désordre des trois protéines composant le complexe Hélicase-Primase d'EBV obtenu par le script D-score. A : sous-unité accessoire (BBLF2/3), B : hélicase (BBLF4), C : primase (BSLF1).**

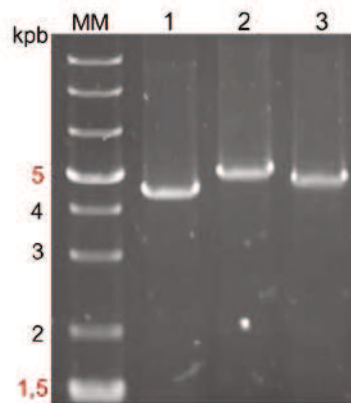
Cette étude préliminaire montre que les protéines du complexe Hélicase-Primase d'EBV sont prédites pour être relativement bien ordonnées et suggère qu'une étude par cristallographie aux rayons X, semble réalisable sur ces protéines entières. Au sein de la famille des herpèsvirus, aucune structure d'un complexe H-P n'a été obtenue jusqu'à ce jour. Nous avons donc décidé d'étudier les protéines entières afin de pouvoir les décrire une à une, mais aussi pouvoir observer leurs interactions lors de la formation du complexe. Pour réaliser ces objectifs, nous avons procédé au clonage de chaque protéine dans un baculovirus différent. Le clonage et la production des protéines par ce système d'expression sont décrits dans la section suivante.

### 3. Expression des protéines individuelles

#### 3.1. Clonage des gènes dans les vecteurs baculovirus

Les clonages des trois gènes BSLF1, BBLF2/3 et BBLF4 dans trois vecteurs baculovirus ont été réalisés en suivant le protocole Bac-to-Bac® (Invitrogen). Dans un premier temps les gènes sont insérés dans le plasmide pFastBac™HTB permettant d'avoir une étiquette hexahistidine en N-terminal de chaque protéine. Une étape supplémentaire a été nécessaire pour le clonage du gène

BBLF2/3 dans le plasmide pFastBac™HTB. *In vivo*, l'ARNm de BBLF2/3 est épissé pour retirer un intron de 129 pb. Le gène contenant seulement les deux exons codant directement pour la protéine BBLF2/3 a été obtenu grâce à une délétion par PCR. Une fois les vecteurs plasmidiques obtenus, les gènes sont transposés dans trois bacmides différents. La transfection des bacmides en cellules d'insectes a permis d'obtenir trois baculovirus contenant chacun un gène du complexe H-P d'EBV. La vérification de la présence des trois gènes d'intérêt dans les trois bacmides a été réalisée par PCR.



**Figure 47 : Contrôle par PCR de la présence des gènes BBLF2/3, BSLF1 et BBLF4 dans les bacmides.** L'amplification a été permise, après purification des bacmides, par les amorces M13, ajoutant ainsi 2300 pb à taille attendue de chaque gène. 1 : M13 + BBLF2/3 = 2300 + 2130 = 4430 pb, 2 : M13 + BSLF1 = 2300 + 2625 = 4925 pb, 3 : M13 + BBLF4 = 2300 + 2430 = 4730 pb. MM : marqueur de poids moléculaire.

Les baculovirus permettent la production individuelle des trois protéines BSLF1, BBLF2/3 et BBLF4 par infection des cellules d'insectes.

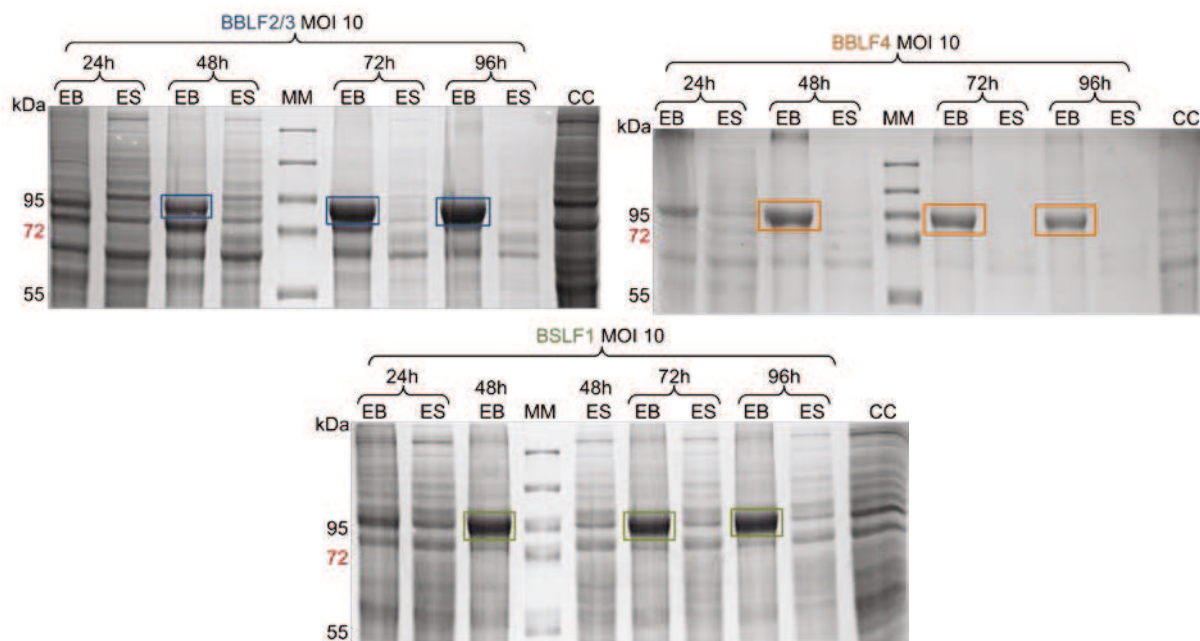
### 3.2. Production individuelle des protéines BSLF1, BBLF2/3 et BBLF4 en cellules adhérentes

Les tests d'expression permettent d'estimer la production optimale des protéines BSLF1, BBLF2/3 et BBLF4, en fonction de la MOI et du temps d'infection. Le temps d'infection optimal est déterminé grâce à une production à petite échelle dans des plaques 6 puits. Une plaque 6 puits est réalisée par temps d'infection avec un puits pour chaque MOI. Les résultats de ces expériences sont détaillés pour chaque protéine.

- **Test d'expression de BBLF2/3 :** Le profil d'expression révèle la surexpression d'une protéine avec une masse proche de 80 kDa qui concorde avec la masse théorique de BBLF2/3 (78 kDa). La bande de BBLF2/3 est présente seulement dans la fraction brute. Une forte surexpression est observée pour un temps d'infection de 72 h (Figure 48).

- **Test d'expression de BBLF4** : Cette expérience réalisée sur la protéine BBLF4 montre une bande de surexpression à environ 95 kDa, en accord avec la masse attendue pour cette protéine (90 kDa). BBLF4 est surexprimée dès 48 h et le niveau reste stable. Tout comme BBLF2/3, la protéine BBLF4 semble seulement présente dans la fraction brute du test (Figure 48).
- **Test d'expression de BSLF1** : Les tests d'expression de BSLF1 sont très similaires à ceux obtenus pour BBLF4. La surexpression de la protéine est visible dès 48 h et le niveau d'expression n'augmente pas pour les temps supérieurs. La bande de surexpression a une masse d'environ 95 kDa, concordant avec la valeur théorique de 98 kDa attendue pour BSLF1 (Figure 48).

Pour chaque protéine, seul le test à la MOI 10 est présenté (Figure 48), car c'est pour cette valeur de MOI que les surexpressions sont les plus marquées. La surexpression de BBLF2/3 est bien visible à partir de 72 h et se maintient à 96 h. Les protéines BBLF4 et BSLF1 sont surexprimées dès 48 h et le niveau d'expression de ces protéines reste stable pour les temps supérieurs (72 h et 96 h). Ainsi un temps de surexpression de 72 h semble optimal pour la production des trois protéines.



**Figure 48 : Profil d'expression des trois protéines du complexe Hélicase-Primase.** Infection des cellules Hi5 à une MOI 10 pour des temps d'infection de 24 à 96 h, observation du profil d'expression sur gels SDS-PAGE 10% colorés au bleu de Coomassie. MM : Marqueur de poids moléculaire, EB : Extrait Brut, ES : Extrait Soluble, CC : Contrôle Cellulaire.

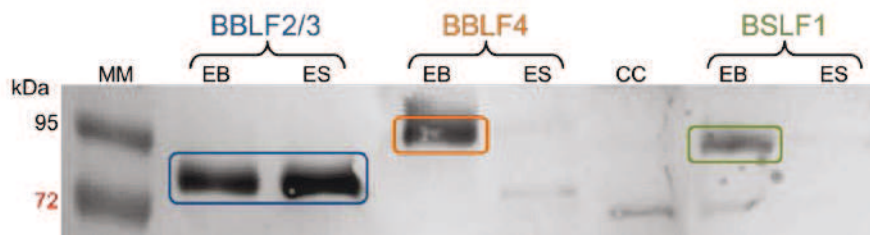
Les tests d'expression par « Time course » ont permis de déterminer les conditions optimales d'infection des 3 baculovirus : MOI 10 pendant 72h. Les trois baculovirus possèdent les mêmes paramètres d'infection, facilitant ainsi la mise en place des co-infections décrites plus tard.

### 3.3. Identification des protéines surexprimées

#### 3.3.1. Western Blot anti étiquette hexahistidine

Les tailles observées sur les gels SDS-PAGE suite aux tests d'expression concordent avec les masses théoriques des trois protéines, mais cette information n'est pas suffisante pour affirmer qu'il s'agit réellement des protéines d'intérêt. Les trois protéines possèdent une étiquette hexahistidine en N-terminal permettant une purification facilitée des protéines. Cette étiquette peut être détectée sur Western Blot pour aider à l'identification des protéines.

L'étiquette polyhistidine des trois protéines détectée par Western Blot avec un anticorps anti-étiquette hexahistidine permet d'estimer la production et la solubilité des protéines (Figure 49). BBLF4 et BSLF1 sont bien surexprimées, mais aucun signal n'est observé dans la fraction soluble indiquant que ces protéines sont à priori insolubles. La protéine BBLF2/3 est soluble car nous observons un signal aussi bien dans fraction brute que dans fraction soluble. Si on compare cette observation avec celles données par le test d'expression (Figure 48), on en déduit que seule une fraction de l'expression de BBLF2/3 est soluble. La protéine BSLF1 est détectée sur ce Western Blot, mais l'intensité de la bande étant plus faible que pour les deux autres protéines, nous avons entrepris son identification par séquençage N-terminal.



**Figure 49 : Western Blot anti étiquette polyhistidine.** Détection de BBLF2/3 dans la fraction soluble et brute. BSLF1 et BBLF4 sont seulement détectées dans la fraction brute. EB : Extrait Brut, ES : Extrait soluble, CC : contrôle cellulaire, MM : Marqueur de poids moléculaire.

### 3.3.2. Séquençage N-terminal

Pour confirmer l'identité de la bande attribuée à BSLF1, un séquençage N-terminal (plateforme IBS, J-P Andrieu) a été effectué pour cette protéine. La séquence SYHHHHH a été identifiée, elle correspond à l'étiquette polyhistidine de BSLF1.

Les techniques d'identification menées par Western Blot et séquençage N-terminal ont confirmé la surexpression des trois protéines du complexe H-P par le système cellules d'insectes/baculovirus. Cette observation est en accord avec les résultats des précédentes études pour la production des protéines du complexe H-P d'EBV par le système cellules d'insectes/baculovirus [91, 94](Tableau 10).

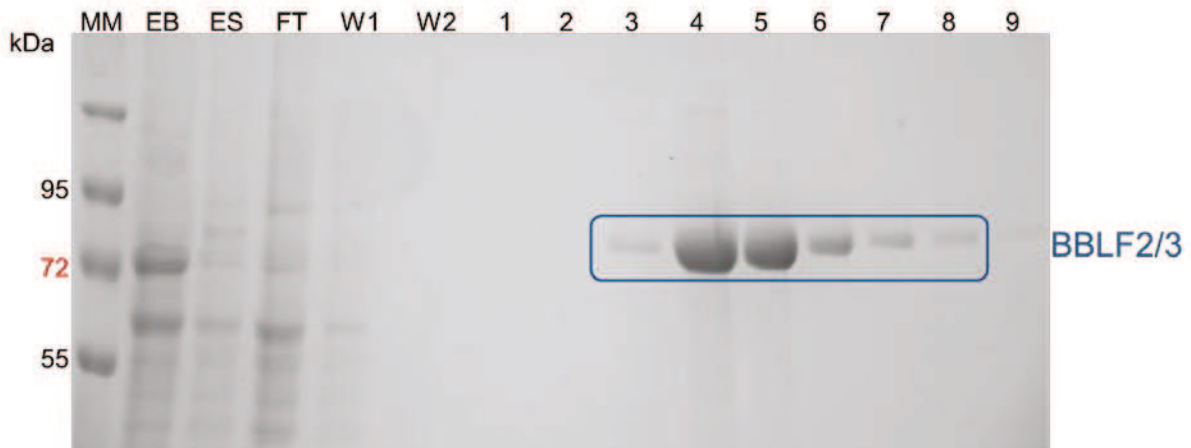
**Tableau 10 : Tableau récapitulatif des niveaux de surexpression et de solubilité des trois protéines du complexe H-P d'EBV.**

	Surexpression	Solubilité
BSLF1	+	-
BBLF4	++	-
BBLF2/3	+++	++

## 4. Purification par affinité sur colonne de nickel

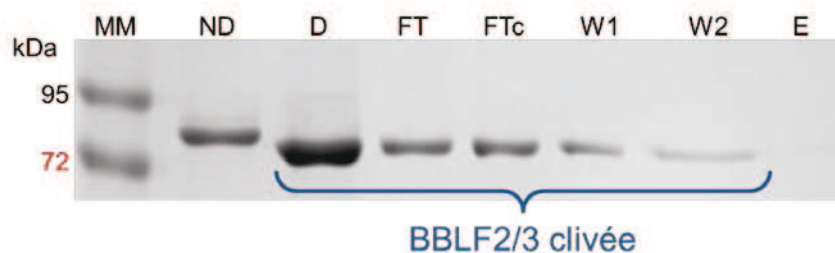
### 4.1. BBLF2/3

La protéine BBLF2/3 est surexprimée (Figure 48) sous forme soluble (Figure 49). Pour caractériser la structure ou l'activité enzymatique d'une protéine, celle-ci doit être pure et suffisamment concentrée. BBLF2/3 a été purifiée par affinité sur colonne de nickel, grâce à l'étiquette hexahistidine située en N-terminal. Une seule étape de purification sur colonne de nickel permet d'obtenir BBLF2/3 avec une pureté relativement élevée (Figure 50). Le tampon final dans lequel est conservé la protéine est composé de 20 mM Tris pH7,5 ; 250 mM NaCl ; 10 % glycérol.



**Figure 50 : Profil d'élution de la protéine BBLF2/3 d'une colonne de nickel**, visualisé sur Gel SDS-PAGE 8% coloré au bleu de Coomassie. MM : Marqueur de poids moléculaire, EB : Extrait Brut, ES : Extrait Soluble, FT : « FlowThrough », W1 : Lavage 1, W2 : Lavage 2, 1-9 : Fractions d'élution.

L'étiquette hexahistidine situé en N-terminal de la protéine possède un site de clivage pour la protéase TEV. L'efficacité du clivage est observée directement sur gel SDS-PAGE, la bande de BBLF2/3 est décalée vers une masse plus faible après digestion sur la nuit par la protéase TEV (Figure 51). Chaque puits du gel contient le même volume de solution protéique sauf pour le puits ND ayant un volume de dépôt plus faible. On remarque que la bande de protéine digérée (Figure 51, puits D) a une intensité plus forte que la protéine en sortie de la seconde colonne d'affinité (Figure 51, puits FT). Cette baisse d'intensité indique une perte de protéines lors du passage sur la deuxième colonne d'affinité probablement due à une interaction non spécifique de BBLF2/3 clivée sur la colonne de nickel. Le tampon de la protéine lors du passage sur la seconde colonne de nickel était composé de 20 mM Tris-HCl pH 7,5 ; 250 mM NaCl. Les bandes présentes dans les puits FT et FTc (Figure 51) possèdent une intensité équivalente. Cette constatation révèle des difficultés pour la concentration de BBLF2/3.



**Figure 51 : Digestion de l'étiquette hexahistidine de BBLF2/3 par la protéase TEV.** Gel SDS-PAGE 8%, coloré au bleu de Coomassie. MM : Marqueur de poids moléculaire, ND : non digérée, D : digérée, FT : « Flowthrough », FTc : « Flowthrough » concentré, W1 : lavage 1, W2 : Lavage 2, E : Elution.

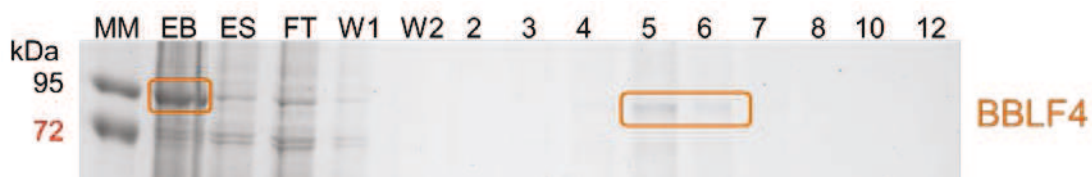


La protéine BBLF2/3 s'exprime très bien par le système cellules d'insectes/baculovirus. La phase de purification par affinité sur colonne de nickel permet d'obtenir une protéine pure et en grande quantité. La purification permet d'obtenir 100 µg de BBLF2/3 à partir de 1g de cellules d'insectes. Une perte de protéine est constatée après clivage de l'étiquette de purification et pendant l'étape de concentration.

## 4.2. BBLF4 et BSLF1

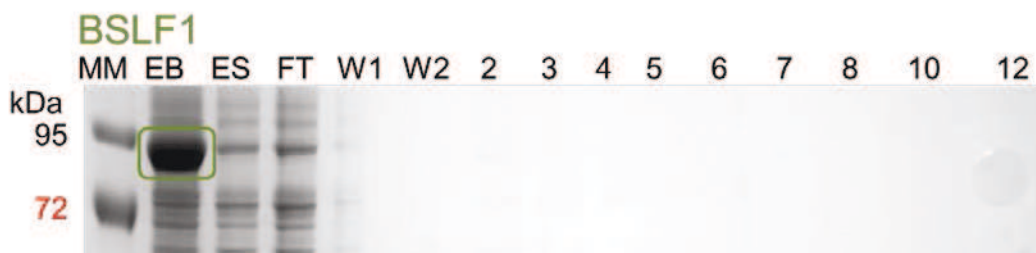
Les profils d'expression (Figure 48) et le Western Blot (Figure 49), semblent indiquer que les protéines BSLF1 et BBLF4 sont surexprimées, mais pas solubles. Pour vérifier cette constatation les protéines ont été produites à grande échelle et purifiées par affinité sur colonne de nickel. Les profils d'élution de chaque protéine sont visualisés sur gel SDS-PAGE (Figure 52 : BBLF4, Figure 53 : BSLF1).

- **La protéine BBLF4** : La bande présente dans la fraction brute (Figure 52, puits EB), confirme la surexpression de BBLF4. L'observation d'une bande de faible intensité à 95 kDa dans les fractions d'élution 5 et 6, indique que BBLF4 est faiblement soluble. La production à grande échelle de cette protéine permet d'isoler une petite fraction soluble de la protéine.



**Figure 52 : Profil d'élution de la protéine BBLF4 d'une colonne de nickel**, visualisé sur gel SDS-PAGE 8% coloré au bleu de Coomassie. MM : Marqueur de poids moléculaire, EB : Extrait Brut, ES : Extrait Soluble, FT : Flowthrough, W1 : Lavage 1, W2 : Lavage 2, 2-12 : Fractions d'élution.

- **La protéine BSLF1** : Une bande de forte intensité est observée à 95 kDa dans le puits de la fraction brute, confirmant ainsi la surexpression de la protéine BSLF1. Contrairement à la protéine BBLF4, BSLF1 n'est pas visualisée dans les fractions d'élution, lors de la production protéique à grande échelle (Figure 53). Ces résultats confirment les conclusions faites à partir des expressions à petite échelle : BSLF1 est surexprimée mais insoluble.



**Figure 53 : Profil d'éluion de la protéine BSLF1 d'une colonne de nickel**, visualisé sur gel SDS-PAGE 8%, coloré au bleu de Coomassie. MM : Marqueur de poids moléculaire, EB : Extrait Brut, ES : Extrait Soluble, FT : Flowthrough, W1 : Lavage 1, W2 : Lavage 2, 2-12 : Fractions d'éluion.

Les productions protéiques menées à petite et grande échelle indiquent que BSLF1 est surexprimée, mais sous une forme insoluble.

La purification de BBLF4 à partir d'une production à grande échelle a permis de préciser qu'une fraction de cette protéine est soluble. Des ajustements en concentration saline et en pH n'ont pas permis d'améliorer la solubilité des protéines. Nous avons aussi essayé d'optimiser les conditions de purification notamment en ajoutant au tampon du  $MgCl_2$ , car les activités hélicase et primase sont dépendantes de la présence d'ions divalents, mais la solubilité des deux protéines n'a pas été améliorée. Ainsi les deux protéines sont bien produites par le système cellules d'insectes/baculovirus, sous forme insoluble pour BSLF1 et très faiblement soluble pour BBLF4.

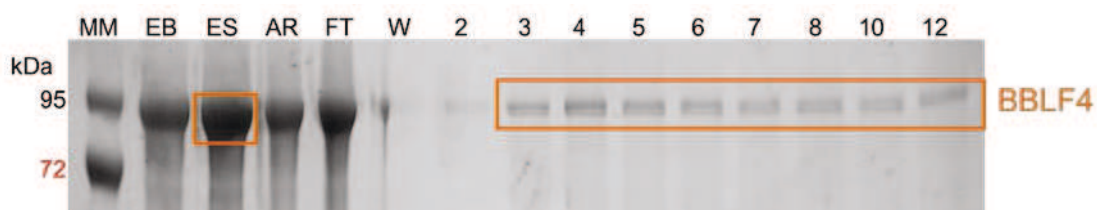
## 5. Purification des protéines BBLF4 et BSLF1 en conditions dénaturantes et renaturation

Les purifications traditionnelles par affinité ont montré que les protéines BSLF1 et BBLF4 sont insolubles. Il est possible de purifier les protéines même insolubles en utilisant des conditions dénaturantes. Après la dénaturation des protéines, on essaie de les replier dans leur état natif par dialyse. Pour cette approche nous avons utilisé le kit « Protein Refolding ® » (Novagen).

### 5.1. L'hélicase (BBLF4)

La lyse en conditions dénaturantes permet d'observer la protéine BBLF4 dans la fraction soluble (Figure 54, puits ES). Le repliement de la protéine est effectué par des dialyses successives, avant la purification par affinité sur colonne de nickel. La protéine BBLF4 est présente dans plusieurs fractions d'éluion (Figure 54) et l'intensité des bandes est plus forte que lors de la purification en conditions non dénaturantes (Figure 52). Par contre au moment de l'éluion on observe la protéine dans l'ensemble des fractions. Ce profil particulier d'éluion peut être révélateur de la présence d'agrégats solubles. Un agrégat possède plusieurs étiquettes pouvant interagir avec la colonne,

rendant plus difficile son élution malgré la présence de 500 mM d'imidazol dans le tampon. Il semblerait que la phase de repliement par dialyse provoque la formation d'agrégats solubles de la protéine. Les phases de dialyse du protocole « Protein Refolding ® » ont été modifiées pour limiter la formation d'agrégats solubles au cours du repliement de BBLF4. Nous avons ajouté au tampon de dialyse les acides aminés arginine et glutamate (R/E). La présence des acides aminés R/E dans un tampon permet d'améliorer la solubilité et la stabilité des protéines ; elle permet aussi de limiter l'agrégation pendant les étapes de concentration et d'augmenter la stabilité à long-terme de la protéine [168, 169]. La présence des deux acides aminés n'a pas permis un meilleur repliement de BBLF4, le gel après purification donne les mêmes résultats que ceux observés sur la Figure 54.



**Figure 54 : Profil d'élution de la protéine BBLF4 d'une colonne de nickel après lyse des corps d'inclusion et repliement de BBLF4 par dialyse.** Observation sur gel SDS-PAGE 8% coloré au bleu de Coomassie. EB : Extrait Brut, ES : Extrait Soluble, AR : Après repliement, W : Lavage, 2-12 : Fractions d'élution.

## 5.2. La primase (BSLF1)

Tout comme pour la protéine BBLF4, la lyse en conditions dénaturantes permet d'obtenir BSLF1 dans la fraction soluble. La protéine renaturée ne se fixe pas à la colonne de nickel, elle est complètement éluée par le premier lavage. Le repliement de la protéine, au cours des dialyses, rend peut être l'étiquette de purification inaccessible provoquant l'élution de la protéine dès le premier lavage. Les dialyses de BSLF1 ont aussi été essayées en présence du tampon R/E. Mais le profil de purification est resté inchangé malgré de ce nouveau tampon.

Les purifications menées en conditions dénaturantes ont permis de solubiliser les protéines BBLF4 et BSLF1. BSLF1 n'a pas pu être purifiée par affinité après repliement, contrairement à BBLF4. Mais la présence de BBLF4 dans toutes les fractions d'élutions sur le profil de purification laisse à penser que cette protéine est majoritairement agrégée après renaturation. Les difficultés de repliement rencontrées pour ces deux protéines proviennent sûrement de leur grandes tailles (BBLF4 : 809 aa, BSLF1 : 874 aa), de plus elles nécessitent peut-être la présence des autres protéines du complexe H-P ou d'un cofacteur pour un bon repliement *in vitro*.

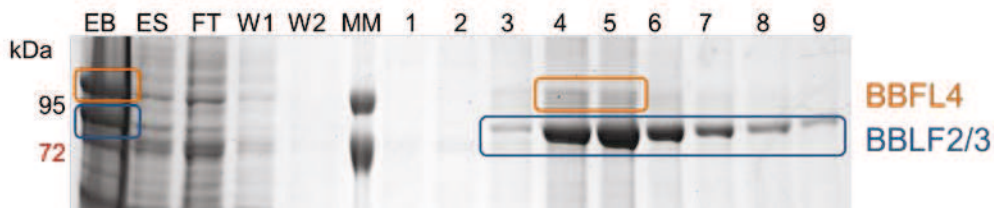
En conclusion, les productions individuelles des trois protéines du complexe H-P d'EBV ont montré que la sous unité accessoire se purifie en grande quantité et avec une pureté élevée par affinité sur colonne de nickel. Les protéines BSLF1 et BBLF4 sont majoritairement insolubles. L'utilisation de conditions dénaturantes a permis de les solubiliser, mais une renaturation efficace n'est pas possible. Le caractère insoluble observé pour ces protéines peut-être dû à l'absence d'un ou plusieurs partenaires.

## **6. Co-production par co-infection avec les baculovirus codant chacun pour une protéine**

Dans la section suivante je vais analyser les résultats des différentes co-infections réalisées afin d'obtenir le complexe Hélicase-Primase d'EBV sous forme soluble. Nous avons déjà mentionné l'importance de la MOI lors d'une infection des cellules d'insectes par le baculovirus. Lors d'une co-infection, la MOI peut être interprétée de deux façons différentes. Soit chaque virus est utilisé à sa propre MOI, approche que nous appelons : « MOI stricte ». Soit les virus sont utilisés à une MOI plus faible pour que la somme des différentes MOI soit égale à celle optimale pour un seul virus, cette approche est nommée : « MOI cumulée ». Ces deux approches ont été testées lors les co-infections.

### **6.1. Co-production BBLF4 + BBLF2/3**

La co-infection des cellules d'insectes par les baculovirus de BBLF2/3 et BBLF4 a été réalisée à une MOI stricte de 10 pour chaque virus (production à grande échelle). La Figure 55 représente le profil de purification d'affinité sur colonne de nickel de la co-infection. Les fractions d'élution contiennent une bande très majoritaire proche de 80 kDa, correspondant à BBLF2/3 et une seconde bande minoritaire à environ 95 kDa, correspondant à la protéine BBLF4. Les bandes de BBLF4 et BBLF2/3 dans la fraction brute possèdent des intensités équivalentes. Ainsi la faible quantité de BBLF4 élue de la colonne ne provient pas d'une faible expression de cette protéine, mais de son caractère peu soluble. Nous avons aussi réalisé des infections avec un ratio de baculovirus plus élevé pour BBLF4 afin de favoriser sa production. Cette approche n'améliore pas la production de BBLF4.



**Figure 55 : Profil d'élution d'une colonne de nickel de la co-infection BBLF2/3 et BBLF4**, visualisé sur gel SDS-PAGE 8%, coloré au bleu de Coomassie. EB : Extrait Brut, ES : Extrait Soluble, FT : Flowthrough, W1 : Lavage 1, W2 : Lavage 2, MM : Marqueur de poids moléculaire, 1-9 : Fraction d'élution.

La production simultanée en cellules d'insectes des protéines BBLF2/3 et BBLF4 ne favorise pas la forme soluble de BBLF4. La présence de BBLF2/3 seule ne permet pas le bon repliement de BBLF4 *in vivo*.

## 6.2. Co-production de BBLF4 +BSLF1

Les co-infections par les baculovirus BSLF1 et BBLF4 ont été réalisées avec une MOI stricte et cumulée. Lors de la purification par affinité aucune des deux protéines n'est présente dans les fractions d'élution. La présence des deux protéines *in vivo*, ainsi que l'utilisation de MOI stricte et cumulée lors de l'infection ne sont pas des paramètres suffisants pour produire les deux protéines sous leurs formes solubles.

## 6.3. Co-production de BBLF2/3 + BSLF1

Le dernier couple de co-production testé est composé de BBLF2/3 + BSLF1. Cette production a été réalisée par co-infection à MOI stricte et cumulée. La fraction brute de la co-production présente une seule bande majoritaire proche de 80 kDa correspondant à BBLF2/3 (Figure 56). La poursuite de la purification révèle des bandes de très faible intensité, à environ 80 kDa, dans les fractions d'élution. Ces bandes correspondent à la taille de BBLF2/3, mais il est étonnant que ces bandes aient une intensité aussi faible.



**Figure 56 : Analyse de la co-expression des protéines BSLF1 et BBLF2/3**, sur gels SDS-PAGE 8% coloré au bleu de Coomassie. **Gel de gauche** : Analyse des fractions brutes et solubles. **Gel de droite** : Profil d'élution des protéines BSLF1 et BBLF2/3 d'une colonne de nickel. MM : Marqueur de poids moléculaire, C : MOI cumulée, S

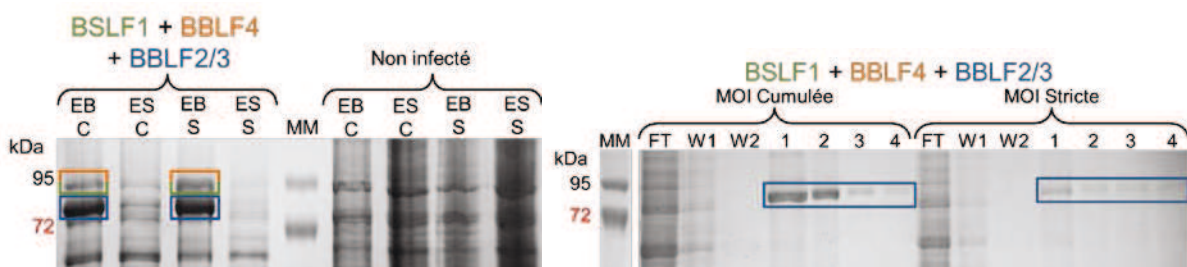
: MOI stricte, Ct : Contrôle non infecté, EB : Extrait brut, ES : Extrait soluble, FT : Flowthrough, W1 : Lavage 1, W2 : Lavage 2, 1-4 : Fraction d'élution.

Ce résultat est très étonnant, car il semblerait que BSLF1 soit très peu exprimée lorsqu'elle est co-produite avec BBLF2/3 et que cette co-production retient une grande partie de BBLF2/3 dans la fraction insoluble. Une explication à ce résultat n'est pas évidente, mais on pourrait penser que BSLF1 est produite en faible quantité et qu'elle séquestre la protéine BBLF2/3 dans la fraction insoluble ou que l'étiquette de purification est masquée par l'interaction entre BSLF1 et BBLF2/3 ne permettant pas la purification de cette protéine par affinité.

Les co-infections avec deux sous unités du complexe n'ont pas permis d'obtenir les protéines BSLF1 et BBLF4 dans la fraction soluble. La solubilité de BBLF2/3 ne permet pas d'entraîner BSLF1 ou BBLF4 dans la fraction soluble. Les protéines BBLF4 et BSLF1 nécessitent peut être la formation du complexe entier *in vivo* pour être soluble.

## 7. Triple co-infection

L'infection triple a été menée à petite échelle avec des MOI strictes ou cumulées (Figure 57). On distingue très bien, dans les extraits bruts cumulés et stricts, la présence de BBLF2/3 (80 kDa). Une seconde bande à 95 kDa est présente, elle peut correspondre à l'hélicase et/ou la primase. Les fractions d'élution, de la triple production sur colonne de nickel, contiennent seulement la bande à 80 kDa correspondant à BBLF2/3. Les bandes de BBLF2/3, éluées donc solubles, sont plus intenses pour l'infection à MOI cumulée que stricte. Il semblerait donc que l'utilisation d'une MOI cumulée lors d'une co-infection permette une meilleure production protéique. Avec une MOI stricte, les cellules font face à une multiplicité d'infection plus importante pouvant impacter sur leur capacité de production protéique.



**Figure 57 : Analyse de la triple production BSLF1, BBLF4 et BBLF2/3 à partir d'une MOI stricte ou cumulée** sur gel SDS-PAGE 8% coloré au bleu de Coomassie. **Gel de gauche** : Observation des fractions brutes et solubles infectées et non infectées. **Gel de droite** : Profil d'élution d'une colonne de nickel de la triple infections BSLF1, BBLF4 et BBLF2/3. EB : Extrait Brut, ES : Extrait Soluble, C : MOI cumulée, S : MOI stricte, FT : Flowthrough, W1 : Lavage 1, W2 : Lavage 2, MM: Marqueur de poids moléculaire, 1-9 : Fraction d'élution.

Les protéines BSLF1 et BBLF4 restent insolubles même lors du test de triple infection. La formation du complexe est peut être seulement réalisable si les trois protéines sont synthétisées simultanément dans la cellule d'insecte. La co-infection par des baculovirus individuels n'implique pas obligatoirement une infection concomitante de la cellule par les trois virus. Il y a peut-être un trop faible pourcentage de cellules triplement infectées.

En conclusion, l'approche d'expression par gène individuel a permis de produire et de purifier BBLF2/3. Les protéines BBLF4 et BSLF1 sont bien surexprimées, mais BBLF4 est très peu soluble et BSLF1 pas du tout. Ainsi seule BBLF4 a pu être observée dans la fraction d'élution de la colonne de nickel, mais en quantité beaucoup trop faible pour pouvoir la caractériser. Cette approche n'a pas été poursuivie, nous avons choisi d'utiliser la technique MultiBac pour pouvoir produire le complexe entier avec un seul baculovirus codant simultanément pour les trois gènes d'intérêts (approche développée dans le prochain chapitre).

## **II. Etude biophysique de BBLF2/3**

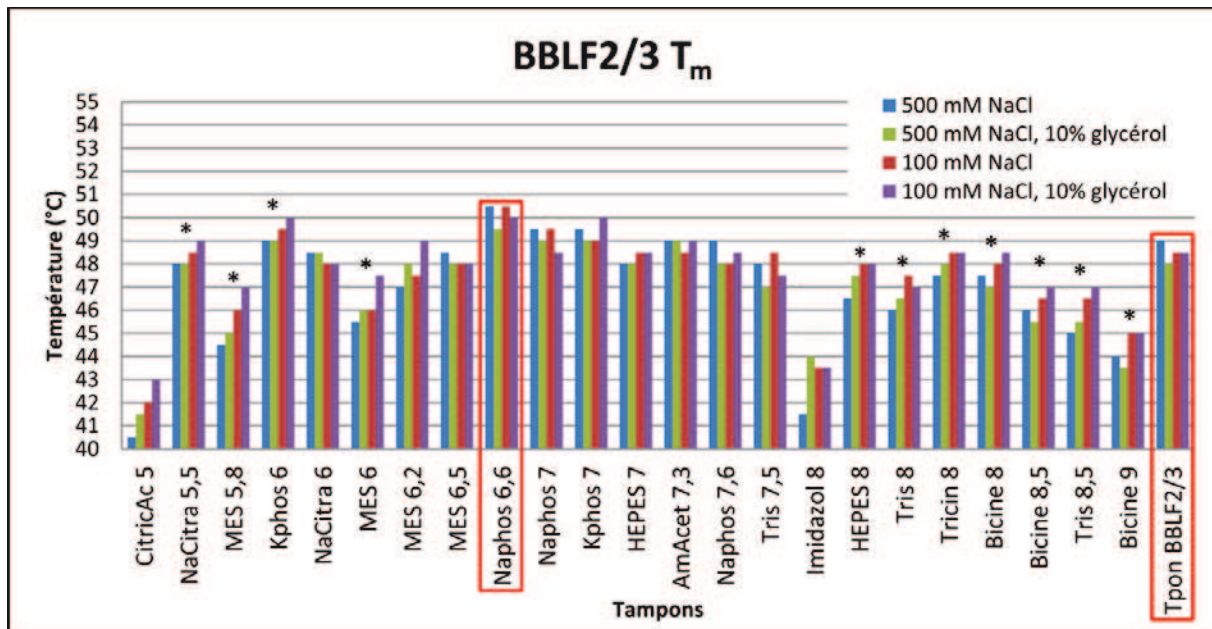
### **1. Optimisation du tampon de purification**

Les premiers tests de purification de BBLF2/3 (Figure 51) ont montré une perte de protéine lors de la seconde colonne d'affinité après clivage de l'étiquette de purification par la protéase TEV et pendant la phase de concentration. Pour limiter la perte protéique, on a recherché un tampon permettant de stabiliser la protéine sans étiquette hexahistidine.

Les tests de stabilité de BBLF2/3 ont été effectués par la technique Thermofluor à partir de 24 tampons différents (23 tampons + tampon spécifique de BBLF2/3). Ces 24 tampons (pH 5 à 9) ont été testés dans 4 conditions différentes, en présence de 100 mM ou 500 mM de NaCl et avec ou sans glycérol (10 %). Le criblage des 24 tampons dans les 4 conditions en sel et glycérol a donné une plage de  $T_m$  allant de 41°C à 50,5°C (Figure 58). La présence ou l'absence de glycérol n'influe pas sur la valeur du  $T_m$  de la protéine. Pour les deux concentrations en sel, il semble que la concentration à 100 mM NaCl permette d'obtenir des valeurs de  $T_m$  plus élevées de l'ordre de 1 à 2°C (Figure 58, astérisque). Cette faible augmentation n'est pas réellement significative, mais elle indique tout de même que BBLF2/3 semble être légèrement plus stable avec des tampons ayant une concentration saline inférieure à 500 mM. Le tampon utilisé pour la purification de BBLF2/3 (Figure 58 Tpon BBLF2/3 : 20 mM Tris HCl pH 7,5 ; 250 mM NaCl) a un  $T_m$  de 48,5°C. Cette valeur de  $T_m$  indique une stabilité acceptable de la protéine. Par ailleurs seul le tampon phosphate pH 6,6 (Naphos 6,6) permet d'avoir un  $T_m$  plus élevé : 50,5°C. Il est admis qu'une augmentation de 2°C du  $T_m$  est significative et

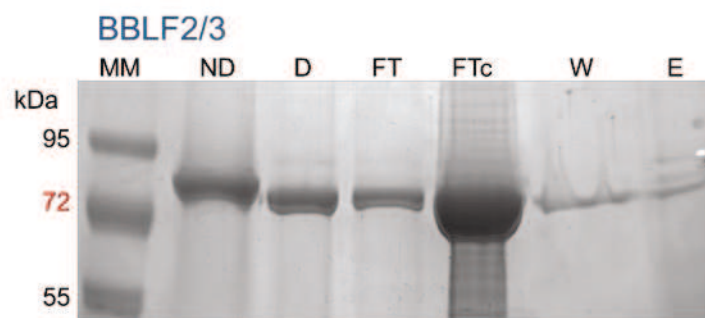


traduit une meilleure stabilité de la protéine. Le tampon phosphate pH 6,6 permet donc d'augmenter la stabilité de BBLF2/3. Mais ce tampon possède un désavantage important car il forme très facilement des cristaux de sel lors des tests de cristallisation. Il est donc plus intéressant de conserver le tampon d'origine de la protéine en vue des tests futurs de cristallisation.



**Figure 58 : Valeurs des  $T_m$  de BBLF2/3 dans 24 tampons** à 100 et 500 mM NaCl en présence ou absence de glycérol (10 %). Tpon BBLF2/3 : 20 mM Tris HCl pH 7,5 ; 250 mM NaCl, tampon d'origine de la protéine.

La seconde phase critique de perte de protéines est lors de la phase de concentration par centrifugation. Différentes marques de concentrateur ont été essayées, mais aucune n'a permis de concentrer efficacement BBLF2/3. On a pu constater par Thermofluor que la stabilité de BBLF2/3 dans son tampon de purification est correcte. Ce test confirme que la présence de glycérol ne modifie pas sa stabilité. Nous avons donc essayé de purifier BBLF2/3 en présence de 10 % de glycérol. Cette modification apportée au tampon permet de concentrer efficacement la protéine par centrifugation à 15°C et 4000 rpm (centrifugeuse de paillasse 5840 Eppendorf). La perte de protéines lors du passage sur la seconde colonned'affinité peut venir d'une interaction non spécifique de la protéine clivée avec le nickel. Pour limiter ces interactions non spécifiques, nous avons ajouté 15 mM d'imidazol au tampon de dessalage de BBLF2/3. Lors de l'application de la protéine clivée sur la seconde colonne de nickel, l'imidazol présent dans le tampon possède une affinité pour le nickel plus forte que la protéine clivée et permettra donc de limiter les interactions non spécifiques de la protéine. La présence d'imidazol dans le tampon ne perturbe pas la concentration de la protéine par centrifugation (Figure 59).



**Figure 59 : Profil d'élution de BBLF2/3 d'une colonne de nickel après clivage de l'étiquette hexahistidine sur gel SDS-PAGE 10%, coloré au bleu de Coomassie.** MM : Marqueur de poids moléculaire, ND : Non digérée, D : Digérée, FT : « Flowthrough », FTc : « Flowthrough » concentré, W : Lavage, E : Elution.

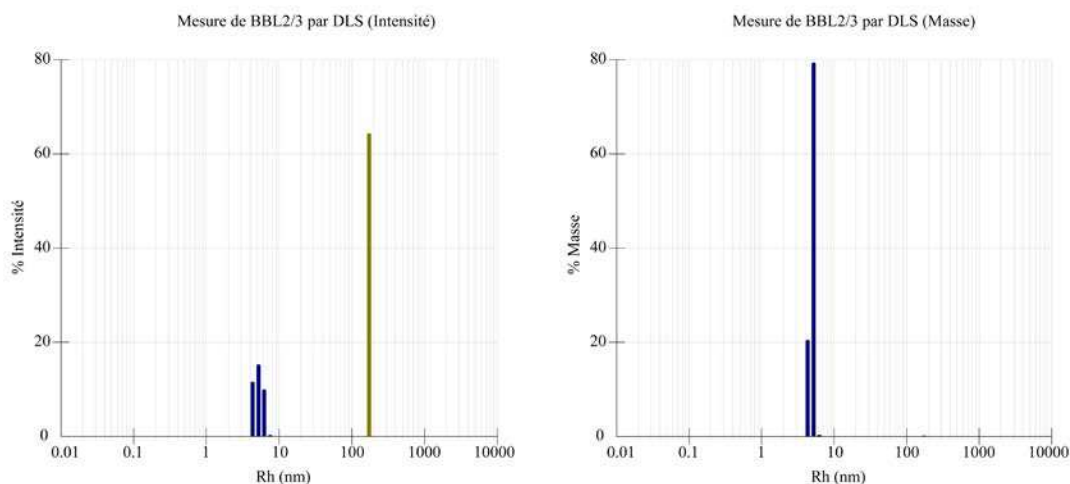
Les modifications apportées au tampon de purification de BBLF2/3 a permis de limiter la perte de protéines après clivage de l'étiquette de purification et de la concentrer efficacement par centrifugation. Le nouveau tampon de BBLF2/3 (20 mM Tris HCl pH 7,5 ; 250 mM NaCl ; 10 % glycérol ; 15 mM imidazol) permet d'obtenir jusqu'à 2 mg de protéine à partir de 5 g de cellules (21 flasques 300 cm<sup>2</sup> obtenus en 3 semaines de production), pour une concentration finale de 5 mg/mL.

## 2. Caractérisation biophysique par chromatographie d'exclusion de taille, MALLS et DLS

### 2.1. Diffusion dynamique de lumière (DLS, Dynamic Light Scattering)

La technique de diffusion dynamique de lumière permet de mesurer le rayon hydrodynamique de particules en solution. En assimilant la particule à une sphère cette technique permet aussi le calcul approximatif de sa masse. De plus cette technique permet d'estimer la polydispersité de la solution protéique de BBLF2/3.

L'étude de BBLF2/3 par DLS donne un rayon hydrodynamique de l'ordre de 5 nm, correspondant à une taille d'environ 145 à 150 kDa (Figure 60, %Masse). Une seconde taille de 200 nm est mesurée sur cet échantillon (Figure 60, %Intensité). Cette valeur correspond à la diffusion de lumière par des agrégats. Le signal fort de l'agrégat réduit la précision des mesures des particules de plus petites tailles. Ces histogrammes indiquent que la quasi-totalité de l'échantillon possède une taille de 5 nm. Cette information est très importante car la monodispersité d'une protéine est une des exigences pour la cristallisation des protéines.

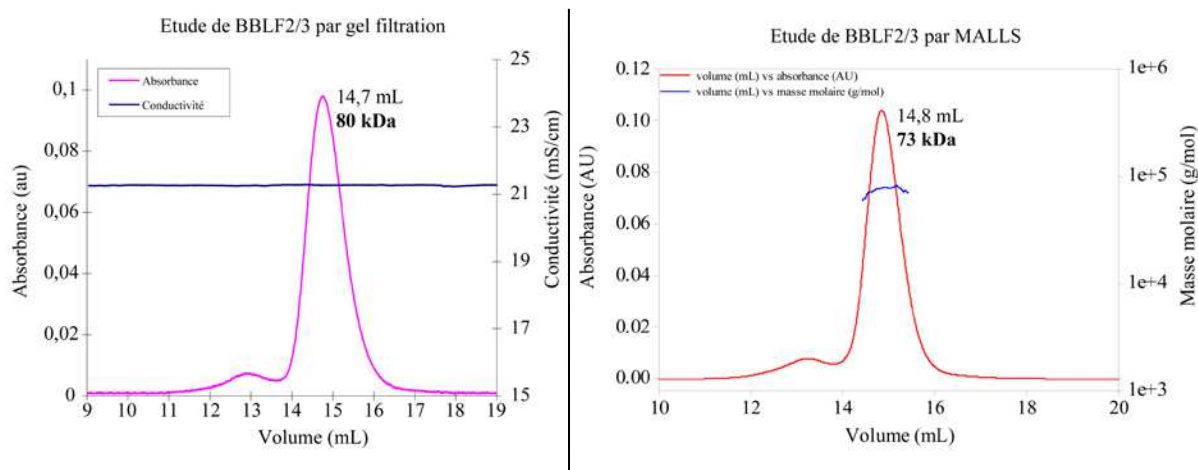


**Figure 60 : Histogrammes représentant le pourcentage d'intensité (à gauche) et du pourcentage de masse (à droite) en fonction du rayon d'hydrodynamique (Rh) de BBLF2/3.** La mesure a été réalisée avec une solution de BBLF2/3 à 1,5 mg/mL dans du tampon de dessalage (20 mM Tris HCl pH 7,5 ; 250 mM NaCl, 10 % glycérol ; 15 mM imidazol).

L'étude par DLS a révélé que la solution protéique de BBLF2/3 est globalement monodisperse avec une taille des particules de l'ordre de 5 nm suggérant plutôt une forme dimérique de BBLF2/3. Cette hypothèse doit être vérifiée par des expériences complémentaires.

## 2.2. Chromatographie d'exclusion de taille et MALLS

Pour vérifier l'état d'oligomérisation de BBLF2/3 en solution nous avons procédé à des analyses par chromatographie d'exclusion de taille. Cette technique permet d'évaluer la masse de la protéine, sa multimérisation et sa pureté. Une première mesure a été réalisée sur un système basse pression relié à la colonne SuperDex200 10/300. Le volume d'élution de BBLF2/3 est de 14,7 mL. D'après la calibration de la colonne ce volume d'élution correspond à une protéine globulaire ayant une masse d'environ 80 kDa. Cette valeur est très proche de la valeur théorique de 78 kDa prédite à partir de la séquence protéique de BBLF2/3. Pour une détermination plus précise de la masse de BBLF2/3, nous avons couplé la colonne SuperDex200 10/300 au système MALLS. Le MALLS n'utilise pas le volume d'élution de la protéine pour calculer sa masse, mais l'intensité de diffusion de la lumière et la concentration obtenue par réfractométrie. La masse calculée n'est pas influencée par la forme de la protéine. Le volume d'élution (14,8 mL) est quasiment identique à celui obtenu précédemment (14,7 mL), mais la masse de BBLF2/3 est plus faible que précédemment, 73 kDa. La taille mesurée expérimentalement est plus petite que celle calculée théoriquement à partir de la séquence en acides aminés de BBLF2/3 (78 kDa). Ainsi les deux approches montrent que BBLF2/3 est monomérique en solution avec une taille de 73 kDa.



**Figure 61 : Profil d'élution de BBLF2/3 d'une colonne S200 sur un système Basse Pression (à gauche) et MALLS (à droite).** Pour les deux mesures injection de 1,5 mg/mL de BBLF2/3 (20 mM Tris HCl pH 7,5 ; 250 mM NaCl ; 10% glycérol).

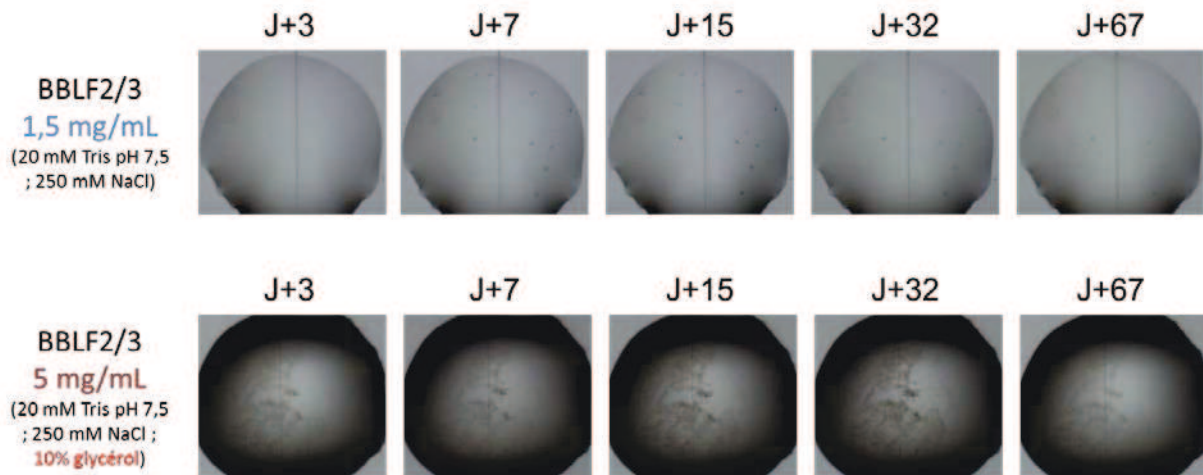
On peut remarquer que les deux chromatogrammes présentent un épaulement à gauche du pic principal. Aucune masse n'a pu être calculée pour ce pic avec le MALLS, car le signal détecté par le réfractomètre est trop faible. En se basant sur la calibration de la colonne et le volume d'élution de ce pic, on obtient une masse de 150 kDa. Ce pic est toujours présent sur les chromatogrammes, même si on réinjecte sur la colonne les fractions correspondant au pic principal. Ainsi la solution de BBLF2/3 semble contenir une large majorité de protéines monomériques avec une faible fraction de protéines dimériques. La différence obtenue pour la mesure de la taille de BBLF2/3 entre le DLS et le MALLS peut être expliquée par une forme peu compacte de BBLF2/3, car le volume d'élution surestime sa taille, et la précision de la mesure par DLS est réduite par la présence des agrégats. Le résultats obtenu grâce au MALLS est indépendant de la forme de la protéine et permet donc de mieux décrire l'état d'oligomérisation de BBLF2/3.

Les techniques biophysiques précédentes ont permis de décrire BBLF2/3 en solution. Cette protéine est monomérique avec une masse de 73 kDa (MALLS). La faible polydispersité de la solution permet d'envisager les tests de cristallisation (DLS).

### III. Tests de cristallisation de BBLF2/3

Les tests précédents ont montré que le tampon le plus adapté pour purifier BBLF2/3 était composé de 20 mM Tris HCl pH 7,5 ; 250 mM NaCl ; 10 % glycérol. Il a été possible d'obtenir BBLF2/3 à une concentration de 1,5 mg/mL en absence de glycérol, concentration suffisante pour des tests de cristallisation préliminaires. La plateforme de cristallisation (HTX Lab) a permis de réaliser un premier crible de conditions de cristallisation avec cet échantillon. Sur les 576 conditions testées 40% des gouttes étaient encore claires après 6 inspections (67 j). Une condition de cristallisation (100 mM

Hépès-Na pH 7,5 ; 100 mM Ammonium Sulfate ; 10% PEG4000) a fait apparaître au fur et à mesure des inspections des objets de forme carrée dans une goutte claire. Ces formes ont atteint une taille maximale à l'inspection 4 (15 j) et ont progressivement disparu au cours des inspections suivantes (Figure 62).



**Figure 62 : Clichés des inspections à J+3, J+7, J+15, J+32 et J+67 de BBL2/3 à 1,5 et 5 mg/mL.** Le réservoir contient 100 mM Hépès-Na pH 7,5 ; 100 mM Ammonium Sulfate ; 10 % PEG4000.

Des conditions de cristallisation ont été criblées pour échantillon à 1 mg/mL en présence de 10 % de glycérol (HTX Lab) afin d'estimer l'influence du glycérol sur la cristallisation de BBLF2/3. A cette concentration 95 % des gouttes étaient encore claires à la 6<sup>ème</sup> inspection. Soit la concentration de 1 mg/mL est trop basse pour initier la cristallisation, soit le glycérol ralentit la cinétique de cristallisation. La goutte correspondant à la condition de cristallisation précédemment décrite s'est collée aux parois du puits, rendant impossible toute interprétation. Le dernier test de cristallisation mené pour BBLF2/3 a été réalisé avec une solution protéique à 5 mg/mL avec la présence de 10 % de glycérol. A cette concentration seulement 10 % des gouttes étaient encore claires à la 6<sup>ème</sup> inspection. La condition de cristallisation décrite précédemment présente un précipité tout au long des 6 inspections.

Le fait que les formes disparaissent au cours du temps, laissent à penser qu'elles sont sûrement dues à une dégradation de la protéine d'intérêt au cours du temps. Le glycérol est connu pour avoir un effet important sur la cinétique de cristallisation, en ralentissant voir empêchant l'apparition de cristaux. Cette condition n'a pas été plus étudiée car elle n'a pas été reproduite à plus forte concentration.

## IV. Synthèse des résultats pour les protéines produites par des baculovirus porteurs d'un seul gène d'intérêt

L'expression des protéines du complexe H-P d'EBV par des baculovirus exprimant un seul gène d'intérêt n'a pas apporté les résultats escomptés. La primase et l'hélicase sont bien produites par ce système d'expression, mais restent dans la fraction insoluble. Des approches par purification en conditions dénaturantes ont permis d'obtenir les deux protéines dans la fraction soluble. Les phases de repliement des protéines n'ont pas permis de les renaturer sous forme soluble et en quantité suffisante. Le repliement des protéines *in vitro* est très empirique et la difficulté augmente quand les protéines ont une taille élevée (BBLF4 : 95kDa et BSLF1 : 100kDa). Aucune approche testée n'a permis de favoriser la solubilité de l'hélicase et de la primase.

L'utilisation de baculovirus portant un seul gène d'intérêt a été plus concluante pour l'étude de la sous unité accessoire (BBLF2/3). La protéine se produit très bien par cette approche et permet d'avoir la protéine sous forme soluble. Cette protéine est majoritairement monomérique en solution avec une masse de 73kDa. Pour caractériser la structure de cette protéine nous avons essayé de la cristalliser, une condition semblait prometteuse mais n'a pas pu être explorée plus en détail. Les prédictions d'ordre et de désordre suggèrent que la protéine possède une alternance de régions ordonnées et désordonnées. L'expression et la production de ces régions ordonnées de BBLF2/3 auraient probablement permis d'obtenir des cristaux. Cependant cette protéine ayant peu d'homologie avec les autres sous-unités accessoires des Herpèsvirus et n'étant pas prédite pour avoir une fonction enzymatique, j'ai préféré me concentrer sur le clonage permettant d'avoir les trois protéines portées par un seul baculovirus. Le clonage avec un baculovirus unique permet d'être sûr de produire simultanément les trois protéines du complexe dans une même cellule, favorisant ainsi la formation du complexe *in vivo* et permettant probablement la production du complexe soluble.

## CHAPITRE II :

Etude du complexe H-P  
d'EBV produit par un  
baculovirus portant les trois  
gènes d'intérêts



# I. Production du complexe par un seul baculovirus

La co-infection des cellules d'insecte par plusieurs baculovirus ne garantit pas que les trois virus infectent simultanément une même cellule. Pour éviter ce problème et obtenir la formation du complexe H-P d'EBV *in vivo*, nous avons décidé de cloner les trois gènes dans un seul et même baculovirus. Pour ce faire nous avons utilisé la technique appelée MultiBac.

## 1. Expression des protéines par le système MultiBac

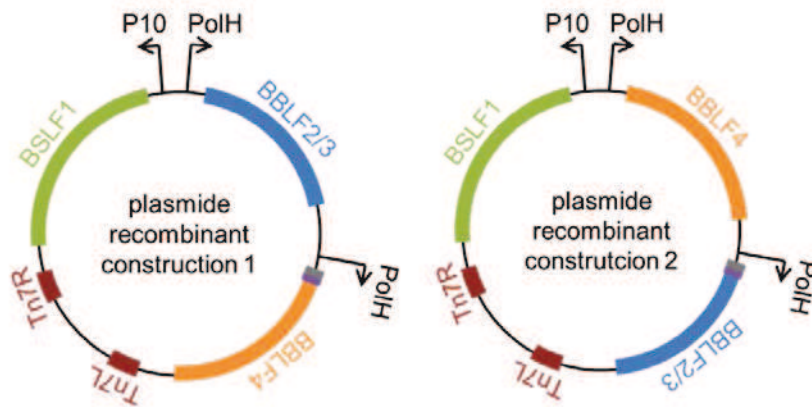
### 1.1. Le système MultiBac

Le système MultiBac, développé au sein de l'EMBL Grenoble par l'équipe d'Imre BERGER [166] permet d'insérer plusieurs gènes d'intérêt au sein d'un même bacmide. C'est une stratégie sophistiquée qui exige beaucoup de rigueur lors des différentes étapes du clonage. En plus des gènes d'intérêt, le bacmide porte un gène codant pour une protéine fluorescente : la YFP (« Yellow Fluorescent Protein »). C'est un gène rapporteur qui permet de suivre la production de protéine au cours de l'infection. La YFP est produite en même temps que les trois protéines du complexe clonées dans le bacmide. La mesure de la fluorescence de la YFP permet de suivre le niveau de production protéique au cours de l'infection. La fluorescence atteint normalement un plateau à 72 ou 96h post-infection. Lorsque ce maximum de production est atteint, on estime que les protéines d'intérêt sont aussi à leur maximum d'expression et les cellules sont alors récupérées pour les futures purifications.

### 1.2. La stratégie de clonage

Le complexe étant composé de trois protéines, plusieurs stratégies pour le positionnement de l'étiquette de purification étaient possibles. Les résultats que nous avons obtenus sur la protéine BBLF2/3 et les résultats décrits dans la littérature pour le complexe H-P d'HSV-1 [167] nous ont conduit à réaliser deux constructions pour le clonage MultiBac. La construction 1 (BSLF1 + BBLF2/3 + BBLF4his) possède l'étiquette de purification sur la protéine BBFL4 (similaire à l'étude réalisée chez HSV-1), la construction 2 (BSLF1 + BBLF4 + BBLF2/3his) possède l'étiquette hexahistidine sur la protéine BBLF2/3 (en accord avec nos résultats). La stratégie que nous avons mise en place pour le clonage est la même pour les deux constructions.

Les gènes non étiquetés ont été clonés dans un plasmide donneur (pUCDM), ce plasmide est ensuite ligué avec les pFastBac™HTB de BBLF4 (construction 1) et BBLF2/3 (construction 2), formant un plasmide recombinant contenant les trois gènes du complexe H-P (Figure 63).

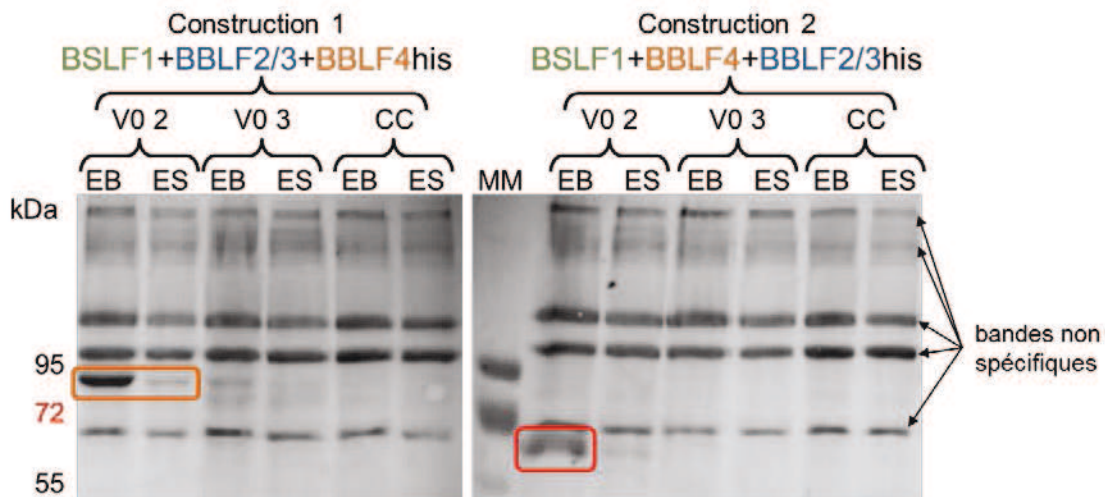


**Figure 63 : Représentation schématique des plasmides recombinants des constructions 1 et 2 contenant les trois gènes du complexe H-P d'EBV.**

Les trois gènes sont ensuite transposés dans un bacmide modifié codant pour le gène rapporteur YFP. Les bacmides recombinants sont transfectés dans les cellules d'insectes pour produire les baculovirus des constructions 1 et 2.

### 1.3. Test d'expression et de purification

La transfection des bacmides EmBacY1 et 2 dans les cellules SF21, permet la production des premières générations des baculovirus (V0) de chaque construction. Chaque baculovirus produit une protéine portant une étiquette hexahistidine, qui a été utilisée pour contrôler les produits de la transfection par Western Blot (Figure 64).



**Figure 64 : Test d'expression des baculovirus des constructions 1 et 2, détection de l'étiquette hexahistidine par Western Blot. MM : Marqueur de poids moléculaire. V0 2 et V0 3 : Virus issu des transfections numéro 2 et 3. CC : Contrôle Cellulaire. EB : Extrait Brut. ES : Extrait Soluble.**

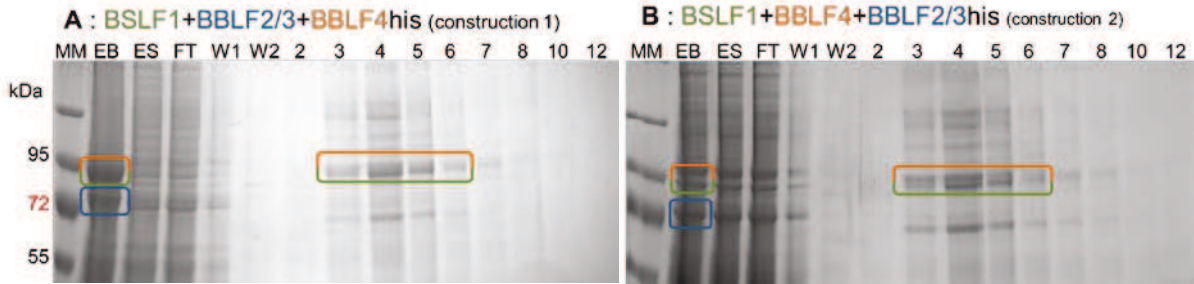
Pour la construction 1 (BSLF1 + BBLF2/3 + BBLF4his) une étiquette hexahistidine est détectée à 95 kDa. Cette taille correspond à la protéine BBLF4his. On remarque une forte intensité de la bande dans la fraction brute et la présence d'une bande de faible intensité dans la fraction soluble, indiquant que la protéine est bien surexprimée et qu'une partie de cette expression est soluble. Ce résultat est très encourageant pour la suite des expériences.

Pour la construction 2 (BSLF1 + BBLF4 + BBLF2/3his) une bande est observée avec le virus 2 (V0 2) avec une taille faisant moins de 70 kDa (Figure 64, rectangle rouge) ce qui est moins que la taille attendue. L'étiquette est présente sur BBLF2/3, or il a été montré précédemment que cette protéine possède une masse de 73 kDa par MALLS (Figure 61) et migre sur gel SDS-PAGE au-dessus du marqueur de taille 72 kDa (Figure 50). Le clonage du gène BBLF2/3 dans le système MultiBac a été réalisé à partir du pFastBac™HTB obtenus lors de la création des baculovirus à gène unique. La vérification des intermédiaires de clonage des deux approches par séquençage ont permis de valider la présence du gène complet de BBLF2/3. La présence de cette bande sur le gel est probablement due à une interaction non spécifique de l'anticorps anti-étiquette hexahistidine avec un contaminant tandis que la vraie bande de BBLF2/3 reste sûrement sous le seuil de détection.

Le Western Blot réalisé sur les cellules des deux transfections a permis de détecter l'étiquette hexahistidine de la construction 1 et montrant un profil non attendu pour la construction 2. Les deux baculovirus V0 2 des constructions 1 et 2 ont été sélectionnés pour l'amplification virale et procéder à des tests d'expression à plus grande échelle. Ces expériences permettront d'éclaircir le résultat observé pour le virus V0 2 de la construction 2.

## **2. Purification du complexe H-P d'EBV**

L'amplification des baculovirus des constructions 1 et 2 est effectuée sur des volumes de cellules de 25 mL. Les virus sont récoltés 48 h post infection, la culture cellulaire est maintenue jusqu'à obtention du plateau de fluorescence pour la YFP. Une purification par affinité sur colonne de nickel est réalisée à partir des cellules ayant servies à l'amplification virale.



**Figure 65 : Profils de purification par affinité sur colonne de nickel des constructions 1 et 2. A :** construction 1 BSLF1 + BBLF2/3 + BBLF4his. **B :** construction 2 BSLF1 + BBLF4 + BBLF2/3his. Gels SDS-PAGE à 8%, colorés au bleu de Coomassie. MM : marqueur de poids moléculaires. EB : Extrait Brut. ES : Extrait Soluble. FT : Flowthrough. W1 : Lavage 1. W2 : Lavage 2. Puits 2 à 12 : Fraction d'élution.

Les purifications des constructions 1 et 2, respectivement BSLF1 + BBLF2/3 + BBLF4his (Figure 65 A) et BSLF1 + BBLF4 + BBLF2/3his (Figure 65 B) donnent des profils équivalents sur gels SDS-PAGE.

## 2.1. Construction 1 (BSLF1 +BBLF2/3 + BBLF4his)

- **Analyse de la fraction brute :** La construction 1 possède l'étiquette hexahistidine sur BBLF4. La fraction brute contient 2 fortes bandes à 80 et 95 kDa. La bande à 80 kDa (rectangle bleu puits EB, Figure 65 A) correspond probablement à la sous-unité accessoire du complexe. La seconde bande à 95 kDa (rectangle orange et vert puits EB, Figure 65 A) peut être la primase et/ou l'hélicase du complexe. On peut déjà affirmer que le baculovirus de la construction 1 permet de produire au moins deux des protéines du complexe H-P d'EBV.
- **Analyse des fractions d'élution :** Elles révèlent une bande majoritaire à 95 kDa. Même si BSLF1 est proche de cette taille, il est plus probable que cette bande corresponde à BBLF4 car c'est la protéine portant l'étiquette hexahistidine. Mais on ne peut pas exclure que cette bande est seulement composée de BBLF4, il se peut que les protéines BBLF4 et BSLF1 migrent avec des tailles très proches ne permettant de les séparer sur le gel SDS-PAGE. Aucune bande n'est détectée vers 80 kDa, suggérant que la protéine BBLF2/3 n'est pas soluble avec ces conditions d'expression ou de purification. Une bande est présente sous le marqueur de taille 72 kDa, correspondant probablement à une protéine contaminante.

## 2.2. Construction 2 (BSLF1 + BBLF4 + BBLF2/3his)

- **Analyse de la fraction brute** : L'étude de l'extrait brut de la construction 2 révèle la présence de trois bandes, une à 80 et deux à environ 95 kDa. La première bande située à 80 kDa correspond probablement à la protéine BBLF2/3 (rectangle bleu puits EB, Figure 65 B). Les deux autres bandes situées de part et d'autre du marqueur de taille 95 kDa, correspondent probablement aux protéines BSLF1 et BBLF4. Il n'est pas possible d'attribuer chaque bande à sa protéine respective avec ce gel comme seul support. On pourrait attribuer ces bandes en se basant sur les masses théoriques des deux protéines 98 kDa pour BSLF1 et 90kDa pour BBLF4. Mais les travaux de Yokoyama & al (1999) [94] sur le complexe H-P d'EBV ont révélé que BSLF1 et BBLF4 migrent à des niveaux très proches, respectivement 89 et 90 kDa. L'attribution des bandes sera réalisée lors de l'identification des protéines. Par contre on peut affirmer que le baculovirus de la construction 2 permet d'exprimer les trois protéines du complexe H-P d'EBV.
- **Analyse des fractions d'élution** : Elles contiennent une double bande à proximité du marqueur de taille 95 kDa. Ces deux bandes correspondent sûrement aux protéines BSLF1 et BBLF4 du complexe (rectangle orange et vert, Figure 65 B). Tout comme pour la construction 1, on distingue une bande en dessous du marqueur de taille 72 kDa, suggérant la présence d'un contaminant. En revanche on ne détecte pas de bande à environ 80 kDa pouvant correspondre à la protéine BBLF2/3. Ce résultat est étonnant car c'est la protéine qui porte l'étiquette de purification. En regardant attentivement les puits FT et W1 de la Figure 65 B, on remarque qu'elles contiennent toutes deux la protéine BBLF2/3. Ainsi il semblerait que l'étiquette hexahistidine ne soit pas accessible avec la construction 2, ne permettant pas de purifier la protéine BBLF2/3. L'hélicase et la primase sont présentes dans les fractions d'élution, car ces deux protéines sont surexprimées et doivent interagir de façon non spécifique sur la colonne de nickel.

Les informations données par ces deux gels sont très intéressantes. Premièrement, la protéine BBLF2/3 est présente dans les fractions brutes de chaque purification mais pas dans les fractions d'élution. Son étiquette de purification est peut-être masquée par l'interaction avec les autres protéines du complexe, ne permettant pas sa purification par affinité. Deuxièmement, on observe les bandes pouvant correspondre à l'hélicase et la primase dans les fractions d'élution, indiquant que ces deux protéines sont produites sous forme soluble par le système d'expression MultiBac. Mais l'information principale est apportée par le gel de la construction 2, la présence de la double bande à 95 kDa indique qu'il est possible de co-purifier BSLF1 et BBLF4. Ainsi une seule

protéine étiquetée sur les trois semble être suffisante pour envisager la purification du complexe entier sur colonne d'affinité. Pour valider les résultats suggérés par ces gels, il est indispensable d'optimiser les conditions de purification de chaque construction afin d'obtenir les trois protéines du complexe sous forme soluble.

### 3. Optimisation des conditions de lyse

#### 3.1. Test de tampons de lyse en plaque 96 puits

Un crible de tampons de lyse ou de purification en plaque 96 puits a été développé au sein de l'institut. Dans le cadre du complexe H-P d'EBV, nous avons utilisé cette technique afin de trouver les conditions de lyse permettant une bonne solubilisation des trois protéines, et facilitant ainsi la purification de grande quantité de protéine. Dans ce but, nous avons établi une matrice de solubilité permettant de tester différents paramètres : 8 conditions de pH, 2 types de sels et 8 additifs différents (Tableau 11).

**Tableau 11 : Tableau détaillant les 96 conditions de tampons de lyse utilisés pour les tests de solubilisation des protéines produites par les constructions MultiBac 1 et 2.**

	1	2	3	4	5	6	7	8	9	10	11	12
A	Acide citric pH 5,5 NaCl 150mM	Acide citric pH 5,5 NaCl 500mM	Acide citric pH 5,5 NaCl 2M	Acide citric pH 5,5 SulfAmo 500mM	Acide citric pH 5,5 NaCl 500mM 10% glycerol	Acide citric pH 5,5 NaCl 500mM 25mM Glucose	Acide citric pH 5,5 NaCl 500mM 0,1% tritonx100	Acide citric pH 5,5 NaCl 500mM 250mM NDSB	Acide citric pH 5,5 NaCl 500mM 10mM Glycine	Acide citric pH 5,5 NaCl 500mM 50mM Imidazole	Acide citric pH 5,5 NaCl 500mM 2mM MgCl2	Acide citric pH 5,5 NaCl 500mM 0,5M UREA
B	MES pH 6 NaCl 150mM	MES pH 6 NaCl 500mM	MES pH 6 NaCl 2M	MES pH 6 SulfAmo 500mM	MES pH 6 NaCl 500mM 10% glycerol	MES pH 6 NaCl 500mM 25mM Glucose	MES pH 6 NaCl 500mM 0,1% tritonx100	MES pH 6 NaCl 500mM 250mM NDSB	MES pH 6 NaCl 500mM 10mM Glycine	MES pH 6 NaCl 500mM 50mM Imidazole	MES pH 6 NaCl 500mM 2mM MgCl2	MES pH 6 NaCl 500mM 0,5M UREA
C	MES pH 6,5 NaCl 150mM	MES pH 6,5 NaCl 500mM	MES pH 6,5 NaCl 2M	MES pH 6,5 SulfAmo 500mM	MES pH 6,5 NaCl 500mM 10% glycerol	MES pH 6,5 NaCl 500mM 25mM Glucose	MES pH 6,5 NaCl 500mM 0,1% tritonx100	MES pH 6,5 NaCl 500mM 250mM NDSB	MES pH 6,5 NaCl 500mM 10mM Glycine	MES pH 6,5 NaCl 500mM 50mM Imidazole	MES pH 6,5 NaCl 500mM 2mM MgCl2	MES pH 6,5 NaCl 500mM 0,5M UREA
D	HEPES pH 7 NaCl 150mM	HEPES pH 7 NaCl 500mM	HEPES pH 7 NaCl 2M	HEPES pH 7 SulfAmo 500mM	HEPES pH 7 NaCl 500mM 10% glycerol	HEPES pH 7 NaCl 500mM 25mM Glucose	HEPES pH 7 NaCl 500mM 0,1% tritonx100	HEPES pH 7 NaCl 500mM 250mM NDSB	HEPES pH 7 NaCl 500mM 10mM Glycine	HEPES pH 7 NaCl 500mM 50mM Imidazole	HEPES pH 7 NaCl 500mM 2mM MgCl2	HEPES pH 7 NaCl 500mM 0,5M UREA
E	HEPES pH 7,5 NaCl 150mM	HEPES pH 7,5 NaCl 500mM	HEPES pH 7,5 NaCl 2M	HEPES pH 7,5 SulfAmo 500mM	HEPES pH 7,5 NaCl 500mM 10% glycerol	HEPES pH 7,5 NaCl 500mM 25mM Glucose	HEPES pH 7,5 NaCl 500mM 0,1% tritonx100	HEPES pH 7,5 NaCl 500mM 250mM NDSB	HEPES pH 7,5 NaCl 500mM 10mM Glycine	HEPES pH 7,5 NaCl 500mM 50mM Imidazole	HEPES pH 7,5 NaCl 500mM 2mM MgCl2	HEPES pH 7,5 NaCl 500mM 0,5M UREA
F	Tris pH 8 NaCl 150mM	Tris pH 8 NaCl 500mM	Tris pH 8 NaCl 2M	Tris pH 8 SulfAmo 500mM	Tris pH 8 NaCl 500mM 10% glycerol	Tris pH 8 NaCl 500mM 25mM Glucose	Tris pH 8 NaCl 500mM 0,1% tritonx100	Tris pH 8 NaCl 500mM 250mM NDSB	Tris pH 8 NaCl 500mM 10mM Glycine	Tris pH 8 NaCl 500mM 50mM Imidazole	Tris pH 8 NaCl 500mM 2mM MgCl2	Tris pH 8 NaCl 500mM 0,5M UREA
G	Tris pH 8,5 NaCl 150mM	Tris pH 8,5 NaCl 500mM	Tris pH 8,5 NaCl 2M	Tris pH 8,5 SulfAmo 500mM	Tris pH 8,5 NaCl 500mM 10% glycerol	Tris pH 8,5 NaCl 500mM 25mM Glucose	Tris pH 8,5 NaCl 500mM 0,1% tritonx100	Tris pH 8,5 NaCl 500mM 250mM NDSB	Tris pH 8,5 NaCl 500mM 10mM Glycine	Tris pH 8,5 NaCl 500mM 50mM Imidazole	Tris pH 8,5 NaCl 500mM 2mM MgCl2	Tris pH 8,5 NaCl 500mM 0,5M UREA
H	Tris pH 9 NaCl 150mM	Tris pH 9 NaCl 500mM	Tris pH 9 NaCl 2M	Tris pH 9 SulfAmo 500mM	Tris pH 9 NaCl 500mM 10% glycerol	Tris pH 9 NaCl 500mM 25mM Glucose	Tris pH 9 NaCl 500mM 0,1% tritonx100	Tris pH 9 NaCl 500mM 250mM NDSB	Tris pH 9 NaCl 500mM 10mM Glycine	Tris pH 9 NaCl 500mM 50mM Imidazole	Tris pH 9 NaCl 500mM 2mM MgCl2	Tris pH 9 NaCl 500mM 0,5M UREA

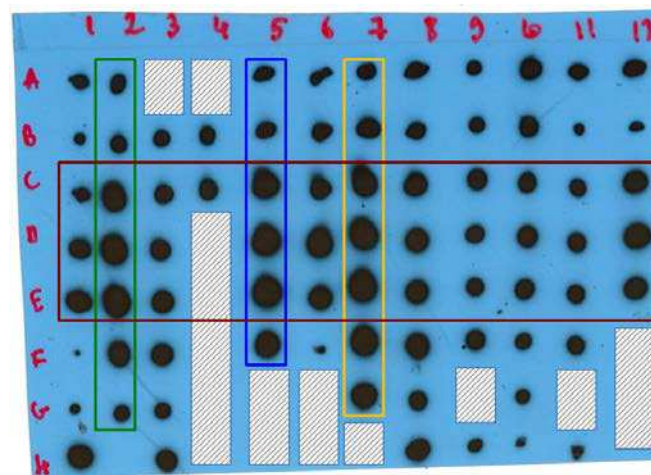


Pour réaliser ce test, les baculovirus des constructions 1 et 2 ont été utilisés pour des productions à grande échelle. Un total de 400 mL de cellules infectées ( $1.10^6$  cellules/mL) permet de réaliser un test de solubilité en plaque 96 puits. Après lyse par choc thermique et centrifugation les surnageants de chaque solution sont analysés par Dot Blot avec un anticorps anti étiquette hexahistidine.

### 3.2. Optimisation des conditions de solubilisation de la construction 1 (BSLF1 + BBLF2/3 + BBLF4his)

#### 3.2.1. Dot Blot du test de solubilisation en plaque 96 puits

Le Dot Blot (Figure 66) représente les résultats que nous avons obtenus pour la construction 1. Une analyse rapide des résultats montre quatre conditions semblant augmenter la solubilité de la protéine BBLF4his. Le rectangle de couleur rouge indique une plage de trois pH (6,5 ; 7 et 7,5) permettant une forte détection de la protéine BBLF4his. La concentration en sel la plus intéressante, repérée par un rectangle vert, semble être à 500mM NaCl. Deux additifs montrent un fort signal de BBLF4his, le glycérol (rectangle bleu) et le Triton-X100 (rectangle jaune). Cette expérience simple et rapide a permis de définir quatre conditions influençant la solubilité de la protéine BBLF4his. Comme ce test détecte seulement l'étiquette de purification, on ne peut pas être sûr que l'ensemble du complexe a été solubilisé.



**Figure 66 : Dot Blot de la construction 1 : BSLF1 + BBLF2/3 + BBLF4his dans 96 conditions de lyse.** La protéine BBLF4his est détectée par un anticorps anti étiquette hexahistidine. Les rectangles de couleurs repèrent les conditions intéressantes du test, **rouge** : le pH (6,5 à 7,5), **vert** : 500 mM NaCl, **bleu** : 10% glycérol et **gris** : 0,1% Triton-X100. Les rectangles blancs hachurés indiquent les puits avec des problèmes de dépôt de l'extrait soluble.

Cette expérience confirme avec certitude que l'on peut produire l'hélicase sous forme soluble grâce à la technique d'expression MultiBac.



### 3.2.2. Optimisation des tampons de purifications sur colonne de nickel

Les résultats du test en plaque 96 puits ont permis de définir 5 conditions de tampons de lyse. Les tampons vont permettre d'estimer l'efficacité et la synergie des différents composants pour la solubilisation du complexe. Le détail de chaque tampon de lyse est donné dans le Tableau 12. La purification sur colonne de nickel nécessite deux tampons supplémentaires pour le lavage de la colonne et l'élution de la protéine. Ces deux tampons possèdent les mêmes composants que le tampon utilisé pour la lyse cellulaire, avec en plus 1 M NaCl pour le tampon de lavage et 500 mM d'imidazol pour le tampon d'élution.

**Tableau 12 : Détails des tampons utilisés pour la purification par affinité sur colonne de nickel.** Création des tampons à partir des composants identifiés par le test de solubilité en plaque 96 puits.

Nom	Tampon	Sel	Additif 1	Additif 2
Tampon II	20 mM Hepes pH 7,6	500 mM NaCl		
Tampon III	20 mM Hepes pH 7,6	500 mM NaCl	10% Glycérol	
Tampon IV	20 mM Hepes pH 7,6	500 mM NaCl		0,1% Triton X-100
Tampon V	20 mM Hepes pH 7,6	500 mM NaCl	10% Glycérol	0,1% Triton X-100
Tampon VI	20 mM Hepes pH 7,6	500 mM NaCl	10% Glycérol	0,1% Triton X-100 <i>juste pour la lyse</i>
Affinement des conditions en présence de détergent				
Tampon VIa	20 mM HépeS pH 7,6	500 mM NaCl	10% Glycérol	0,44% $\beta$ -OG <i>juste pour la lyse</i>
Tampon VIb	20 mM HépeS pH 7,6	500 mM NaCl	10% Glycérol	1,2% $\beta$ -OG <i>juste pour la lyse</i>
Tampon VIIa	20 mM Hepes pH 7,6	500 mM NaCl	10% Glycérol	0,005% Triton X-100 <i>juste pour la lyse</i>
Tampon VIIb	20 mM Hepes pH 7,6	500 mM NaCl	10% Glycérol	0,01% Triton X-100 <i>juste pour la lyse</i>

A ce stade des résultats nous ne pouvons pas relier avec certitude chaque bande avec sa protéine respective. Pour simplifier l'analyse et la description des gels SDS-PAGE suivants, je distingue les trois bandes par une dénomination : P1, P2 et P3. La protéine 1 ou P1 correspond à la bande ayant le plus haut poids moléculaire, la protéine 2 (P2) à la bande intermédiaire et la protéine 3 (P3) à la protéine ayant le plus petit poids moléculaire. Les tampons II à IV nous permettent d'estimer indépendamment l'impact des différents composants identifiés par le test en plaque 96 puits. Les tampons VIa, VIb, VIIa et VIIb permettent d'affiner le type de détergent à utiliser pour une lyse efficace. Pour pouvoir comparer de façon reproductible les différentes purifications, nous avons utilisé un seul culot cellulaire divisé en trois volumes égaux pour réaliser les purifications.

**Tableau 13 : Analyse des protéines de la construction 1 (BSLF1 + BBLF2/3 + BBLF4his) éluées d'une colonne de nickel en utilisant 9 tampons différents.** Les gels SDS-PAGE 8% sont colorés au bleu de Coomassie. P1, P2 et P3 : Les protéines du complexe H-P d'EBV. F1, F2 : fractions d'éluion 1 et 2. → : Présence de la protéine. →→ : Présence probable de la protéine. ∅ : Absence de la protéine. ?P1/?P2 : Présence de la protéine 1 et/ou 2.

Tampons	Profil d'éluion sur colonne de nickel	Observations
<b>Tampons II</b> (seulement sel)		-Présence des protéines P1 et P2 à proximité du marqueur de taille 95 kDa -Intensité de la bande P3 est très faible
<b>Tampon III</b> (sel + glycérol)		-Bande proche de 95 kDa présence des protéines P1 et/ou P2 -Pas de bande P3 observée vers 80 kDa
<b>Tampon IV</b> (sel + détergent)		-Bande proche de 95 kDa présence des protéines P1 et/ou P2 -Pas de bande P3 observée vers 80 kDa
<b>Tampon V</b> (sel + glycérol + détergent)		-Présence des trois bandes, P1 et P2 proche de 95 kDa et P3 vers 80 kDa
<b>Tampon VI</b> (sel + glycérol + détergent) <i>détergent dans le tampon de lyse</i>		-Présence des trois bandes, P1 et P2 proche de 95 kDa et P3 vers 80 kDa
<b>Tampon VIa</b> (sel + glycérol + détergent) <i>β-OG (0,44% m/v) dans le tampon de lyse</i>		-Forte intensité des bandes P1 et P2 à 95 kDa -Absence de la protéine P3
<b>Tampon VIb</b> (sel + glycérol + détergent) <i>β-OG (1,2% m/v) dans le tampon de lyse</i>		-Bande proche de 95 kDa présence des protéines P1 et/ou P2 -Absence de la protéine P3
<b>Tampon VIIa</b> (sel + glycérol + détergent) <i>Triton X-100 (0,005% m/v) dans le tampon de lyse</i>		-Forte intensité des trois bandes, P1 et P2 vers 95 kDa et P3 proche de 80 kDa
<b>Tampon VIIb</b> (sel + glycérol + détergent) <i>Triton X-100 (0,01% m/v) détergent que pour la lyse</i>		-Forte intensité des trois bandes, P1 et P2 vers 95 kDa et P3 proche de 80 kDa

Les tampons II, III et IV permettent de tester indépendamment l'effet du sel, du glycérol et du Triton X-100 sur la solubilisation des protéines du complexe H-P. Les tampons III (sel + glycérol) et IV (sel + détergent) présentent une seule bande vers 95 kDa dans les fractions d'élution de la colonne de nickel (Tableau 13). Cette bande peut contenir les protéines P1 et/ou P2. La protéine P3 n'est pas présente dans les fractions d'élution. Le tampon II présente un profil d'élution différents des deux tampons précédents. On distingue clairement la présence deux bandes vers 95 kDa et une bande de très faible intensité vers 80 kDa. Ainsi c'est le tampon II qui permet une meilleure purification du complexe. Ce résultat est étonnant car ce tampon contient seulement la présence de 500 mM NaCl. Les tampons III et IV contiennent respectivement du glycérol et du Triton-X100 en plus des 500 mM NaCl. La présence de ces composants devrait faciliter la purification du complexe car le glycérol permet de stabiliser les protéines alors que le Triton-X100 favorise leur solubilité.

Les composants (NaCl, glycérol et Triton X-100) ayant permis une détection forte de la protéine BBLF4 lors du test en plaque 96 puits ne permettent pas l'obtention des trois protéines dans les fractions d'élution lorsqu'ils sont utilisés séparément.

Les tampons V et VI ont été réalisés pour évaluer si ces différents composants peuvent avoir un effet synergique permettant la purification simultanée des trois protéines. La différence entre ces deux tampons dépend du Triton X-100. Pour la condition V il est présent dans tous les tampons utilisés pour la purification, alors que dans la condition VI il est juste présent dans le tampon de lyse. Le Triton X-100 forme facilement des micelles qui sont pénalisants pour l'étude structurale des protéines. Ces deux conditions de tampons vont permettre d'estimer si l'utilisation du Triton-X100 seulement pour la lyse est suffisant pour une bonne purification du complexe.

Ces deux tampons permettent d'observer clairement les bandes P1, P2 et P3 dans la fraction d'élution F2 de la colonne de nickel. Le sel, le glycérol et le détergent possède un effet synergique pour la purification des protéines du complexe H-P. On remarque même que les bandes obtenues avec la condition VI ont une intensité plus importante que celles de la condition V. C'est une observation très intéressante car l'absence de détergent dans les tampons de lavage et d'élution n'altère pas la purification des trois protéines. Ainsi le Triton X-100 peut être éliminé dès la phase de purification.

Le complexe H-P entier peut être purifié à partir de la construction MultiBac 1 (BSLF1 + BBLF2/3 + BBLF4his), avec un tampon composé de sel (NaCl 500 mM), de glycérol (10 %), d'un détergent (Triton X-100 0,1%) et d'un pH fixé à 7,6 (20 mM Hépès).

### 3.2.3. Affinement des conditions d'utilisation des détergents

Dans le test en plaque 96 puits, les additifs sont utilisés à des concentrations relativement élevées afin de « forcer » leur rôle et voir plus facilement les effets qu'ils engendrent sur la ou les protéines d'intérêts. Jusqu'à présent nous avons utilisé le Triton X-100 à 0,1% (m/v), à cette concentration nous sommes largement au-dessus de sa concentration micellaire critique (CMC) de 0,02% (m/v). Afin d'optimiser plus précisément le tampon de lyse, deux nouvelles concentration en Triton X-100 ont été testées, 0,01% (<CMC) et 0,005% (<<CMC). Le Triton X-100 forme très facilement des micelles puisque sa CMC est très basse, de plus les micelles formées possèdent une taille de 90 kDa équivalente aux protéines d'intérêts. Un autre détergent non ionique, l'Octylglucoside ( $\beta$ -OG) a été testé. Il possède deux avantages comparé au Triton X-100, une CMC de 0,73% (m/v) et des micelles de 8kDa. Comme ce tensio-actif possède des micelles de petites tailles, nous avons testé deux concentrations, une supérieure (1,2 %) à sa CMC et l'autre inférieure (0,44%). Quatre nouveaux tampons ont été réalisés pour évaluer l'efficacité de ces détergents sur la purification des protéines du complexe (Tableau 12).

Les tampons VIa et VIb ont pour détergent le  $\beta$ -OG, lorsque sa concentration (0,44%) est inférieure à sa CMC, les bandes P1 et P2 sont observées dans les fractions d'élution de la colonne de nickel et possèdent une forte intensité. En revanche lorsque la concentration (1,2%) de ce détergent est supérieure à sa CMC, on n'observe plus qu'une seule bande vers 95 kDa pouvant représenter P1 et/ou P2. Quelques soit la concentration utilisé par ce détergent, on n'observe pas de bande correspondant à la protéine P3 (80 kDa). Ainsi les propriétés du détergent  $\beta$ -OG ne semblent pas appropriées pour permettre la purification du complexe H-P d'EBV.

Les tampons VIIa et VIIb contenant du Triton X-100 permettent d'obtenir les trois bandes P1, P2 et P3 dans les fractions d'élution de la colonne de nickel. Il n'est pas évident de déterminer quelle concentration en Triton X-100 est la plus adaptée pour la purification du complexe. Il semblerait que l'intensité des bandes éluées avec le tampon VIIb (0,01% Triton X-100) est plus forte qu'en présence du tampon VIIa. De plus la concentration en détergent du tampon VIIb est plus proche de sa CMC, accentuant plus fortement les propriétés de déstabilisation des bicouches lipidiques et donc favorisant la purification des protéines du complexe. Nous avons finalement choisi d'utiliser le tampon VIIb pour la purification des protéines du complexe H-P d'EBV. Pour plus de simplicité il sera nommé tampon VII dans la suite du manuscrit.

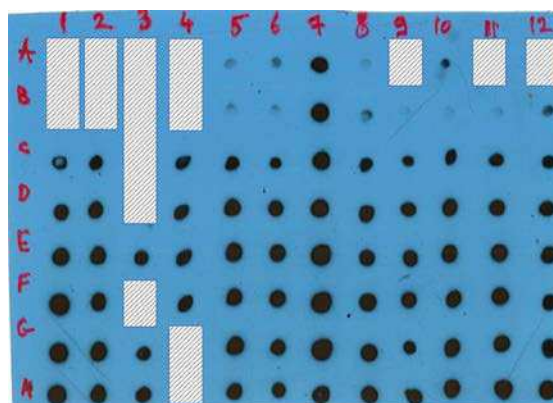
La phase d'optimisation des tampons de purification a permis de montrer que nous pouvons purifier les trois protéines du complexe H-P d'EBV par affinité sur colonne de nickel. La construction 1

ne possède qu'une seule étiquette hexahistidine présente sur la protéine BBLF4. La présence des deux autres protéines du complexe au sein des fractions d'élution indique qu'elles interagissent avec BBLF4. Une information sur la stœchiométrie du complexe peut être déduite de ces gels, car on observe que les trois bandes protéiques possèdent une intensité très similaire. Cette observation suggère que le complexe suit une stœchiométrie de type 1/1/1, en accord avec les résultats de Yokoyama & al (1999) [94]. Ils ont aussi identifié chaque protéine du complexe expérimentalement, par analogie avec leurs résultats on peut relier les bandes à leur protéines respectives, bande P1 : BBLF4his, bande P2 : BSLF1 et bande P3 : BBLF2/3.

### 3.3. Optimisation des conditions de solubilisation de la construction 2 (BSLF1 + BBLF4 + BBLF2/3his)

#### 3.3.1. Dot Blot du test de solubilisation en plaque 96 puits

La même expérience a été réalisée avec la seconde construction : BSLF1+BBLF4+BBLF2/3his. Le Dot Blot (Figure 67) donne des résultats beaucoup moins clairs que pour la première construction. Aucune condition ne présente une forte détection de la protéine BBLF2/3his. Malgré tout il semble se dégager de ce test que les faibles pH ne sont pas adaptés pour la solubilisation du complexe lors de la lyse. Deux conditions semblent augmenter sensiblement la solubilité de la protéine BBLF2/3his, la condition à 150 mM NaCl (Figure 67, colonne 1) et la condition avec le détergent Triton-X100 à 0,1% (Figure 67, colonne 7). Le Triton-X100 est une condition favorable pour la solubilisation des deux protéines portant l'étiquette de purification. En revanche les conditions en sel, favorisant la solubilité de chaque protéine, sont très différentes. Il n'est pas si étonnant que deux protéines appartenant au même complexe aient des comportements différents en fonction de la force ionique du tampon.

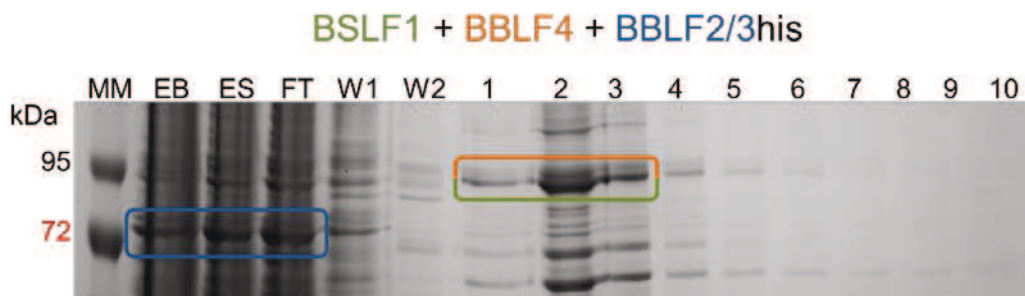


**Figure 67 : Dot Blot de la construction 2 : BSLF1 + BBLF4 + BBLF2/3his dans 96 conditions de lyse.** La protéine BBLF2/3his est détectée par un anticorps anti étiquette hexahistidine. Les rectangles blancs hachurés indiquent les puits pour lesquels aucune fraction soluble n'a pu être récupérée.

Le test en plaque 96 puits est beaucoup moins parlant que celui obtenu pour la construction 1. Il met tout de même légèrement en avant les conditions 150 mM NaCl et 0,1% Triton X-100. Ces deux composants ont été utilisés lors d'une purification sur colonne de nickel pour estimer leur potentiel de solubilisation des protéines du complexe H-P.

### 3.3.2. Optimisation des tampons de purification sur colonne de nickel

Une purification a été réalisée en lysant les cellules d'insectes dans le tampon : 50 mM Hépès pH 7,5 ; 150 mM NaCl ; 0,01% Triton X-100 ; 15 mM imidazol. Les tampons de lavage et de d'élution sont les mêmes que ceux utilisés pour la purification de BBLF2/3 seule. Les fractions d'élution contiennent un grand nombre de bandes en plus de celles attendues. Les bandes entourées par un rectangle orange et vert (Figure 68) peuvent correspondre à BSLF1et/ou BBLF4. Il n'y a toujours pas la présence d'une bande à 80 kDa pouvant correspondre à BBLF2/3 dans les fractions d'élution. On observe la présence de BBLF2/3 dans les puits, EB, ES et FT, indiquant que la protéine est produite et soluble, mais qu'elle ne reste pas fixée à la colonne de nickel. Un clivage à la protéase TEV a été réalisé sur les fractions d'élution, mais aucune bande n'a révélé une réduction de taille. Toutes ces informations semblent indiquer que l'étiquette de purification de la protéine BBLF2/3 n'est plus accessible une fois le complexe formé. BBLF2/3 interagirait avec au moins un de ses partenaires par son extrémité N-terminale.



**Figure 68 : Profil d'élution de la construction 2 (BSLF1 + BBLF4 + BBLF2/3his) sur une colonne de nickel.** Le gel SDS-PAGE 8% est coloré au bleu de Coomassie. MM : Marqueur de poids moléculaires. EB : Extrait brut. ES : Extrait soluble. FT : « Flowthrough ». W1 : Lavage1. W2 : Lavage2. 1 – 10 : Fractions d'élution.

Malgré les modifications apportées au tampon de lyse, le profil d'élution de la colonne d'affinité n'a pas été amélioré. La construction 2 semble bien exprimer la protéine BBLF2/3 mais son étiquette n'est pas accessible pour permettre sa purification. Nous avons donc décidé de ne pas poursuivre les tests sur la construction 2 et de porter tous nos efforts sur les tests réalisés avec la construction 1 (BSLF1 + BBLF2/3 + BBLF4his).

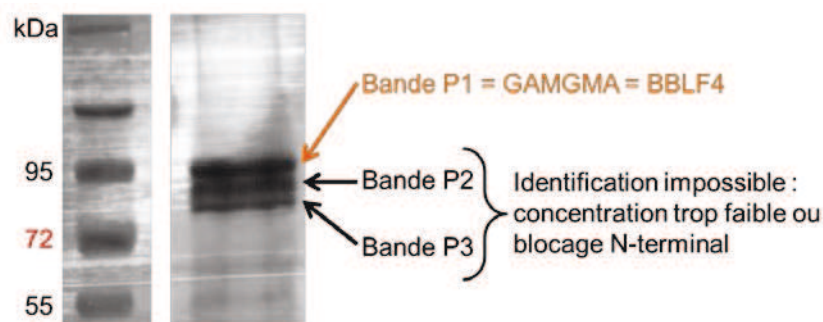
## II. Confirmation des protéines produites par la construction MultiBac 1

Nous avons pu déterminer le système d'expression et les conditions de purification permettant d'obtenir le complexe H-P d'EBV sous forme soluble après purification par affinité. Une identification des trois bandes présentes sur gel SDS-PAGE est nécessaire.

### 1. Identification par séquençage N-terminal

La nomenclature des bandes P1, P2 et P3, établie précédemment est conservée pour la description des résultats de séquençage N-terminal (Figure 69). L'échantillon transmis à la plateforme de séquençage a été digéré par la protéase TEV afin de retirer l'étiquette hexahistidine présente sur l'hélicase. Des trois bandes soumises, seule la bande P1 a pu être séquencée. La séquence identifiée dans la bande P1 est GAMGMA. Elle correspond à la séquence de la protéine BBLF4 après clivage de l'étiquette de purification par la protéase TEV, les résidus GAMG correspondent aux résidus du site de clivage TEV restant après action de la protéase et les résidus MA sont les deux premiers résidus de la séquence de BBLF4. Ce résultat concorde avec la détection de BBLF4 par Western Blot anti étiquette polyhistidine. Cette identification permet de relier la bande P1 observée sur gel SDS-PAGE à l'hélicase du complexe d'EBV.

Aucune séquence n'a pu être identifiée pour les bandes P2 et P3. Ce résultat négatif est probablement dû à une modification en N-terminal, bloquant la réaction chimique nécessaire à l'identification.

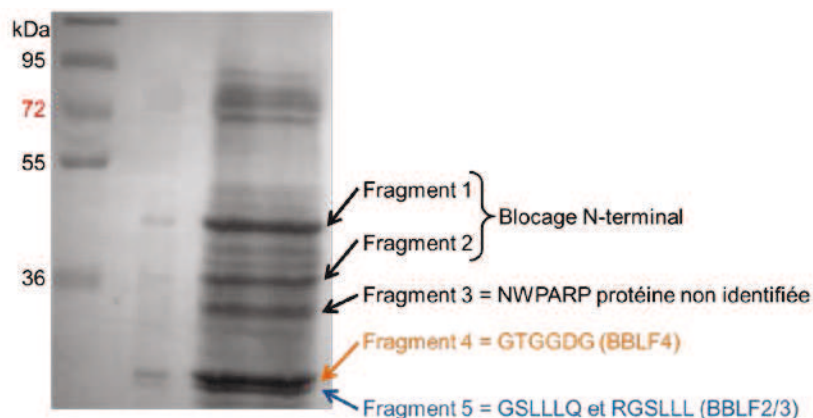


**Figure 69 : Membrane PVDF colorée au bleu portant les bandes utilisées pour le séquençage N-terminal.** La bande P1 contient la séquence GAMGMA, attribuée à la protéine BBLF4. Les deux autres bandes P2 et P3 n'ont pas pu être identifiées.

Pour résoudre le problème du blocage N-terminal, les protéines du complexe ont été digérées à la trypsine pour créer des fragments protéines BBLF2/3 et BSLF1. La digestion ménagée à la trypsine a permis de créer 5 fragments de tailles différentes à celle des protéines d'origine (Figure



70). Les fragments 1 et 2 n'ont pas pu être séquencés car leurs extrémités N-terminales sont bloquées. Ces deux fragments portent sûrement les extrémités d'origine des protéines BSLF1 et/ou BBLF2/3. Le fragment 3 contient les six résidus formant la séquence NWPARP, cette séquence n'est retrouvée dans aucune des protéines formant le complexe H-P d'EBV. Le fragment 4 présent sur la membrane PVDF possède la séquence GTGGDG qui est portée par BBLF4 à partir du résidu 399, ce qui confirme la production sous forme soluble de BBLF4. Enfin la cinquième et dernière bande contient deux séquences, GSLLLQ et RGSLLL commençant respectivement aux résidus 361 et 360 de la protéine BBLF2/3.



**Figure 70 : Profil de digestion ménagée du complexe H-P d'EBV par la Trypsine.** Mise en avant des 5 fragments choisis pour le séquençage N-terminale.

Cette approche par digestion ménagée a permis de surmonter le problème du blocage en N-terminal de BBLF2/3, mais malheureusement nous n'avons toujours pas confirmé que BSLF1 fait partie des trois bandes que nous observons sur gel SDS-PAGE. La séquence identifiée pour le fragment 3 ne correspond à aucune des trois protéines du complexe. Ce résultat est un peu déroutant, car il indique la présence d'une protéine contaminante. Si ce fragment provient de la bande P2, cela révèle que nous ne surexprimons pas la protéine BSLF1. Cette conclusion n'est pas définitive car la digestion à la trypsine a été réalisée sur un échantillon en sortie de colonne de nickel. A ce niveau de la purification, il y a encore la présence de protéines contaminantes qui auraient pu générer ce fragment.

Le séquençage N-terminal a permis d'identifier les bandes correspondantes aux protéines BBLF4 (bande P1) et BBLF2/3 (bande P3) du complexe H-P d'EBV. L'incertitude sur la présence de la protéine BSLF1 persiste.

## 2. Identification par spectrométrie de masse

La technique utilisée pour l'identification des protéines en spectrométrie de masse est appelée « Fingerfoot Printing ». Chaque bande P1, P2 et P3 est digérée au sein du gel par la trypsine. Les peptides générés par la protéase sont extraits du gel puis analysés par un spectromètre MALDI-TOF. On obtient alors une carte massique expérimentale pour chaque bande P1, P2 et P3. Chaque carte de peptides expérimentale est comparée avec les cartes théoriques générée *in silico* à partir des séquences de chaque protéine du complexe H-P d'EBV.

- **Identification de la bande P1** : La carte peptidique expérimentale de P1 est composée de 41 peptides. L'utilisation du moteur de recherche d'empreintes peptidiques Mascot, a permis de faire correspondre P1 à la protéine BBLF4. Treize peptides générés expérimentalement à partir de la bande P1 sont identifiés dans la séquence de BBLF4 et représentent une couverture de 15% de la séquence de BBLF4.
- **Identification de la bande P2** : La carte peptidique expérimentale de la bande P2 contient 59 peptides. L'utilisation de Mascot a permis d'aligner 16 peptides sur la séquence de la protéine BSLF1, correspondant à une couverture de 16%. Pour vérifier ce résultat très intéressant, la carte de peptides expérimentaux de P2 a été comparée avec celles générées à partir des protéines correspondant à la taxonomie « virus » de la base de données SwissProt. La protéine identifiée dans cette taxonomie correspond bien à la protéine BSLF1.
- **Identification de la bande P3** : La carte expérimentale issue de P3 dénombre 49 peptides, 9 de ces peptides ont été attribués à la séquence de BBLF2/3 par le logiciel Mascot. Les 9 peptides recouvrent 14% de la séquence de BBLF2/3.

<b>BSLF1</b>				
1	MSAPVVIKAL	VASNTDIAEA	ILDAILSRPD	EGFRLFCLCH NASPLHHVAG
51	SLVELQLHLP	<b>KKRLTSQSRC</b>	GLVLTLLHLP	EEAFPFRL <b>GL TPLTADRLST</b>
101	<b>YLDR</b> AGALRS	LTPLVELLTL	SAKKQPQDA	RGRVAVLRPK IVGCLRRIYR
151	VNISARWFIS	TFGSHEAQFV	LVTAAAYFWG	IPCTIETLAH LTELFTSESG
201	QSLAAVTSLA	ELGEVFGSSA	WAEQTEAFAH	FAHEKLRDSD REIR <b>AVARTI</b>
251	<b>DAYR</b> GRLPLA	SADLVRYVYL	AHAQCFNEGT	FKRYSQLTSM GEIGCLPSGG
301	VVLPSLLDRG	FAEHMRTYFT	<b>RETYLAEHVR</b>	VQQLKIRMPEP POPYTWDPPD
351	DDGLMRAWAG	LSVDVARE <b>LV ELAR</b> WHADEG	PTYPTLQGF	LCLAGQATCR
401	GQWNP <b>KEQFL</b>	<b>PPTVLR</b> RVQR	LPVFLCHFAD	RHYFVMTAAD PFSSHLAEVV
451	STPTNCRLPD	TCLTRALSYT	PVYYSQNSLS	EQLFVSR <b>HEY FNPRL</b> PVCNL
501	VLDLCLKIKG	APWSLEEIYD	LCRTVRRVVL	RLMRRLGPPVS <b>RAHPVYFFKS</b>
551	ACPPADPDNM	EDVLPFCICT	GKLGFR <b>VITP LPR</b> GHAIVGT	SAVQGFVSVL
601	QKLMGLTACL	RRMRHKIKEI	GAPLFDGCVY	HAGRCIR <b>LPH TYK</b> VDRGGGL
651	<b>SRQLR</b> LFVCH	PEEEDKHSYV	KNALNIQNL	HSHLVGWPA PKTFCYHIAD
701	DGR <b>DYLIQRT</b>	RETLPPTVEN	VCAMIEGHLG	LDLVAWVSSC IWPSLMSTLA
751	TAVPEDKFPQ	FLHVTFEQTG	PNLVQVCHAR	GRNFACL RHT HRASSKNVRV
801	FLVLYYTSQA	ITVTFMSQCF	AGR <b>CGANQPT AHFSISVPAS</b>	<b>RI</b> INRAEASQ
851	DSTTSQLARR	<b>RDRQDGSFSE TLP</b>		

Couverture de 15% de la séquence.

<b>BBLF2/3</b>				
1	MMETPAESVR	AR <b>VSSVT</b> FYN	<b>VTQTAGR</b> WWA	IWVVGIVPIK REDVETLIVV
51	QACQPPLGGS	LEPPVVNAPS	TTELNFLRWE	RELRRSGGLI AMLADAAEKD
101	LFDLSEFRTRD	RRLLSAARVE	DEQGLIFQPL	FPAQVVCQSC SGDDGRDQQP
151	PPVDGFGSEM	EGEQTCPHAQ	RHSESPGQLD	VYIRTPRGDV FTYSTETPDD
201	PSPVPFRDIL	RPVTYEVDLV	SSDGATGRGG	DARRHRVSLK ILEPAGGFES
251	WLVNSWSMAG	GGLYAFRLSI	YASCYANHRG	TKPIFYLLDP ELCPPGSDFFQ
301	PYVPGFPFLP	IHYVGRARPA	FWRHAPHSEG	LLLLDLNLGV SGTPLADALL
351	GLDARSGQRR	GSLLLQQIWP	PTRK <b>EINPRH</b>	<b>VCTREGGEGG GEDETTVVGR</b>
401	AEATAILEAD	ATWWLYELAR	CHLSARGAPV	<b>GTPDGGGQAR DAQTWLR</b> ALH
451	<b>RYGTS</b> DTTRA	<b>LGGLY</b> TAVTR	VLLHAAADLG	LTWAYADEFI LGFVAP TSAH
501	PSEEP LAQAF	LQGV <b>KDSEDA</b>	<b>SR</b> LDRDVMGG	EATVARRHIR VKARRGPGL
551	LMAIFQGDLY	VGGCREHSGP	FLVWHEAFSW	TLDQLAARPE ADKAPP SHDH
601	LLTLVRDLTR	RLAPGRRNR	FWALPRAWLQ	RLRRAGLRLS GSHVCLLDKD
651	GARPAPCQTA	TEHGLSPTAY	FREIMAFLLD	VISALHPGYT IPMEITR <b>ETD</b>
701	<b>LLMTVLSLF</b>			

Couverture de 14% de la séquence

<b>BBLF4</b>				
1	<b>MAEE</b> PRAPEA	<b>LSSTF</b> MLNMT	<b>SDASVRR</b> IVR	<b>RIGTLARRV</b> <b>QQLPDMETFS</b>
51	<b>PEFDE</b> ELSEP	PFLPFSAYVI	TGTAGAGKST	SVSCLHHTMD CLVTGATTVA
101	AQNLSQTLRA	YCPTVYSAFG	FKSR <b>HINMTQ</b>	<b>RVSSHGRSTD AALEELQR</b> RD
151	LAKYWPVLS	IAAEFRRTKP	<b>RGLYSGVSGP</b>	<b>AFEVLR</b> DMHQ GQLWTTNVIV
201	VDEAGTSLVH	ILTAVVFCYW	FFNAWLRTPL	YRRGRIPCIV CVGSPTQ TDA
251	FQSSFSHETQ	VNKIRECDNI	LTFVLGNPRA	<b>ATYVDVAR</b> NW ALFINNKRCT
301	DVQFGHLMKT	LEYGLELSPD	ILAYVDRFVV	PRAAIMDPAQ YVGWTR <b>LFLS</b>
351	<b>HAEVK</b> TFLT	LHATLKTAGQ	GRAARGTGGD	GGGVMTMFTCP VECEVFLDPL
401	AQYKTLVGLP	GLTAHTWLQK	NYARLGNYSQ	FADQDMVPVG TEQDEERV <b>KV</b>
451	<b>TYNVTYVKHS</b>	<b>SVSVNCK</b> TKK	SICGYTGTFG	DFMDTLEADS FVEAHGHEQP
501	EYVYSFLARL	IYGGIYAFSH	GGHSLCENGE	YVAELGAVPL PGRTWDPEVT
551	AGMELGELPL	EVAWDGERS <b>P</b>	<b>AAVFYAR</b> VLA	PPAANSAPLC SLLNIYNDLR
601	AYFRQCLDVA	VRYGGREFRD	LPFCTFTNNM	LIRDNIEFTS DEPLLHGLLD
651	YASTTENYTL	LGYTHLNVFF	GIRGK <b>QQPQD</b>	<b>AGSSR</b> MPRLM VKDEAGFVCC
701	LEHNTNKLYE	<b>TIEDK</b> SLNLC	SIRDYGISSK	<b>LAMTI</b> AKAQQ LSLNKVAICF
751	GSHR <b>NIKPGH</b>	<b>VYVALSR</b> ARH	SNCVVMDRNP	LSEMITGEGN PASGYIVDAL
801	KNSRALLVY			

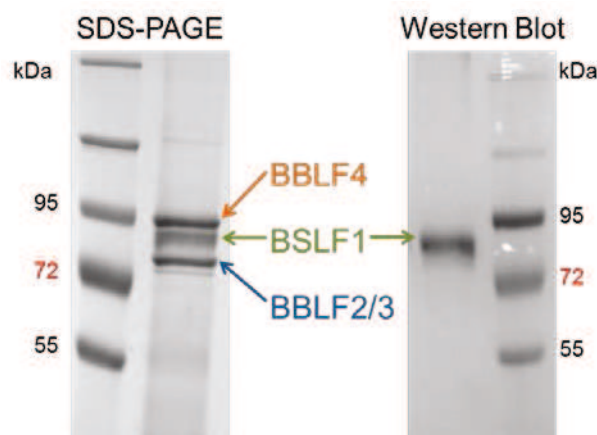
Couverture de 15% de la séquence

Figure 71 : Couverture des séquences protéiques des protéines BSLF1, BBLF2/3 et BBLF4 par les peptides identifiés expérimentalement en rouge.

La spectrométrie de masse a permis d'attribuer la bande P2 à la protéine BSLF1. Elle confirme que l'utilisation d'un seul baculovirus contenant les trois gènes d'intérêt, permet la production des trois protéines sous forme soluble et la formation du complexe *in vivo*.

### 3. Identification par Western Blot

En parallèle des tests effectués par spectrométrie de masse, nous avons pu obtenir des anticorps anti BSLF1 générés par l'équipe du Dr Tatsuya TSURUMI. Cet anticorps a été utilisé dans la publication montrant une stœchiométrie de type 1/1/1 du complexe H-P d'EBV [94]. Cette anticorps est spécifique de BSLF1 et ne provoque donc pas de réaction croisée sur les deux autres protéines du complexe. En comparant un profil des trois protéines obtenu par SDS-PAGE et Western Blot anti BSLF1, nous avons pu confirmer que la bande intermédiaire correspond bien à la protéine BSLF1.



**Figure 72 : Détection de la protéine BSLF1 par un anticorps anti BSLF1 par Western Blot. A gauche,** un gel SDS-PAGE coloré au bleu de Coomassie. **A droite,** un Western Blot anti BSLF1. La comparaison des deux profils permet d'attribuer la bande intermédiaire à la protéine BSLF1.

L'utilisation de l'anticorps anti BSLF1 a permis de franchir un pas important dans la thèse en confirmant que la protéine BSLF1 est bien présente dans les trois bandes observées sur gel SDS-PAGE.

Les résultats des différentes techniques d'identification protéique ont permis de confirmer la production des trois protéines du complexe. L'utilisation d'un baculovirus exprimant les trois protéines simultanément permet d'obtenir le complexe ternaire sous forme soluble. Les optimisations pour l'obtention d'un complexe pur et en grande quantité sont poursuivies afin de pouvoir réaliser des investigations structurales et enzymatiques.

### **III. Analyse du complexe par chromatographie d'exclusion de taille**

La chromatographie d'exclusion de taille va nous permettre de juger si le complexe que nous purifions est utilisable pour les analyses structurales par cristallographie.

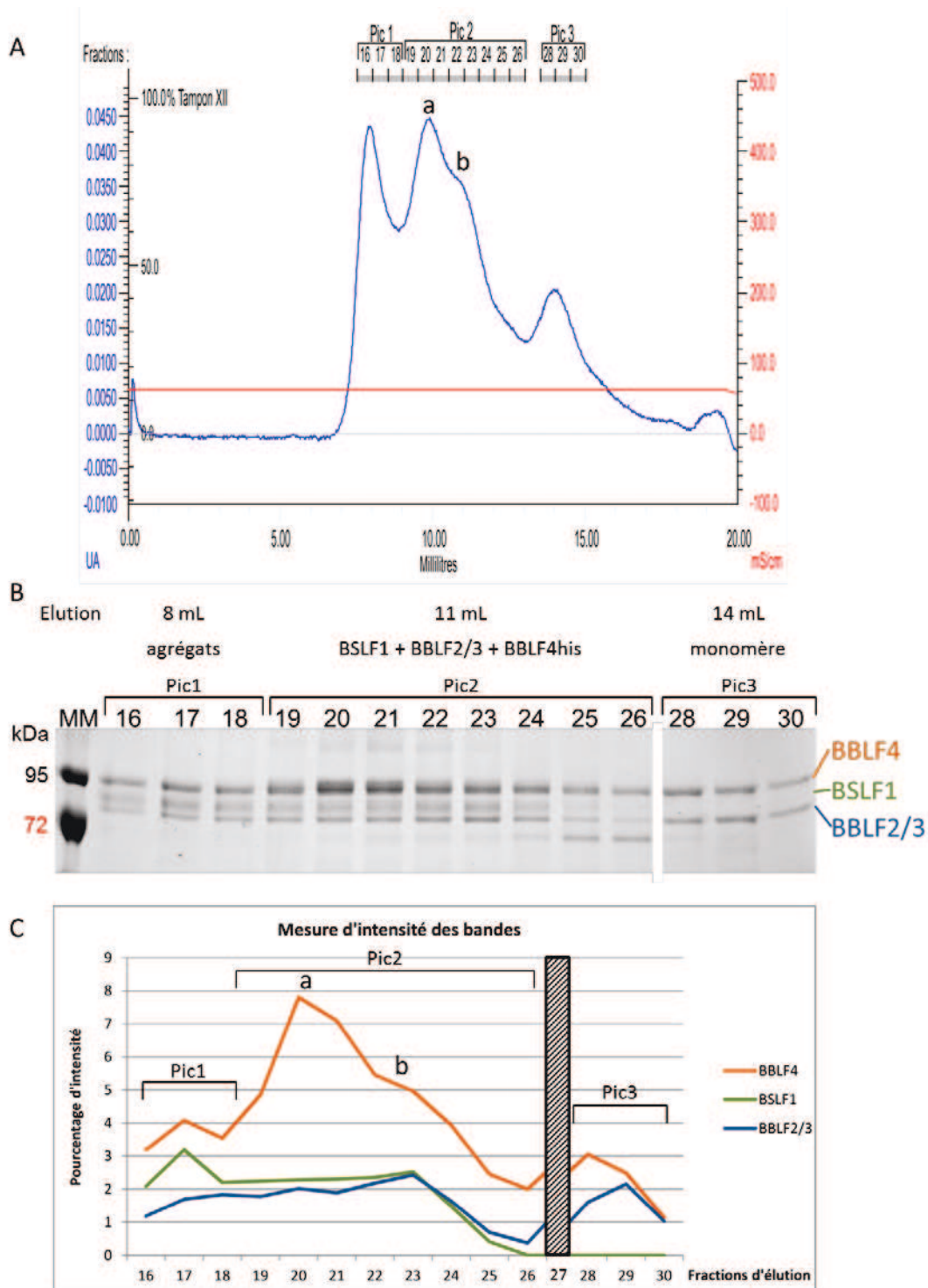
#### **1. Tampon optimisé pour la purification par exclusion de taille**

A l'issue des optimisations précédentes nous avons sélectionné le tampon VII (50mM Hépès pH 7,6 ; 500mM NaCl ; 10 % Glycérol) pour conserver le complexe. Les tampons VII (lyse, lavage, élution, conservation) permettent d'avoir un profil de purification sur colonne de nickel présentant les trois bandes du complexe avec des intensités similaires. L'utilisation de la chromatographie d'exclusion de taille permet d'accroître la pureté du complexe et d'obtenir des informations préliminaires sur sa taille et la qualité de l'échantillon. Les premiers tests de purification par exclusion de taille à partir du tampon VII ont révélé que l'échantillon était majoritairement agrégé. Nous avons alors essayé d'optimiser les conditions de tampon pour réduire l'agrégation du complexe. Dans un premier temps nous avons essayé de voir si le changement de la concentration en sel ou du pH du tampon pouvait réduire l'agrégation. Un tampon présentant une concentration en sel de 250 mM NaCl ou une valeur de pH de 6,6 déstabilise complètement le complexe, favorisant les formes monomériques des protéines BBLF2/3 et BBLF4 et rendant difficile la détection de la protéine BSLF1. Les conditions trouvées par l'optimisation sur colonne de nickel sont indispensables pour la stabilité du complexe. Aux conditions précédemment trouvées, 50 mM arginine/glutamate ont été ajoutés pour limiter l'agrégation et favoriser la solubilité des protéines [168, 169].

Ce nouveau tampon (tampon XII) permet de réduire l'agrégation et de favoriser la forme soluble du complexe. Le premier pic (fractions 16, 17 et 18) correspond à un volume d'élution de 8 mL (Figure 73, A). La colonne SuperDex200 ayant un volume mort de 8 mL, les fractions du pic 1 correspondent à de la protéine agrégée. Le profil SDS-PAGE des fractions composant ce pic révèle la présence des trois bandes des protéines du complexe (Figure 73, B). Ainsi une partie de l'échantillon après purification sur colonne de nickel reste agrégée malgré la présence des acides aminés R/E.

Le deuxième pic du chromatogramme (fractions 19, 20, 21, 22, 23, 24, 25 et 26) est hors volume mort et représente donc les protéines solubles de notre échantillon. Le volume d'élution de 11 mL correspond à des protéines ayant une masse proche de 600 kDa d'après la calibration de la colonne. Le gel SDS-PAGE indique que les fractions du pic 2 contiennent les trois protéines du complexe. Le pic 2 possède un épaulement et peut donc être décrit comme un double pic (Figure 73,

a et b). La partie « a » du pic 2 présente un excès de BBLF4 vis-à-vis des deux autres protéines d'après la mesure d'intensité des bandes du gel SDS-PAGE (Figure 73, C). La partie « b » du pic se rapproche plus d'une stœchiométrie de type 1/1/1 comme décrite dans la littérature. La masse théorique du complexe est de 260 kDa, les 600 kDa donnés par le volume d'élution équivalent à 2,5 complexes, mais on ne peut faire une relation aussi simpliste car la calibration de la colonne est réalisée par des protéines globulaires. La forme des protéines influe fortement sur leur volume d'élution. Une protéine allongée aura un temps d'élution lent et sa masse sera surestimée. La forme des protéines composant le complexe et la forme du complexe lui-même ne sont pas connues, il n'est donc pas possible de décrire la stœchiométrie des complexes formant le pic 2.



**Figure 73 : Purification du complexe H-P par chromatographie d'exclusion de taille (SuperDex200 10/300) avec le tampon n°XII. A : chromatogramme de l'absorbance en fonction du volume d'élution. B : Observation des fractions d'intérêt sur gel SDS-PAGE 8% coloré au bleu de Coomassie. C : Mesure d'intensité des bandes présentes sur le gel SDS-PAGE. MM : Marqueur de poids moléculaire. 16-30 : Fractions d'élution.**



Le troisième et dernier pic (fractions 28, 29, 30) correspond à un volume 14 mL indiquant l'élution de protéines ayant une masse proche de 100 kDa. D'après le gel SDS-PAGE les fractions du pic 3 contiennent seulement les protéines BBLF4 et BBLF2/3. La taille des protéines formant ce pic est de l'ordre de 100 kDa, indiquant que les protéines n'interagissent pas entre elles et sont éluées indépendamment.

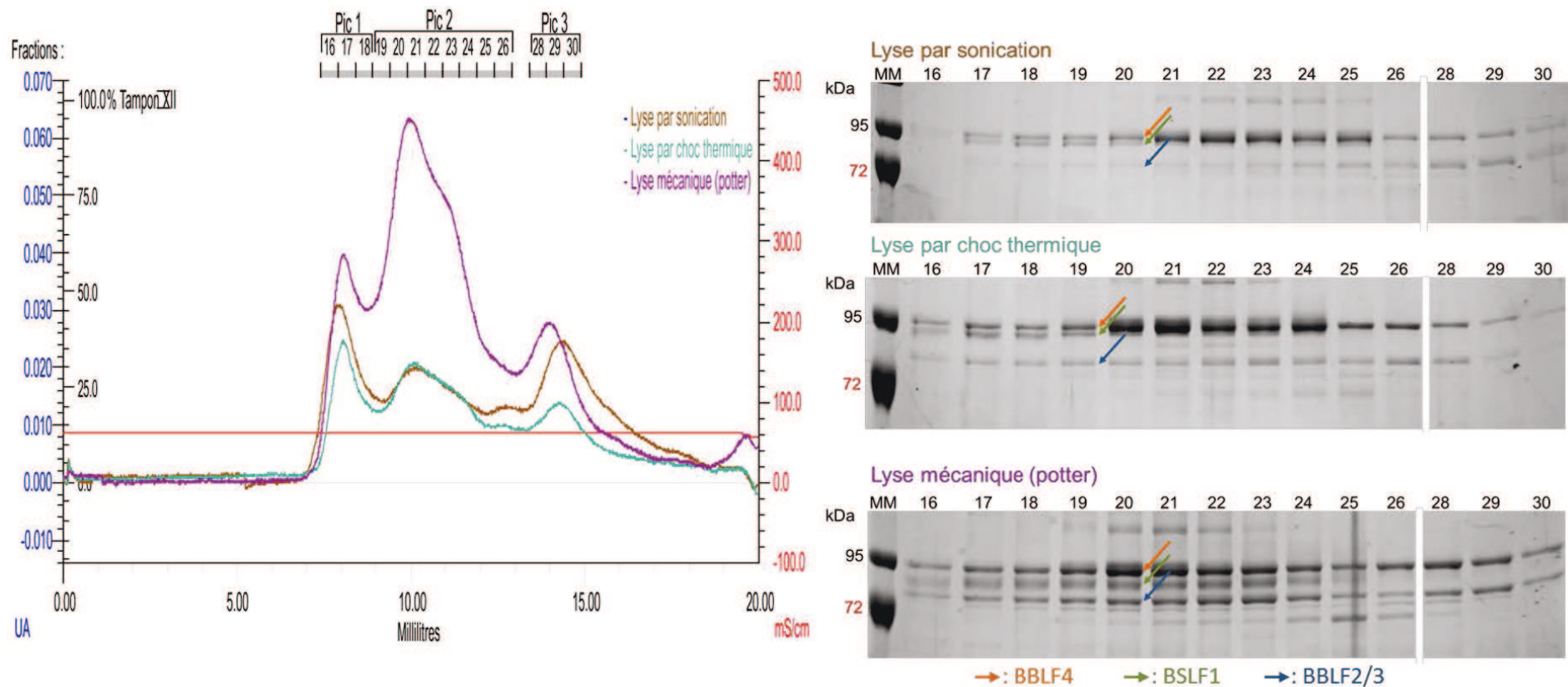
Une chromatographie d'exclusion de taille couplée au MALLS-RI a été réalisée afin de déterminer avec plus de précision la masse du complexe et pouvoir estimer la composition du pic 2. De façon standard 50  $\mu$ L d'échantillon est injecté sur ce système. Le pic d'absorbance correspondant au complexe soluble donne un très faible signal en réfractométrie et ne permet pas le calcul d'une masse précise, seule une approximation peut être réalisée. Le pic 2 présent sur les chromatogrammes contient une espèce protéique faisant entre 240 et 330 kDa. Cette approximation possède un  $\Delta$ Masse pouvant correspondre soit à un complexe avec une stœchiométrie de type 1/1/1 ou avec une forme oligomérique de l'hélicase au sein du complexe. Malgré l'utilisation de la technique MALLS nous ne pouvons pas conclure sur la stœchiométrie exacte du complexe.

Ces tests de différents tampons nous ont permis de confirmer que les conditions de base, en sel et pH, déterminées lors des premières phases d'optimisation stabilisent le complexe. La présence de 50 mM R/E permet de purifier une plus grande quantité de complexe soluble. Nous avons aussi ajouté à ce tampon 5 % de glycérol pour faciliter la concentration du complexe après l'étape de chromatographie d'exclusion, ce composant supplémentaire n'a pas d'influence sur le comportement du complexe. Le tampon XII (20mM Hépès pH 7,6 ; 500mM NaCl ; 50mM R/E et 5 % Glycérol) semble ne pas être suffisant pour séparer les deux formes solubles du complexe. Aucune condition n'a permis de favoriser une forme du complexe par rapport à l'autre.

Malgré les efforts d'optimisation de tampon, les purifications n'ont pas été vraiment reproductibles, la quantité d'agrégats et de protéines monomériques varie fortement d'une purification à l'autre. Nous avons alors suspecté que ces variations pouvaient être dues à la technique de lyse utilisée.

## **2. Conditions de lyse cellulaire**

Les différentes purifications que nous avons pu faire avec le tampon XII ont révélé des profils très variables. Pour juger si la technique de lyse est à l'origine des variations, nous avons comparé l'efficacité de trois techniques de lyse : la lyse par ultrason, la lyse par choc thermique et la lyse mécanique (potter).



**Figure 74 : Etude des techniques de lyse du complexe dans le tampon n°XII. A gauche** : chromatogramme de l'absorbance en fonction du volume d'élution pour les trois méthodes de lyse, ultrason (marron), choc thermique (vert) et mécanique (violet). Observation des fractions d'intérêt sur gel SDS-PAGE 8% coloré au bleu de Coomassie, **Haut droite** : lyse par ultrason. **Milieu droite** : lyse par choc thermique. **Bas droite** : lyse mécanique. MM : Marqueur de poids moléculaire. 16-30 : Fractions d'élution.

En fonction de la technique de lyse, les profils de chromatographie d'exclusion sont différents. Les purifications ont été réalisées à partir de la même masse de culot cellulaire pour les conditions de lyse par ultrason et de lyse par choc thermique. La condition utilisant le potter a été obtenue avec 20% de cellules en plus que les conditions précédentes. On constate que la lyse par ultrason induit plus d'agrégation (pic 1) et de protéines monomériques (pic 3). Le profil obtenu par choc thermique est relativement similaire à celui par ultrason, l'absorbance du pic 2 est équivalente pour les deux techniques de lyse. La dernière technique de lyse, le potter permet d'augmenter fortement l'absorbance du pic 2 et obtenir ainsi le complexe soluble en quantité bien supérieure par rapport aux autres techniques de lyse. Les gels SDS-PAGE des trois purifications sont très similaires. La technique de lyse mécanique avec le potter permet de purifier en quantité beaucoup plus importante le complexe et donc de pouvoir envisager la réalisation d'études structurales.

Quelle que soit la technique de lyse utilisée, le pic 2 présente toujours deux populations du complexe. N'ayant trouvé aucune solution permettant de séparer ces deux formes nous avons décidé de poursuivre les démarches structurales et voir si nous pouvons avoir une meilleure description des formes de complexes composants le pic 2.

## **IV. Etudes structurales du complexe H-P**

### **1. Tests de cristallisation**

L'objectif premier de la thèse étant d'étudier le complexe par cristallographie et diffraction aux rayons X, nous avons entrepris des tests de cristallisation. Comme pour les essais réalisés pour BBLF2/3, nous avons utilisé la plateforme de cristallisation HTX Lab pour le complexe entier. Tous les tests de cristallisation ont été réalisés avec les fractions concentrées du pic 2 observées sur le profil de purification par exclusion de taille.

Dans un premier temps nous avons soumis un échantillon à 0,6 mg/mL dans le tampon XII (20 mM Hépès pH 7,6 ; 500 mM NaCl ; 50 mM R/E ; 5% glycérol). La concentration de l'échantillon est faible pour des tests de cristallisation, mais plus l'échantillon à cristalliser possède un poids moléculaire élevé plus il est possible qu'il cristallise à basse concentration. Pour cet échantillon 6 plaques de cristallisation (576 conditions) ont été réalisées et placées dans un robot à 20°C pour le suivi de la cristallisation. Sur les 3 mois de suivi de la cristallisation, 100 % des gouttes sont restées claires. Il faut donc impérativement augmenter la concentration du complexe pour les tests de cristallisation.

Un second échantillon a été utilisé pour des essais de cristallisation, cette fois ci à une concentration de 1,3 mg/mL. A cette concentration le volume n'est pas suffisant pour pouvoir réaliser les 6 plaques standard de cristallisation. Nous avons choisi de réaliser la plaque nommée JCSG (Joint Center for Structural Genomics). Le crible JCSG a été établi à partir d'une étude statistique regroupant 67 conditions de cristallisation ayant permis la formation de cristaux et la résolution de la structure par diffraction aux rayons X. Les 67 conditions ont été étendues à 96 en complétant la gamme de pH et de précipitants. La plaque JCSG est idéale pour les premiers tests de cristallisation sur une protéine. La plaque a été placée dans le robot à 20°C pour le suivi de la cristallisation. Après 3 mois d'inspection 35% des gouttes sont encore claires. La concentration de 1,3 mg/mL est suffisante pour observer la formation de précipité ou de dénaturation des protéines du complexe. Aucun cristal n'a pu être observé, il faudrait pouvoir réaliser des tests en ayant une concentration en complexe plus élevée ou une préparation protéique plus homogène.

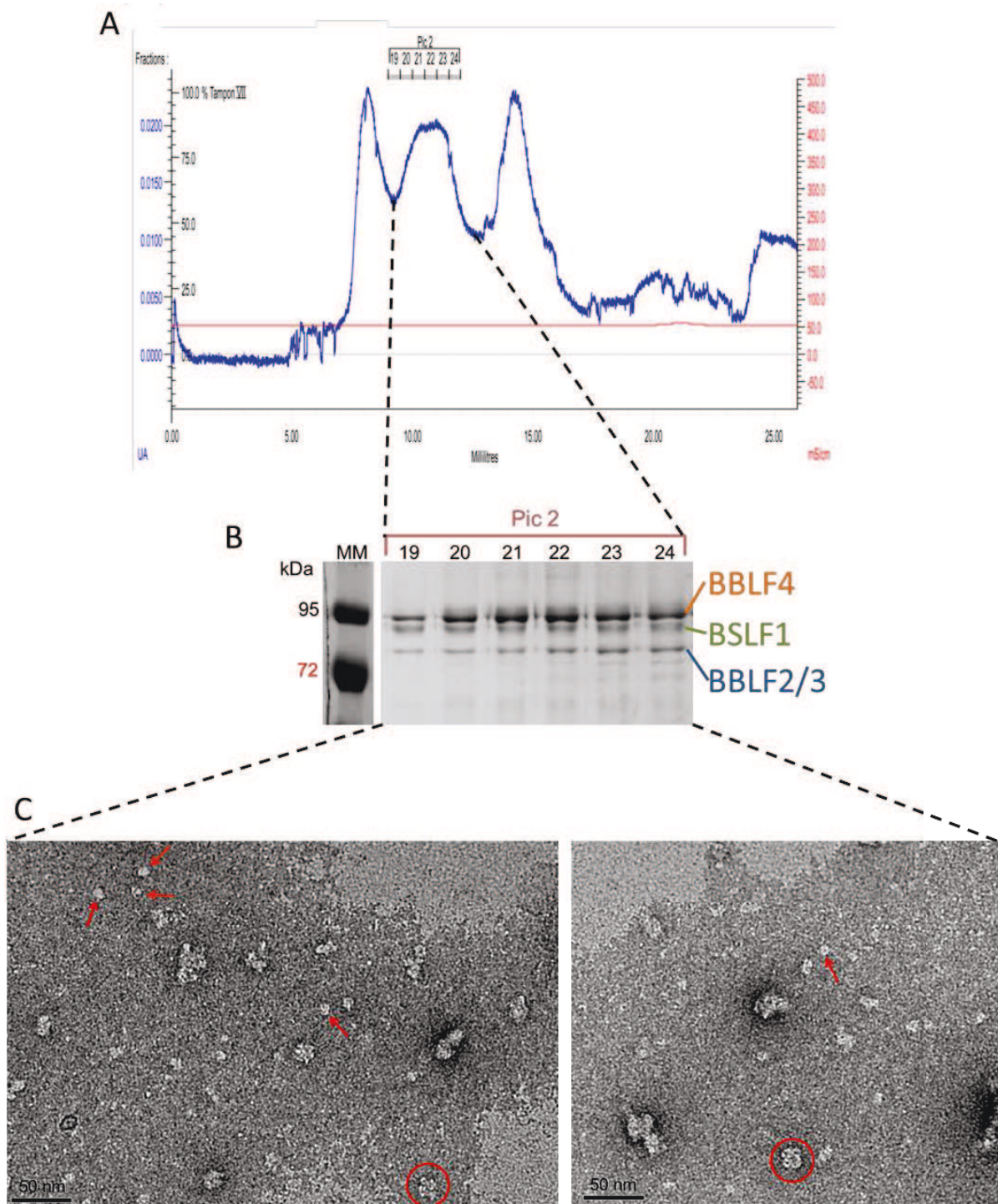
## **2. Etudes du complexe par microscopie électronique**

### **2.1. Etudes préliminaires**

Les concentrations finales obtenues pour le complexe sont faibles et rendent difficiles la recherche de condition de cristallisation. La microscopie électronique demande beaucoup moins de matériel pour pouvoir observer le complexe en coloration négative.

Dans un premier temps les analyses par microscopie électronique ont été utilisées comme contrôle qualité de l'échantillon après purification. Ces expériences ont été réalisées lors des premiers tests de purification par exclusion de taille avec le tampon VII. Les clichés observés ont révélé une forte hétérogénéité de l'échantillon et la présence d'agrégats. Malgré la qualité faible des échantillons certaines formes particulières de type anneaux ou tétramères ont été observées (Figure 75, C).

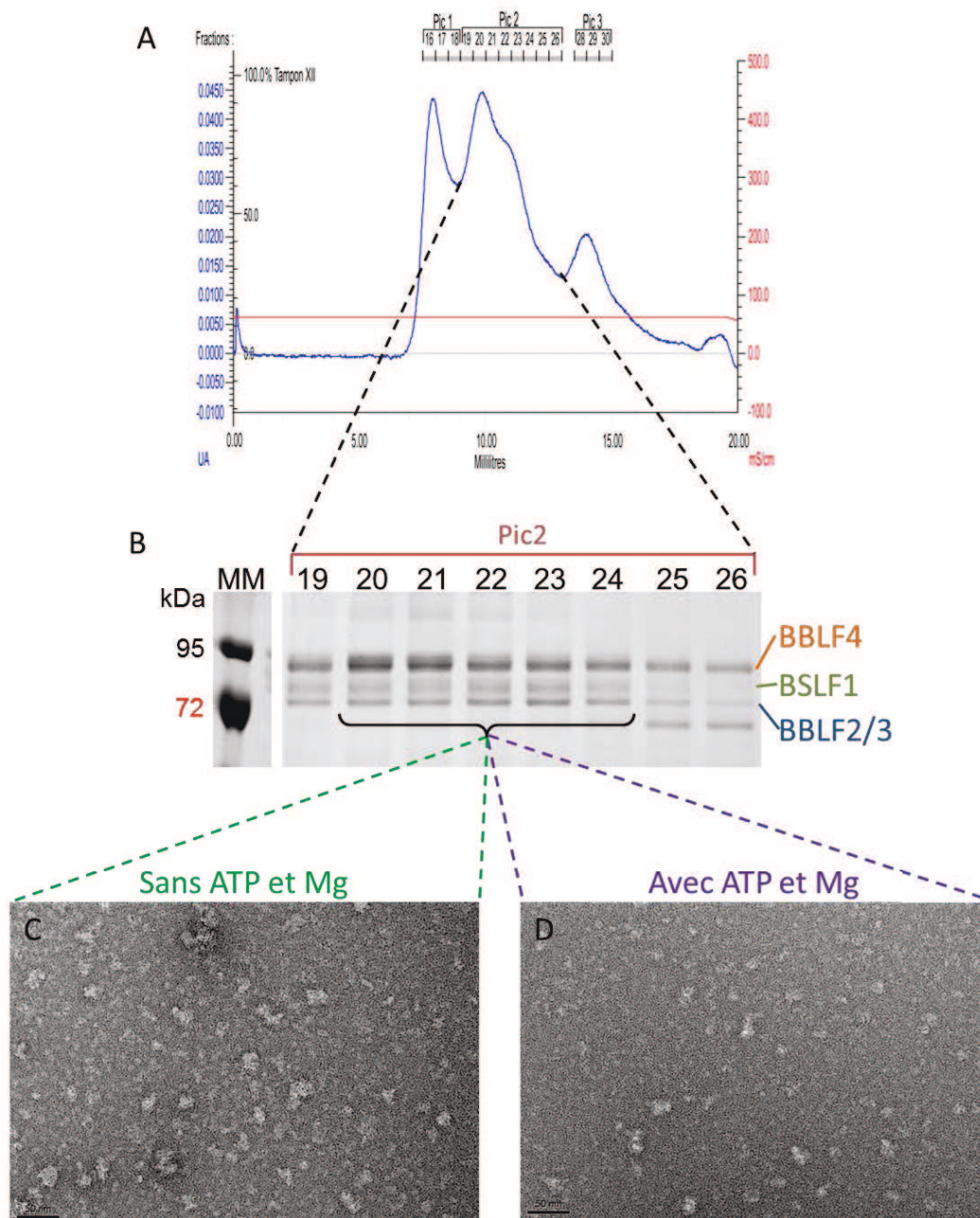
Le complexe H-P est composé de trois protéines avec une stœchiométrie théorique de type 1/1/1. Les formes observées sur les clichés de microscopie électronique (Figure 75, C) ne permettent pas de vérifier cette hypothèse. Lors de l'optimisation du tampon par chromatographie d'exclusion de taille, nous avons mis en avant une forte agrégation du complexe qui est visible sur les clichés. L'hétérogénéité présente sur les photos n'est pas seulement due à l'agrégation de l'échantillon, mais aussi à la présence d'une population non homogène du complexe au sein des fractions du pic 2. La qualité de l'échantillon doit être optimisé afin d'observer des formes plus régulières du complexe.



**Figure 75 : Clichés de microscopie électronique du pic 2 de gel filtration après concentration. A :** Profil de purification par exclusion de taille du complexe dans le tampon VII. **B :** Fractions composant le pic 2 visualisé sur gel SDS-PAGE 8% coloré au bleu de coomassie. **C :** Clichés de microscopie électronique en coloration négative (2% m/v acétate d'uranyle).



L'utilisation du tampon XII permet de réduire la présence de très gros agrégats, mais présente toujours une forte hétérogénéité (Figure 76, C). En ajoutant de l'ATP et du Mg dans le tampon XII après purification, il y a toujours la présence d'agrégats, mais globalement l'échantillon est moins hétérogène (Figure 76, D). Les clichés ont été réalisés à partir des fractions du pic 2 de la chromatographie d'exclusion de taille (Figure 76, A), les fractions n'ont pas été concentrées pour limiter l'aggrégation des échantillons.



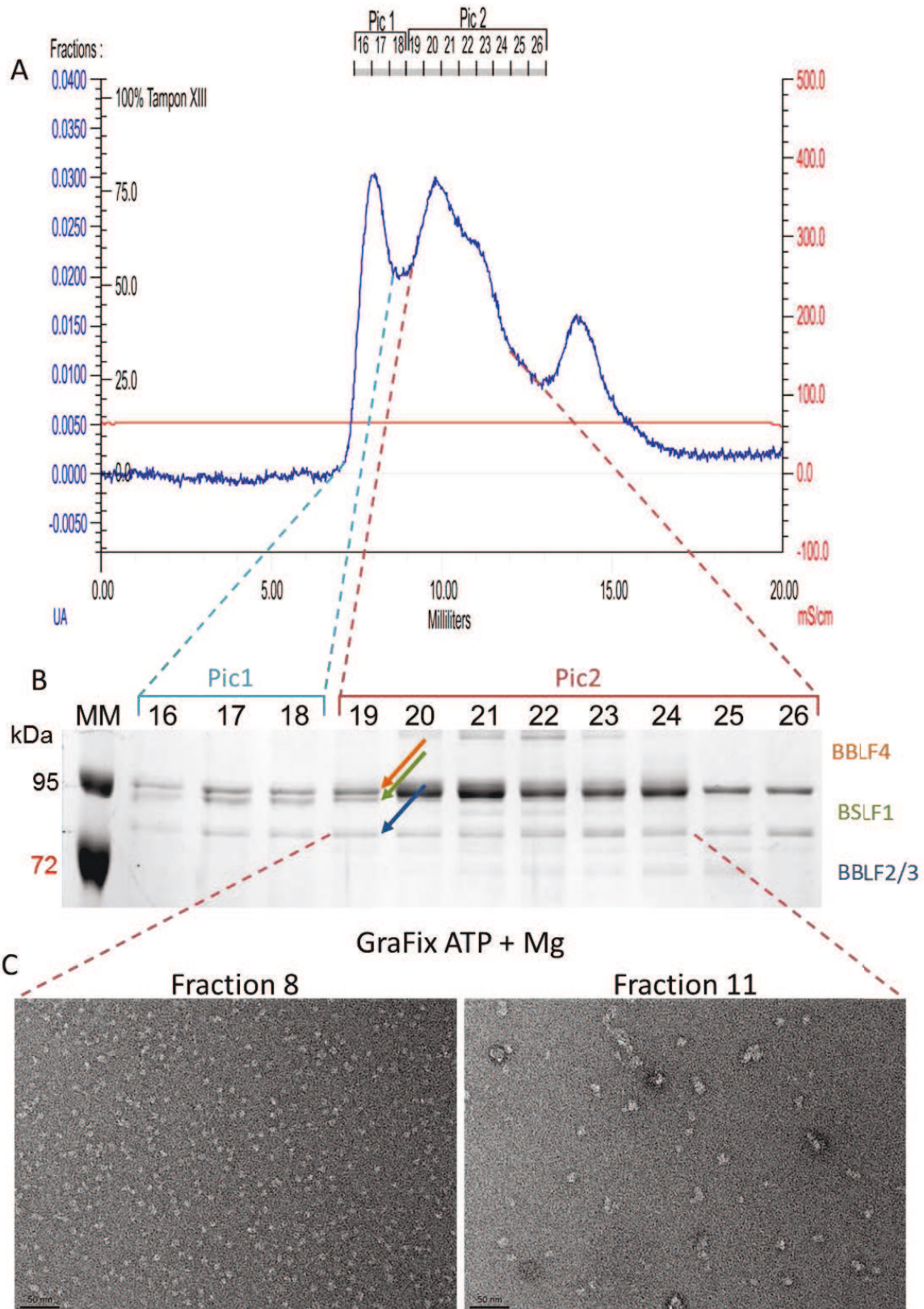
**Figure 76 : Clichés de microscopie électronique en transmission du complexe H-P en coloration négative (2% m/w acétate d'uranyle). A :** Profil de purification par exclusion de taille du complexe dans le tampon XII. **B :** Fractions composant le pic 2 visualisé sur gel SDS-PAGE 8% coloré au bleu de coomassie. **C :** Observation du complexe par ME dans le tampon XII. **D :** Observation du complexe par ME dans le tampon XII, ajout de 1 mM ATP et 5 mM MgCl<sub>2</sub> après la purification par exclusion de taille.

L'ATP et le Mg semble réduire un peu l'hétérogénéité de l'échantillon. Ces molécules sont indispensables pour les activités des protéines de type hélicase et primase et peuvent permettre de figer la structure du complexe. Les observations du complexe par microscopie électronique révèlent la persistance d'une hétérogénéité importante et de la présence d'agrégats au sein des échantillons malgré la purification par exclusion de taille et l'optimisation des tampons. Ces constatations peuvent être dues à la présence de deux formes solubles du complexe au sein des fractions du pic 2, mais aussi à une stabilité faible du complexe, supportant mal la procédure de coloration négative. La coloration peut induire des changements de composition ou de conformation augmentant l'hétérogénéité de l'échantillon.

## **2.2. GraFix**

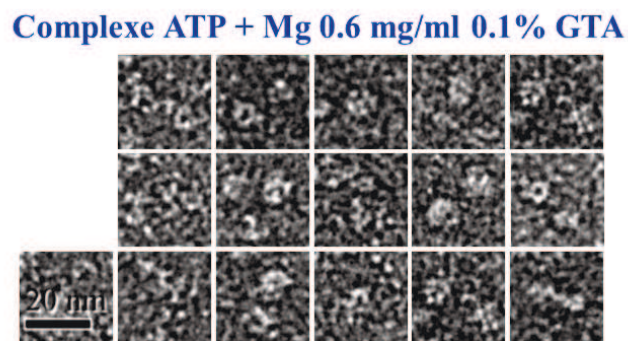
La technique GraFix permet de stabiliser les complexes en utilisant un gradient d'agent fixateur (glutaraldéhyde) pendant une ultra-centrifugation de l'échantillon. Nous avons appliqué cette technique au complexe obtenu après purification par exclusion de taille dans le tampon XII en présence ou absence d'ATP (1 mM) et  $Mg^{2+}$  (5 mM). L'échantillon déposé sur le gradient GraFix est à une concentration de 0,6 mg/mL. Cette approche a permis d'augmenter fortement l'homogénéité des échantillons et d'éliminer les agrégats. Les échantillons en présence ou absence d'ATP et Mg sont très similaires après l'étape de GraFix, seules les conditions en présence d'ATP et Mg sont détaillées dans les Figure 77 et Figure 78.





**Figure 77 : Etude par microscopie électronique à transmission du complexe H-P après GraFix. A :** Profil de chromatographie d'exclusion de taille du complexe H-P dans le tampon XII + 1 mM ATP + 5mM MgCl<sub>2</sub>. **B :** Gel SDS-PAGE 8%, coloré au bleu de Coomassie, des fractions présentes dans les pics 1 et 2 du chromatogramme. **C :** Observation par ME des fractions 8 et 11 du gradient GraFix.

Le gel SDS-PAGE de la purification d'exclusion de taille révèle que le pic 2 possède une variabilité importante de la stœchiométrie des protéines le composant (Figure 77, B). Nous sommes toujours en présence de plusieurs formes solubles du complexe. Les échantillons observés en microscopie électronique après le GraFix (Figure 77, C) révèlent une qualité d'échantillon bien supérieure à ceux observés directement après purification par exclusion de taille. La fraction 8 montre un échantillon avec une hétérogénéité très réduite et il n'y a plus du tout de présence d'agrégats. La fraction 11 révèle des particules de taille plus élevée, probablement dû à une fixation excessive induite par la forte concentration en glutaraldéhyde dans cette partie du gradient. L'utilisation de la technique GraFix apporte une réelle amélioration de la qualité de l'échantillon, sa faible hétérogénéité permet même d'isoler quelques particules (Figure 78).



**Figure 78 : Détails de particules isolées pour la condition en présence d'ATP et Mg.**

La qualité de l'échantillon a été fortement améliorée, mais elle n'est pas encore optimale pour pouvoir effectuer une analyse plus approfondie des particules isolées car il y a la persistance d'une faible hétérogénéité. Deux raisons peuvent expliquer cette persistance de l'hétérogénéité des échantillons malgré l'utilisation de la technique GraFix. Premièrement, le pic 2 utilisé pour le GraFix contient toujours les deux formes solubles du complexe pouvant participer à cette hétérogénéité. Deuxièmement, le gradient de glycérol et de GTA étant réalisé manuellement, il se peut qu'il ne permette pas de séparer avec une grande précision les différentes formes du complexe présentes dans l'échantillon. L'utilisation d'un gradient continu permettrait peut-être une meilleure séparation des deux formes du complexe et donc améliorer un peu plus la qualité de l'échantillon.

La technique GraFix permet une avancée importante pour l'étude structurale du complexe. Quelques optimisations sont encore à réaliser pour pouvoir faire une analyse des particules isolées. La réalisation d'une reconstruction de l'enveloppe du complexe dépendra de la présence d'une symétrie dans la forme du complexe.

## V. Etudes des activités enzymatiques du complexe

### 1. Test ATPase

Les mesures de l'activité ATPase du complexe ont été réalisées à l'aide du kit « ADP-Glo Kinase » (Promega). Le test permet de mesurer le pourcentage d'ATP converti en ADP pendant la réaction ATPasique. Pour établir les conditions de tampon du test ATPase nous nous sommes inspirés des conditions utilisées pour le complexe H-P d'HSV [171]. La réaction est réalisée dans un volume de 25  $\mu$ L pendant 30 min à 37°C. Une courbe étalon réalisée par des rapports variables ATP/ADP (50  $\mu$ M final) permet d'évaluer le pourcentage de conversion d'ATP en ADP pendant la réaction ATPasique.

Les conditions testées pour l'activité ATPase sont détaillées dans le Tableau 14, avec leurs pourcentages de conversion d'ATP en ADP respectifs.

**Tableau 14 : Détails des conditions utilisées pour le test ATPasique**, réalisé sur le complexe (BSLF1 + BBLF2/3 + BBLF4his) et leur valeur de conversion de l'ATP en ADP. Réaction 30 min à 37°C. ADNsb : ADN simple brin.

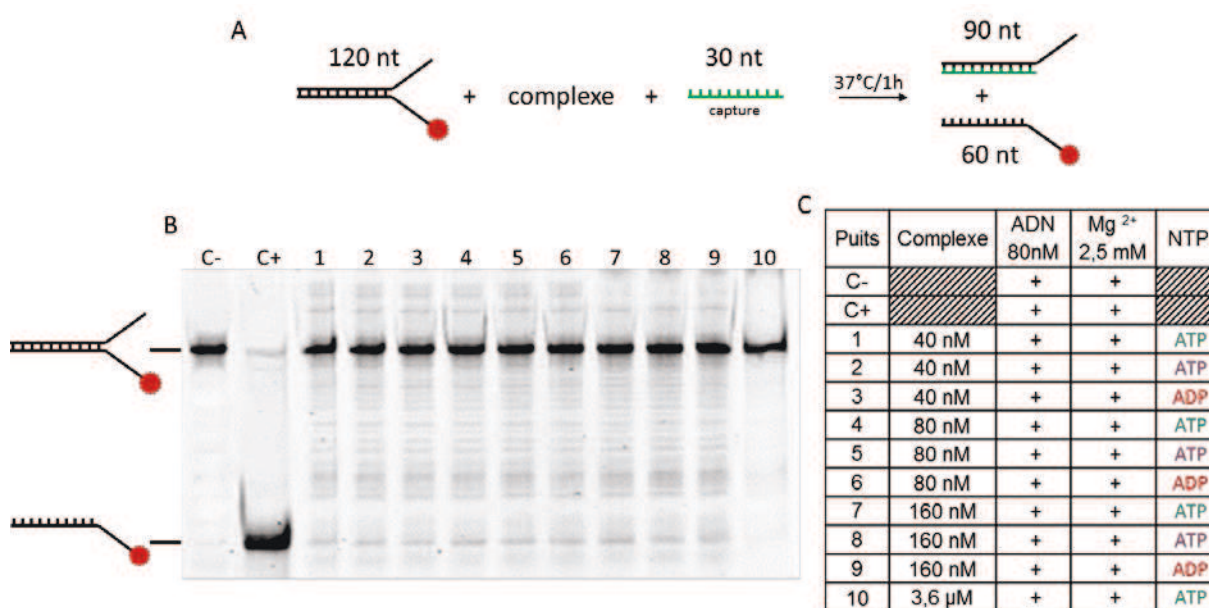
Conditions expérimentales	% conversion ATP en ADP
1 : Complexe (1 $\mu$ M)	0
2 : Complexe (410 nM) + ADNsb (200 nM)	0
3 : Complexe (410 nM) + ADNsb (1 $\mu$ M)	0
4 : Complexe (410 nM) + ATP (50 $\mu$ M)	38
5 : Complexe (410 nM) + ATP (50 $\mu$ M) + ADNsb (200 nM)	44
6 : Complexe (410 nM) + ATP (50 $\mu$ M) + ADNsb (1 $\mu$ M)	42
7 : Complexe (1 $\mu$ M) + ATP (50 $\mu$ M) + ADNsb (200 nM)	70
8 : Complexe (1 $\mu$ M) + ATP (50 $\mu$ M) + ADNsb (1 $\mu$ M)	65

Les pourcentages de conversion d'ATP en ADP pour les conditions 4 (sans ADNsb), 5 (ADNsb 200 nM) et 6 (ADNsb 1  $\mu$ M) sont respectivement de 38%, 44% et 42%. Ces valeurs révèlent que l'activité ATPasique mesurée est indépendante de la présence d'ADN simple brin. Pour bien observer l'action ATPasique, l'ATP est en excès (120 fois) vis-à-vis du complexe pour qu'il ne soit pas le facteur limitant de la réaction. Il semblerait que nous mesurons seulement l'activité basale du complexe et non l'hydrolyse permettant le déplacement le long d'un ADN simple brin. L'activité du complexe est très faible car elle représente l'hydrolyse de 0,03 pmol d'ATP pour 1 pmol de complexe par seconde.

Ce test préliminaire révèle que le complexe H-P que nous purifions possède une activité ATPase basale qui n'est pas influencée par la présence d'ADN. Les investigations futures permettront de déterminer s'il possède seulement une activité basale comme montré dans ce test ou s'il est réellement actif mais dans des conditions différentes de celles utilisées actuellement.

## 2. Tests hélicase

Des tests hélicase ont été réalisés sur le complexe H-P d'EBV (BSLF1 + BBLF2/3 + BBLF4his). La réalisation des tests a été inspirée de ceux ayant permis de mettre en évidence l'activité hélicase du complexe UL5/UL52/UL8 de HSV-1 [97, 108]. Aucune activité hélicase spécifique du complexe n'a pu être observée, que ce soit en utilisant des substrats courts possédant une fourche (Figure 79) ou des substrats composés du plasmide M13mp18 présentant une partie hybridée avec un bout flottant en 5' (non montré).



**Figure 79 : Réaction hélicase avec des concentrations croissantes de complexe (BSLF1 + BBLF2/3 + BBLF4his) en présence d'un substrat ADN court de type fourche, de Mg<sup>2+</sup> et d'ATP ou d'ADP.** **A** : Principe du test d'activité hélicase. **B** : gel contrôle de la réaction hélicase (polyacrylamide non dénaturant 20 %). **C** : Les concentrations utilisées pour le test sont les suivantes : ADN de type fourche (80 nM), MgCl<sub>2</sub> (2,5 mM), ATP et ADP à 1 mM. Deux ATP sont utilisés dans ce test, **ATP** : obtenue à partir d'une poudre d'ATP (Sigma Aldrich), **ATP** : ATP ultra pur du kit « ADP-Glo Kinase » (Promega), **ADP** : ADP ultra pur du kit « ADP-Glo Kinase ».

Les réactions hélicase sont très difficiles à mettre en œuvre car elles sont très sensibles aux composants présents lors du test et aux types de substrats utilisés. Le test ATPase précédent a permis d'observer seulement une activité ATPasique basale du complexe. Cette activité est sûrement insuffisante pour permettre l'ouverture d'un duplex d'ADN de type fourche.

Partie IV :  
Discussion  
et  
Perspectives

## I. Discussion

Au début de ma thèse les informations disponibles dans la littérature sur la réplication du génome viral d'EBV portaient essentiellement sur l'identification des protéines composant le complexe H-P et les interactions possibles avec d'autres protéines répliquatives ou cellulaires. Les études menées par Fixman & al (1992 et 1995) ont permis de décrire l'importance des protéines BBLF4 et BSLF1 pour la réplication du génome viral d'EBV [50]. Les expériences menées sur la protéine BBLF2/3 ont permis de relier cette protéine au complexe H-P [51]. Par analogie avec le complexe H-P d'HSV-1 [172], il était supposé que les protéines BBLF4, BSLF1 et BBLF2/3 interagissaient ensemble pour former le complexe H-P d'EBV. La localisation cellulaire des protéines BBLF4, BSLF1 et BBLF2/3 réalisée par Gao & al (1998) [90] révèle une co-localisation des trois protéines dans le noyau de la cellule infectée suggérant que ces protéines forment bien un complexe. Cette étude a aussi révélé que le complexe H-P interagissait avec le facteur de transcription ZEBRA par l'intermédiaire des protéines BBLF4 et BBLF2/3. Les interactions entre les protéines BBLF4, BSLF1 et BBLF2/3 ont été confirmées par les expériences de co-immunoprécipitation réalisées par Yokoyama & al (1999) [94]. Dans cette publication, ils décrivent que chaque protéine interagit avec ces deux autres partenaires. D'après ces résultats et les connaissances sur le complexe H-P d'HSV-1, ils ont suggéré que le complexe H-P d'EBV avait probablement une stœchiométrie de type 1/1/1. Les interactions formant le complexe H-P étant bien décrites, des expériences complémentaires ont été réalisées afin de définir si ce complexe pouvait interagir avec d'autres protéines répliquatives. Gao & al (1998) avaient déjà montré que le complexe était capable d'interagir avec la protéine ZEBRA grâce aux protéines BBLF4 et BBLF2/3, favorisant la présence de ce complexe pré-répliquatif au niveau des foyers de réplication au sein du noyau de la cellule infectée [90, 127]. Des études complémentaires réalisées par Fujii & al (2000) [91] ont permis de décrire l'interaction du complexe H-P avec la sous-unité catalytique de la polymérase (BALF5). L'interaction a été observée *in vitro* par co-immunoprécipitation à partir de cellules de la lignée B95-8 et avec des cellules Hi5 transfectées par des baculovirus exprimant les protéines du complexe H-P et BALF5. Cette étude révèle que BALF5 interagissait avec chacune des protéines composant le complexe H-P. Il a aussi été montré que le complexe H-P peut interagir avec la protéine cellulaire ZBRK1 et son co-répresseur KAP-1 par l'intermédiaire de la protéine BBLF2/3. Cette étude semble montrer que BBLF2/3 contribue à la réplication lytique d'EBV aussi par le recrutement de protéines cellulaires en relation avec la réplication. Toutes ces études ont permis de décrire le complexe H-P d'EBV et son implication au sein du complexe de réplication. Aucune information structurale ou enzymatique n'était connue pour le

complexe H-P d'EBV au début de ma thèse, c'était un sujet attractif permettant préciser un peu plus le mécanisme répliatif d'EBV.

Lorsque j'ai débuté ma thèse les protéines BBLF4, BSLF1 et BBLF2/3 n'avaient pas encore été étudiées par l'équipe. J'ai choisi d'utiliser le système d'expression baculovirus/cellules d'insectes car il avait été utilisé pour les expériences décrivant le complexe H-P d'EBV. La réalisation de trois baculovirus codant chacun pour une protéine du complexe H-P m'a permis d'étudier la protéine accessoire BBLF2/3. J'ai pu décrire que cette protéine est monomérique en solution (MALLS et DLS) et possède une masse de 73 kDa (MALLS). La pureté de la protéine et la quantité produite m'ont permis de réaliser des tests de cristallisation. Une condition particulière a permis l'apparition de cristaux transitoires. La disparition des cristaux au cours du temps peut-être due à plusieurs facteurs : soit la protéine présente dans le cristal est protéolysée au cours du temps, soit les conditions permettant l'apparition des cristaux ne correspondent pas à la composition finale de la goutte à l'équilibre. Le script de recherche d'ordre et de désordre « D-score » réalisé sur cette protéine avait révélé qu'elle était composée d'une alternance de régions ordonnées et désordonnées, pouvant expliquer la présence des cristaux transitoires. J'ai donc décidé de me concentrer sur la production du complexe entier, car les possibilités d'études sur le complexe sont plus nombreuses que sur la sous-unité accessoire seule. Et les informations obtenues pourraient permettre de décrire le mécanisme de répliation d'EBV.

La production des protéines BBLF4 et BSLF1 solubles n'a pas été possible par cette stratégie d'expression. J'ai tenté d'obtenir ces deux protéines en réalisant des co-infections, mais elles sont restées insolubles, je n'ai donc pas réussi à reproduire les expériences de co-infections réalisées dans les études précédentes [90, 91, 94, 127, 129]. L'explication de ce résultat peut provenir des méthodes de détection des protéines du complexe. Au cours de mon travail, les protéines produites sont visualisées par coloration au bleu de coomassie sur gel SDS-PAGE, alors que les études de co-immunoprécipitation utilisent des anticorps spécifiques pour chaque protéine du complexe H-P permettant une détection beaucoup plus sensible. De plus Yokoyama & al (1999) [94] estiment que les protéines surexprimées représentent seulement 10% des protéines totales de chaque infection et considèrent que seulement 5% des protéines surexprimées sont solubles. Ces considérations peuvent expliquer les difficultés rencontrées pour la production sous forme soluble des protéines BSLF1 et BBLF4. Les productions des trois protéines sont réalisées en vue d'une analyse structurale par cristallographie, ce qui implique une exigence en terme de pureté et de concentration. Ces conditions sont difficilement réalisables avec seulement 5 % de protéines d'intérêts solubles. Pour la première stratégie d'expression des protéines individuelles ; les tampons de lyses et de purification



n'étaient pas optimisés, ce qui peut expliquer les problèmes de solubilité des protéines BBLF4 et BSLF1. Les études réalisées avant le début de ma thèse n'ont pas abouti jusqu'aux études enzymatiques ou structurales du complexe, ce qui pourrait indiquer une incapacité des auteurs à produire le complexe sous une forme suffisamment pure pour entreprendre ces investigations.

L'approche utilisant trois baculovirus codant pour chaque protéine du complexe ne m'ayant pas permis de produire le complexe entier, j'ai alors décidé de cloner les trois gènes d'intérêts dans un seul et même baculovirus par la technique de clonage MultiBac. Cette nouvelle stratégie d'expression en cellules d'insectes a été développée à l'EMBL. J'ai ainsi pu profiter de nombreux conseils pour mener à bien ce clonage. Cette stratégie de clonage avec le système MultiBac n'avait pas encore été réalisée au sein de l'équipe. Ce clonage demande une rigueur et un contrôle de toutes les étapes pour parvenir à insérer plusieurs gènes dans un seul bacmide. La difficulté fut d'autant plus grande que je devais insérer trois gènes de tailles importantes, posant notamment des problèmes dans les choix des enzymes de restrictions à utiliser pour le clonage. L'obtention des deux constructions MultiBac décrites dans les résultats a demandé 6 mois de travail et a permis au final de produire le complexe H-P sous forme soluble. La recherche des conditions optimales de lyse a été facilitée par l'utilisation de la matrice de solubilité en plaque 96 puits. Cette approche développée au sein de l'institut m'a permis de trouver rapidement des conditions favorisant la forme soluble de la protéine portant l'étiquette de purification. Des deux constructions réalisées par le clonage MultiBac, seule celle portant l'étiquette de purification sur l'hélicase a été utilisée pour les investigations structurales et enzymatiques. La construction portant l'étiquette de purification sur la sous-unité accessoire a été abandonnée car la protéine BBLF2/3 n'était pas détectée dans les fractions d'élution après purification par affinité. Un parallèle entre cette observation et la co-infection avec BSLF1 et BBLF2/3 peut être réalisé : dans les deux cas on détecte très faiblement ou pas du tout la protéine BBLF2/3his dans les fractions d'élution de la colonne de nickel. Ainsi ces deux informations suggèrent que l'étiquette de purification située en N-terminal de BBLF2/3 n'est pas accessible. Il semblerait donc que le N-terminal de BBLF2/3 interagisse sur BSLF1. Cette hypothèse doit être vérifiée car elle pourrait permettre de préciser les domaines d'interactions du complexe décrit par Yokoyama & al (1999) [94].

Le travail que j'ai réalisé sur le complexe entier ayant l'étiquette de purification sur l'hélicase, a permis de décrire avec précision les conditions d'expression et de purification permettant l'obtention du complexe soluble. Pour confirmer que l'on purifiait bien les trois protéines j'ai procédé à leur identification. J'ai combiné diverses techniques d'identification (séquençage N-terminal, spectrométrie de masse et Western Blot) pour pouvoir identifier les trois protéines du complexe.

L'identité des protéines confirmée et les conditions d'expression et de purification établies, m'ont permis d'envisager la caractérisation structurale et enzymatique du complexe. Lors des purifications par exclusion de taille, il semblerait que deux formes solubles du complexe soient présentes. La première forme semble être composée des trois protéines du complexe mais avec un excès de BBLF4, la seconde forme est beaucoup plus proche de la stœchiométrie de type 1/1/1 décrite par Yokoyama & al (1999) [94] (MALLS). Les résultats de Chen & al (2011) [123] pourraient expliquer la forme soluble du complexe présentant un excès de BBLF4. L'étude qu'ils ont menée sur le complexe H-P d'HSV-1 a révélé que deux complexes H-P pouvaient former des complexes d'ordre supérieur, au moins deux complexes, en fonction de la concentration. Nous observons un excès de BBLF4 probablement dû à la présence de l'étiquette de purification. La protéine BBLF4 étant sur représentée, elle pourrait s'oligomériser sur le complexe de base ayant une stœchiométrie de type 1/1/1. L'optimisation des conditions de purification par exclusion de taille ne m'ont pas permis de séparer les deux formes du complexe et j'ai décidé de débiter les investigations structurales et enzymatiques du complexe.

Les 672 tests de cristallisation réalisés avec le complexe entier n'ont pas permis d'observer de cristaux. L'absence de cristaux peut être due à la présence des deux états du complexe dans la solution protéique. Afin de juger de l'homogénéité des échantillons, nous avons utilisé la microscopie électronique. Une hétérogénéité des échantillons persistait malgré l'amélioration des conditions de purification, j'ai alors décidé d'utiliser la technique GraFix pour stabiliser le complexe. Cette méthode utilise un gradient de glycérol et de glutaraldéhyde permettant de séparer le complexe en fonction de sa taille et stabiliser les différentes tailles par fixation au glutaraldéhyde. Cette approche a fortement amélioré la qualité des échantillons en réduisant l'hétérogénéité. La persistance d'une faible hétérogénéité ne permet pas d'isoler des particules et d'envisager une description de l'enveloppe du complexe. Pour pouvoir continuer les investigations structurales il faut tenter de séparer les deux formes solubles du complexe ou arriver à favoriser une des deux formes.

En parallèle des études structurales nous avons abordé l'aspect fonctionnel du complexe. Nous avons pu mesurer une activité ATPase du complexe H-P d'EBV de l'ordre de : 1 pmol de complexe pouvant hydrolyser 0,03 pmol d'ATP en une seconde. Cette valeur est indépendante de la présence d'un ADN simple brin. Ce résultat est inférieur de celui obtenu pour le complexe H-P d'HSV-1 en absence d'ADN : 0,3 pmol d'ATP hydrolysé par 1 pmol [107]. On observe un facteur d'ordre dix entre l'activité ATPase du complexe H-P d'EBV et celui d'HSV-1 en absence d'ADN. En revanche les activités du complexe H-P d'HSV-1 [104, 107] et de la Dda [140] sont nettement supérieures en présence d'un substrat d'ADN simple brin, elles sont respectivement de 2,5 à 10 pmol et 150 pmol

d'ATP hydrolysé par 1 pmol d'enzyme en une seconde. Nous mesurons probablement une activité basale du complexe d'EBV, suggérant que le complexe n'est pas stimulé par la présence d'un ADN simple brin. Il faudrait effectuer d'autres mesures pour estimer l'importance du substrat ADN sur l'activité ATPase du complexe H-P d'EBV.

J'ai tenté également de mesurer l'activité hélicase du complexe. Les expériences ont été élaborées à partir des tests réalisés pour le complexe d'HSV-1 [97, 108]. Le test permet de suivre l'apparition de forme simple brin de l'ADN après ouverture d'un duplex d'ADN par l'hélicase. Quel que soit le substrat ADN utilisé, brin 5' flottant sur le plasmide M13mp18 ou un ADN double brin présentant une fourche, aucune activité hélicase n'a pu être observée. Malgré un grand nombre de tests réalisés en faisant varier, les substrats ADN et leur concentration, la concentration du complexe H-P, la présence ou non de Mg, différents type d'ATP, je n'ai jamais pu observer une activité hélicase. Cette absence d'activité peut être due à l'activité ATPase faible portée par le complexe qui n'est peut-être pas suffisante pour permettre le déplacement du complexe sur l'ADN ou l'ouverture du duplex.

## II. Perspectives

Au cours de ma thèse, j'ai réussi à établir les conditions d'expression et de purification qui permettent de poursuivre l'étude structurale et enzymatique du complexe H-P d'EBV. Mais il est aussi important de s'intéresser à la protéine BBLF2/3 qui est exploitable pour les caractérisations structurales.

La condition ayant permis l'apparition de cristaux transitoires mériterait une investigation complémentaire. Il serait intéressant de reproduire manuellement ces conditions de cristallisation et de les affiner. Le tampon de la protéine ne présentait pas de glycérol lorsque les cristaux ont été observés. L'utilisation de glycérol facilite la concentration de l'échantillon par centrifugation et est nécessaire pour l'obtention d'une concentration élevée de BBLF2/3. En revanche le glycérol ralentit la cinétique de la cristallisation et modifie la concentration en eau dans la goutte. Il serait donc intéressant d'évaluer si la présence ou absence de glycérol influe sur la cristallogénèse.

L'obtention de cristaux exploitables étant assez aléatoire nous pourrions également envisager une étude par diffractions aux petits angles (SAXS) afin d'obtenir l'enveloppe basse résolution de la protéine. Cette approche permettrait de décrire la conformation de BBLF2/3.

Le problème majeur rencontré lors la purification du complexe est la présence des deux populations solubles du complexe. Les deux formes ne facilitent pas les investigations structurales et

enzymatiques. Il faut donc poursuivre les étapes de purification afin de pouvoir séparer les deux populations de complexe. La stabilité du complexe est dépendante du pH (7,6) et de la concentration en sel (500 mM NaCl), il n'est donc pas possible d'utiliser des colonnes de type échangeuses d'ions pour affiner la purification. Il serait intéressant de voir s'il est possible de séparer les deux complexes sur un gel natif. Si la séparation est possible par cette méthode, il sera alors possible d'utiliser la technique d'élution continue par électrophorèse (« prep-cell ») pour purifier les deux formes du complexe.

Cependant la séparation des deux formes du complexe semble un objectif relativement difficile à atteindre par des techniques de purification. Les autres facteurs importants influençant les problèmes rencontrés actuellement sont la stabilité et la solubilité du complexe. La présence d'autres protéines répliquatives interagissant potentiellement avec le complexe H-P pourrait permettre de le stabiliser et peut-être de favoriser qu'une seule forme soluble. Les protéines, BALF2 (protéine d'adhésion à l'ADN simple brin), BALF5 (sous-unité catalytique de la polymérase) et ZEBRA (facteur de transcription) sont disponibles au laboratoire, il serait donc relativement facile d'étudier l'influence de ces protéines sur la stabilité du complexe. Il serait aussi possible de vérifier les interactions mises en avant par des expériences de co-localisation cellulaire et de co-immunoprécipitation (BALF2 [90], BALF5 [91], ZEBRA [127]). Si l'addition, post-purification, de ces protéines n'apporte pas d'amélioration pour le comportement du complexe H-P, il peut être envisageable de créer de nouvelles constructions MultiBac comportant chaque protéine en plus du complexe H-P.

Concernant l'aspect fonctionnel du complexe, nous avons pu observer une activité ATPase basale indépendante du substrat ADN. Pour vérifier si ce résultat préliminaire est une caractéristique du complexe d'EBV, il faudrait répéter la mesure avec un plus grand nombre de substrats de type double brin bouts francs ou en présence d'une fourche pour vérifier si le taux de conversion d'ATP en ADP est modifié ou non. Ces conditions permettront d'estimer si la valeur ATPase basale observée est due à un substrat non adapté. D'autres tests peuvent être réalisés afin d'affiner les caractéristiques de l'activité ATPase. Une condition en présence d'EDTA permettrait d'évaluer l'importance des ions divalents, comme le  $Mg^{2+}$ , pour l'activité du complexe. L'utilisation d'une gamme d'ATP permettrait de vérifier si le complexe a besoin d'une concentration minimale en ATP pour avoir une activité maximale.

Les perspectives vis-à-vis du test hélicase sont plus délicates à décrire car nous n'avons pas trouvé de conditions permettant d'observer cette activité. Il se peut que l'absence d'activité hélicase soit due à la faible activité ATPase. Les tests mis en place permettent d'observer une activité hélicase

seulement si les deux brins sont complètement séparés. Si le complexe présente une faible activité hélicase, elle n'est peut-être pas suffisante pour permettre la séparation complète du double brin. Une publication très récente révèle qu'il est possible de suivre une activité hélicase incomplète en utilisant les propriétés de l'intercalant fluorescent de l'ADN « SYBR Green I » [173]. Cet intercalant est spécifique des substrats double brin, sa fluorescence décroît fortement en présence d'un ADN simple brin. Comme la fluorescence est maximale en début de réaction, une activité hélicase partielle pourra quand même être observée par une décroissance de la fluorescence. Cette approche pourrait être testée afin d'estimer si le complexe H-P d'EBV possède une faible activité hélicase, de plus elle est peu coûteuse et facilement réalisable. Il faudrait aussi poursuivre les tests hélicases en affinant encore les conditions car les hélicases ont besoin de conditions très strictes pour être actives. L'absence d'un seul composant ou une concentration non adaptée peut bloquer l'activité hélicase du complexe.

# Partie V : Conclusion

## **Conclusion :**

Le travail réalisé au cours de ma thèse a permis de produire le complexe H-P d'EBV sous forme soluble et de pouvoir débiter des analyses structurales et enzymatiques. L'avancement actuel du projet permet d'étudier la structure et les activités du complexe. Les résultats à venir contribueront à la description et la compréhension du rôle du complexe H-P d'EBV. Ce futur permet aussi d'envisager des informations complémentaires pour le mécanisme de réplication des herpesvirus en ayant les caractéristiques d'un complexe H-P d'un  $\gamma$ -herpès virus qui pourront être directement comparées à celles déjà connues pour les  $\alpha$ -herpèsvirus. Personnellement je suis déçu de ne pas pouvoir poursuivre les investigations sur le complexe H-P d'EBV. Mais le travail que j'ai réalisé durant ma thèse pose les bases de l'expression et de la purification du complexe facilitant les investigations structurales et enzymatiques du complexe H-P d'EBV.



# Bibliographie

1. Epstein, M.A., Achong, B.G., and Barr, Y.M., *Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma*. *Lancet*, 1964. **1**(7335): p. 702-3.
2. Epstein, M.A., Henle, G., Achong, B.G., and Barr, Y.M., *Morphological and Biological Studies on a Virus in Cultured Lymphoblasts from Burkitt's Lymphoma*. *The Journal of experimental medicine*, 1965. **121**: p. 761-70.
3. Strauss, J.H. and Strauss, E.G., *Viruses and human disease* 2002, San Diego etc.: Academic Press. VII, 383 p.
4. Young, L.S. and Rickinson, A.B., *Epstein-Barr virus: 40 years on*. *Nature reviews. Cancer*, 2004. **4**(10): p. 757-68.
5. Niederman, J.C., Miller, G., Pearson, H.A., Pagano, J.S., and Dowaliby, J.M., *Infectious mononucleosis. Epstein-Barr-virus shedding in saliva and the oropharynx*. *The New England journal of medicine*, 1976. **294**(25): p. 1355-9.
6. Tattevin, P., Cremieux, A.C., Descamps, D., and Carbon, C., *Transfusion-related infectious mononucleosis*. *Scandinavian journal of infectious diseases*, 2002. **34**(10): p. 777-8.
7. Gratama, J.W., Oosterveer, M.A., Zwaan, F.E., Lepoutre, J., Klein, G., and Ernberg, I., *Eradication of Epstein-Barr virus by allogeneic bone marrow transplantation: implications for sites of viral latency*. *Proceedings of the National Academy of Sciences of the United States of America*, 1988. **85**(22): p. 8693-6.
8. Meyohas, M.C., Marechal, V., Desire, N., Bouillie, J., Frottier, J., and Nicolas, J.C., *Study of mother-to-child Epstein-Barr virus transmission by means of nested PCRs*. *Journal of virology*, 1996. **70**(10): p. 6816-9.
9. Junker, A.K., Thomas, E.E., Radcliffe, A., Forsyth, R.B., Davidson, A.G., and Rymo, L., *Epstein-Barr virus shedding in breast milk*. *The American journal of the medical sciences*, 1991. **302**(4): p. 220-3.
10. Niedobitek, G., Agathangelou, A., Herbst, H., Whitehead, L., Wright, D.H., and Young, L.S., *Epstein-Barr virus (EBV) infection in infectious mononucleosis: virus latency, replication and phenotype of EBV-infected cells*. *The Journal of pathology*, 1997. **182**(2): p. 151-9.
11. Nicolas, J.S., J., *Virologie Médicale, Herpesviridae : Virus d'Epstein-Barr*. Presses Universitaires de Lyon, 2002.
12. Miller, N. and Hutt-Fletcher, L.M., *Epstein-Barr virus enters B cells and epithelial cells by different routes*. *Journal of virology*, 1992. **66**(6): p. 3409-14.
13. White, J.M., *Viral and cellular membrane fusion proteins*. *Annual review of physiology*, 1990. **52**: p. 675-97.
14. Fingerroth, J.D., Weis, J.J., Tedder, T.F., Strominger, J.L., Biro, P.A., and Fearon, D.T., *Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2*. *Proceedings of the National Academy of Sciences of the United States of America*, 1984. **81**(14): p. 4510-4.
15. Kirschner, A.N., Sorem, J., Longnecker, R., and Jardetzky, T.S., *Structure of Epstein-Barr virus glycoprotein 42 suggests a mechanism for triggering receptor-activated virus entry*. *Structure*, 2009. **17**(2): p. 223-33.
16. Mullen, M.M., Haan, K.M., Longnecker, R., and Jardetzky, T.S., *Structure of the Epstein-Barr virus gp42 protein bound to the MHC class II receptor HLA-DR1*. *Molecular cell*, 2002. **9**(2): p. 375-85.
17. Hutt-Fletcher, L.M., *Epstein-Barr virus entry*. *Journal of virology*, 2007. **81**(15): p. 7825-32.
18. Tugizov, S.M., Berline, J.W., and Palefsky, J.M., *Epstein-Barr virus infection of polarized tongue and nasopharyngeal epithelial cells*. *Nature medicine*, 2003. **9**(3): p. 307-14.
19. Wang, X. and Hutt-Fletcher, L.M., *Epstein-Barr virus lacking glycoprotein gp42 can bind to B cells but is not able to infect*. *Journal of virology*, 1998. **72**(1): p. 158-63.
20. Wang, X., Kenyon, W.J., Li, Q., Mullberg, J., and Hutt-Fletcher, L.M., *Epstein-Barr virus uses different complexes of glycoproteins gH and gL to infect B lymphocytes and epithelial cells*. *Journal of virology*, 1998. **72**(7): p. 5552-8.

21. Hadinoto, V., Shapiro, M., Sun, C.C., and Thorley-Lawson, D.A., *The dynamics of EBV shedding implicate a central role for epithelial cells in amplifying viral output*. PLoS pathogens, 2009. **5**(7): p. e1000496.
22. Borza, C.M. and Hutt-Fletcher, L.M., *Alternate replication in B cells and epithelial cells switches tropism of Epstein-Barr virus*. Nature medicine, 2002. **8**(6): p. 594-9.
23. Adams, A. and Lindahl, T., *Epstein-Barr virus genomes with properties of circular DNA molecules in carrier cells*. Proceedings of the National Academy of Sciences of the United States of America, 1975. **72**(4): p. 1477-81.
24. Rowe, M., Lear, A.L., Croom-Carter, D., Davies, A.H., and Rickinson, A.B., *Three pathways of Epstein-Barr virus gene activation from EBNA1-positive latency in B lymphocytes*. Journal of virology, 1992. **66**(1): p. 122-31.
25. Kutok, J.L. and Wang, F., *Spectrum of Epstein-Barr virus-associated diseases*. Annual review of pathology, 2006. **1**: p. 375-404.
26. Chaudhuri, B., Xu, H., Todorov, I., Dutta, A., and Yates, J.L., *Human DNA replication initiation factors, ORC and MCM, associate with oriP of Epstein-Barr virus*. Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(18): p. 10085-9.
27. Ritzi, M., et al., *Complex protein-DNA dynamics at the latent origin of DNA replication of Epstein-Barr virus*. Journal of cell science, 2003. **116**(Pt 19): p. 3971-84.
28. Humme, S., Reisbach, G., Feederle, R., Delecluse, H.J., Bousset, K., Hammerschmidt, W., and Schepers, A., *The EBV nuclear antigen 1 (EBNA1) enhances B cell immortalization several thousandfold*. Proceedings of the National Academy of Sciences of the United States of America, 2003. **100**(19): p. 10989-94.
29. Levitskaya, J., et al., *Inhibition of antigen processing by the internal repeat region of the Epstein-Barr virus nuclear antigen-1*. Nature, 1995. **375**(6533): p. 685-8.
30. Cohen, J.I., Wang, F., Mannick, J., and Kieff, E., *Epstein-Barr virus nuclear protein 2 is a key determinant of lymphocyte transformation*. Proceedings of the National Academy of Sciences of the United States of America, 1989. **86**(23): p. 9558-62.
31. Abbot, S.D., et al., *Epstein-Barr virus nuclear antigen 2 induces expression of the virus-encoded latent membrane protein*. Journal of virology, 1990. **64**(5): p. 2126-34.
32. Marshall, D. and Sample, C., *Epstein-Barr virus nuclear antigen 3C is a transcriptional regulator*. Journal of virology, 1995. **69**(6): p. 3624-30.
33. Zhao, B. and Sample, C.E., *Epstein-barr virus nuclear antigen 3C activates the latent membrane protein 1 promoter in the presence of Epstein-Barr virus nuclear antigen 2 through sequences encompassing an spi-1/Spi-B binding site*. Journal of virology, 2000. **74**(11): p. 5151-60.
34. Harada, S. and Kieff, E., *Epstein-Barr virus nuclear protein LP stimulates EBNA-2 acidic domain-mediated transcriptional activation*. Journal of virology, 1997. **71**(9): p. 6611-8.
35. Rickinson, A.B.a.K., E, *Epstein-Barr virus*, in *Fields Virology*, D.M. Knipe, Howley, P.M., Griffin, B.D., Lamb, R.A., Martin, M.A., Roizman, B. and Straus, S.E., Editor 2007, Lippincott Williams et Wilkins: Philadelphia. p. 2655-2700.
36. Hatzivassiliou, E.G., Kieff, E., and Mosialos, G., *Constitutive CD40 signaling phenocopies the transforming function of the Epstein-Barr virus oncoprotein LMP1 in vitro*. Leukemia research, 2007. **31**(3): p. 315-20.
37. Merchant, M., Caldwell, R.G., and Longnecker, R., *The LMP2A ITAM is essential for providing B cells with development and survival signals in vivo*. Journal of virology, 2000. **74**(19): p. 9115-24.
38. Takada, K., *Role of EBER and BARF1 in nasopharyngeal carcinoma (NPC) tumorigenesis*. Seminars in cancer biology, 2012. **22**(2): p. 162-5.
39. Samanta, M., Iwakiri, D., and Takada, K., *Epstein-Barr virus-encoded small RNA induces IL-10 through RIG-I-mediated IRF-3 signaling*. Oncogene, 2008. **27**(30): p. 4150-60.

40. Feederle, R., Linnstaedt, S.D., Bannert, H., Lips, H., Bencun, M., Cullen, B.R., and Delecluse, H.J., *A viral microRNA cluster strongly potentiates the transforming properties of a human herpesvirus*. PLoS pathogens, 2011. **7**(2): p. e1001294.
41. Fahmi, H., Cochet, C., Hmama, Z., Opolon, P., and Joab, I., *Transforming growth factor beta 1 stimulates expression of the Epstein-Barr virus BZLF1 immediate-early gene product ZEBRA by an indirect mechanism which requires the MAPK kinase pathway*. Journal of virology, 2000. **74**(13): p. 5810-8.
42. Luka, J., Kallin, B., and Klein, G., *Induction of the Epstein-Barr virus (EBV) cycle in latently infected cells by n-butyrate*. Virology, 1979. **94**(1): p. 228-31.
43. Petosa, C., Morand, P., Baudin, F., Moulin, M., Artero, J.B., and Muller, C.W., *Structural basis of lytic cycle activation by the Epstein-Barr virus ZEBRA protein*. Molecular cell, 2006. **21**(4): p. 565-72.
44. Dickerson, S.J., Xing, Y., Robinson, A.R., Seaman, W.T., Gruffat, H., and Kenney, S.C., *Methylation-dependent binding of the Epstein-Barr virus BZLF1 protein to viral promoters*. PLoS pathogens, 2009. **5**(3): p. e1000356.
45. Bhende, P.M., Seaman, W.T., Delecluse, H.J., and Kenney, S.C., *BZLF1 activation of the methylated form of the BRLF1 immediate-early promoter is regulated by BZLF1 residue 186*. Journal of virology, 2005. **79**(12): p. 7338-48.
46. Adamson, A.L., Darr, D., Holley-Guthrie, E., Johnson, R.A., Mauser, A., Swenson, J., and Kenney, S., *Epstein-Barr virus immediate-early proteins BZLF1 and BRLF1 activate the ATF2 transcription factor by increasing the levels of phosphorylated p38 and c-Jun N-terminal kinases*. Journal of virology, 2000. **74**(3): p. 1224-33.
47. Strockbine, L.D., et al., *The Epstein-Barr virus BARF1 gene encodes a novel, soluble colony-stimulating factor-1 receptor*. Journal of virology, 1998. **72**(5): p. 4015-21.
48. Tarbouriech, N., Ruggiero, F., de Turenne-Tessier, M., Ooka, T., and Burmeister, W.P., *Structure of the Epstein-Barr virus oncogene BARF1*. Journal of molecular biology, 2006. **359**(3): p. 667-78.
49. Salek-Ardakani, S., Arrand, J.R., and Mackett, M., *Epstein-Barr virus encoded interleukin-10 inhibits HLA-class I, ICAM-1, and B7 expression on human monocytes: implications for immune evasion by EBV*. Virology, 2002. **304**(2): p. 342-51.
50. Fixman, E.D., Hayward, G.S., and Hayward, S.D., *trans-acting requirements for replication of Epstein-Barr virus ori-Lyt*. Journal of virology, 1992. **66**(8): p. 5030-9.
51. Fixman, E.D., Hayward, G.S., and Hayward, S.D., *Replication of Epstein-Barr virus oriLyt: lack of a dedicated virally encoded origin-binding protein and dependence on Zta in cotransfection assays*. Journal of virology, 1995. **69**(5): p. 2998-3006.
52. Zimmermann, J. and Hammerschmidt, W., *Structure and role of the terminal repeats of Epstein-Barr virus in processing and packaging of virion DNA*. Journal of virology, 1995. **69**(5): p. 3147-55.
53. Henson, B.W., Perkins, E.M., Cothran, J.E., and Desai, P., *Self-assembly of Epstein-Barr virus capsids*. Journal of virology, 2009. **83**(8): p. 3877-90.
54. Germi, R., Effantin, G., Grossi, L., Ruigrok, R.W., Morand, P., and Schoehn, G., *Three-dimensional structure of the Epstein-Barr virus capsid*. The Journal of general virology, 2012. **93**(Pt 8): p. 1769-73.
55. Sixbey, J.W., Nedrud, J.G., Raab-Traub, N., Hanes, R.A., and Pagano, J.S., *Epstein-Barr virus replication in oropharyngeal epithelial cells*. The New England journal of medicine, 1984. **310**(19): p. 1225-30.
56. Klein, G., Svedmyr, E., Jondal, M., and Persson, P.O., *EBV-determined nuclear antigen (EBNA)-positive cells in the peripheral blood of infectious mononucleosis patients*. International journal of cancer. Journal international du cancer, 1976. **17**(1): p. 21-6.

57. Tierney, R.J., Steven, N., Young, L.S., and Rickinson, A.B., *Epstein-Barr virus latency in blood mononuclear cells: analysis of viral gene transcription during primary infection and in the carrier state*. Journal of virology, 1994. **68**(11): p. 7374-85.
58. Hislop, A.D., Taylor, G.S., Sauce, D., and Rickinson, A.B., *Cellular responses to viral infection in humans: lessons from Epstein-Barr virus*. Annual review of immunology, 2007. **25**: p. 587-617.
59. Bertrand, S., et al., *Variant translocation in a non endemic case of Burkitt's lymphoma: t(8;22) in an Epstein-Barr virus negative tumour and in a derived cell line*. European journal of cancer, 1981. **17**(5): p. 577-84.
60. Bernheim, A., Berger, R., and Lenoir, G., *Cytogenetic studies on African Burkitt's lymphoma cell lines: t(8;14), t(2;8) and t(8;22) translocations*. Cancer genetics and cytogenetics, 1981. **3**(4): p. 307-15.
61. Niedobitek, G., et al., *Heterogeneous expression of Epstein-Barr virus latent proteins in endemic Burkitt's lymphoma*. Blood, 1995. **86**(2): p. 659-65.
62. Gerrard, M., et al., *Excellent survival following two courses of COPAD chemotherapy in children and adolescents with resected localized B-cell non-Hodgkin's lymphoma: results of the FAB/LMB 96 international study*. British journal of haematology, 2008. **141**(6): p. 840-7.
63. Dolcetti, R. and Boiocchi, M., *Epstein-Barr virus in the pathogenesis of Hodgkin's disease*. Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie, 1998. **52**(1): p. 13-25.
64. Kapatai, G. and Murray, P., *Contribution of the Epstein Barr virus to the molecular pathogenesis of Hodgkin lymphoma*. Journal of clinical pathology, 2007. **60**(12): p. 1342-9.
65. Kuppers, R., Engert, A., and Hansmann, M.L., *Hodgkin lymphoma*. The Journal of clinical investigation, 2012. **122**(10): p. 3439-47.
66. Yu, M.C. and Yuan, J.M., *Epidemiology of nasopharyngeal carcinoma*. Seminars in cancer biology, 2002. **12**(6): p. 421-9.
67. Busson, P., Ooka, T., and Corbex, M., *[Nasopharyngeal carcinomas and Epstein-Barr virus: from epidemiology and detection to therapy]*. Medecine sciences : M/S, 2004. **20**(4): p. 453-7.
68. Klein, G., Giovanella, B.C., Lindahl, T., Fialkow, P.J., Singh, S., and Stehlin, J.S., *Direct evidence for the presence of Epstein-Barr virus DNA and nuclear antigen in malignant epithelial cells from patients with poorly differentiated carcinoma of the nasopharynx*. Proceedings of the National Academy of Sciences of the United States of America, 1974. **71**(12): p. 4737-41.
69. Raab-Traub, N., *Epstein-Barr virus in the pathogenesis of NPC*. Seminars in cancer biology, 2002. **12**(6): p. 431-41.
70. Cochet, C., et al., *Expression of the Epstein-Barr virus immediate early gene, BZLF1, in nasopharyngeal carcinoma tumor cells*. Virology, 1993. **197**(1): p. 358-65.
71. Lee, A.W., et al., *Treatment results for nasopharyngeal carcinoma in the modern era: the Hong Kong experience*. International journal of radiation oncology, biology, physics, 2005. **61**(4): p. 1107-16.
72. Shibata, D. and Weiss, L.M., *Epstein-Barr virus-associated gastric adenocarcinoma*. The American journal of pathology, 1992. **140**(4): p. 769-74.
73. Camargo, M.C., et al., *Determinants of Epstein-Barr virus-positive gastric cancer: an international pooled analysis*. British journal of cancer, 2011. **105**(1): p. 38-43.
74. Young, L.S. and Murray, P.G., *Epstein-Barr virus and oncogenesis: from latent genes to tumours*. Oncogene, 2003. **22**(33): p. 5108-21.
75. Carbone, A., Cesarman, E., Spina, M., Gloghini, A., and Schulz, T.F., *HIV-associated lymphomas and gamma-herpesviruses*. Blood, 2009. **113**(6): p. 1213-24.
76. Greenspan, J.S., Greenspan, D., Lennette, E.T., Abrams, D.I., Conant, M.A., and Petersen, V.H., *Oral viral leukoplakia--a new AIDS-associated condition*. Advances in experimental medicine and biology, 1985. **187**: p. 123-8.

77. Ma, S.D., et al., *A new model of Epstein-Barr virus infection reveals an important role for early lytic viral protein expression in the development of lymphomas*. Journal of virology, 2011. **85**(1): p. 165-77.
78. Sugino, A., Hirose, S., and Okazaki, R., *RNA-linked nascent DNA fragments in Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America, 1972. **69**(7): p. 1863-7.
79. Wu, C.A., Nelson, N.J., McGeoch, D.J., and Challberg, M.D., *Identification of herpes simplex virus type 1 genes required for origin-dependent DNA synthesis*. Journal of virology, 1988. **62**(2): p. 435-43.
80. Tsurumi, T., *EBV replication enzymes*. Current topics in microbiology and immunology, 2001. **258**: p. 65-87.
81. Pari, G.S. and Anders, D.G., *Eleven loci encoding trans-acting factors are required for transient complementation of human cytomegalovirus oriLyt-dependent DNA replication*. Journal of virology, 1993. **67**(12): p. 6979-88.
82. Russo, J.J., et al., *Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8)*. Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(25): p. 14862-7.
83. Honess, R.W. and Roizman, B., *Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins*. Journal of virology, 1974. **14**(1): p. 8-19.
84. Stow, N.D., Hammarsten, O., Arbuckle, M.I., and Elias, P., *Inhibition of herpes simplex virus type 1 DNA replication by mutant forms of the origin-binding protein*. Virology, 1993. **196**(2): p. 413-8.
85. Sun, R., Lin, S.F., Gradoville, L., Yuan, Y., Zhu, F., and Miller, G., *A viral gene that activates lytic cycle expression of Kaposi's sarcoma-associated herpesvirus*. Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(18): p. 10866-71.
86. Gawn, J.M. and Greaves, R.F., *Absence of IE1 p72 protein function during low-multiplicity infection by human cytomegalovirus results in a broad block to viral delayed-early gene expression*. Journal of virology, 2002. **76**(9): p. 4441-55.
87. Cohen, J.I., Fauci, A.S., Varmus, H., and Nabel, G.J., *Epstein-Barr virus: an important vaccine target for cancer prevention*. Science translational medicine, 2011. **3**(107): p. 107fs7.
88. Hammerschmidt, W. and Sugden, B., *Identification and characterization of oriLyt, a lytic origin of DNA replication of Epstein-Barr virus*. Cell, 1988. **55**(3): p. 427-33.
89. Rennekamp, A.J. and Lieberman, P.M., *Initiation of Epstein-Barr virus lytic replication requires transcription and the formation of a stable RNA-DNA hybrid molecule at OriLyt*. Journal of virology, 2011. **85**(6): p. 2837-50.
90. Gao, Z., Krithivas, A., Finan, J.E., Semmes, O.J., Zhou, S., Wang, Y., and Hayward, S.D., *The Epstein-Barr virus lytic transactivator Zta interacts with the helicase-primase replication proteins*. Journal of virology, 1998. **72**(11): p. 8559-67.
91. Fujii, K., et al., *The Epstein-Barr virus pol catalytic subunit physically interacts with the BBLF4-BSLF1-BBLF2/3 complex*. Journal of virology, 2000. **74**(6): p. 2550-7.
92. Wang, Y., Tang, Q., Maul, G.G., and Yuan, Y., *Kaposi's sarcoma-associated herpesvirus ori-Lyt-dependent DNA replication: dual role of replication and transcription activator*. Journal of virology, 2006. **80**(24): p. 12171-86.
93. Prichard, M.N., Jairath, S., Penfold, M.E., St Jeor, S., Bohlman, M.C., and Pari, G.S., *Identification of persistent RNA-DNA hybrid structures within the origin of replication of human cytomegalovirus*. Journal of virology, 1998. **72**(9): p. 6997-7004.
94. Yokoyama, N., et al., *Assembly of the Epstein-Barr virus BBLF4, BSLF1 and BBLF2/3 proteins and their interactive properties*. The Journal of general virology, 1999. **80** ( Pt 11): p. 2879-87.

95. McMahon, T.P. and Anders, D.G., *Interactions between human cytomegalovirus helicase-primase proteins*. *Virus research*, 2002. **86**(1-2): p. 39-52.
96. McLean, G.W., Abbotts, A.P., Parry, M.E., Marsden, H.S., and Stow, N.D., *The herpes simplex virus type 1 origin-binding protein interacts specifically with the viral UL8 protein*. *The Journal of general virology*, 1994. **75 ( Pt 10)**: p. 2699-706.
97. Falkenberg, M., Bushnell, D.A., Elias, P., and Lehman, I.R., *The UL8 subunit of the heterotrimeric herpes simplex virus type 1 helicase-primase is required for the unwinding of single strand DNA-binding protein (ICP8)-coated DNA substrates*. *The Journal of biological chemistry*, 1997. **272**(36): p. 22766-70.
98. Barnard, E.C., Brown, G., and Stow, N.D., *Deletion mutants of the herpes simplex virus type 1 UL8 protein: effect on DNA synthesis and ability to interact with and influence the intracellular localization of the UL5 and UL52 proteins*. *Virology*, 1997. **237**(1): p. 97-106.
99. Ahn, J.H. and Hayward, G.S., *The major immediate-early proteins IE1 and IE2 of human cytomegalovirus colocalize with and disrupt PML-associated nuclear bodies at very early times in infected permissive cells*. *Journal of virology*, 1997. **71**(6): p. 4599-613.
100. Burkham, J., Coen, D.M., and Weller, S.K., *ND10 protein PML is recruited to herpes simplex virus type 1 prereplicative sites and replication compartments in the presence of viral DNA polymerase*. *Journal of virology*, 1998. **72**(12): p. 10100-7.
101. Wu, F.Y., Ahn, J.H., Alcendor, D.J., Jang, W.J., Xiao, J., Hayward, S.D., and Hayward, G.S., *Origin-independent assembly of Kaposi's sarcoma-associated herpesvirus DNA replication compartments in transient cotransfection assays and association with the ORF-K8 protein and cellular PML*. *Journal of virology*, 2001. **75**(3): p. 1487-506.
102. Calder, J.M. and Stow, N.D., *Herpes simplex virus helicase-primase: the UL8 protein is not required for DNA-dependent ATPase and DNA helicase activities*. *Nucleic acids research*, 1990. **18**(12): p. 3573-8.
103. Dodson, M.S. and Lehman, I.R., *Association of DNA helicase and primase activities with a subassembly of the herpes simplex virus 1 helicase-primase composed of the UL5 and UL52 gene products*. *Proceedings of the National Academy of Sciences of the United States of America*, 1991. **88**(4): p. 1105-9.
104. Crute, J.J., Mocarski, E.S., and Lehman, I.R., *A DNA helicase induced by herpes simplex virus type 1*. *Nucleic acids research*, 1988. **16**(14A): p. 6585-96.
105. Graves-Woodward, K.L. and Weller, S.K., *Replacement of gly815 in helicase motif V alters the single-stranded DNA-dependent ATPase activity of the herpes simplex virus type 1 helicase-primase*. *The Journal of biological chemistry*, 1996. **271**(23): p. 13629-35.
106. Falkenberg, M., Elias, P., and Lehman, I.R., *The herpes simplex virus type 1 helicase-primase. Analysis of helicase activity*. *The Journal of biological chemistry*, 1998. **273**(48): p. 32154-7.
107. Chen, Y., Carrington-Lawrence, S.D., Bai, P., and Weller, S.K., *Mutations in the putative zinc-binding motif of UL52 demonstrate a complex interdependence between the UL5 and UL52 subunits of the human herpes simplex virus type 1 helicase/primase complex*. *Journal of virology*, 2005. **79**(14): p. 9088-96.
108. Chono, K., et al., *ASP2151, a novel helicase-primase inhibitor, possesses antiviral activity against varicella-zoster virus and herpes simplex virus types 1 and 2*. *The Journal of antimicrobial chemotherapy*, 2010. **65**(8): p. 1733-41.
109. Graves-Woodward, K.L., Gottlieb, J., Challberg, M.D., and Weller, S.K., *Biochemical analyses of mutations in the HSV-1 helicase-primase that alter ATP hydrolysis, DNA unwinding, and coupling between hydrolysis and unwinding*. *The Journal of biological chemistry*, 1997. **272**(7): p. 4623-30.
110. Zhu, L.A. and Weller, S.K., *The six conserved helicase motifs of the UL5 gene product, a component of the herpes simplex virus type 1 helicase-primase, are essential for its function*. *Journal of virology*, 1992. **66**(1): p. 469-79.



111. Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J., *Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold*. The EMBO journal, 1982. **1**(8): p. 945-51.
112. Fry, D.C., Kuby, S.A., and Mildvan, A.S., *ATP-binding site of adenylate kinase: mechanistic implications of its homology with ras-encoded p21, F1-ATPase, and other nucleotide-binding proteins*. Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(4): p. 907-11.
113. Story, R.M. and Steitz, T.A., *Structure of the recA protein-ADP complex*. Nature, 1992. **355**(6358): p. 374-6.
114. Copeland, W.C. and Wang, T.S., *Mutational analysis of the human DNA polymerase alpha. The most conserved region in alpha-like DNA polymerases is involved in metal-specific catalysis*. The Journal of biological chemistry, 1993. **268**(15): p. 11028-40.
115. Braithwaite, D.K. and Ito, J., *Compilation, alignment, and phylogenetic relationships of DNA polymerases*. Nucleic acids research, 1993. **21**(4): p. 787-802.
116. Ilyina, T.V., Gorbalenya, A.E., and Koonin, E.V., *Organization and evolution of bacterial and bacteriophage primase-helicase systems*. Journal of molecular evolution, 1992. **34**(4): p. 351-7.
117. Dracheva, S., Koonin, E.V., and Crute, J.J., *Identification of the primase active site of the herpes simplex virus type 1 helicase-primase*. The Journal of biological chemistry, 1995. **270**(23): p. 14148-53.
118. Mendelman, L.V., Beauchamp, B.B., and Richardson, C.C., *Requirement for a zinc motif for template recognition by the bacteriophage T7 primase*. The EMBO journal, 1994. **13**(16): p. 3909-16.
119. Pavletich, N.P. and Pabo, C.O., *Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å*. Science, 1991. **252**(5007): p. 809-17.
120. Agarwal, K., Baek, K.H., Jeon, C.J., Miyamoto, K., Ueno, A., and Yoon, H.S., *Stimulation of transcript elongation requires both the zinc finger and RNA polymerase II binding domains of human TFIIIS*. Biochemistry, 1991. **30**(31): p. 7842-51.
121. Lee, S.J., Zhu, B., Akabayov, B., and Richardson, C.C., *Zinc-binding domain of the bacteriophage T7 DNA primase modulates binding to the DNA template*. The Journal of biological chemistry, 2012. **287**(46): p. 39030-40.
122. Carrington-Lawrence, S.D. and Weller, S.K., *Recruitment of polymerase to herpes simplex virus type 1 replication foci in cells expressing mutant primase (UL52) proteins*. Journal of virology, 2003. **77**(7): p. 4237-47.
123. Chen, Y., Bai, P., Mackay, S., Korza, G., Carson, J.H., Kuchta, R.D., and Weller, S.K., *Herpes simplex virus type 1 helicase-primase: DNA binding and consequent protein oligomerization and primase activation*. Journal of virology, 2011. **85**(2): p. 968-78.
124. Marsden, H.S., et al., *The catalytic subunit of the DNA polymerase of herpes simplex virus type 1 interacts specifically with the C terminus of the UL8 component of the viral helicase-primase complex*. Journal of virology, 1997. **71**(9): p. 6390-7.
125. Falkenberg, M., Lehman, I.R., and Elias, P., *Leading and lagging strand DNA synthesis in vitro by a reconstituted herpes simplex virus type 1 replisome*. Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(8): p. 3896-900.
126. Stengel, G. and Kuchta, R.D., *Coordinated leading and lagging strand DNA synthesis by using the herpes simplex virus 1 replication complex and minicircle DNA templates*. Journal of virology, 2011. **85**(2): p. 957-67.
127. Liao, G., Wu, F.Y., and Hayward, S.D., *Interaction with the Epstein-Barr virus helicase targets Zta to DNA replication compartments*. Journal of virology, 2001. **75**(18): p. 8792-802.

128. El-Guindy, A., Heston, L., and Miller, G., *A subset of replication proteins enhances origin recognition and lytic replication by the Epstein-Barr virus ZEBRA protein*. PLoS pathogens, 2010. **6**(8): p. e1001054.
129. Liao, G., Huang, J., Fixman, E.D., and Hayward, S.D., *The Epstein-Barr virus replication protein BBLF2/3 provides an origin-tethering function through interaction with the zinc finger DNA binding protein ZBRK1 and the KAP-1 corepressor*. Journal of virology, 2005. **79**(1): p. 245-56.
130. Jongeneel, C.V., Bedinger, P., and Alberts, B.M., *Effects of the bacteriophage T4 dda protein on DNA synthesis catalyzed by purified T4 replication proteins*. The Journal of biological chemistry, 1984. **259**(20): p. 12933-8.
131. Gauss, P., Park, K., Spencer, T.E., and Hacker, K.J., *DNA helicase requirements for DNA replication during bacteriophage T4 infection*. Journal of bacteriology, 1994. **176**(6): p. 1667-72.
132. Jongeneel, C.V., Formosa, T., and Alberts, B.M., *Purification and characterization of the bacteriophage T4 dda protein. A DNA helicase that associates with the viral helix-destabilizing protein*. The Journal of biological chemistry, 1984. **259**(20): p. 12925-32.
133. Ma, Y., Wang, T., Villemain, J.L., Giedroc, D.P., and Morrical, S.W., *Dual functions of single-stranded DNA-binding protein in helicase loading at the bacteriophage T4 DNA replication fork*. The Journal of biological chemistry, 2004. **279**(18): p. 19035-45.
134. Behme, M.T. and Ebisuzaki, K., *Characterization of a bacteriophage T4 mutant lacking DNA-dependent ATPase*. Journal of virology, 1975. **15**(1): p. 50-4.
135. Brister, J.R., *Origin activation requires both replicative and accessory helicases during T4 infection*. Journal of molecular biology, 2008. **377**(5): p. 1304-13.
136. Perumal, S.K., Raney, K.D., and Benkovic, S.J., *Analysis of the DNA translocation and unwinding activities of T4 phage helicases*. Methods, 2010. **51**(3): p. 277-88.
137. Byrd, A.K. and Raney, K.D., *Increasing the length of the single-stranded overhang enhances unwinding of duplex DNA by bacteriophage T4 Dda helicase*. Biochemistry, 2005. **44**(39): p. 12990-7.
138. Byrd, A.K. and Raney, K.D., *Protein displacement by an assembly of helicase molecules aligned along single-stranded DNA*. Nature structural & molecular biology, 2004. **11**(6): p. 531-8.
139. Eoff, R.L. and Raney, K.D., *Intermediates revealed in the kinetic mechanism for DNA unwinding by a monomeric helicase*. Nature structural & molecular biology, 2006. **13**(3): p. 242-9.
140. He, X., et al., *The T4 phage SF1B helicase Dda is structurally optimized to perform DNA strand separation*. Structure, 2012. **20**(7): p. 1189-200.
141. Saikrishnan, K., Powell, B., Cook, N.J., Webb, M.R., and Wigley, D.B., *Mechanistic basis of 5'-3' translocation in SF1B helicases*. Cell, 2009. **137**(5): p. 849-59.
142. Lee, S.S. and Lehman, I.R., *Unwinding of the box I element of a herpes simplex virus type 1 origin by a complex of the viral origin binding protein, single-strand DNA binding protein, and single-stranded DNA*. Proceedings of the National Academy of Sciences of the United States of America, 1997. **94**(7): p. 2838-42.
143. He, X. and Lehman, I.R., *Unwinding of a herpes simplex virus type 1 origin of replication (Ori(S)) by a complex of the viral origin binding protein and the single-stranded DNA binding protein*. Journal of virology, 2000. **74**(12): p. 5726-8.
144. Buisson, M., Geoui, T., Flot, D., Tarbouriech, N., Rensing, M.E., Wiertz, E.J., and Burmeister, W.P., *A bridge crosses the active-site canyon of the Epstein-Barr virus nuclease with DNase and RNase activities*. Journal of molecular biology, 2009. **391**(4): p. 717-28.
145. Sele, C., Gabel, F., Gutsche, I., Ivanov, I., Burmeister, W.P., Iseni, F., and Tarbouriech, N., *Low-resolution structure of vaccinia virus DNA replication machinery*. Journal of virology, 2013. **87**(3): p. 1679-89.

146. Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J., and Russell, R.B., *Protein disorder prediction: implications for structural proteomics*. *Structure*, 2003. **11**(11): p. 1453-9.
147. Ward, J.J., McGuffin, L.J., Bryson, K., Buxton, B.F., and Jones, D.T., *The DISOPRED server for the prediction of protein disorder*. *Bioinformatics*, 2004. **20**(13): p. 2138-9.
148. Prilusky, J., et al., *FoldIndex: a simple tool to predict whether a given protein sequence is intrinsically unfolded*. *Bioinformatics*, 2005. **21**(16): p. 3435-8.
149. Linding, R., Russell, R.B., Neduva, V., and Gibson, T.J., *GlobPlot: Exploring protein sequences for globularity and disorder*. *Nucleic acids research*, 2003. **31**(13): p. 3701-8.
150. Dosztanyi, Z., Csizmok, V., Tompa, P., and Simon, I., *IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content*. *Bioinformatics*, 2005. **21**(16): p. 3433-4.
151. Shimizu, K., Hirose, S., and Noguchi, T., *POODLE-S: web application for predicting protein disorder by using physicochemical features and reduced amino acid set of a position-specific scoring matrix*. *Bioinformatics*, 2007. **23**(17): p. 2337-8.
152. Hirose, S., Shimizu, K., Kanai, S., Kuroda, Y., and Noguchi, T., *POODLE-L: a two-level SVM prediction system for reliably predicting long disordered regions*. *Bioinformatics*, 2007. **23**(16): p. 2046-53.
153. Hirose, S., Shimizu, K., and Noguchi, T., *POODLE-I: Disordered Region Prediction by Integrating POODLE Series and Structural Information Predictors Based on a Workflow Approach*. *In silico biology*, 2010. **10**(3): p. 185-91.
154. Ishida, T. and Kinoshita, K., *PrDOS: prediction of disordered protein regions from amino acid sequence*. *Nucleic acids research*, 2007. **35**(Web Server issue): p. W460-4.
155. Yang, Z.R., Thomson, R., McNeil, P., and Esnouf, R.M., *RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins*. *Bioinformatics*, 2005. **21**(16): p. 3369-76.
156. Vullo, A., Bortolami, O., Pollastri, G., and Tosatto, S.C., *Spritz: a server for the prediction of intrinsically disordered regions in protein sequences using kernel machines*. *Nucleic acids research*, 2006. **34**(Web Server issue): p. W164-8.
157. Vucetic, S., Brown, C.J., Dunker, A.K., and Obradovic, Z., *Flavors of protein disorder*. *Proteins*, 2003. **52**(4): p. 573-84.
158. Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., Brown, C.J., and Dunker, A.K., *Predicting intrinsic disorder from amino acid sequence*. *Proteins*, 2003. **53 Suppl 6**: p. 566-72.
159. Obradovic, Z., Peng, K., Vucetic, S., Radivojac, P., and Dunker, A.K., *Exploiting heterogeneous sequence properties improves prediction of protein disorder*. *Proteins*, 2005. **61 Suppl 7**: p. 176-82.
160. Rost, B. and Liu, J., *The PredictProtein server*. *Nucleic acids research*, 2003. **31**(13): p. 3300-4.
161. Gasteiger E., H.C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A., *The Proteomics Protocols Handbook*, ed. J.M. Walker 2005: Humana Press.
162. Kosugi, S., Hasebe, M., Matsumura, N., Takashima, H., Miyamoto-Sato, E., Tomita, M., and Yanagawa, H., *Six classes of nuclear localization signals specific to different binding grooves of importin alpha*. *The Journal of biological chemistry*, 2009. **284**(1): p. 478-85.
163. Kosugi, S., Hasebe, M., Tomita, M., and Yanagawa, H., *Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs*. *Proceedings of the National Academy of Sciences of the United States of America*, 2009. **106**(25): p. 10171-6.
164. Baer, R., et al., *DNA sequence and expression of the B95-8 Epstein-Barr virus genome*. *Nature*, 1984. **310**(5974): p. 207-11.
165. Maniatis, T., Fritsch, E.F., and Sambrook, J., *Molecular Cloning: A Laboratory Manual* 1987: Cold Spring Harbor Laboratory.

166. Fitzgerald, D.J., Berger, P., Schaffitzel, C., Yamada, K., Richmond, T.J., and Berger, I., *Protein complex expression by using multigene baculoviral vectors*. *Nature methods*, 2006. **3**(12): p. 1021-32.
167. Schreiner, U., Theune, M., Althof, F., Kehm, E., and Knopf, C.W., *One-step column purification of herpes simplex virus 1 helicase-primase subcomplex using C-terminally his-tagged UL5 subunit*. *Virus genes*, 2009. **39**(1): p. 19-29.
168. Golovanov, A.P., Hautbergue, G.M., Wilson, S.A., and Lian, L.Y., *A simple method for improving protein solubility and long-term stability*. *Journal of the American Chemical Society*, 2004. **126**(29): p. 8933-9.
169. Blobel, J., et al., *Protein loop compaction and the origin of the effect of arginine and glutamic acid mixtures on solubility, stability and transient oligomerization of proteins*. *European biophysics journal : EBJ*, 2011. **40**(12): p. 1327-38.
170. Kastner, B., et al., *GraFix: sample preparation for single-particle electron cryomicroscopy*. *Nature methods*, 2008. **5**(1): p. 53-5.
171. Kleymann, G., et al., *New helicase-primase inhibitors as drug candidates for the treatment of herpes simplex disease*. *Nature medicine*, 2002. **8**(4): p. 392-8.
172. Dodson, M.S., Crute, J.J., Bruckner, R.C., and Lehman, I.R., *Overexpression and assembly of the herpes simplex virus type 1 helicase-primase in insect cells*. *The Journal of biological chemistry*, 1989. **264**(35): p. 20835-8.
173. Siddiqui, S., Khan, I., Zarina, S., and Ali, S., *Use of the SYBR Green dye for measuring helicase activity*. *Enzyme and microbial technology*, 2013. **52**(3): p. 196-8.



## **Résumé :**

Le virus Epstein-Barr (EBV) est un gamma herpèsvirus humain infectant plus de 95 % de la population mondiale. Lorsque la primo-infection a lieu pendant l'adolescence ou à l'âge adulte, elle peut induire la mononucléose infectieuse (MNI), cette maladie est le plus souvent bénigne. EBV est aussi associé à un certain nombre de cancers de type lymphome (lymphomes de Burkitt et d'Hodgkin) et de type carcinome (carcinomes gastriques et indifférenciés du rhinopharynx). L'importance des protéines de latence du virus dans l'apparition des tumeurs a été très étudiée. Des études récentes montrent que les protéines lytiques d'EBV sont aussi très importantes pour l'apparition et le développement des tumeurs. Le complexe Hélicase-Primase (H-P) du virus herpès simplex 1 (HSV-1, alpha herpèsvirus) est la cible de nouveaux antiviraux. Les activités ATPase, hélicase et primase du complexe d'HSV-1 ont été largement étudiées, mais aucune information structurale du complexe H-P n'est disponible actuellement pour un membre des herpèsvirus humains. Nous avons entrepris l'étude du complexe H-P d'EBV (BBLF4 : hélicase, BSLF1 : primase et BBLF2/3 : sous-unité accessoire) afin de caractériser sa structure et les activités qu'il porte. Nous avons pu établir les conditions d'expression et de purification du complexe et débiter des études structurales et enzymatiques préliminaires. Nous avons pu observer une activité ATPase basale du complexe indépendante de la présence d'un substrat ADN simple brin. Nous observons deux formes solubles du complexe lors des purifications, une présentant probablement une stœchiométrie proche de 1/1/1 et une seconde forme ayant surement un excès de la protéine Hélicase (BBLF4). Ces premiers résultats apportent des informations nouvelles pour le complexe H-P d'EBV et doivent être poursuivis afin de les confirmer et de pouvoir les comparer avec ceux déjà connus pour le complexe H-P d'HSV-1.

**Mots clés : EBV, complexe Hélicase-Primase, BBLF4, BSLF1, BBLF2/3, ATPase, stœchiométrie.**

## **Summary:**

Epstein-Barr Virus (EBV) is a human herpesvirus largely present worldwide. When primary infection occurs during adolescence or adulthood, it could cause infectious mononucleosis (IM). This disease is most of the time minor. EBV is also associated with several cancers like lymphomas (Burkitt's lymphoma and Hodgkin's lymphoma) or carcinomas (gastric carcinoma and nasopharyngeal carcinoma). Latent proteins of the virus are largely studied and are important for apparition of tumors. Recent studies show that lytic proteins are also important for tumor apparition and progression. The Helicase-Primase complex (H-P) of herpes simplex 1 (HSV-1), a well-known herpèsvirus, is a target for new antiviral drugs. ATPase, helicase and primase activities of HSV-1 complex are well studied, but no information is available for the structure of the H-P complex of human herpesvirus. We studied the H-P complex of EBV (BBLF4: helicase, BSLF1: primase and BBLF2/3: accessory subunit) to characterize its structure and enzymatic activities. We describe the expression and purification conditions and begin preliminary studies on structure and activities of the H-P complex. We show a basal ATPase activity that is DNA single strand independent. We were able to purify two forms of the H-P complex, the first has a stoichiometry close to 1/1/1 and the second one has an excess of helicase protein (BBLF4). These preliminary results on H-P complex of EBV have to be validated with other experiments before being compared to information already known for the HSV-1 complex.

**Key words : EBV, Helicase-Primase complex, BBLF4, BSLF1, BBLF2/3, ATPase, stoichiometry.**