

# Remerciements

Le travail présenté ici a été réalisé en grande partie à l'Institut Universitaire Européen de la Mer à Plouzané avec le soutien financier du Ministère de l'Enseignement Supérieur et de la Recherche.

Je tiens tout d'abord à remercier Daniel Prieur, mon Directeur de Thèse, de m'avoir donné l'opportunité de faire une thèse sur un sujet passionnant, de m'avoir fait confiance tout au long de ces années et de m'avoir toujours encouragée dans mes travaux, notamment durant la correction de ce manuscrit. Je lui suis également très reconnaissante de m'avoir permis de participer à une campagne océanographique dans le Pacifique sud au cours de laquelle j'ai eu le privilège d'observer des cheminées hydrothermales actives derrière le hublot du Nautile à - 2 583 m de profondeur. Je n'oublie pas non plus la formidable expérience de ma participation au congrès international « Thermophiles 2005 » qui s'est tenu en Australie.

Je voudrais également remercier Marc Le Romancer, mon Responsable Scientifique, pour la dynamique qu'il a insufflé à cette thèse en suggérant que certaines manipulations soient réalisées dans les laboratoires spécialistes des domaines concernés. Je lui suis aussi très reconnaissante de la qualité des corrections qu'il a apportées à ce manuscrit.

Je remercie sincèrement Marc Nadal et Christa Schleper d'avoir accepté d'être rapporteurs. Je suis très touchée qu'ils aient accepté de juger ce travail.

Je remercie également Mickael DuBow et Jean-Michel Rossignol pour avoir consenti à apporter leur vision de virologue dans l'évaluation de ce travail.

La liste est longue mais il me tient particulièrement à cœur de remercier toutes les personnes qui m'ont consacré une partie de leur temps pour me former à différentes techniques :

- Jean-Louis Birrien qui a rendu possible les expériences menées sous pression hydrostatique. Merci Jean-Lou pour ta bonne humeur et ton paternalisme si touchant.
- Thierry Bouvier qui m'a permis de réaliser des essais d'ultrafiltration tangentielle au sein de son laboratoire de Montpellier.
- Gaël Erauso qui, de part son expérience sur les plasmides, a pu me guider judicieusement dans les différentes étapes de l'étude moléculaire du virus PAV1.
- Didier Flament qui a fait preuve de patience et d'un optimisme sans faille au cours de l'étude de la protéine 899 de PAV1.
- Claire Geslin (mon ancienne maîtresse) qui m'a appris tout ce qu'elle sait sur la culture des Thermococcales et la « chasse » aux virus hydrothermaux. Merci Clear d'avoir toujours été là !

- Ghislaine Henneke qui m'a initiée aux tests d'incorporation de nucléotides marqués.

- Tristan Montier qui s'est tout de suite impliqué avec enthousiasme dans le projet d'hybridation *in situ* actuellement en cours.

- Marc Nadal qui m'a enseigné les subtilités du surenroulement de l'ADN et surtout qui m'a accueillie à bras ouvert pendant plusieurs dans son laboratoire de Versailles. Merci Marc pour ton soutien si réconfortant.

- Sophie Quevillon-Cheruel qui m'a donné des conseils avisés concernant les essais de solubilisation de la protéine 899.

- Gérard Siquin qui a en charge la gestion du microscope électronique en transmission de l'Université de Bretagne Occidentale et qui a bien voulu s'investir dans le projet d'hybridation *in situ* en cours de réalisation.

- Jean Sourimant qui a toujours rendu disponible le microscope électronique en transmission de la Station Biologique de Roscoff dès que j'en avais besoin.

- Jean-Marc Verbavatz qui a réalisé des essais de cryofracture sur les particules de type virale portées par la souche KAZA et qui m'a accueillie très gentiment pendant quelques jours dans son laboratoire du CEA de Saclay.

Je voulais aussi remercier Nadège Bienvenu pour l'aide si précieuse qu'elle m'a apportée pendant cette thèse. Merci Nad pour ton enthousiasme à faire toujours de nouvelles choses qui n'aboutissent pas nécessairement et pour ton amitié qui est un soutien moral de tous les jours.

Je remercie aussi Astrid Hoste pour avoir bravé les odeurs de *Pyrococcus abyssi* et réalisé certaines expériences complémentaires concernant le surenroulement de l'ADN de PAV1. Je voulais également remercier Astrid, ainsi que Florence Garnier et Sharon Larmony, pour ces merveilleuses semaines que j'ai passées dans leur laboratoire.

Un énorme merci à Evelyne Marguet, Nicolas Soler et Patrick Forterre pour m'avoir permis de réaliser des cultures d'hyperthermophiles en anaérobiose dans leur laboratoire, pour m'avoir hébergée plusieurs fois lors de mes séjours parisiens et pour leur accueil toujours si chaleureux.

J'adresse un grand merci aux deux stagiaires que j'ai eu l'honneur d'encadrer : Krsitin Krieger et Brice Mvou Lekogo. Merci aussi à Aziz Lies pour le test en épifluorescence.

Je tiens aussi à remercier les deux experts virologues qui ont accepté de me faire part de leur point de vue concernant la VLP-TBPP : le Pr Chastel (Université de Bretagne Occidentale - Brest) et le Pr Ackermann (Université de Laval - Québec).

Il me tient à cœur d'adresser ma gratitude aux organismes prestataires de service qui m'ont été d'une grande aide au cours de mes travaux. Il s'agit d'une part de

la Souchotèque de Bretagne avec ses deux égéries, Nadège Bienvenu et Valérie Cueff-Gauchard, et d'autre part, du Plateau d'Analyse Protéomique par Séquençage et Spectrométrie de Masse d'INRA de Jouy-en-Josas avec Céline Henry qui a toujours été d'une grande disponibilité.

Je profite aussi de ces pages pour adresser mes plus vifs remerciements aux membres du LM2E qui n'ont pas déjà été cités : Anne, Adeline, Audrey, Caroline, Charlotte, Eloïse, Erwan, Floriane, Françoise, Hélène, Joël, Joëlle, Josiane, JP, Julien, Lolo, Lucile, Marianne, Mathieu, Patricia, Raja, Totof, Sophie, Steph, et Zyao. Des remerciements tout spéciaux à Krikri et Nat pour leur amitié, leur générosité et leur bonne humeur.

Au cours de ces cinq ans passés à Brest j'ai rencontré des gens formidables qui chacun à leur manière ont rendu mon quotidien plus joyeux. Un grand merci donc : à Anne-So (rencontrée à l'autre bout du monde), Denise (ma troisième grand-mère), Elise (la plus grande fan de Daby Touré), Eric (le Gérard d'Aboville de Plouzané) Gregouille (la fripouille), Guillaume (le plus grand fan de Scary Movie), Gwen (ma poissonnière préférée), Lol et Cédric (les mélomanes nantais), Julien (la terreur de Gandalf), Justine et Luc (vive la Finlande à l'Eurovision), Mich (le plus charmant des draculas), Mickael (le plus américain d'entre nous), Olive (I need you), Sab (ma Thérèse préférée) Sandrine (la plus fervente admiratrice de Gandalf) et Seb (l'inventeur du « Je ne sais pas, il faut le voir porté »). Merci aussi à : Adeline, Anne-So, Annick, Annick, Antoine, Aurélie, Elisabeth, Emilie, Eric, Estérine, Gene, Gwen, Isa, Joëlle, Marc, Mado, Rose-Marie, Sonia et Xav.

Ces remerciements touchent à leur fin mais je voudrais avant cela remercier les amis non brestois qui m'ont soutenu dans cette aventure : Amandine, Cédric, Christophe, Cissou, Didier et Martine, Floflo, Fred, Guillemette, Guy et Yvette, Hélène et Nico, Ingrid, Lolo, Lucie, Marie, M Klein, la famille Oden, Pierrot, Sandrine D, Sandrine M, Sandra et Julien, Sil et Song.

Un énorme merci à ma Célinette qui est toujours là pour moi depuis bientôt dix ans et à qui je dois énormément !

Je voulais aussi sincèrement remercier toute ma famille pour le soutien sans faille qu'elle m'a toujours témoigné au cours de cette thèse et plus particulièrement ma douce Mamoun, mon adorable Baboun, mon irremplaçable Tamalytye et mon Brother préféré.

Enfin, merci à toi Jean-Marc pour ton soutien, ton infinie patience et pour toutes les choses merveilleuses que tu m'apportes.



# THESE

Présentée pour l'obtention du grade de

**Docteur de l'Université de Bretagne Occidentale**

Ecole Doctorale des Sciences de la Mer et du Littoral

(Mention Microbiologie Marine)

**MELUSINE GAILLARD**

**Etude moléculaire de PAV1,**

**le premier virus hyperthermophile marin.**

**Recherche et caractérisation préliminaire de particules de type virale  
chez les archées de l'ordre de *Thermococcales* inféodées aux sources  
hydrothermales marines profondes.**

Soutenue le 28 juin 2006 devant le jury composé de :

Pr. Mickael DuBow	Examineur
Dr. Marc Le Romancer	Examineur
Pr. Marc Nadal	Rapporteur
Pr. Daniel Prieur	Directeur de thèse
Pr. Jean-Michel Rossignol	Examineur
Pr. Christa Schleper	Rapporteur



# Sommaire

<b>Introduction bibliographique : « Les virus de procaryotes »</b>	<b>1</b>
<b>I. Généralités sur les virus de procaryotes</b>	<b>1</b>
1. Découverte	1
2. Terminologie	1
3. Classification	1
a. Classification générale des virus	1
b. Classification des virus de procaryotes	3
4. Cellules hôtes	5
5. Relation à l'hôte et production virale	6
a. Les différents types de relation hôte – virus	6
b. L'influence de la pression hydrostatique	7
6. Génomique, origine et évolution	7
a. La structure génétique des populations virales	7
b. La notion de génomes en mosaïque	8
c. Les théories de l'origine et de l'évolution	9
<b>II. Les techniques de détection et de quantification des virus de procaryotes</b>	<b>11</b>
1. La technique des plages de lyse	11
2. La microscopie électronique en transmission (MET)	12
3. La microscopie à épifluorescence (ME)	13
4. La cytométrie en flux	14
5. Les techniques d'amplification de gène	15
a. La Ppolymerase Chain Reaction (PCR) et la Reverse Transcription- Polymerase Chain Reaction (RT-PCR)	15
b. La PCR et la RT-PCR quantitatives	15
6. La métagénomique	16
<b>III. Les virus de procaryotes dans les environnements extrêmes</b>	<b>17</b>
1. Les virus des milieux hypersalins	18
a. La diversité globale	18
b. Les virus modèles	18
2. Les virus des lacs alcalins	24
3. Les virus du désert	24
4. Les virus des glaces polaires	25
5. Les virus des sédiments de subsurface	25
6. Les virus des sources chaudes terrestres acides	27
a. La diversité globale	27
b. Les virus modèles	27
7. Les virus des sources hydrothermales marines	39
a. La diversité globale	39
b. PAV1, le virus modèle	39
<b>IV. Pourquoi étudier les virus de procaryotes ?</b>	<b>40</b>
1. Pour un enjeu en biodiversité	40
2. Pour un enjeu en écologie marine	41
a. Au niveau de la mortalité microbienne	41
b. Au niveau des cycles biogéochimiques	42
c. Au niveau des flux de gènes	43

3. Pour un enjeu en biotechnologie moléculaire	44
a. En utilisant leur cycle lysogène	44
b. En utilisant leur acide nucléique	44
c. En utilisant leurs enzymes	45
4. Pour un enjeu en thérapeutique	46
a. En tant qu'« antibiotiques »	46
b. Pour la synthèse d'anticorps spécifiques	46
c. En tant que vecteurs pour la thérapie génique	47

## **Introduction au travail de thèse** **49**

## **Matériel et Méthode** **53**

### **I. Echantillons bruts** **53**

### **II. Culture des *Thermococcales*** **53**

1. Milieu utilisé	53
1. Culture en milieu liquide à pression atmosphérique	54
2. Culture en milieu liquide sous pression hydrostatique	54
3. Isolement sur milieu solide à pression atmosphérique	54
a. Réalisation de la « sous-couche »	55
b. Réalisation de la « sur-couche »	55
c. Incubation et repiquage des colonies	55

### **III. Recherche de nouveaux virus dans des cultures de *Thermococcales*** **56**

1. La microscopie électronique en transmission (MET)	56
a. Cultures d'enrichissement et isolats observés	56
b. Protocole	57
2. La Polymerase Chain Reaction (PCR)	57
a. Isolats analysés	57
b. Protocole	58

### **IV. Description d'un système hôte - particules de type viral chez les *Thermococcales*** **59**

1. Détermination de l'identité de la souche KAZA	59
2. Détermination de la croissance de la souche KAZA et de la production des deux VLP	59
3. Détermination des conditions optimales de pression hydrostatique pour la production de la VLP-TBPP	60
4. Recherche de la VLP-TBPP dans différentes souches de <i>Thermococcales</i>	60
5. Recherche d'un acide nucléique chez les deux particules de type viral	60
a. Concentration des particules	60
b. Essai d'extraction d'ADN	61
c. Essai d'extraction d'ARN	61
d. Technique du « cracking »	62
e. Test en épifluorescence	62
6. Purification des deux particules de type viral	63
a. Ultracentrifugation et gradient de densité	63
b. Précipitation et gradient de densité	63
c. Ultrafiltration tangentielle	63
7. Etude préliminaire de la structure de la VLP-TBPP	64
a. Tests de sensibilité	64
b. Recherche des protéines majoritaires	64

<b>V. Etude du génome de PAV1</b>	<b>64</b>
1. Détermination du nombre de copies du génome viral dans la cellule	64
2. Détermination du surenroulement de l'ADN viral	66
3. Recherche de méthylations sur le génome viral et le chromosome archéen	66
4. Recherche d'un intermédiaire de réplication simple brin	67
5. Cartographie des transcrits du virus	67
<b>VI. Etude des protéines de PAV1</b>	<b>69</b>
1. Electrophorèse de protéines en conditions dénaturantes (SDS-PAGE) et coloration au bleu de Coomassie	69
2. Surproduction de la protéine 899 de PAV1	69
a. Couples d'amorces et vecteurs d'expression utilisés	69
b. Clonage de l' <i>orf</i> 899 en vecteur d'expression	70
c. Surexpression de l' <i>orf</i> 899	71
d. Séparation des fractions contenant les protéines solubles et insolubles	72
3. Essais d'amélioration de la production de la protéine 899 sous sa forme soluble	72
a. Abaissement de la température de culture	72
b. Co-production de la protéine 899 avec des chaperonnes	72
c. Production <i>in vitro</i> de la protéine 899	73
4. Extraction et solubilisation des corps d'inclusion	73
5. Renaturation de la protéine 899	74
a. Recherche d'un tampon de renaturation	74
b. Dialyse de la protéine 899	74
6. Essais de mise en évidence d'une implication de la protéine 899 dans le métabolisme de l'ADN viral	75
a. Recherche d'une activité ADN polymérase chez la protéine 899	75
b. Recherche d'interactants pour la protéine 899	75
7. Recherche de la protéine 899 dans son hôte naturel, la souche GE23 de <i>P. abyssi</i>	76
8. Identification d'une protéine majoritaire de PAV1	76

## **Résultats et Discussion** **77**

### **I. Préambule - Données apportées par l'analyse de la séquence du virus PAV1** **77**

### **II. Etude moléculaire du virus PAV1** **78**

1. Etude du génome de PAV1	79
a. Détermination du nombre de copies du génome viral	79
b. Etude du surenroulement de l'ADN viral dans la cellule	81
c. Recherche de méthylations sur le génome viral et le chromosome de l'hôte	85
d. Etude du mode de réplication	87
e. Etude des transcrits	88
3. Etude de deux protéines de PAV1	91
a. Etude de la protéine 899	91
➤ Pourquoi étudier la protéine 899 ?	91
➤ Surexpression de l' <i>orf</i> 899	91
➤ Essais de production de la protéine 899 sous sa forme soluble	93
➤ Extraction et solubilisation des corps d'inclusion	96
➤ Renaturation de la protéine 899	96
➤ Essais de mise en évidence d'une implication de la protéine 899 dans le métabolisme de l'ADN viral	98
➤ Recherche de la protéine 899 native dans son hôte naturel	100
b. Etude d'une protéine majoritaire du virus	101

<b>III Recherche de nouveaux virus chez les <i>Thermococcales</i> des sources hydrothermales océaniques profondes</b>	<b>102</b>
1. Préambule	102
2. Recherche de VLP dans des cultures d'enrichissement de <i>Thermococcales</i>	103
3. Recherche de particules de type viral dans des isolats de <i>Thermococcales</i>	108
a. Par microscopie électronique en transmission	108
b. Par Polymerase Chain Reaction	109
<b>IV. Caractérisation préliminaire de deux particules de type viral portées par une souche de <i>Thermococcales</i></b>	<b>110</b>
1. Etude de la souche hôte KAZA	110
2. Description morphologique de la particule de type viral filamenteuse	111
3. Etude de l'influence de la pression hydrostatique sur la VLP-TBFP	112
4. Etude de la production de la VLP-TBFP et de la VLP-TBPP	114
5. Recherche d'un acide nucléique chez la VLP-TBFP et la VLP-TBPP	118
6. Purification de la VLP-TBFP et de la VLP-TBPP	119
7. Etude préliminaire de la structure de la VLP-TBPP	120
<b>Conclusion générale et Perspectives</b>	<b>122</b>
<b>Références Bibliographiques</b>	<b>126</b>

# Liste des Figures et Tableaux

## Liste des Figures :

- Figure 1** : Morphologie des différentes familles de virus de procaryotes. p 4
- Figure 2** : Distribution des virus dans les domaines bactérien et archéen. p 5
- Figure 3** : Schéma expliquant la théorie de l'évolution des virus et des cellules défendue par P. Forterre. p 10
- Figure 4** : Schéma du protocole utilisé dans la technique des plages de lyse. p 11
- Figure 5** : Photographies de bactériophages après coloration négative et observés en MET prises dans les années 1940 et 1970. p 12
- Figure 6** : Photographie en épifluorescence de procaryotes (gros points) et de leurs virus (petits points) marqués au SYBR Green I. p 14
- Figure 7** : Photographie en MET de différents virus isolés de milieux hypersalins représentant les principales morphologies rencontrées. p 19
- Figure 8** : Photographie en MET des particules de type viral observées dans les échantillons de sable du désert du Sahara et étant assimilées aux familles des *Myoviridae*, des *Podoviridae* et des *Siphoviridae*. p 26
- Figure 9** : Photographie au MET des isolats 1a, 11b et 21c isolés d'échantillons de glace prélevés en Arctique au large de la côte est du Groenland. p 26
- Figure 10** : Photographie en MET de différents virus isolés de sources chaudes acides terrestres et représentant les principales morphologies rencontrées. p 28
- Figure 11** : Photographie en MET de différentes particules de type viral observées au niveau de sources chaudes acides terrestres et possédant des morphologies originales par rapport aux virus décrits. p 28
- Figure 12** : Comparaison du génome de SSV1 avec trois autres *Fusellovirus* : SSV2, SSVRH et SSVK1. p 31
- Figure 13** : Carte des transcrits des virus SIRV1 et SIRV2. p 35
- Figure 14** : Photographies en MET de PAV1 : après purification, attaché à un débris cellulaire, allongé à l'extrême et après traitement au chloroforme. p 40
- Figure 15** : Place des Thermococcales dans l'arbre phylogénétique du vivant non enraciné et basé sur la comparaison des séquences des ADNr 16S et 18S. p 50
- Figure 16** : Origine géographique des échantillons bruts utilisés. p 53
- Figure 17** : Origine géographique des différentes cultures d'enrichissement et isolats observés en MET. p 56
- Figure 18** : Origine géographique des différents isolats analysés par PCR. p 58
- Figure 19** : Schéma du génome de PAV1 récapitulant les principales informations livrées par l'annotation. p 78
- Figure 20** : Signal de l'hybridation des sondes « PAV1 » et « 16S GE23 » sur l'ADN total de la souche GE23. Profil généré et intensités calculées par le logiciel NIH Image à partir des signaux d'hybridation. p 79
- Figure 21** : Mode d'action de la topoisomérase I eucaryote, de la topoisomérase A d'*Escherichia coli* et de la reverse gyrase de *Sulfolobus*. p 82
- Figure 22** : Profils électrophorétiques des ADN de PAV1 et de pTZ18 : en natif, après action de la topoisomérase I eucaryote, de la topoisomérase A d'*Escherichia coli* et de la reverse gyrase de *Sulfolobus*. p 82
- Figure 23** : Schéma explicatif de la migration des topoisomères dans un gel en deux dimensions en présence de BET dans la deuxième dimension. p 82
- Figure 24** : Profils électrophorétiques de l'ADN de PAV1 après migration en deux dimensions : à l'état natif, après incubation avec la topoisomérase I eucaryote et après incubation avec la reverse gyrase de *Sulfolobus*. p 84

<b>Figure 25</b> : Profils électrophorétiques l'ADN de PAV1 et de l'ADN total de la souche GE23 natifs et digérés par <i>Bst</i> NI, <i>Dpn</i> I, <i>Mbo</i> I et <i>Scr</i> FI.	p 86
<b>Figure 26</b> : Signal donné par l'hybridation des sondes spécifiques du brin + et du brin – sur l'ADN de PAV1.	p 88
<b>Figure 27</b> : Cartographie des transcrits de PAV1 avec les promoteurs et les terminateurs de la transcription déterminés par l'analyse de séquence.	p 89
<b>Figure 28</b> : Profils électrophorétiques des protéines contenues dans les fractions solubles et insolubles obtenues 2 h après induction à l'IPTG des cellules HMS 174 contenant: le vecteur pET 11d non recombinant ou le vecteur pET 11d contenant l' <i>orf</i> 899; le vecteur pET 28a non recombinant ou le vecteur pET 28a+ contenant l' <i>orf</i> 899.	p 92
<b>Figure 29</b> : Profils électrophorétiques des protéines contenues dans les fractions solubles et insolubles obtenues après induction à l'IPTG des cellules HMS 174 portant le vecteur pET 28a+ contenant l' <i>orf</i> 899 et cultivées à 37°C, 30°C, 26°C ou 18°C.	p 94
<b>Figure 30</b> : Profil électrophorétique des protéines obtenues après expression <i>in vitro</i> pratiquée sur le plasmide pEX1-GW/ <i>lacZ</i> contenant le gène de la $\beta$ -galactosidase, le vecteur pET 28a+ non recombinant ou le vecteur pET 28a+ contenant l' <i>orf</i> 899.	p 95
<b>Figure 31</b> : Profil électrophorétique des protéines obtenues après extraction puis après solubilisation des corps d'inclusion.	p 96
<b>Figure 32</b> : Densité optiques mesurées à 340 nm pour la protéine 899 mélangée à chacun des 80 tampons de renaturation aux temps t = 0 et t = 52 h.	p 97
<b>Figure 33</b> : Profils électrophorétiques des protéines contenues dans les fractions après dialyse contre les tampons 1H, 2H, 7H ou 9H et dans le surnageant après centrifugation des fractions dialysées	p 98
<b>Figure 34</b> : Profil d'interaction protéine 899 / PCNA obtenu avec le Biacore X.	p 99
<b>Figure 35</b> : Profil électrophorétique des protéines contenues dans : la fraction insoluble des cellules HMS 174 portant le vecteur pET 28a+ contenant l' <i>orf</i> 899 ainsi que dans la fraction soluble de la souche GE23 de <i>Pyrococcus abyssi</i> .	p 100
<b>Figure 36</b> : Profil électrophorétique des protéines issues des virions de PAV1 purifiés.	p 101
<b>Figure 37</b> : Alignement des séquences de la protéine prédite et de la protéine séquencée expérimentalement.	p 101
<b>Figure 38</b> : Photographies des VLPS ne se maintenant pas dans la culture d'enrichissement après plusieurs repiquages observées en MET après coloration négative.	p 104
<b>Figure 39</b> : Photographies des VLPS se maintenant dans la culture d'enrichissement après plusieurs repiquages observées en MET après coloration négative.	p 105
<b>Figure 40</b> : Photographie en microscopie optique à contraste de phase de la souche KAZA à pression atmosphérique et à 35 MPa.	p 110
<b>Figure 41</b> : Cinétiques de croissance de la souche KAZA à pression atmosphérique et à 350 MPa.	p 111
<b>Figure 42</b> : Photographies de la VLP-TBFP entière et de fragments observés en MET après coloration négative.	p 112
<b>Figure 43</b> : Photographies des morphologies de la VLP-TBPP : le type « tête-queue », de type haltère, de type à extensions bipolaires et de type bourgeonnant observées en MET après coloration négative.	p 113
<b>Figure 44</b> : Cinétiques de croissance de la souche KAZA et de production de la VLP-TBFP avec traitement à la mitomycine C à pression atmosphérique.	p 114
<b>Figure 45</b> : Cinétiques de la souche KAZA et de production de la VLP-TBPP avec traitement à la mitomycine C à 35 MPa.	p 115
<b>Figure 46</b> : Production des la VLP-TBPP à différentes pressions hydrostatiques avec et sans traitement à la mitomycine C.	p 116

- Figure 47 :** Production de la VLP-TBPP par différentes souches de Thermococcales cultivées à 15 MPa avec et sans traitement à la mitomycine C. p 117
- Figure 48 :** Photographies après marquage au SYBR Green I du témoin négatif (eau milliQ), du témoin positif (PAV1), de la VLP-TBFP et de la VLP-TBPP observés en microscopie à épifluorescence. p 118
- Figure 49 :** Photographies de la VLP-TBPP de type haltère et de type « tête-queue » traitée à la protéinase K et observées en MET après coloration négative. p 120
- Figure 50 :** Profil électrophorétique des protéines contenues dans la fraction obtenue après purification de la VLP-TBPP. p 121

### Liste des Tableaux :

- Tableau 1 :** Principales caractéristiques de l'écologie virale des différents environnements extrêmes. p 17
- Tableau 2 :** Principales caractéristiques des virus décrits isolés de milieux hypersalins. p 19
- Tableau 3 :** Principales caractéristiques des virus décrits isolés de sources chaudes acides terrestres. p 29
- Tableau 4 :** Séquences des différents couples d'amorces dessinés sur le génome de PAV1. p 68
- Tableau 5 :** Détail des constructions sélectionnées. p 69
- Tableau 6 :** Composition des tampons de renaturation testés p 74
- Tableau 7 :** Mesures de la radioactivité (en coups par minute) incorporée après incubation de la matrice avec la protéine 899 dialysée contre les tampons 1H, 2H, 7H et 9H. p 98
- Tableau 8 :** Caractéristiques des VLPs ne se maintenant pas dans la culture d'enrichissement après plusieurs repiquages. p 104
- Tableau 9 :** Caractéristiques des VLPs se maintenant dans la culture d'enrichissement après plusieurs repiquages. p 105
- Tableau 10 :** Protéines assignées aux différentes bandes protéiques obtenues après électrophorèse en SDS-PAGE de la VLP-TBPP purifiée. p 121



# Liste des Abréviations

**ADN** acide désoxyribonucléique  
**ADNc** acide désoxyribonucléique complémentaire  
**ADNr 16S** ou **18S** acide désoxyribonucléique codant pour la petite sous-unité du ribosome  
**ARN** acide ribonucléique  
**ARNm** acide ribonucléique messenger  
**ATP** adénosine triphosphate  
**BET** bromure d'éthidium  
**CHES** acide 2(cyclohexylamino)ethanesulfonique  
**CsCl** chlorure de césium  
**Ci** curie  
**CIR** ride centro-Indienne  
**DIG** digoxygénine  
**DNase** désoxyribonucléase  
**dNTP** désoxyribonucléotide triphosphate  
**DTT** 1,4-dithithreitol  
**[<sup>3</sup>H]dTTP** désoxythymidine triphosphate tritiée  
**EDTA** éthylène diamine-tétraacétate  
**EPR** ride est-Pacifique  
**g, mg, µg** gramme, milligramme, microgramme  
**h** heure  
**ICTV** international committee on taxonomy of viruses  
**IgG** immunoglobuline G  
**IPTG** isopropyl-βd-thiogalactopyranoside  
**kDa** kilodalton  
**L, mL, µL** litre, millilitre, microlitre  
**LB** Luria-Bertani  
**m, cm, µm, nm** mètre, centimètre, micromètre, nanomètre  
**M, mM** molaire, millimolaire  
**Maldi-Tof** « matrix assisted laser desorption ionization - time of flight »  
**MAR** ride médio-Atlantique  
**MES** acide 2-morpholineoethanesulfonique  
**MET** microcopie électronique en transmission  
**min** minute  
**mmol** millimole  
**MPa** mégapascal  
**nt** nucléotide  
**orf** « open reading frame »  
**p** poids  
**PAGE** « polyacrylamide gel electrophoresis »  
**pb, kpb** paire de bases, kilopaire de bases  
**PCNA** « poliferating cell nuclear antigen »  
**PCR** polymerase chain reaction  
**PEG** polyéthylenglycol  
**PIPES** piperazine-N,N'-bis(2-ethanesulfonic acid)  
**RNase** ribonucléase  
**RT-PCR** reverse transcription-polymerase chain reaction  
**s** seconde  
**SDS** sulphate dodécylsulfate  
**TAE** tris-HCl/acetate/EDTA  
**TE** tris-HCl/EDTA  
**U** unité  
**v** volume  
**V** volt  
**VLP** "virus-like particle"  
**W** watt



# Introduction Bibliographique

## « Les Virus de Procaryotes »

### I. Généralités sur les virus de procaryotes

#### 1. Découverte

L'existence des virus de procaryotes fut révélée au début du XXème siècle par deux scientifiques. L'anglais F. W. Twort décrivit le premier, en 1915, une lyse chez les microcoques (Twort *et al.*, 1915). Deux ans plus tard, le canadien F. d'Hérelle publia le même phénomène chez les shigelles (D'Hérelle, 1917).

Les études, menées depuis maintenant près d'un siècle, ont permis de mettre en lumière l'extraordinaire diversité morphologique et génomique de ces virus. Ils forment actuellement le groupe viral le plus important et constituent également le groupe d'entités biologiques le plus abondant de la planète (Ackermann, 2001 ; Rohwer, 2003). Ils sont présents dans des environnements extrêmement variés, en relation étroite avec l'écologie de leurs hôtes (sols, océans, sources hydrothermales, sédiments marins, tractus digestifs de différents animaux, etc). Ainsi, à l'heure actuelle, plus de 5 000 virus de procaryotes ont été visualisés au microscope électronique et environ 150 sont décrits chaque année (Ackermann, 2001).

#### 2. Terminologie

Les virus infectant les procaryotes sont répartis en deux catégories suivant la nature de leur hôte (bactérie ou archéobactérie). En raison du fait qu'ils ont souvent un pouvoir lytique, les termes bactériophage et phage ont été initialement utilisés pour nommer les virus de bactéries. En ce qui concerne les archées, le mot archéophage existe mais il est peu usité. En effet, les scientifiques spécialistes du domaine parlent plutôt de virus d'archéobactéries pour souligner les différences existant vis-à-vis des bactériophages.

#### 3. Classification

##### a. Classification générale des virus

Le premier système de classement des virus date des années 1920. Il reposait sur les symptômes occasionnés par l'agent pathogène chez son hôte (animal, homme ou plante). Puis, avec l'apparition du microscope électronique, c'est la forme et la taille du virion qui devinrent les critères majeurs de classification trente ans plus tard (Fauquet, 2002).

En 1962, A. Lwoff proposa un système répartissant les virus en sous-phyla, classes, ordres, sous-ordres et familles. La hiérarchisation se fit alors suivant des critères bien particuliers, à savoir la nature de l'acide nucléique (ADN ou ARN) et sa structure (monocaténaire, bicaténaire, linéaire ou circulaire), la symétrie de la capsid (binaire, cubique ou hélicoïdale), la présence ou l'absence d'enveloppe et le mode de réplication (Lwoff *et al.*, 1962).

En 1971, D. Baltimore proposa une autre solution qui consistait à classer les virus en fonction de la stratégie de réplication, directement liée à la structure du génome. Ainsi, six groupes viraux furent définis : le groupe I (virus à ADN double brin), le groupe II (virus à ADN simple brin positif), le groupe III (les virus à ARN double brin), le groupe IV (les virus à ARN simple brin positif), le groupe V (les virus à ARN simple brin négatif), le groupe VI (les Rétrovirus) (Baltimore, 1971). Un septième groupe correspondant aux *Reversiviruses* fut ajouté ensuite.

La classification actuelle des virus repose sur un mélange des classifications de A. Lwoff et D. Baltimore (Büchen-Osmond, 2003a). Elle est régie par le Comité International de Taxonomie des Virus (ICTV) créé en 1966 (<http://www.ncbi.nlm.gov/ICTVdb/>). Celui-ci émet régulièrement des rapports faisant état de la classification virale réactualisée (8<sup>th</sup> ICTV Report, 2005). Il a également élaboré une base de données universelle référençant tous les virus décrits (Büchen-Osmond, 2003b).

La nomenclature utilisée en taxonomie virale est très codifiée : les noms d'ordres se terminent par *-virales*, ceux de familles par *-viridae* et ceux de genres par *-virus*. Dans le monde viral, la notion d'espèce a souvent été soumise à la controverse. Depuis 1991, l'ICTV a retenu comme définition celle de M. H. van Regenmortel : « L'espèce virale est une classe de virus polythétiques qui constituent une lignée répliquative et qui occupent une niche écologique particulière. » (Van Regenmortel, 1999 ; Büchen-Osmond, 2003a).

A l'heure où le nombre de génomes viraux disponibles ne cesse d'augmenter, il apparaît nécessaire de les prendre en compte pour élaborer une nouvelle taxonomie. Ainsi, des virus ayant le même type d'acide nucléique et la même morphologie sont souvent génétiquement très éloignés (Nelson, 2004). C'est pourquoi, J. G. Lawrence et ses collaborateurs proposent un système de classification dit « réticulé » où la séquence génomique serait prise en compte. Les groupes seraient élaborés en fonction de similarités résultant d'un ou plusieurs mécanismes dits « de cohésion » comme la recombinaison. Ainsi, un même virus pourrait appartenir à des groupes différents du fait de la multiplicité des origines de son génome (Lawrence *et al.*, 2002) Un autre système de classifi-

cation, basé sur des comparaisons du protéome viral, a également été proposé par F. Rohwer et R. Edwards (Rohwer & Edwards, 2002).

### **b. Classification des virus de procaryotes**

Les virus de procaryotes sont répartis en 14 familles, auxquelles il faut ajouter un nombre restreint de nouvelles propositions en cours d'homologation (<http://www.ncbi.nlm.gov/ICTVdb/>). Sur le seul critère morphologique, quatre grands groupes peuvent être définis (cf Figure 1) :

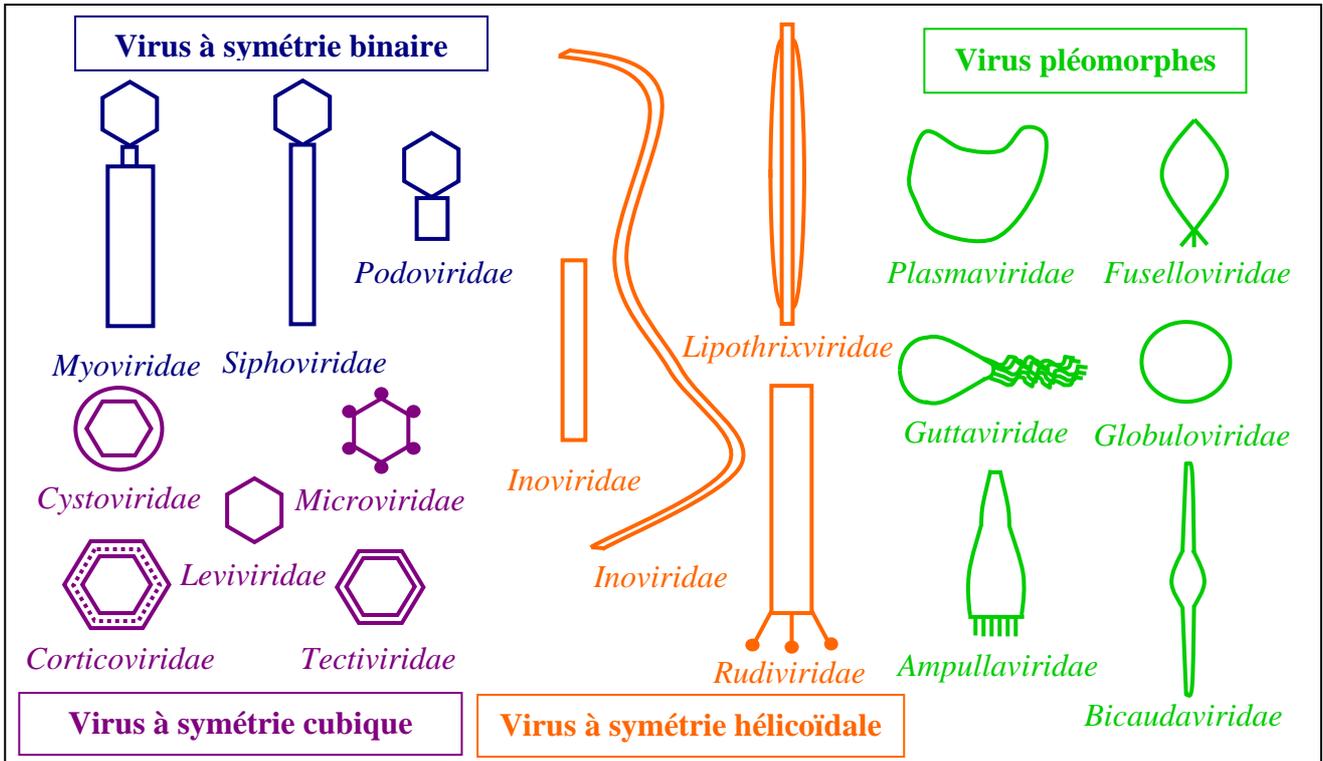
- les virus à symétrie binaire (ou virus caudés). Ils ont comme caractéristique principale de posséder une queue permanente, contractile ou non suivant les cas. Celle-ci possède généralement des organelles de fixation sous forme de fibres ou de crampons terminaux. Leur capsid (ou tête) est à symétrie icosaédrique. Ces virus ne sont pas enveloppés et ont un génome à ADN bicaténaire linéaire. Cette catégorie concentre 96 % des virus de procaryotes, répartis en trois grandes familles : les *Myoviridae*, les *Siphoviridae* et les *Podoviridae*. Ces dernières constituent l'ordre des *Caudovirales*, le seul ordre défini chez les virus de procaryotes (Ackermann, 2001).

- les virus à symétrie cubique. Ils ont une capsid icosaédrique mais pas de queue. Certains possèdent une structure lipidique qui est le plus souvent sous forme d'enveloppe mais qui peut aussi être insérée dans le core. Ce groupe qui représente seulement 2 % des virus de procaryotes, correspond cependant à cinq familles : les *Microviridae* (ADN monocaténaire circulaire), les *Corticoviridae* (ADN bicaténaire circulaire), les *Tectiviridae* (ADN bicaténaire linéaire), les *Leviviridae* (ARN monocaténaire linéaire) et les *Cystoviridae* (ARN bicaténaire linéaire) (Ackermann, 2001).

- les virus à symétrie hélicoïdale. Ils ont la forme de filaments de longueur variable qui sont soit flexueux soit rigides. Ils ne représentent que 1,7 % du nombre total de virus de procaryotes et regroupent trois familles : les *Inoviridae*, les *Lipothrixviridae* et les *Rudiviridae*. Les deux dernières regroupent uniquement des virus infectant les archéobactéries. Seules les *Lipothrixviridae* sont enveloppés. Les *Inoviridae* ont un génome à ADN monocaténaire circulaire alors que les *Lipothrixviridae* ainsi que les *Rudiviridae* ont un génome à ADN bicaténaire linéaire (Ackermann, 2001).

- les virus dits pléomorphes. Ils ne possèdent pas de capsid véritable mais une structure de type enveloppe. Cela ne concerne que 0,3 % des virus de procaryotes mais six familles sont concernées : les *Plasmaviridae*, les *Fuselloviridae*, les *Guttaviridae*, les *Globuloviridae*, les *Ampullaviridae* et les *Bicaudaviridae*. Ces cinq dernières regroupent des virus qui infectent uniquement les archéobactéries. Les *Fuselloviridae* et les *Bicaudaviridae* ont une forme de fuseau alors que les *Guttaviridae* ont une morphologie de goutte d'eau avec une touffe de fins filaments au pôle le plus poin-

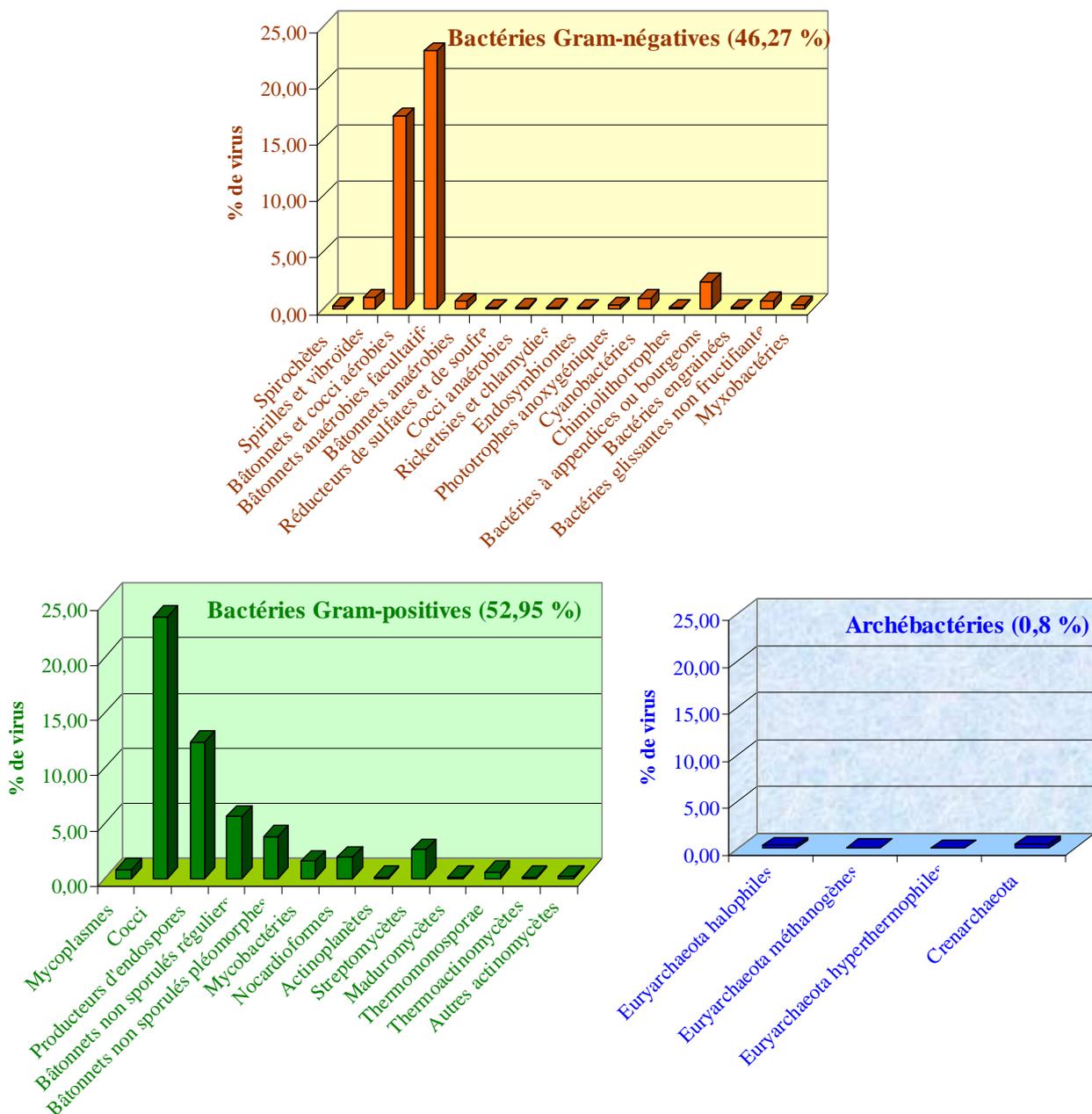
tu. Les *Globuloviridae* sont sphériques et les *Ampullaviridae* ont une forme de bouteille (Ackermann, 2001 ; Arnold *et al.*, 2000b ; Häring *et al.*, 2004, 2005b, 2005c).



**Figure 1 :** Morphologie des différentes familles de virus de procaryotes (d'après Ackermann, 2001 et modifié).

#### 4. Cellules hôtes

La plupart des virus de procaryotes décrits sont des bactériophages et infectent des groupes bactériens très variés (Ackermann, 2001) (cf Figure 2). L'étude des virus d'archéobactéries étant récente, le nombre de spécimens décrits est encore faible (environ une quarantaine) (cf Figure 2). Cependant, des virus ont pu être isolés dans les deux phyla archéens les plus importants, à savoir les *Euryarchaeota* et les *Crenarchaeota* (Dyall-Smith *et al.*, 2003 ; Porter *et al.*, 2005 ; Geslin *et al.*, 2003 ; Prangishvili & Garrett, 2005 ; Ackermann, 2001).



**Figure 2 :** Distribution des virus dans les domaines bactérien et archéen (d'après Ackermann, 2001).

## 5. Relation à l'hôte et production virale

### a. Les différents types de relation hôte - virus

L'infection virale aboutit généralement à la lyse de la cellule infectée et à la libération de nouveaux virions. Dans ce cas, le virus est dit lytique (Ackermann, 2001 ; Maniloff, 2001).

Toutefois, de nombreux virus établissent avec leur hôte un état de lysogénie qui est caractérisé par la persistance du génome viral dans la cellule. Le plus souvent, il est sous forme intégrée dans le génome de l'hôte et appelé provirus. Plus rarement, il est sous forme libre et peut être linéaire comme pour le phage  $\phi$ KO2, ou circulaire comme pour le phage P1 (Casjens *et al.*, 2004, Ackermann, 2001). Dans ces deux cas, le virus est dit tempéré et l'expression de la plupart de ses gènes est réprimée soit par le virus lui-même soit par l'hôte. Le phénomène d'intégration se rencontre chez les virus à symétrie binaire, les *Plasmaviridae*, les *Inoviridae* ainsi que chez les *Fuselloviridae* (Ackermann, 2001 ; Maniloff, 2001).

Les virus tempérés peuvent sortir de leur état de lysogénie sous l'influence de facteurs d'induction qui sont souvent environnementaux (radiations, carence nutritionnelle) et engager alors un cycle lytique. Cependant, il arrive parfois que les nouveaux virions soient excrétés à travers la membrane de la cellule ou libérés par bourgeonnement. Il s'agit alors d'infections permanentes sans mort immédiate de la cellule infectée (Ackermann, 2001 ; Maniloff, 2001).

Il existe un autre état, différent de la lysogénie, qui est appelée pseudolysogénie ou état porteur. Cet état est non inductible et il n'y a pas intégration du génome viral. L'hôte excrète de manière régulière des particules virales, tout en survivant et en se multipliant. Cet état correspond donc plutôt à une situation d'équilibre entre la multiplication cellulaire et la production virale. Dans le cas où l'équilibre est rompu, il y a curage de la cellule hôte ou alors production lytique de virions (Arnold *et al.*, 2000b).

Les mécanismes fins qui caractérisent l'entrée et la sortie du virus sont connus uniquement chez certains bactériophages très étudiés tels que le phage  $\lambda$  ou le phage T4 (Casjens & Hendrix, 2001 ; Matthews, 2001). A l'heure actuelle, aucun travail n'a été mené dans ce sens chez les virus d'archéobactéries.

### **b. L'influence de la pression hydrostatique**

Les études concernant l'influence de la pression hydrostatique sur les virus sont peu nombreuses et elles concernent majoritairement les virus de procaryotes. Elles ont été initiées à la fin des années 1940 par R. C. Foster et ses collaborateurs (Foster *et al.*, 1949). Le scientifique suédois, L. Rutberg, a montré qu'une exposition de la souche K d'*Escherichia coli* à la pression hydrostatique affecte l'immunité des cellules vis-à-vis du phage  $\lambda$  et induit ainsi la production virale (Rutberg, 1965).

La même année, une équipe américaine a rapporté l'inactivation du phage T4 après une exposition de quelques minutes à une pression de 0,4 MPa. Des observations au MET ont montré qu'une partie des virions avaient la queue contractée et avec parfois la tête vide d'ADN. Ces altérations morphologiques ressemblent à ce qui arrive lors du processus d'infection (Solomon *et al.*, 1965).

De nos jours, le traitement en pression hydrostatique est de plus en plus utilisé dans le domaine agro-alimentaire dans le but de préserver les aliments. En effet, un choc hyperbare permet l'inactivation des phages présents dans la matière première ou les produits finis. Les études menées à l'heure actuelle sur l'influence de la pression sur les virus concernent donc plutôt les conditions d'inactivation optimales. (Chen *et al.*, 2004).

Une équipe belge travaille également sur les conséquences de l'exposition aux fortes pressions. Ainsi, elle a mis en évidence l'induction d'une toxine produite par un provirus suite à un choc hyperbare (Aersten *et al.*, 2005).

## **6. Génomique, origine et évolution**

### **a. La structure génétique des populations virales**

Le premier génome viral de procaryotes séquencé a été celui du bactériophage  $\Phi$ X174 (Sanger *et al.*, 1978). A l'heure actuelle, plus de 220 génomes viraux sont disponibles dans les bases de données (<http://www.ncbi.nlm.nih.gov/>). Cependant, la vision actuelle de la structure génétique des virus de procaryotes demeure quelque peu biaisée. En effet, le nombre de génomes séquencés est très restreint comparé à la population totale d'individus qui a été estimée à environ  $10^{31}$  (Hendrix, 2003). Certains gènes présumés homologues ont tellement divergé qu'ils sont difficilement identifiables en tant que tels. Par exemple, les protéines majoritaires de capsid des entérophages P22 et HK97 n'ont aucune similarité de séquences. Cependant, une analyse structurale a révélé que les deux molécules possédaient le même repliement (Jiang *et al.*, 2003).

Il existe des preuves d'échanges génétiques au sein de populations virales telles que celle des phages caudés, même si la fréquence de tels phénomènes est loin d'être constante pour chaque portion de génome (Hendrix *et al.*, 1999). De tels échanges ne sont pas limités à des gènes isolés : il a été mis en évidence l'échange de large blocs de gènes correspondant à plus de la moitié du génome (Lawrence *et al.*, 2002). Ceci peut aboutir à la production de phages très différents au niveau de leur organisation génétique et de leurs propriétés biologiques (Hendrix, 2003).

L'existence de tels processus va à l'encontre d'une hypothèse souvent proposée par le passé et suivant laquelle les phages caudés seraient un ensemble de populations distinctes et assez éloignées. Dans ce cas, les membres d'un groupe restreint (comme les phage lambdaïdes ou les phages du lait) échangeraient fréquemment entre eux mais quasiment jamais avec les membres d'un groupe plus éloigné. Pour R. W. Hendrix, il est plus juste de considérer la population des phages caudés comme un continuum génétique. Ainsi, la structure apparente serait un artefact résultant de l'échantillonnage. De façon plus imagée, la part des phages caudés bien décrits serait à la population totale ce que le grumeau est au pudding (Hendrix, 2003) !

A une échelle plus large, il existe des éléments suggérant une relation de parenté entre les virus procaryotes et eucaryotes. Ainsi le phage  $\phi 6$  (infectant le genre *Pseudomonas*) est un cousin des *Reovirus* eucaryotes en raison de caractéristiques communes telles qu'un génome à ARN double brin segmenté, une capsid double et une ARN polymérase associée au virion (Bamford *et al.*, 2002). De la même manière, il existe des connections évidentes entre les phages de type PRD1 et les *Adenovirus* ainsi qu'entre certains virus procaryotes caudés (bactériens et archéens) et les *Herpesvirus* (Bamford *et al.*, 2002 ; Newcomb *et al.*, 2001). Plus récemment, une homologie de structure chez la protéine majoritaire de capsid a été démontrée pour le virus archéen STIV, le virus eucaryote PBCV-1 et le phage PRD1. De plus, le repliement de cette protéine est très proche de celui observé chez les *Adenovirus* infectant les mammifères (Rice *et al.*, 2004 ; Khayat *et al.*, 2005).

### **b. La notion de génomes en mosaïque**

De telles observations suggèrent que les génomes viraux seraient en fait des mosaïques de « modules » résultant de recombinaisons entre des séquences ancestrales. Les modules sont souvent des gènes individuels mais peuvent aussi être des portions de gènes correspondant à des domaines protéiques (Hendrix, 2003). Plusieurs études comparatives ont été menées, notamment sur des phages lambdaïdes et des mycobactériophages, pour étayer cette théorie (Juhala *et al.*, 2000 ; Pedulla *et al.*, 2003).

Petit à petit, il est apparu que le processus de transfert latéral de gènes résulte en fait d'événements non homologues essentiellement aléatoires en ce qui concerne la position relative des gènes (Pedulla *et al.*, 2003 ; Hendrix, 2002). La remarquable abondance de virus permet de contrebalancer la proportion importante de génomes recombinants non viables impliquée par un tel processus. Ainsi, tous les recombinants de la population sont éliminés sauf ceux qui ne nuisent pas au fonctionnement du phage. Il reste alors seulement les recombinaisons situées à des endroits neutres, ce qui donne l'impression d'un processus dirigé, voir même programmé. Lorsque les mosaïques de modules sont créées par recombinaison non homologue, elles peuvent se réassortir dans la population avec une fréquence élevée de recombinaisons homologues, soit grâce à des gènes homologues, soit en utilisant de courtes régions « délimitées » comme cela est décrit chez certains phages lambdaïdes (Clark *et al.* 2001 ; Hendrix, 2003).

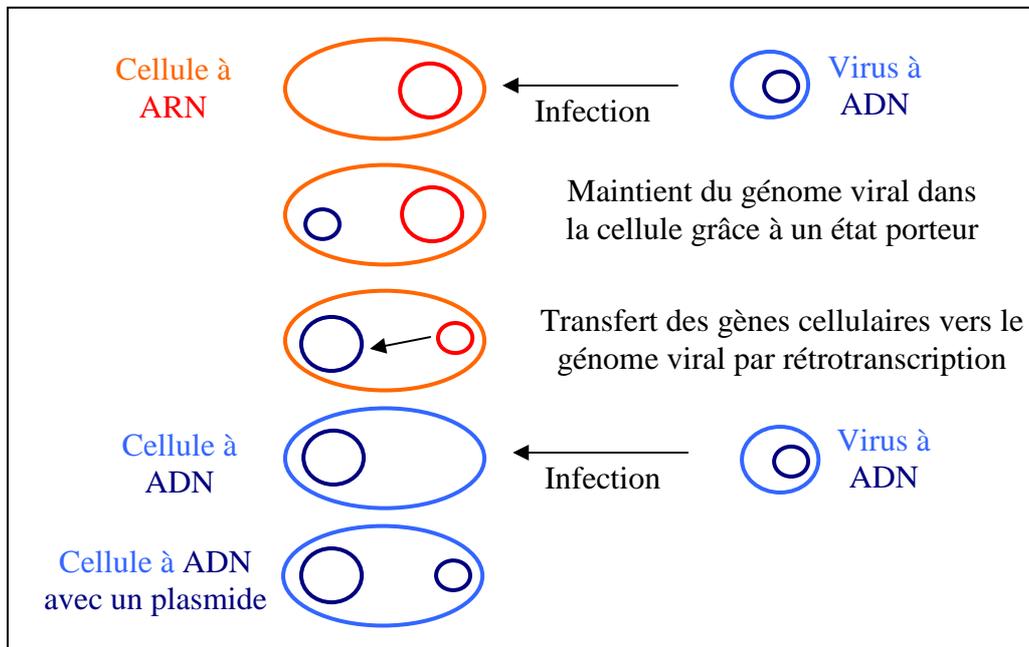
A l'heure actuelle, élaborer des scénarios phylogénétiques à partir de ces nouvelles théories est très tentant mais plein d'incertitudes (Hendrix, 2003). Cependant, certaines concernant l'origine et l'évolution des virus de procaryotes ont été, et sont encore, élaborées.

### **c. Les théories de l'origine et de l'évolution**

Ainsi, deux théories principales s'affrontent actuellement. La première, exposée notamment par H. W. Ackermann, est que les virus descendraient de constituants cellulaires ayant acquis une coque protéique et une répllication indépendante. Cette hypothèse est soutenue par le fait que les phages sont polyphylétiques, ce qui implique plusieurs ancêtres distincts en des temps et des lieux différents. Ainsi, les *Levivirus* ont pu dériver d'ARNm et les *Inovirus* filamenteux peuvent descendre de plasmides ayant acquis une capsid. Les virus caudés, ubiquitaires et présents chez les bactéries et les archées, peuvent très bien être les virus de procaryotes les plus anciens. En effet, ils auraient infecté l'ancêtre commun aux bactéries et aux archées avant la séparation des deux domaines. Une fois apparu, ils ont pu évoluer par mutations ponctuelles et réarrangements génomiques (Ackermann, 2001).

Dans la deuxième théorie, défendue notamment par P. Forterre, les virus auraient transformé l'ARN en l'ADN grâce à des ribonucleotide réductases et des thymidilates synthases. Puis, différents systèmes de répllication, de réparation et de recombinaison de l'ADN auraient été inventés indépendamment dans des lignées virales différentes, et ce à partir d'enzymes initialement impliquées dans le métabolisme de l'ARN. Seule une partie de ces systèmes aurait alors été transférée vers les cellules primordiales à ARN. Ceci explique pourquoi un certain nombre d'activités impliquées dans le métabolisme de l'ADN sont assurées par des protéines non homologues (Forterre, 2005). C'est le cas notamment pour les topoisomérases (Topo IIA et Topo IIB) ou les thymidilates

synthèses (ThyA et ThyX) (Gadelle *et al.*, 2003 ; Myllykallio *et al.*, 2002). Cette théorie permet de comprendre pourquoi beaucoup de protéines virales impliquées dans le métabolisme de l'ADN n'ont pas d'homologues cellulaires (Forterre, 2005).



**Figure 3 :** Schéma expliquant la théorie de l'évolution des virus et des cellules défendue par P. Forterre (d'après Forterre, 2005).

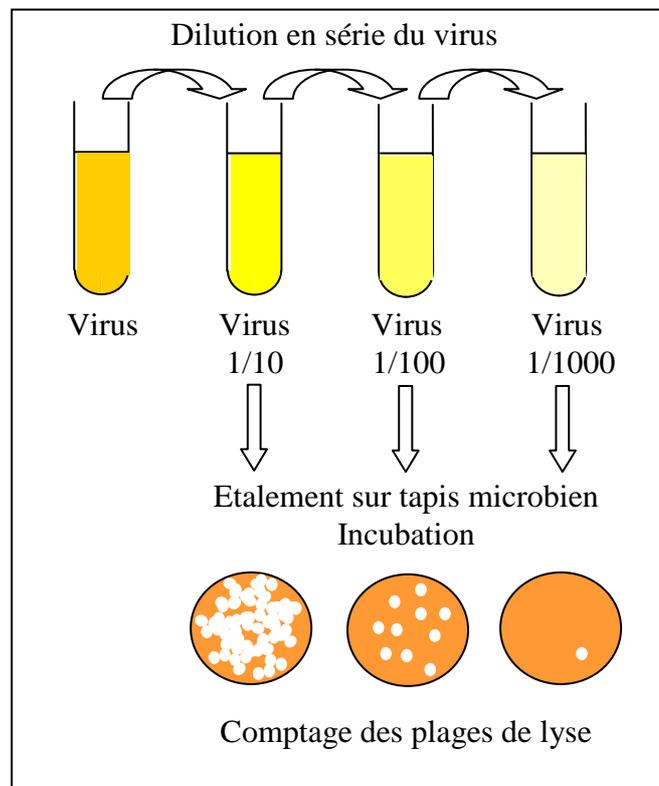
L'hypothèse concernant le processus de transfert est la suivante : un virus à ADN infecte une cellule à ARN et se maintient grâce à un état porteur. Des gènes cellulaires sont progressivement transférés par rétrotranscription vers le génome viral qui persiste sous forme plasmidique. Puis, ce dernier supprime complètement le génome cellulaire à ARN et devient le chromosome cellulaire à ADN. L'infection d'une cellule ADN par un virus ADN donne lieu suivant le même principe à une cellule ADN comportant un chromosome et un plasmide (cf Figure 3). Il est alors envisageable que le virus recrute des membranes intracellulaires pour former une membrane autour de son génome. Ce phénomène a déjà été observé chez les gros virus à ADN double brin pour former leur enveloppe. Ainsi, les virus seraient à l'origine de l'apparition du noyau cellulaire eucaryote (Forterre, 2005).

## II. Les techniques de détection et de quantification des virus de procaryotes

Depuis presque un siècle, les moyens de détection et de quantification des virus de procaryotes se sont peu à peu améliorés et diversifiés. Seules les techniques les plus utilisées sont décrites ci-après.

### 1. La technique des plages de lyse

La mise en évidence des premiers phages a été possible grâce à la lyse qu'ils provoquaient dans une culture de bactéries (Twort *et al.*, 1915 ; D'Hérelle, 1917). Cette propriété a été utilisée très tôt pour déterminer le titre viral des phages lytiques. La technique dite des plages de lyse consiste à étaler une suspension virale très diluée à la surface d'un tapis bactérien dans une boîte de Pétri contenant un milieu nutritif solide. Après incubation, il apparaît des trous dans le tapis bactérien qui sont la résultante de la destruction en cascade des cellules (cf Figure 4) Ces derniers sont appelées plages de lyse et permettent après comptage d'estimer le nombre de virus en pfu (plaque forming unit).mL<sup>-1</sup> (Signer, 1967).



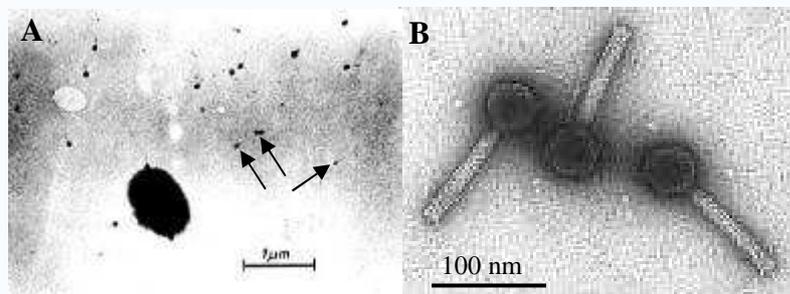
**Figure 4 :** Schéma du protocole utilisé dans la technique des plages de lyse.

Cette technique présente l'avantage d'être simple et peu onéreuse. Cependant, comme son nom l'indique, elle n'est applicable qu'aux virus lytiques. C'est pour cette raison que d'autres techniques ont été développées par la suite.

## 2. La microscopie électronique en transmission (MET)

Le principe de la microscopie électronique en transmission a été mis au point en 1931 par M. Knoll et E. Ruska. Ce dernier reçut le prix Nobel de physique en 1986 pour cette invention. Cette technique de microscopie basée sur le principe de diffraction des électrons permet un facteur de grossissement d'un million. Son utilisation en virologie a été initiée par H. Ruska, E. Pfankuch et G. A. Kausche en 1940 (Ruska, 1940 ; Pfankuch & Kausche, 1940).

Les virus sont adsorbés sur une grille métallique recouverte d'un film de carbone puis colorés négativement. Une solution contenant un agent contrastant (tétroxyde d'osmium, acétate d'uranyle, citrate de plomb, acide phosphotungstique, molybdate d'ammonium) est ajoutée sur la grille. Celui-ci va se fixer préférentiellement au bord des échantillons adsorbés. En raison de sa forte masse atomique, le contrastant dévie les électrons. Ainsi le virus apparaît plus foncé que ce qui l'entoure (cf Figure 5).



**Figure 5 :** Photographies de bactériophages après coloration négative et observés en MET prises dans les années 1940 (A) (d'après Ruska, 1940) et 1970 (B) (d'après Inman *et al.*, 1976).

Cette technique présente l'avantage de permettre une observation directe des particules. Ainsi, la morphologie et par conséquent certains critères de classification peuvent être déterminés. La limite de détection est de  $10^4$  particules.mL<sup>-1</sup> (Ewert & Paynter, 1980). Cependant, c'est une technique extrêmement lourde et coûteuse si le nombre d'échantillons analysés est élevé. Par ailleurs, elle requiert un observateur très expérimenté. De plus, il a été prouvé que les comptages effectués ne sont pas toujours très reproductibles et que le titre viral déterminé est souvent en deçà de la réalité. (Hennes & Suttle, 1995 ; Bettarel *et al.*, 2000).

### 3. La microscopie à épifluorescence (ME)

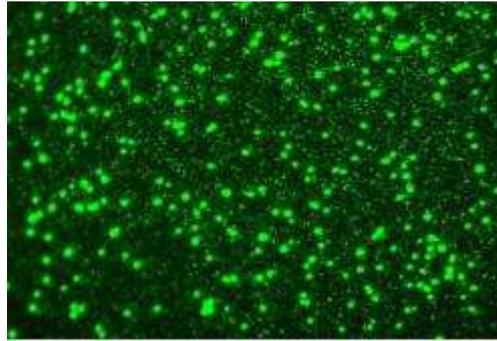
Le coût élevé et la faible disponibilité des MET ont été les principales causes du développement de la microscopie à épifluorescence. Dans cette technique dérivée de celle utilisée pour le comptage des bactéries, l'acide nucléique des virus, concentrés par ultrafiltration, est marqué grâce à un fluorochrome caractérisé notamment par la durée de persistance de sa fluorescence appelée temps de « fading » (cf Figure 6). L'observation se fait sur un microscope à épifluorescence (Weinbauer & Suttle, 1997).

Il existe trois types principaux de fluorochromes utilisables pour cette technique. Le DAPI (ou 4',6'-diamino-2-phenylindole), couramment employé pour détecter les bactéries depuis les années 1980, a été le premier utilisé pour marquer le compartiment viral (Suttle *et al.*, 1990). La méthode au DAPI présente l'avantage d'être très rapide (environ 30 min). Cependant, en raison de la faible intensité de fluorescence et du temps de « fading » relativement court (94 sec), le comptage des échantillons se révèle être assez difficile (Porter & Feig, 1980 ; Weinbauer & Suttle, 1997).

Le Yo-Pro-1 {4-[3-méthyl-2,3-dihydro-(benzo-1,3-oxazole)-2-méthylméthylène]-1-(3'-triméthyl ammoniopropyl)-quinolinium diiode} est un colorant à base de cyanine. Contrairement au DAPI, il produit une fluorescence intense et très stable car son temps de « fading » oscille entre 210 et 244 sec. De plus, le protocole a été amélioré et permet maintenant une incubation de 4 min au lieu de 48 h (Hennes & Suttle, 1995 ; Xenopoulos & Bird, 1997 ; Bettarel *et al.*, 2000).

Le SYBR Green I est arrivé plus récemment sur le marché. Sa fluorescence est aussi intense que le Yo-Pro-1 mais en revanche son temps de « fading » est plus court que celui du DAPI (49 sec). Pour pallier cela, il est donc nécessaire d'appliquer une solution dite « anti-fading ». Son principal inconvénient réside dans la diminution de la précision des comptages lorsque le nombre de particules virales devient trop important ( $> 5.10^7$  particules.mL<sup>-1</sup>) (Bettarel *et al.*, 2000). Cependant, il présente tout de même l'avantage de ne pas marquer les débris présents dans l'échantillon (Noble & Fuhrman, 1998).

Dans certaines études, il s'avère que les comptages obtenus grâce à la microscopie à épifluorescence sont surestimés. En effet, les fragments d'ADN et d'ARN libres sont également marqués et indiscernables d'un « vrai » virus. Ils produisent donc un phénomène d'interférence qui fausse les comptages. Une des solutions envisagée pour pallier ce problème est un traitement de l'échantillon à la DNase et à la RNase. Les résultats restent mitigés en raison du fait que ces enzymes semblent également altérer les capsides virales (Hara *et al.*, 1991 ; Bettarel *et al.*, 2000).



**Figure 6 :** Photographie en épifluorescence de procaryotes (gros points) et de leurs virus (petits points) marqués au SYBR Green I (d'après le site <http://fuhrmanlab.usc.edu>).

Une étude récente a montré que la conservation de l'échantillon a également un rôle prépondérant dans les résultats obtenus en microscopie à épifluorescence. En effet, la fixation au glutaraldhéryde ou au formaldéhyde suivie d'un stockage à 4°C engendre une perte de 72 % des particules après 16 jours. Pour éviter une telle dégradation, il apparaît nécessaire de colorer juste après la fixation puis de stocker la préparation à - 20°C. Il est également possible de plonger l'échantillon dans l'azote liquide après fixation puis de le conserver à - 86°C jusqu'à l'analyse (Wen *et al.*, 2004).

#### **4. La cytométrie en flux**

L'apparition d'une nouvelle génération de fluorochromes colorant spécifiquement les acides nucléiques combinée à l'amélioration des techniques de cytométrie en flux a permis l'obtention d'un nouveau moyen de détection et de quantification des virus. Ce système présente le grand avantage d'être complètement adapté à l'analyse d'un très grand nombre d'échantillons grâce à l'automatisation du procédé. De plus, la cytométrie en flux permet de mettre en évidence différentes populations de virus (Hercher *et al.*, 1979 ; Marie *et al.*, 1999).

Les fluorochromes utilisés sont le SYBR Green I et II, le TOTO-1, le YOYO-1, le PicoGreen, l'OligoGreen et le SYBR Gold. Plusieurs études ont comparé ces différents colorants et le SYBR Gold apparaît comme le meilleur. En effet, il produit une fluorescence beaucoup plus stable mais tout aussi intense que celle du SYBR Green I (Marie *et al.*, 1996, 1997 ; Chen *et al.*, 2001).

Les comptages obtenus sont très similaires à ceux effectués en microscopie à épifluorescence. Aucune corrélation n'a pu être mise en évidence entre l'intensité de fluorescence et la taille des génomes viraux. Cependant, il apparaît que les petits virus à ARN sont difficilement détectables (Marie *et al.*, 1999 ; Brussaard *et al.*, 2000 ; Chen *et al.*, 2001).

## 5. Les techniques d'amplification de gène

Au contraire des techniques culturales et microscopiques, l'amplification de gènes apporte une preuve indirecte de la présence ou de l'absence de tel ou tel virus.

### a. La Polymerase Chain Reaction (PCR) et la Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

Les premières techniques d'amplification de gène qui furent utilisées pour la détection des virus sont la PCR (pour les virus à ADN) et la RT-PCR (pour les virus à ARN) (Saiki *et al.*, 1985 ; Brüssow *et al.*, 1994 ; Lee *et al.*, 1989 ; Griffin *et al.*, 1999). Dans la PCR, un gène cible est amplifié à l'aide d'amorces spécifiques et d'une polymérase thermostable (souvent la Taq polymérase) au cours de plusieurs cycles comprenant trois phases (dénaturation, hybridation, élongation). Pour la RT-PCR, une réverse transcription est réalisée au préalable pour obtenir de l'ADN complémentaire à partir de l'ARN viral grâce à une réverse transcriptase.

Ces techniques moléculaires présentent l'avantage d'être très sensibles car elles permettent de détecter les génomes viraux, même s'il n'y a que quelques copies dans l'échantillon analysé. Par ailleurs, le choix judicieux des amorces utilisées autorise une identification rapide du ou des virus présents. De plus, les provirus, invisibles en MET ou en ME car ils ne produisent pas de virions, sont alors détectables.

Le point crucial de ces techniques reste le gène ciblé. En effet, il n'existe aucun équivalent chez les virus de l'ADNr 16S retrouvé chez tous les procaryotes. Ainsi, des gènes différents sont sélectionnés suivant les virus étudiés. Une région de 1,5 kpb connue pour être présente chez tous les phages de *Bacteroides fragilis* HSP40 a ainsi été utilisée pour détecter ce type de virus dans des échantillons environnementaux (Puig *et al.*, 2000). La séquence codant pour la protéine majeure de capsid commune à tous les phages de type T4 a, par ailleurs, été retenue pour étudier la phylogénie de plusieurs de ces virus provenant de différents milieux aquatiques (Filée *et al.*, 2005).

Le principal inconvénient réside dans le fait que les conditions d'expérimentations doivent être optimisées pour chaque type de virus sous peine d'obtenir des amplifications aspécifiques ou des résultats faussement négatifs.

### b. La PCR et la RT-PCR quantitatives

Plus récemment, la PCR et la RT-PCR quantitative ont été développées dans le but de quantifier les virus et de suivre leur développement. Cette technologie est basée sur la détection et la quantification d'une molécule fluorescente dont l'émission est directement proportionnelle à la

quantité d'amplicons générés pendant la réaction de PCR ou RT-PCR. Il existe 2 types de marquage : les agents se liant à l'ADN double brin (comme le SYBR Green I) et les sondes fluorescentes (comme les sondes Taqman) (Heid *et al.*, 1996 ; Edelman & Barletta, 2003 ; Gibson *et al.*, 1996 ; Shin & Sobsey, 2003).

Le SYBR Green I ne devient fluorescent que s'il se lie à un acide nucléique. Il est donc possible d'inclure cette molécule à faible dose dans une PCR et de suivre l'augmentation des produits de PCR directement en mesurant la fluorescence dans la gamme d'émission du SYBR Green I (maximum à 550 nm). Cela présente l'avantage d'éviter de concevoir une sonde de marquage pour chaque gène ciblé. En revanche, la fluorescence mesurée peut en partie provenir du marquage des produits d'amplification non spécifiques.

La sonde Taqman est un oligonucléotide spécifique d'un morceau interne à la séquence amplifiée. Elle est marquée en 5' par un fluorophore appelé « reporter » et en 3' par un autre type de fluorophore appelé « quencher ». Le spectre d'émission du reporter chevauche le spectre d'excitation du quencher. L'émission du reporter est atténuée par la proximité du quencher. Si la sonde est dégradée par l'activité exonucléasique 5'→3' de la Taq polymérase, les fluorophores s'éloignent et l'émission du reporter sera augmentée. L'augmentation du signal correspondant à la fluorescence du reporter est proportionnel au nombre de copies polymérisées à chaque cycle de PCR. Le principal avantage dans l'utilisation de sondes Taqman est la spécificité du signal obtenu. Il est même possible de suivre plusieurs gènes en même temps. Cependant, cette technique est onéreuse en raison du prix élevé des sondes et du fait qu'il faut autant de sondes que de gènes étudiés.

## 6. La métagénomique

Cette approche, relativement récente, permet de détecter des séquences génomiques virales dans un environnement donné. Le compartiment viral est tout d'abord isolé, puis une banque LASL (Linker-Amplified Shotgun Library) est créée (Rohwer *et al.*, 2001). Ce type de banque permet de contourner les problèmes causés lors d'un clonage classique par les nucléotides modifiés, les gènes viraux létaux et la faible quantité d'ADN contenue dans les échantillons environnementaux. Les séquences obtenues sont alors analysées pour déterminer quels sont les virus représentés et dans quelles proportions (Breitbart *et al.*, 2003 ; Edwards & Rohwer, 2005). Etant donné la masse d'informations récupérées, la plupart (environ 75 %) des séquences obtenues ne trouvent aucun homologue dans les différentes bases de données. De nombreux virus sont ainsi découverts mais reste à l'heure actuelle incultivés (Breitbart *et al.*, 2002, 2003).

La difficulté majeure de cette technique est d'arriver à isoler correctement le compartiment viral sans le contaminer ni perdre trop de matériel. Cependant, une fois ce point résolu, la métagénomique présente l'avantage de permettre l'obtention d'informations concernant la diversité des virus incultivés. La métagénomique permet donc d'évaluer de manière globale la composition ainsi que la structure des communautés virales de l'environnement (Edwards & Rohwer, 2005).

### III. Les virus de procaryotes dans les environnements extrêmes

L'étude bibliographique qui suit est en grande partie reprise dans une revue intitulée « *Viruses in extreme environments* » qui a été déposée à la revue « *Reviews in Environmental Science and Biotechnology* » (cf Annexe 6). La définition exacte d'un environnement extrême n'existe pas mais il paraît relativement correct de dire qu'il s'agit d'un environnement où les contraintes physiques et chimiques sont telles qu'elles confinent aux limites de la vie. Par ailleurs, il apparaît que la virosphère est quasiment omniprésente sur la planète (Suttle, 2005a). Ainsi, elle a même réussi à coloniser ces environnements *a priori* hostiles (cf Tableau 1).

**Tableau 1** : Principales caractéristiques de l'écologie virale des différents environnements extrêmes

Habitat	Abondance	Morphologies	Hôtes
Milieus hypersalins	$10^7$ à $10^9$ .mL <sup>-1</sup>	citron, tête-queue, sphérique	archées appartenant aux genres <i>Halobacterium</i> , <i>Haloferax</i> , <i>Haloarcula</i> , <i>Natrialba</i> , <i>Halorubrum</i>
Lacs alcalins	$10^8$ à $10^9$ .mL <sup>-1</sup>	nd	bactéries apparentées à l'espèce <i>Idiomarina baltica</i>
Sables du désert du Sahara	nd	tête-queue	nd
Glaces de l'Arctique	$9.10^6$ à $3.10^8$ .mL <sup>-1</sup>	tête-queue	bactéries apparentées aux espèces <i>Shewanella frigidimarina</i> , <i>Flavobacterium hiibernum</i> et <i>Colwellia psycherythrae</i>
Sédiments de subsurface	$> 10^9$ .g <sup>-1</sup>	nd	nd
Sources chaudes acides terrestres	nd	citron, filament, bâtonnet, goutte d'eau sphérique, fuseau à extensions bipolaires, bouteille, icosaédrique	archées appartenant aux genres <i>Sulfolobus</i> , <i>Acidianus</i> , <i>Thermoproteus</i> , <i>Desulfurolobus</i> , <i>Pyrobaculum</i>
Sources hydrothermales marines profondes	$1,45.10^5$ à $9,9.10^7$ .mL <sup>-1</sup>	citron, filament, bâtonnet, cuillère, fuseau à extensions bipolaires	archées appartenant à l'ordre des <i>Thermococcales</i>

nd : non déterminé

## 1. Les virus des milieux hypersalins

### a. La diversité globale

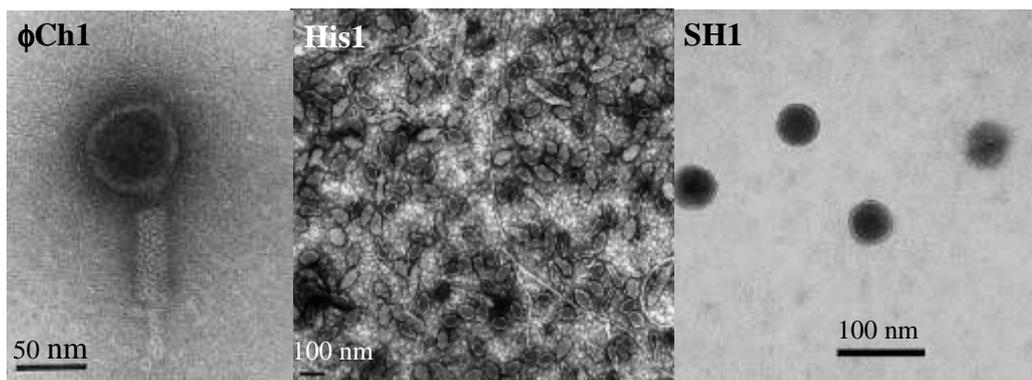
Les milieux hypersalins renferment une quantité importante de virus (de  $10^7$  à  $10^9$ .mL<sup>-1</sup>) mais jusqu'à présent, le nombre de spécimens décrits en laboratoire reste faible (environ 15) (cf Figure 7 et Tableau 2). Ces derniers possèdent en majorité une morphologie tête-queue. Cependant, des observations directes en MET réalisées dans le cadre d'étude de la diversité ont montré que c'est plutôt la forme fuseau qui domine chez les particules virales des milieux hypersalins. Il existe donc une différence entre les résultats apportés par les caractérisations de systèmes hôte – virus et les études de la diversité virale dans l'environnement, peut-être due aux hôtes choisis qui ne correspondent pas aux espèces dominantes dans l'habitat naturel (Guixa-Boixareu *et al.*, 1996 ; Oren *et al.*, 1997 ; Dyll-Smith *et al.*, 2003).

### b. Les virus modèles

#### ➤ $\phi$ H

Le virus  $\phi$ H, le mieux étudié de sa catégorie, infecte cinq souches de l'espèce archéenne *Halobacterium salinarium*.  $\phi$ H appartient à la famille des *Myoviridae*. Sa tête est un polyèdre de 64 nm de diamètre et il possède une queue contractile de 170 nm de long. Trois protéines constitutives majoritaires ont pu être mises en évidence par électrophorèse en SDS-PAGE. Leurs tailles sont respectivement de 22, 53 et 80 kDa. Par ailleurs, la concentration en sel du milieu dans lequel se trouve le virus semble primordiale. Ainsi, en dessous d'une concentration en NaCl ou en KCl de 3 M, la particule perd son infectiosité alors que les concentrations requises sont moindres en présence de 0,1 M de Mg<sup>2+</sup> (Schnabel *et al.*, 1982a ; Dyll-Smith *et al.*, 2003).

$\phi$ H est un virus tempéré qui peut produire des plages de lyses chez la souche R1 (dont la mutation caractéristique est une absence de vacuoles gazeuses) ainsi que chez la souche sauvage mais avec une efficacité réduite d'un facteur  $10^3$ - $10^4$ . L'apparition de particules se produit 5,5 h après le début de l'infection, alors que la lyse n'est observée qu'au bout de 7 h (Schnabel *et al.*, 1982a). Lorsque le génome viral est présent dans la cellule, cette dernière devient résistante à toute surinfection (Reiter *et al.*, 1988).



**Figure 7 :** Photographie en MET de différents virus isolés de milieux hypersalins représentant les principales morphologies rencontrées (d'après Dyall-Smith *et al.*, 2003).

**Tableau 2 :** Principales caractéristiques des virus décrits isolés de milieux hypersalins.

Nom du virus	Morphologie et taille du virion	Taille du génome*	Physiologie du virus	Références bibliographiques
Hh1	Tête-queue 60 nm ; 100 nm	37,6 kpb	Tempéré	Pauling <i>et al.</i> , 1982 Rohrmann <i>et al.</i> , 1983
Hh 3	Tête-queue 75 nm ; 50 nm	29,6 kpb	Tempéré	Pauling <i>et al.</i> , 1982 Rohrmann <i>et al.</i> , 1983
Ja1	Tête-queue 90 nm ; 150 nm	230 kpb	nd	Wais <i>et al.</i> , 1975
Hs1	Tête-queue 50 nm ; 120 nm	nd	Lytique	Torsvik & Dundas, 1974 Torsvik & Dundas 1980
φH	Tête-queue 64 nm ; 170 nm	59 kpb	Tempéré	Schnabel <i>et al.</i> , 1982a Stolt & Zillig, 1992
φN	Tête-queue 55 nm ; 80 nm	56 kpb	nd	Vogelsang-Wenke & Oesterhelt, 1988
S45	Tête-queue 40 nm ; 70 nm	nd	Lytique	Daniels & Wais, 1984
S5100	Tête-queue Taille nd	nd	nd	Daniels & Wais, 1990
B10	Tête-queue Taille nd	nd	nd	Dyall-Smith <i>et al.</i> , 2003
HF1	Tête-queue 58 nm ; 94 nm	77,7 kpb	Lytique	Nuttall & Dyall-Smith., 1993
HF2	Tête-queue 58 nm ; 94 nm	75,9 kpb	Lytique	Nuttall & Dyall-Smith, 1993
His1	Fuseau 74 nm x 44 nm	14,9 kpb	Etat porteur	Bath & Dyall-Smith., 1998
His2	Sphérique 62 nm	16,2 kpb	Etat porteur	Bath <i>et al.</i> , 2006
φCh1	Tête-queue 70 nm ; 130 nm	58,5 kpb	Tempéré	Witte <i>et al.</i> , 1997 Klein <i>et al.</i> , 2002
SH1	Polyèdre 55 nm	30,9 kpb	Lytique	Porter <i>et al.</i> , 2005

\* : Tous les génomes des virus des milieux hypersalins sont constitués d'une molécule d'ADN double brin linéaire.

nd : non déterminé

Le génome de  $\phi$ H est un ADN linéaire bicaténaire de 59 kpb caractérisé par un pourcentage en G + C de 65. Ses extrémités sont redondantes, ce explique pourquoi sous forme de provirus, il adopte une structure circulaire. Par ailleurs, cela suggère que l'ADN serait synthétisé sous forme de concatémères, encapsidés au fur et à mesure dans les têtes à partir d'une séquence nommée pac (Schnabel *et al.*, 1982b, Schnabel & Zillig, 1984).

Les profils de restrictions du génome obtenus avec les endonucléases *ClaI*, *BglII* et *BglIII* génèrent des bandes surnuméraires peu intenses qui disparaissent lorsque le virus est purifié grâce à plusieurs « plaque-tests » successifs. Ceci suggère l'existence d'au moins deux variants du virus. Après étude plus poussée de ces résultats, ce sont en fait 8 variants (notés 1 à 8) qui ont pu être mis en évidence. Leurs génomes diffèrent par des insertions, des délétions ou des inversions d'un fragment d'ADN d'une taille minimum de 11 kpb nommé segment L. Les inversions ont lieu à une fréquence élevée chez les variants qui portent un élément de 1,8 kpb, appelé ISH1.8, au niveau des régions flanquantes du segment L. Ce dernier peut se circulariser sous forme d'un plasmide dénommé  $\phi$ HL et être impliqué dans l'immunité de certaines souches vis-à-vis des différents variants de  $\phi$ H. Il peut aussi s'intégrer au génome viral par recombinaison homologue et contribuer aux réarrangements d'ADN chez *Halobacterium salinarum* (Schnabel *et al.*, 1982b, 1984 ; Schnabel, 1984a, 1984b).

La transcription chez  $\phi$ H, lorsqu'il entame son cycle lytique, est dépendante du temps. Au cours de la phase précoce, il a pu être mis en évidence au niveau du segment L un transcrit majoritaire de 2 000 nt. Celui-ci est ensuite séparé en deux molécules qui font respectivement 1 100 nt (notée T<sub>4</sub>') et 700 nt (notée T<sub>4</sub>). Un second transcrit nommé T<sub>1</sub>/T<sub>2</sub>/T<sub>3</sub> démarre à l'extrémité droite du segment L. Ces deux produits apparaissent moins de 30 min après le début de l'infection. Les gènes tardifs, eux, se situent à la gauche du segment L et sont transcrits seulement après 3 h. Ils renferment notamment les trois séquences codant pour les protéines majoritaires qui seraient peut-être organisées en unité de transcription (Stolt & Zillig, 1994).

Le phénomène est légèrement différent lorsque la souche R1 renferme le plasmide  $\phi$ HL. Ainsi, 3 transcrits « d'immunité » ont pu être détectés. Le promoteur d'un transcrit « d'immunité », noté T<sub>6</sub>, est dos-à-dos avec le promoteur du transcrit lytique T<sub>4</sub>. Il a été montré *in vitro* que T<sub>6</sub> code pour un répresseur Rep qui se fixe au niveau du promoteur de T<sub>4</sub>. De plus, les deux autres transcrits « d'immunité », notés T<sub>9</sub> et T<sub>10</sub>, codent pour une protéine Per qui augmente d'un facteur dix l'action de Rep. Il a également été mis en évidence que l'immunité était aussi médiée par un ARN antisens noté T<sub>ant</sub>. En effet, T<sub>ant</sub> démarre 151 pb en aval du point de départ de T<sub>1</sub> et il est transcrit dans le sens inverse à ce dernier. Ainsi, il peut former un duplex avec T<sub>1</sub> et rendre la molécule sensible au

clivage par des RNAses. C'est la première fois que ce type de régulation a pu être mis en évidence chez les archées (Stolt & Zillig, 1993, 1994).

L'étude de la séquence du génome de  $\phi$ H a permis de mettre en évidence de fortes similarités avec  $\phi$ Ch1, un autre halovirus. Néanmoins, à part quelques gènes codant pour des méthylases, il existe très peu d'homologues chez les génomes de bactériophages et de virus eucaryotes (Dyall-Smith *et al.*, 2003).

#### ➤ $\phi$ Ch1

$\phi$ Ch1 a été décrit plus récemment et infecte l'archéobactérie haloalkalophile *Natrialba magadii*. Il appartient à la famille des *Myoviridae* et possède une structure très similaire à celle de  $\phi$ H (cf Figure 7). La queue comporte de petites structures certainement impliquées dans l'adsorption à la cellule hôte. La concentration minimale en NaCl requise est de 2 M, même en présence de  $Mg^{2+}$ . En deçà, les particules se désagrègent (Witte *et al.*, 1997).  $\phi$ Ch1 est un virus tempéré. Dans sa phase lysogène, il est sous forme de prophage intégré au génome de l'hôte. Cependant, il est aussi capable de produire des plages de lyse. La souche porteuse peut néanmoins être curée grâce à des repiquages successifs (Witte *et al.*, 1997).

Des analyses en SDS-PAGE ont permis de mettre en évidence 8 protéines structurales (de 15 à 75 kDa (notées B à I) dont les points isoélectriques sont compris entre 3,3 et 5,2. Ceci est typique des micro-organismes alkalophiles. Par ailleurs, il a pu être déterminé que la neuvième protéine structurale de 80 kDa, nommée A, était en fait la résultante, soit d'un hétéromultimère de plusieurs protéines de  $\phi$ Ch1, soit d'un homomultimère de protéine H. Dans les deux cas, les molécules sont covalamment liées et contribuent à la stabilité de la structure du virion (Witte *et al.*, 1997 ; Klein *et al.*, 2002).

Les virions renferment en plus du génome de  $\phi$ Ch1, un ADN double brin linéaire de 58 kpb *dam* méthylé, des ARNs de différentes tailles (80 à 700 nt). Des expériences d'hybridation ont permis de déterminer qu'il s'agissait d'ARNs codés par l'hôte. C'est la première fois qu'un tel phénomène a pu être observé chez les virus d'archées (Witte *et al.*, 1997 ; Baranyi *et al.*, 2000). Le génome semble être organisé en 3 parties. La région de gauche regroupe des gènes impliqués dans la morphogénèse du virion ou codant pour des protéines structurales. La région centrale comporte les fonctions de répllication, de stabilisation de plasmide et de régulation de l'expression génique. La partie de droite est composée majoritairement de gènes ayant une fonction inconnue mais elle renferme aussi des gènes codant pour des enzymes de restriction et de méthylation de l'ADN. Cette

organisation modulaire est très similaire à celle observée chez les bactériophages caudés, ce qui met en évidence une conservation à travers les différents domaines (Klein *et al.*, 2002).

L'analyse bioinformatique a prédit 98 gènes et une fonction a pu être assignée pour 1/5<sup>ème</sup> d'entre eux. Bien que la plupart n'aient pas de fonction connue, un nombre important comporte tout de même des similarités avec  $\phi$ H, et plus particulièrement avec le segment L. Une partie de ces séquences homologues se trouve être inversée entre les deux virus. Ceci est très certainement la résultante d'une recombinaison site-spécifique au niveau des sites pal1 et pal2 qui bordent le fragment inversé. Ce phénomène pourrait être médié par une recombinase virale codée par int1 et/ou int2 qui présentent des homologies avec l'intégrase du bactériophage  $\lambda$  (Klein *et al.*, 2002). L'ADN de  $\phi$ Ch1 est redondant à ses extrémités. Il peut donc, comme  $\phi$ H, se circulariser. Des analyses de restriction menées sur de l'ADN viral isolé de cellules lysogènes suggèrent un intermédiaire de répllication adoptant cette forme plasmidique. Ceci est en accord avec le fait que plusieurs gènes ont des homologues impliqués dans la répllication et la stabilisation des plasmides. Par exemple, le gène 54 (rep H) ressemble à des séquences faisant partie du réplicon minimal de deux plasmides d'archées halophiles (Klein *et al.*, 2002).

#### ➤ His1

His1 est le premier virus d'halophile décrit qui n'a pas de structure tête-queue. Il infecte différentes archéobactéries : une de l'espèce *Haloarcula hispanica* mais aussi 16 autres souches des genres *Haloferax*, *Halorubrum* et *Halobacterium*. Les virions ont une forme de citron de 74 nm de long par 44 nm de large et sont pourvus d'une queue de 7 nm (cf Figure 7). Cette morphologie ressemble beaucoup à celle des virus de la famille des *Fuselloviridae* tel SSV1. Cependant, comme His1 est très différent au niveau moléculaire des autres virus en forme de fuseau, le genre *Salterprovirus* a été proposé. Très récemment, His2, un halovirus morphologiquement et génétiquement proche de His1, a été isolé et classé dans ce genre (Bath *et al.*, 2006).

His1 est dans un état porteur. L'hôte relargue en continu des particules 4 à 6 h après le début de l'infection sans jamais subir de lyse. Par ailleurs, aucune copie du génome viral ne semble ni être intégrée au chromosome archéen, ni persister sous forme plasmidique (Bath & Dyal-Smith, 1998). Le génome de His1 est un ADN double brin linéaire de 15,5 kpb. Il ne possède pas, comme  $\phi$ H ou  $\phi$ Ch1, d'extrémités redondantes qui permettent d'adopter une forme circulaire. Il ne semble pas non plus posséder le même système d'encapsidation. Enfin, il ne comporte pas de protéine associée à l'extrémité 5' de l'ADN (Bath & Dyal-Smith, 1998). La répllication se ferait selon le mode dit « d'amorçage protéique » qui est assez peu répandu mais rencontré chez certains virus de mammifère-

res (*Adenovirus*), certains bactériophages ( $\phi$ 29, PRD1), certains plasmides linéaires de *Streptomyces* et l'ADN mitochondrial de plante (Salas, 1991 ; Bath *et al.*, 2006).

Le séquençage et l'annotation très récents du génome de His1 ont permis de déterminer 35 gènes. Un d'eux code pour une ADN polymérase putative, présentant 42 % de similarité avec celle trouvée chez His2. Ces dernières possèdent des caractéristiques communes avec les polymérases de type B existant chez certains virus ainsi que dans les mitochondries des plantes et des champignons. Etant donné leurs morphologies proches, la structure identique de leurs génomes et l'homologie de leurs ADN polymérases, His1 et His2 ont certainement un ancêtre commun. Au cours du temps ils ont considérablement divergé, ce qui suggère des événements de recombinaison entre halovirus mais aussi entre les halovirus et leurs hôtes (Bath *et al.*, 2006).

#### ➤ SH1

SH1 est le dernier virus halophile décrit. Il infecte deux espèces archéennes : *Haloarcula hispanica* et *Halorubrum sodomense*. Les particules, polyédriques avec un diamètre de 70 nm, semblent dénuées de queue (cf Figure 7). L'analyse en SDS-PAGE des protéines constitutives du virion a révélé 15 bandes de poids moléculaire compris entre 4 et 185 kDa. Plusieurs d'entre elles forment des complexes protéiques. Une membrane externe fragile entoure la capsidie icosaédrique. Il a également été montré que SH1 renferme des constituants lipidiques sélectionnés et provenant de son hôte (Porter *et al.*, 2005 ; Bamford *et al.*, 2005). SH1 entraîne la lyse de son hôte 5 à 6 h après le début de l'infection. A ce stade, le nombre de particules s'élève à 200 pfu par cellule. Cependant, il arrive que les archées forment des îlots où les cellules ont une paroi épaissie. Ces dernières se sont révélées être résistantes à l'infection (Porter *et al.*, 2005).

Le génome de SH1 est une molécule d'ADN double brin linéaire de 30,9 kpb qui possède à ses extrémités des séquences répétées inversées de 309 pb. Son pourcentage en G+C est de 68,4, ce qui est tout à fait comparable à celui retrouvé chez les organismes halophiles. Le génome a été entièrement séquencé et comporte 56 gènes prédits. Aucun gène de SH1 n'a d'homologues dans les bases de données. Le séquençage N-terminal a permis de relier 8 protéines structurales aux gènes les encodant. Par ailleurs, un gène est homologue à l'ATPase permettant l'empaquetage de l'ADN chez les phages du type PRD1. Ce gène coderait alors pour une protéine de structure et serait la signature d'un mode d'assemblage commun aux deux virus. Ainsi, des capsides vides seraient synthétisées et l'encapsidation aurait lieu ensuite. Une étude structurale détaillée de SH1 devrait permettre d'évaluer la possibilité d'une relation entre ce virus et le phage PRD1 (Bamford *et al.*, 2005).

## 2. Les virus des lacs alcalins

Les lacs alcalins sont, en réalité, un cas très particulier de milieu hypersalin. Ainsi, la composition chimique de l'eau est proche de celle des marais salants mais la quantité de carbonate minéral y est beaucoup plus élevée. La seule étude portant sur les virus de cet environnement a été menée par une équipe californienne dans le lac Mono où la valeur du pH avoisine 10. L'abondance virale varie entre  $1.10^8$  et  $1.10^9$ .mL<sup>-1</sup>, ce qui fait des lacs alcalins un des systèmes aquatiques où la densité de virus est la plus importante. Des électrophorèses en champ pulsé ont permis de déterminer que la taille des génomes viraux varie de 14 à 400 kpb avec une majorité comprise entre 30 et 60 kpb. L'analyse des profils de bandes obtenus a mis en évidence que la communauté virale des eaux profondes forme un groupe distinct de celles des eaux de surface et de demi-surface. Cela montre une forte stratification de la distribution virale entre les eaux anoxiques et les eaux plus oxygénées (Jiang *et al.*, 2004).

Au cours de la même étude, un phage lytique nommé  $\Phi$ Mono1 a été isolé. Il infecte une bactérie dont le plus proche représentant décrit appartient à l'espèce *Idiomarina baltica*. Le génome de ce virus, utilisé comme sonde, a permis de déterminer que les communautés virales étaient soumises à une certaine saisonnalité. En effet, les hybridations réalisées sur un échantillon prélevé en mai 2000 donnent des résultats différents de ceux obtenus sur un échantillon récolté en août 2000 (Jiang *et al.*, 2004).

## 3. Les virus du désert

Une seule étude a été menée par une équipe française sur les virus de ce type d'environnement dans le désert du Sahara. Des échantillons de sable ont été soniqués dans différents milieux de culture puis incubés avec et sans mitomycine C. Des observations en MET réalisées sur les cultures concentrées ont révélé la présence de particules de type viral ayant des morphologies comparables à celle des virus de trois familles différentes : les *Myoviridae*, les *Siphoviridae* et les *Podoviridae* (cf Figure 8). L'analyse électrophorétique en champ pulsé de l'ADN extrait des préparations enrichies en particules suggère la présence de plusieurs génomes viraux dont la taille varie entre 45 et 270 kpb (Prigent *et al.*, 2005).

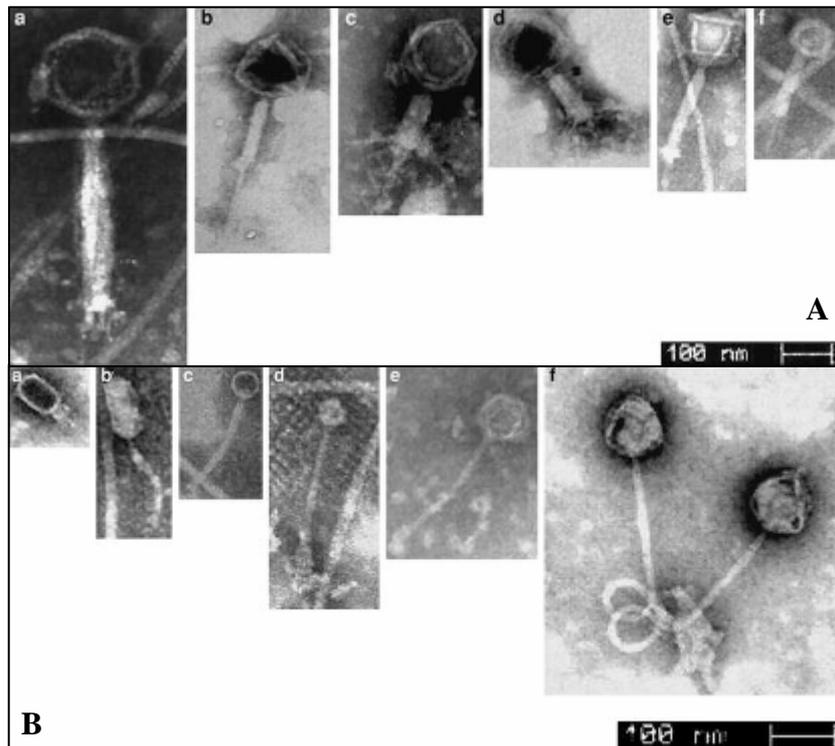
#### 4. Les virus des glaces polaires

Des travaux sont menés sur les virus psychrophiles depuis les années 1990. Une étude, menée sur des glaces prélevées au sud de Resolute dans l'Arctique canadien, a révélé une extraordinaire abondance de particules virales (de  $3,7 \cdot 10^{11}$  à  $4,9 \cdot 10^{12} \cdot \text{m}^{-2}$  ce qui correspond à des valeurs de  $9,0 \cdot 10^6$  à  $1,3 \cdot 10^8 \cdot \text{mL}^{-1}$ ). En effet, elle est 10 à 100 fois supérieure aux comptages réalisés dans la colonne d'eau située en dessous des glaces analysées ( $1,1 \cdot 10^6 \cdot \text{mL}^{-1}$ ) (Maranger *et al.*, 1994). Une autre étude menée près de la côte nord de Ellesmere Island, également dans l'Arctique canadien, a donné des résultats similaires ( $10^{10}$  à  $10^{11}$  virus. $\text{m}^{-2}$ ) (Vincent *et al.*, 2000).

La première observation au MET de phages présents dans de la glace est très récente car elle ne date que de trois ans. Dans cette étude, trois systèmes hôte - virus ont été isolés d'échantillons de glace prélevés en Arctique au large de la côte est du Groenland. L'analyse du gène de l'ADNr 16S des bactéries hôtes montre respectivement des similitudes avec les espèces suivantes : *Shewanella frigidimaria* (96,0 %), *Flavobacterium hibernum* (94,0 %) et *Colwellia psychrerythraea* (98,4 %). Les phages sont actifs jusqu'à 14°C mais leur température optimale pour la formation de plages de lyse se situe plus aux environs de 0°C. Ces virus sont hôte-spécifiques et leur structure tête-queue implique leur appartenance à deux familles différentes : les *Myoviridae* (pour l'isolat 1a) et les *Siphoviridae* (pour les isolats 11b et 21c) (cf Figure 9). Leur génome sont des ADN double brin dont la taille est de 70 kpb pour l'isolat 1a, de 40 à 50 kpb pour l'isolat 21c et de 30 kpb pour l'isolat 11b (Boriss *et al.*, 2003).

#### 5. Les virus des sédiments de subsurface

La plupart des études menées sur les virus des sédiments concerne la couche des dix premiers centimètres (Danovaro *et al.*, 2001, 2002). Une équipe canadienne a réalisé des études en épifluorescence ainsi qu'en MET sur deux échantillons de sédiments prélevés respectivement à 105 et à 118 m sous le plancher océanique près de la côte ouest du Canada. Une importante quantité de virus a pu être mise en évidence car l'abondance moyenne déterminée est supérieure à  $10^9 \text{ g}^{-1}$  de sédiment sec. Etant donné que ces virus semblent être en étroite association avec les bactéries de cet environnement, ils apparaissent comme des acteurs potentiels de la biogéochimie des sédiments marins profonds (Bird *et al.*, 2001).



**Figure 8 :** Photographie en MET des particules de type viral observées dans les échantillons de sable du désert du Sahara et étant assimilées aux familles des *Myoviridae* (A a-f), des *Podoviridae* (B a,b) et des *Siphoviridae* (B c-f) (d'après Prigent *et al.*, 2005).



**Figure 9 :** Photographie au MET des isolats 1a (A), 11b (B) et 21c (C) isolés d'échantillons de glace prélevés en Arctique au large de la côte est du Groenland. La barre représente 100 nm (d'après Boriss *et al.*, 2003).

## 6. Les virus des sources chaudes terrestres acides

### a. La diversité globale

Sur la base du seul critère morphologique, la diversité virale observée au niveau des sources chaudes terrestres paraît bien supérieure à celle des écosystèmes aquatiques mésophiles. En effet, la morphologie des particules observées varie du citron au bâtonnet en passant par la cuillère, la fermeture éclair et la structure tête-queue (cf Figure 11) (Rice *et al.*, 2001 ; Rachel *et al.*, 2002).

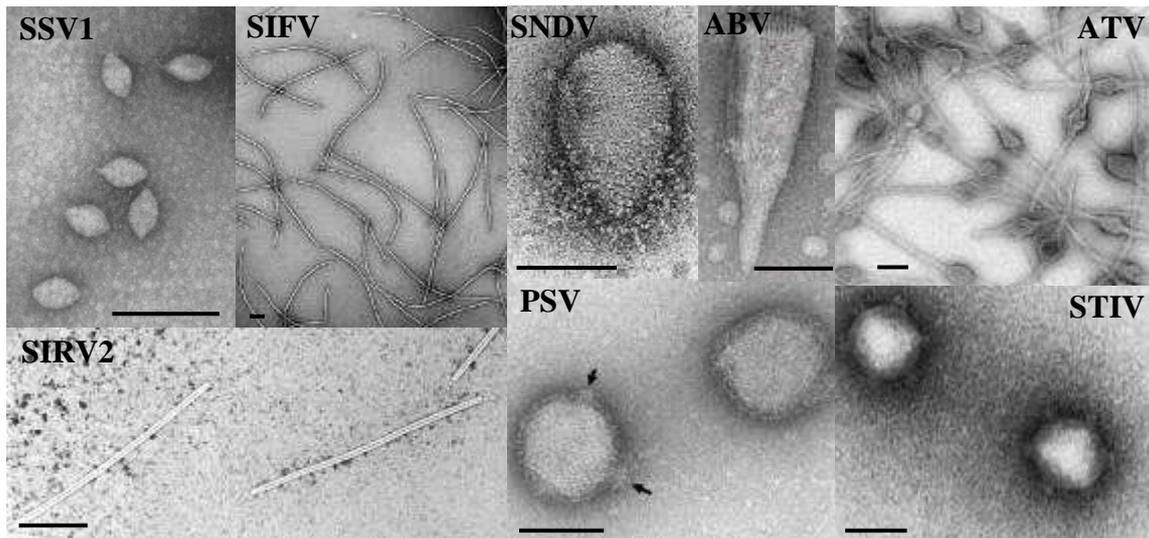
Le travail pionnier mené dans le laboratoire de W. Zillig depuis les années 1980 a donné lieu à la description d'une myriade de virus infectant uniquement des archéobactéries thermophiles terrestres (cf Figure 10 et Tableau 3). Ces derniers ont des morphologies extrêmement variées mais également des génomes très différents nécessitant la création de 7 familles totalement nouvelles : les *Fuselloviridae* (en forme de fuseau), les *Lipothrixviridae* (filamenteux), les *Rudividae* (en forme de bâtonnet), les *Guttaviridae* (en forme de goutte d'eau), les *Globuloviridae* (sphérique), les *Ampullaviridae* (en forme de bouteille) et les *Bicaudaviridae* (en forme de fuseau avec des extensions aux deux pôles) (Prangishvili & Garrett, 2005).

Deux virus sont cependant orphelins étant donné leurs caractéristiques tout à fait originales. Le premier, nommé STIV (pour « *Sulfolobus Turrests and Icosahedral Virus* »), possède une capsidie icosaédrique ayant la même organisation structurale que celles du bactériophage PRD1, du virus d'algue PBCV-1 et de certains *Adenovirus* humains (Rice *et al.*, 2004). Le deuxième, nommé STSV1 (pour *Sulfolobus tengchongensis Spindle Shaped Virus* »), est le plus grand virus en forme de fuseau jamais isolé car il fait 230 nm de long et 107 nm de large avec une queue atteignant en moyenne 68 nm (Xiang *et al.*, 2005).

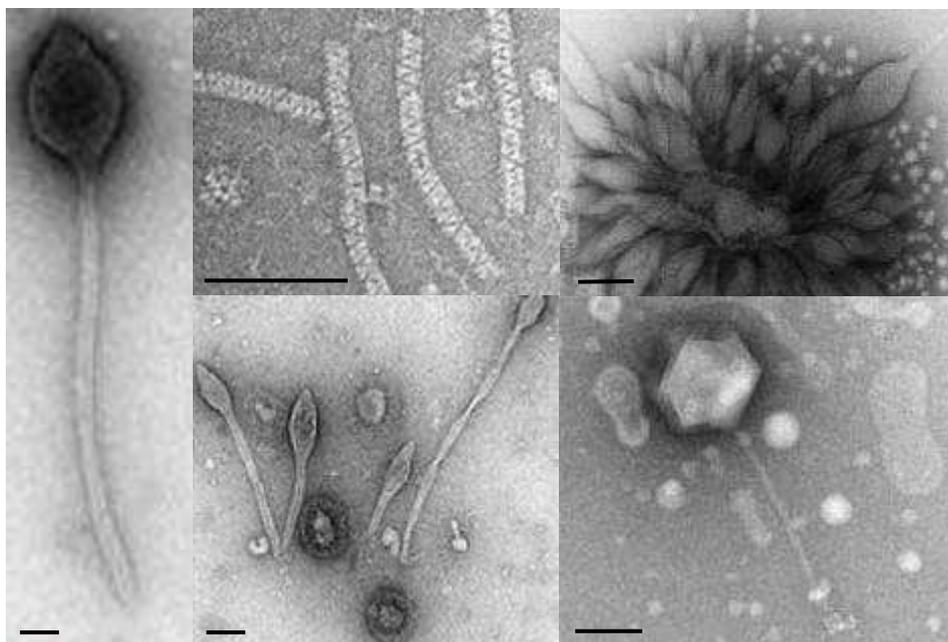
### b. Les virus modèles

#### ➤ SSV1, le *Fusellovirus*

Les membres de la famille des *Fuselloviridae* ont la caractéristique commune d'avoir une forme de fuseau d'une taille moyenne de 100 nm x 60 nm (cf Figure 10). Leur ADN est bicaténaire circulaire, surenroulé positivement et associé à une protéine. Le virion consiste en une enveloppe protéique hydrophobe possédant des fibres caudales à un des pôles. Les *Fusellovirus* sont tous des virus tempérés qui infectent des archéobactéries du phylum des *Crenarchaeota*. Il s'agit le plus souvent du genre *Sulfolobus* mais aussi d'une espèce d'*Acidianus*. Ils ont été isolés de sources hydrothermales terrestres ainsi que de solfatares au Japon, en Islande, en Amérique du Nord et en Russie (Zillig *et al.*, 1998 ; Prangishvili & Garrett, 2005).



**Figure 10 :** Photographie en MET de différents virus isolés de sources chaudes acides terrestres et représentant les principales morphologies rencontrées. La barre représente 100 nm (d'après Martin *et al.*, 1984 ; Arnold *et al.*, 2000a ; Prangishvili *et al.*, 1999 ; Arnold *et al.*, 2000b ; Häring *et al.*, 2004 ; Häring *et al.*, 2005b, 2005c ; Rice *et al.*, 2001).



**Figure 11 :** Photographie en MET de différentes particules de type viral observées au niveau de sources chaudes acides terrestres et possédant des morphologies originales par rapport aux virus décrits. La barre représente 100 nm (d'après Rice *et al.*, 2001 ; Rachel *et al.*, 2002).

Introduction Bibliographique  
« Les Virus de Procaryotes »

**Tableau 3 :** Principales caractéristiques des virus décrits isolés de sources chaudes acides terrestres.

Nom du virus	Morphologie et taille du virion	Type d'acide nucléique	Taille du génome	Famille attribuée	Relation à l'hôte	Références
SSV1	Fuseau 100 nm x 60 nm	ADN double brin circulaire	15,5 kpb	<i>Fuselloviridae</i>	Tempéré	Martin <i>et al.</i> , 1984
SSV2	Fuseau 80 nm x 55 nm	ADN double brin circulaire	14,8 kpb	<i>Fuselloviridae</i>	Tempéré	Prangishvili, 2003
SSV3	Fuseau 80 nm x 55 nm	ADN double brin circulaire	15 kpb	<i>Fuselloviridae</i>	Tempéré	Prangishvili, 2003
SSVK1	Fuseau 90 nm x 60 nm	ADN double brin circulaire	17,4 kpb	<i>Fuselloviridae</i>	Tempéré	Wiedenheft <i>et al.</i> , 2004
SSVRH	Fuseau 90 nm x 60 nm	ADN double brin circulaire	16,5 kpb	<i>Fuselloviridae</i>	Tempéré	Wiedenheft <i>et al.</i> , 2004
SIRV1	Bâtonnet rigide 830 nm x 22 nm	ADN double brin linéaire	32,3 kpb	<i>Rudiviridae</i>	Tempéré	Prangishvili <i>et al.</i> , 1999
SIRV2	Bâtonnet rigide 900 nm x 23 nm	ADN double brin linéaire	35,4 kpb	<i>Rudiviridae</i>	Etat porteur	Prangishvili <i>et al.</i> , 1999
TTV4	Bâtonnet rigide 500 nm x 35 nm	ADN double brin linéaire	17 kpb	<i>Rudiviridae</i>	Lytique	Zillig <i>et al.</i> , 1996
ARV1	Bâtonnet rigide 610 nm x 22 nm	ADN double brin linéaire	24,7 kpb	<i>Rudiviridae</i>	Etat porteur	Vestergaard <i>et al.</i> , 2005
TTV1	Filament flexueux 400 nm x 40 nm	ADN double brin linéaire	16 kpb	$\alpha$ <i>Lipothrixviridae</i>	Tempéré	Janekovic <i>et al.</i> , 1983
SIFV	Filament flexueux 2000 nm x 24 nm	ADN double brin linéaire	40,8 kpb	$\beta$ <i>Lipothrixviridae</i>	Etat porteur	Arnold <i>et al.</i> , 2000a
TTV2	Filament flexueux 1200 nm x 20 nm	ADN double brin linéaire	16 kpb	$\beta$ <i>Lipothrixviridae</i>	Tempéré	Janekovic <i>et al.</i> , 1983
TTV3	Filament flexueux 2500 nm x 30 nm	ADN double brin linéaire	27 kpb	$\beta$ <i>Lipothrixviridae</i>	nd	Janekovic <i>et al.</i> , 1983
AFV1	Filament flexueux 900 nm x 24 nm	ADN double brin linéaire	21 kpb	$\gamma$ <i>Lipothrixviridae</i>	Etat porteur	Bettstetter <i>et al.</i> , 2003
AFV2	Filament flexueux 1100 nm x 24 nm	ADN double brin linéaire	31,8 kpb	$\delta$ <i>Lipothrixviridae</i>	Etat porteur	Häring <i>et al.</i> , 2005a
DAFV	Filament flexueux 2200 nm x 27 nm	ADN double brin linéaire	56 kpb	<i>Lipothrixviridae</i>	nd	Zillig <i>et al.</i> , 1994
SNDV	Goutte d'eau 145 nm x 80 nm	ADN double brin circulaire	20 kpb	<i>Guttaviridae</i>	Etat porteur	Arnold <i>et al.</i> , 2000b
PSV	Sphérique 100 nm	ADN double brin linéaire	28,3 kpb	<i>Globuloviridae</i>	Etat porteur	Häring <i>et al.</i> , 2004
TTSV1	Sphérique	ADN double brin linéaire	20,9 kpb	<i>Globuloviridae</i>	nd	Prangishvili & Garret, 2005
ABV	Bouteille 230 nm x 80 nm	ADN double brin linéaire	23,9 kpb	<i>Ampullaviridae</i>	Etat porteur	Häring <i>et al.</i> , 2005b
ATV	Fuseau bicaudé 1000 nm x 85 nm	ADN double brin circulaire	62,7 kpb	<i>Bicaudaviridae</i>	Tempéré	Häring <i>et al.</i> , 2005c
STIV	Icosaédrique 52 nm	ADN double brin circulaire	17,7 kpb	nd	nd	Rice <i>et al.</i> , 2004
STSV1	Fuseau 230 nm x 107 nm	ADN double brin circulaire	75,3 kpb	nd	Etat porteur	Xiang <i>et al.</i> , 2005

nd : non déterminé

Le virus type de cette famille est SSV1 (pour « *Sulfolobus Spindle-shaped Virus 1* »). C'est le premier isolé (dès 1984) et le mieux décrit. Il infecte la souche B12 de *Sulfolobus solfataricus* isolée au Japon. Son génome fait 15,5 kpb. La production virale est induite par irradiations aux rayons UV ou par traitement à la mitomycine C (Martin *et al.*, 1984).

SSV1, lorsqu'il est dans sa phase lysogène, possède une copie de son génome sous forme plasmidique libre et une autre intégrée. L'intégration se fait niveau d'un ARNt arginine (Reiter *et al.*, 1990) par un processus de recombinaison site spécifique (Muskhelishvili *et al.*, 1993). L'intégrase, encodée par le virus, est également active *in vitro* car elle permet la recombinaison par intégration/excision de molécules d'ADN linéaire (Muskhelishvili., 1994).

L'ADN génomique de SSV1 présente la propriété d'être surenroulé positivement, contrairement à celui des cellules et virus eucaryotes et bactériens. Ceci est dû à la présence chez *Sulfolobus* d'un gène codant pour une topoisomérase ATP-dépendante nommée reverse gyrase. Celle-ci est capable *in vitro* de convertir l'ADN surenroulé négativement ou relâché en un ADN surenroulé positivement (Nadal *et al.*, 1986).

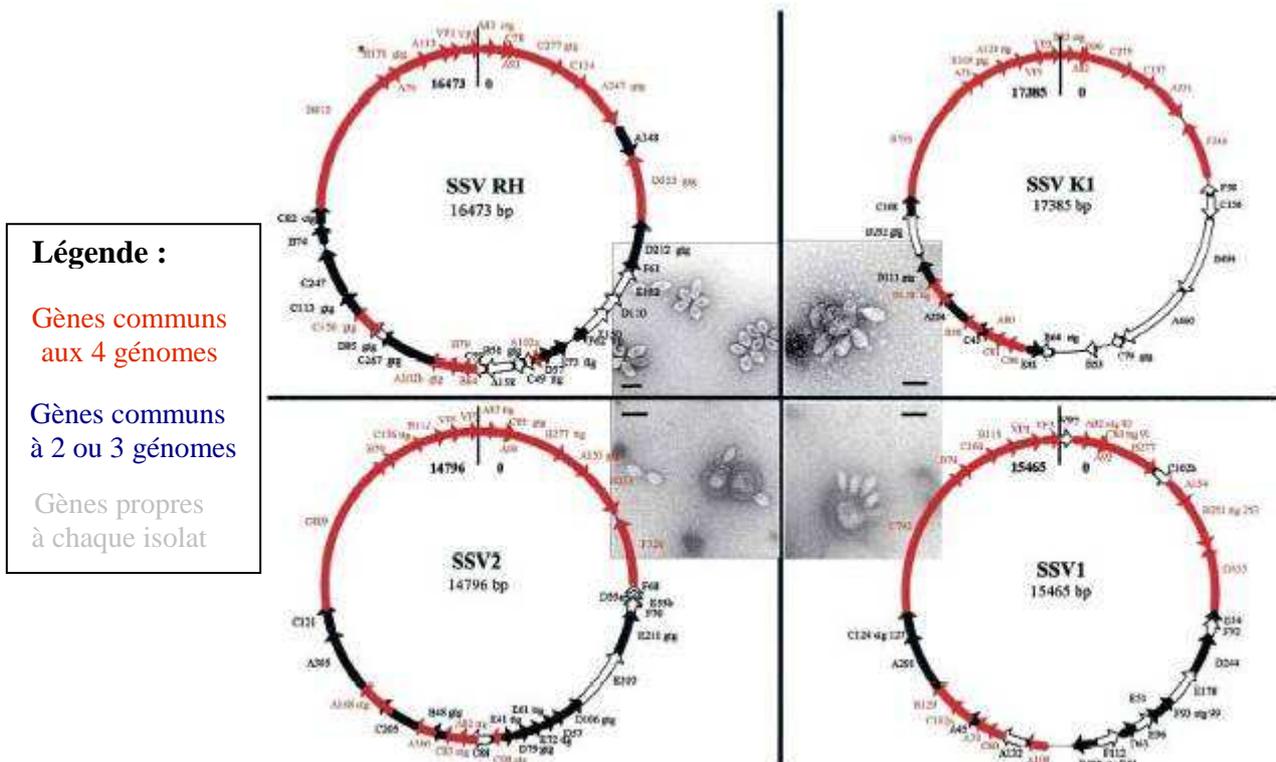
La cartographie des transcrits de SSV1 met en évidence 9 ARNm différents qui sont exprimés en continu. Après induction, deux petits ARNm supplémentaires sont observés (Reiter *et al.*, 1987 ; Palm *et al.*, 1991). En effet, l'induction inactive le répresseur cellulaire d'une petite unité de transcription qui est située entre deux autres unités plus longues et inversées. Dans ce cas, ces dernières se retrouvent en contact ce qui entraîne la réplication de l'ADN viral et permet la formation des virions (Zillig *et al.*, 1996).

Le génome de SSV1 a été entièrement séquencé en 1991 et 34 gènes ont été mis en évidence (Palm *et al.*, 1991). Les seuls gènes correspondant à des fonctions connues sont l'intégrase, les deux protéines de capsid (VP1 et VP3) et la protéine associée à l'ADN (VP2) (Reiter *et al.*, 1987). Quatre gènes (c792, b115, b129 et e96) se sont révélés être indispensables au virus mais leur fonction demeure inconnue (Stedman *et al.*, 1999) et deux autres ont été retrouvés chez des plasmides de la famille pRN (Keeling *et al.*, 1998). Chez ces derniers, ces deux gènes permettent l'encapsidation du plasmide dans de petites particules virales (Arnold *et al.*, 1999).

Une analyse de la composition en acides aminés des différentes protéines virales a permis de mettre en évidence deux catégories bien distinctes. Ainsi, 15 gènes codent pour des protéines comportant des résidus cystéine et sont quasiment tous regroupés (sauf 1) dans une même région qui couvre 6 849 pb. A l'inverse, l'autre portion du génome contient des gènes codant uniquement pour des protéines sans cystéine qui sont notamment VP1, VP2 et VP3. Cette partie du génome pourrait

être impliquée dans l'assemblage du virus et la région « cystéinée » pourrait assurer les fonctions précoces telles que la multiplication virale ainsi que la mise en place et/ou le maintien de la lysogénie. Cette organisation originale fait penser à une fusion de deux génomes à l'évolution différente (Plam *et al.*, 1991).

Lorsque la séquence de SSV1 est comparée à celle d'autres *Fusellovirus* (SSV2 isolé d'Islande, SSVK1 isolé de Russie et SSVRH isolé d'Amérique du nord) il apparaît que l'organisation générale est très similaire (cf Figure 12). Une moitié des gènes semble être très conservée chez les quatre virus, tandis que l'autre moitié ne l'est pas. Il existe aussi des homologies deux à deux ainsi que des gènes propres à chaque virus. Cette diversité génétique serait en faveur de l'existence de différentes souches de *Fusellovirus*. Plusieurs explications sont avancées : isolement géographique, évolution génétique rapide ou pression sélective des caractéristiques physico-chimiques de l'environnement (Stedman *et al.*, 2003 ; Wiedenheft *et al.*, 2004).



**Figure 12 :** Comparaison du génome de SSV1 avec trois autres *Fusellovirus* : SSV2, SSVRH et SSVK1. La barre représente 100 nm sur les photographies en MET des quatre isolats sont positionnées au milieu de la figure (d'après Wiedenheft *et al.*, 2004).

Le gène a291 est homologue au gène 29 de SIFV, un virus archéen de la famille des *Lipothrixviridae* et le gène a45 est homologue au gène 59a du *Lipothrixvirus* AFV1 (Arnold *et al.*, 2000 ; Bettstetter *et al.*, 2003). Deux autres gènes présentent des similarités de séquences avec le génome du *Rudivirus* ARV1 (Vestergaard *et al.*, 2005). Le reste du génome semble très différent des autres virus d'archées et de bactéries (Wiedenheft *et al.*, 2004).

La transcription chez les archées a pu être mieux étudiée grâce à SSV1. En effet, ce virus possède des « TATA-box » ressemblant plus aux promoteurs reconnus par les ARN polymérases eucaryotes que par celle des bactéries. Ceci corrobore les résultats obtenus sur la composition de la sous-unité de l'ARN polymérase archéenne et a permis de prévoir l'existence de « TATA-binding proteins » (TBP) chez les archées. Ces promoteurs ont ensuite été utilisés pour étudier la transcription *in vitro* de l'ARN polymérase archéenne. SSV1 est également à l'origine de deux systèmes de transformation chez *Sulfolobus*. Le premier utilise l'origine putative de réplication du virus et le second est constitué du génome viral entier (Prangishvili, 2001).

➤ SIFV, le *Betalipothrixvirus*

Les membres des *Lipothrixviridae* sont de longs bâtonnets flexueux enveloppés (de 400 nm à 2 000 nm) refermant une molécule d'ADN bicaténaire linéaire (cf Figure 10). Ils infectent quatre genres différents de *Crenarchaeota* : *Thermoproteus*, *Sulfolobus*, *Desulfurolobus* et *Acidianus*. Sur la base des différences concernant les enveloppes, les structures terminales, les propriétés génomiques et les stratégies de réplication, quatre genres ont été créés au sein de cette famille : les *Alphalipothrixvirus*, les *Betalipothrixvirus*, les *Gammalipothrixvirus* et les *Detalipothrixvirus* (Prangishvili & Garret, 2005).

SIFV (pour « *Sulfolobus islandicus Filamentous Virus* ») infecte l'espèce *Sulfolobus islandicus*. C'est un des virus archéens les plus longs car il mesure près de 2 000 nm pour une largeur de 24 nm. L'enveloppe est constituée des phospholipides qui proviennent majoritairement de l'hôte mais un serait spécifique du virus. Ainsi, l'équipement enzymatique de SIFV serait peut-être capable de modifier les lipides de la cellule. Les extrémités de l'enveloppe sont couronnées par 6 fibres protéiques qui peuvent adopter deux positions : ouverte ou fermée. Le modèle proposé pour la structure du core est atypique : l'ADN et les protéines associées formeraient deux hélices et non une seule comme cela existe chez TTV1, un autre *Lipothrixvirus* infectant l'espèce *Thermoproteus tenax* (Arnold *et al.*, 2000a).

SIFV n'entraîne pas la lyse de la cellule mais un retard de croissance. Les virions sont relargués dans le milieu extérieur 6 à 8 h après le début de l'infection. Cependant, la souche peut être

curée par des repiquages successifs. Aucune copie intégrée du génome viral n'a pu être mise en évidence chez son hôte. SIFV est donc dans un état porteur instable (Arnold *et al.*, 2000a).

Son génome, qui fait environ 41 kpb, est presque entièrement séquencé. Seules les extrémités n'ont pas pu être clonées à cause de leur inaccessibilité enzymatique. Soixante quatorze gènes ont été prédits et présentent une bonne couverture de la séquence. Les gènes 34 et 35 codent respectivement pour les protéines de structure HP1 et HP2. Ils seraient transcrits sur un même ARNm polycistronique. Les gènes 7 et 22 comportent des motifs hélicases ATP-dépendantes de type archéen. Trois gènes (41, 48 et 63) présentent 50 % de similarité entre eux et possèdent un motif typique du groupe I des glycosyl transférases. Ils pourraient donc être impliqués dans la synthèse de l'enveloppe lipidique du virus (Arnold *et al.*, 2000a).

Aucune homologie n'a été mise en évidence entre SIFV et TTV1. Par ailleurs, des hybridations ADN/ADN entre SIFV et DAFV (un *Lipothrixvirus* infectant le genre *Desulfurolobus*) donnent une réaction croisée significative. Le gène 29 code pour une protéine homologue à une protéine inconnue de SSV1 (Arnold *et al.*, 2000a). Il a également été montré que SIFV possédait des gènes homologues à des gènes portés par trois *Rudivirus* (SIRV1, SIRV2 et ARV1), un *Gammalipothrixvirus* (AFV1) et un *Deltalipothrixvirus* (AFV2) (Peng *et al.*, 2001 ; Vestergaard *et al.*, 2005 ; Bettstetter *et al.*, 2003 ; Häring *et al.*, 2005a).

#### ➤ SIRV 1 et SIRV2, les *Rudivirus*

Les deux premiers virus décrits appartenant à la famille des *Rudiviridae* sont SIRV1 et SIRV2 (pour « *Sulfolobus islandicus Rod-shaped Virus 1 and 2* »), qui infectent les *Crenarchaeota* du genre *Sulfolobus*. Ils ont une forme de bâtonnets rigides assez longs (830 nm pour SIRV1 et 900 nm pour SIRV2) et ne possèdent pas d'enveloppe (cf Figure 10). A chaque extrémité, la particule possède trois fibres caudales. Le core possède une cavité centrale qui renferme la molécule d'ADN bicaténaire linéaire covalamment liée et associée à une protéine de 15,8 kDa (Prangishvili *et al.*, 1999).

SIRV1 et SIRV2 n'entraînent pas la lyse de leur hôte mais un retard de croissance. Leurs génomes ne sont pas intégrés et la production virale n'est inductible ni par irradiation aux rayons UV, ni par traitement à la mitomycine C. La cellule ne peut être curée par repiquages successifs. Ils sont donc dans un état porteur stable. Par ailleurs, SIRV1 est capable d'induire la production de SSV1 si ce dernier est intégré chez *Sulfolobus solfataricus* (Prangishvili *et al.*, 1999).

Les génomes de SIRV1 et SIRV2 ont été entièrement séquencés en 2001 et font respectivement 32,3 et 35,5 kpb. Ils sont organisés en groupes de séquences très conservées séparés par des séquences non conservées. Ils comportent à leurs extrémités d'importantes régions répétées inversées. Les phénomènes responsables de l'évolution de ces deux génomes sont variés : recombinaison, duplication de gène, transfert horizontal et substitution de gènes (Peng *et al.*, 2001 ; Blum *et al.*, 2001).

Récemment, des variants de SIRV1 ont été mis en évidence. Ils ont des génomes très proches avec tout de même des différences au niveau de l'ordre et la taille des gènes. Un des mécanismes responsables des réarrangements génétiques observés chez les variants est nouveau car il serait médié par un intron archéen de 12 kpb. Il existe également des recombinaisons entre les variants (Peng *et al.*, 2004).

Chez SIRV1, 45 gènes ont été prédits alors que chez SIRV2 ils sont au nombre de 54. Une analyse comparative des deux génomes a permis de déterminer que 44 virus sont communs aux deux virus. De plus, un tiers des gènes de chacun de ces virus possède un homologue chez SIFV. Il a également été mis en évidence des similarités de séquences entre SIRV1, SIRV2, ARV1 (un autre *Rudivirus*), AFV1 (un *Gammalipothrixvirus*) et AFV2 (un *Deltalipothrixvirus*). Ceci laisse à penser que les membres de deux familles, les *Lipothrixviridae* et les *Rudiviridae*, pourraient appartenir à une même superfamille (Peng *et al.*, 2001 ; Vestergaard *et al.*, 2005 ; Bettstetter *et al.*, 2003 ; Häring *et al.*, 2005a).

La fonction de certains gènes a pu être assignée. Ainsi, une ADN polymérase, une dUTPase et une résolvasse de jonction Holliday ont été prédites chez SIRV1 et SIRV2. Le rôle de la résolvasse serait de permettre la résolution des intermédiaires de réplication qui sont en fait des concatémères de chaque brin. Ce phénomène fait penser au mode de réplication qui existe chez certains virus eucaryotes comme les *Poxvirus*. De plus, le génome de ces derniers possède le même type d'organisation (Peng *et al.*, 2001 ; Birkenbihl *et al.*, 2001).

Chez SIRV1 et SIRV2, la transcription démarre en plusieurs endroits sur toute la longueur du génome et sur les deux brins. Les premiers transcrits sont détectables 30 min après le début de l'infection. A ce stade, tous les gènes sont transcrits sauf un et beaucoup le sont sous forme d'ARNm polycistroniques (cf Figure 13). Bien que ce soit le cas pour le gène codant la protéine de capsid, un transcrit monocistronique très abondant est détecté juste avant l'assemblage des virions. De plus, l'expression d'un gène, adjacent à celui-ci, est régulée positivement au cours de la phase tardive. Ceci suggère qu'il pourrait être impliqué dans la régulation du gène codant pour la protéine capsidiale (Kessler *et al.*, 2004).

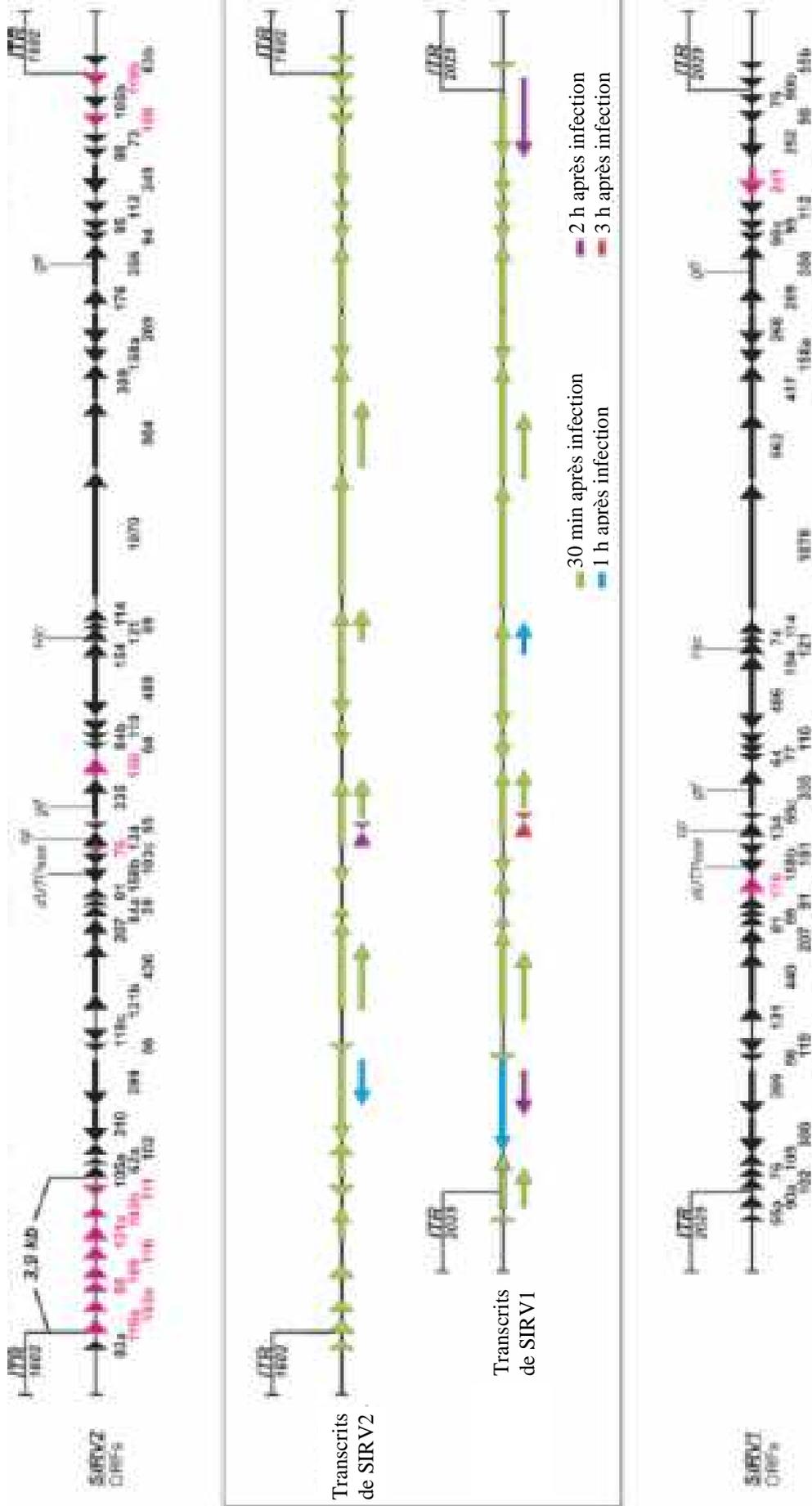


Figure 13: Carte des transcrits des virus SIRV1 et SIRV2 (d'après Kessler *et al.*, 2004).

L'extension d'amorce a permis de déterminer les sites de démarrage de 13 transcrits et les séquences promotrices ont été identifiées. Ces dernières, comme celles de l'hôte, comportent des sites de fixation pour des facteurs de transcription archéens. Cependant, la plupart contiennent aussi un élément consensus spécifique au virus. Ceci suggère l'implication de facteurs de transcription alternatifs (Kessler *et al.*, 2004).

➤ SNDV, le *Guttavirus*

A l'heure actuelle cette famille ne comprend qu'un seul représentant, SNDV (pour « *Sulfolobus neozelandicus Droplet-shaped Virus* »), qui infecte une souche de l'espèce *Sulfolobus neozelandicus*. La particule mesure 110 à 185 nm de long et de 70 à 95 nm de large. Elle a une forme de goutte d'eau avec une touffe de fines fibres caudales à l'extrémité pointue (cf Figure 10). Ces dernières sont peut-être impliquées dans l'adsorption à la S-layer de l'hôte. Le core est recouvert d'une enveloppe dont la structure rappelle la carapace annelée de certains insectes. Trois protéines ont pu être mises en évidence (13, 13,5 et 17,5 kDa) et sont certainement les constituants des fibres, du core et de l'enveloppe (Arnold *et al.*, 2000b).

SNDV n'est pas lytique sauf si la charge virale initiale est trop importante. La production de particules commence 17 h après le début de l'infection. En général, le virus entraîne un retard de croissance chez son hôte qui peut être curé par des repiquages successifs. SNDV est donc dans un état porteur instable. Aucun provirus n'a été détecté dans le chromosome archéen (Arnold *et al.*, 2000b).

Le génome de SNDV est une molécule d'ADN double brin circulaire de 20 kpb qui est *dam*-méthylée. Cette dernière caractéristique est assez exceptionnelle chez les virus de *Crenarchaeota*. Par ailleurs, il a été mis en évidence que le chromosome de l'hôte n'est pas modifié. Ainsi, une méthylase serait codée par le virus lui-même et reconnaîtrait préférentiellement l'ADN hémiméthylé, c'est-à-dire les intermédiaires de réplication du génome viral (Arnold *et al.*, 2000b).

➤ PSV, le *Globulovirus*

La famille des *Globuloviridae* a récemment été proposée à l'ICTV. Son représentant modèle, PSV (pour « *Pyrobaculum Spherical Virus* »), infecte les *Crenarchaeota* des genres *Pyrobaculum* et *Thermoproteus*. Les virions sont sphériques avec un diamètre de 100 nm et plusieurs portent à leur surface de petites protubérances de 15 nm de large (cf Figure 10). Des observations au MET et des analyses de Fourier permettent de penser que l'acide nucléique et les protéines associées sont compactés en une superhélice, elle-même contenue dans une enveloppe. Cette structure est tout à fait inédite dans le monde des virus à ADN (Häring *et al.*, 2004).

Cinq protéines, de 16, 20, 33, 70 et 110 kDa, ont été mises en évidence par électrophorèse en SDS-PAGE. Les trois plus grosses possèdent la même séquence N-terminale et seraient en fait respectivement le monomère, le dimère et le trimère d'une molécule dont la taille réelle est 28 kDa. Cette protéine ne possède pas de résidu cystéine et les deux tiers des acides aminés sont hydrophobes. Ceci contribue à son extrême thermostabilité. Par ailleurs, l'étude des lipides du virus fait apparaître des molécules qui ne sont pas retrouvés chez l'hôte. Ainsi, le virus coderait pour des enzymes capables de modifier les lipides cellulaires (Häring *et al.*, 2004).

PSV n'entraîne pas la lyse de son hôte et des repiquages successifs ne permettent pas de l'éliminer. Ainsi, il serait dans un état porteur comme c'est le cas pour de nombreux virus archéens. (Häring *et al.*, 2004).

Le génome de PSV consiste en une molécule d'ADN double brin linéaire covalamment fermée de 28,3 kpb. L'organisation des 48 gènes prédits est remarquable car ils se situent presque tous sur un seul brin. Ceci rappelle ce qui est observé chez les bactériophages de la famille des *Plasmaviridae*. Les extrémités du génome comportent de longues séquences inversées répétées, comme chez SIRV1 et SIRV2. PSV présente des homologies uniquement avec un de ses hôtes, *Pyrobaculum aerophilum*, mais pas avec les autres virus (Häring *et al.*, 2004).

#### ➤ ABV, l'*Ampullavirus*

La famille des *Ampullaviridae*, comme celle des *Globuloviridae*, doit être soumise à l'approbation de l'ICTV. Elle ne compte pour le moment qu'un seul virus nommé ABV (pour « *Acidianus Bottle-shaped Virus* »). Ce dernier infecte uniquement l'espèce *Acidianus convivator* et entraîne un retard de croissance chez son hôte (Häring *et al.*, 2005b).

Les virions ont une morphologie jamais observée jusqu'alors dans le monde viral. En effet, ils ont une forme de bouteille de 230 nm de long. La largeur des particules varie de 4 nm pour l'extrémité pointue à 75 nm pour l'extrémité opposée. Au niveau de cette dernière, des observations en tomographie ont révélé la présence de fins filaments longs de 20 nm. Ils sont disposés en cercle et chaque virion en possède une vingtaine (cf Figure 10). Ils semblent être impliqués dans la fixation entre virus (Häring *et al.*, 2005b).

Des analyses plus poussées ont permis de déterminer qu'il s'agit en fait d'un « bouchon inséré dans une bouteille sans col ». En réalité, le bouchon est la partie sur laquelle s'enroulent les filaments de nucléoprotéine. Il est certainement impliqué dans l'adsorption du virus et l'injection de l'acide nucléique viral. La bouteille sans col correspond en fait à une enveloppe externe qui recouvre partiellement le bouchon (Häring *et al.*, 2005b).

L'analyse en SDS-PAGE des protéines d'ABV révèle un profil complexe où se distinguent six bandes majoritaires de 15, 22, 65, 66, 70 et 80 kDa. Les études préliminaires menées sur le génome viral montrent qu'il s'agit d'un ADN double brin linéaire de 23,9 kpb (Häring *et al.*, 2005b).

➤ ATV, le *Bicaudavirus*

Les *Bicaudaviridae* correspondent à la dernière famille décrite chez les *Crenarchaeota* et son unique représentant est appelé ATV. Ce virus a été isolé d'une source chaude et acide à Puzzoli en Italie. Il est tempéré et infecte l'archéobactérie de l'espèce *Acidianus convivator*. Ce virus est capable de se développer à l'extérieur de son hôte. En effet, juste après sa sortie de la cellule, la particule a une forme de citron, comme SSV1. En faisant un suivi au cours du temps, il a été mis en évidence l'extension d'une queue à chaque pôle du virion (cf Figure 10). Ce phénomène ne nécessite la présence ni de l'hôte, ni d'une source d'énergie exogène, ni de cofacteurs. Cependant, il n'est observable qu'aux températures physiologiques de l'hôte, c'est-à-dire entre 75°C et 90°C. A 75°C, la particule est complète au bout de 7 jours alors qu'à 90°C il suffit d'une heure (Häring *et al.*, 2005c).

Le génome d'ATV est un ADN double brin circulaire de 62 kpb. L'annotation de ce dernier a permis la mise en évidence de 72 gènes. Environ 40 % ont des homologues dans les bases de données. Ainsi, une intégrase, une nucléase, trois transposases et deux ATPases putatives ont pu être identifiées. De plus, une des neuf protéines structurales identifiées, composée de 800 résidus, pourrait être responsable de la protrusion des queues. En effet, elle possède un motif répété « coiled-coiled » qui est similaire à celui retrouvé chez des protéines pouvant s'assembler en filaments, sans cofacteurs ni source d'énergie, dans les cellules eucaryotes et bactériennes. Après surexpression et purification, il apparaît que cette protéine de 800 acides aminés produit de longs filaments de 2 nm de large ayant des caractéristiques communes à ceux observés au niveau des queues du virion (Häring *et al.*, 2005c).

Ainsi, ATV est, par bien des aspects, un virus hors du commun. C'est le seul virus hyperthermophile qui entraîne la lyse de son hôte. Au contraire de ses comparses, il ne semble pas « craindre » d'être en contact avec le milieu extérieur. Par ailleurs, le fait de développer des extensions uniquement à la température où son hôte est actif peut refléter une stratégie du virus pour multiplier les chances de contact avec son hôte dont la densité cellulaire est faible (Häring *et al.*, 2005c).

## 7. Les virus des sources hydrothermales marines

### a. La diversité globale

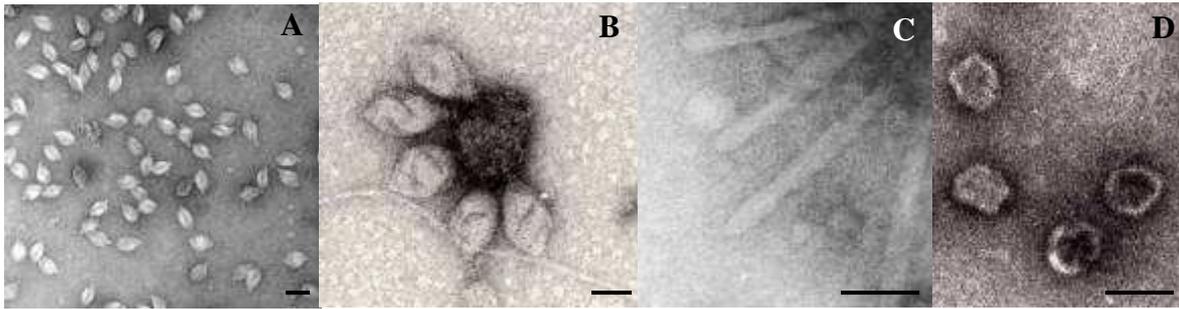
La seule étude traitant de la diversité virale retrouvée au niveau des sources hydrothermales marines a été réalisée dans notre laboratoire. Elle sera développée ultérieurement car elle constitue une partie du travail de thèse présenté dans ce manuscrit. Par ailleurs, une étude récente a été menée sur l'abondance des virus dans ce type d'environnement. Le titre obtenu à partir d'échantillons provenant de trois sites actifs au large de la côte ouest du Canada variait de  $1,45 \cdot 10^5$  à  $9,9 \cdot 10^7 \cdot \text{mL}^{-1}$ . Il apparaît que ces valeurs sont supérieures à celles mesurées dans les océans profonds voisins des sources chaudes (Ortmann & Suttle, 2005).

### b. PAV1, le virus modèle

Hormis la particule de type viral qui a été observée chez l'espèce archéenne *Pyrococcus woesei* (Zillig *et al.*, 1987), un seul virus hyperthermophile marin a été décrit. Il s'agit de PAV1 qui infecte la souche GE23 de l'espèce *Pyrococcus abyssi*. Les virions ont une forme de citron (120 nm x 80 nm) et présentent de petites fibres caudales à un des pôles. Ils peuvent s'allonger à l'extrême comme une baguette, ce qui témoigne de leur forte plasticité. D'après différents tests réalisés notamment avec des solvants organiques, la particule posséderait une enveloppe de nature protéolipidique (cf Figure 14). De plus, des analyses en SDS-PAGE ont permis de mettre en évidence trois protéines majoritaires de 6, 13 et 36 kDa (Geslin *et al.*, 2003).

PAV1 est dans un état porteur stable. En effet, il n'entraîne pas la lyse de son hôte mais il persiste malgré des repiquages successifs. La production de particules est faible (1 000 fois moins que pour SSV1) et n'est pas inductible. Son infectiosité vis-à-vis d'autres souches n'a pas pu être démontrée (Geslin *et al.*, 2003).

Le génome de PAV1 est un ADN double brin circulaire de 18,1 kpb. Il ne s'intègre pas au chromosome archéen mais il est présent dans la cellule sous forme plasmidique. Cette dernière correspond en fait au plasmide pGN23 qui avait été mis en évidence par N. Benbouzid-Rollet et ses collaborateurs 6 ans plus tôt (Benbouzid-Rollet *et al.*, 1997 ; Geslin *et al.*, 2003).



**Figure 14** : Photographies en MET de PAV1 : après purification (A), attaché à un débris cellulaire (B), allongé à l'extrême (C), et après traitement au chloroforme (D). La barre représente 100 nm (d'après Geslin *et al.*, 2003).

## IV. Pourquoi étudier les virus de procaryotes ?

Les virus de procaryotes sont parmi les plus petits constituants de la biosphère, cependant ils constituent le groupe d'organismes le plus abondant de la planète (Casjens, 2005). Leur importance au niveau environnemental ainsi que leurs possibles applications biotechnologiques et thérapeutiques sont en passe de les propulser sur le devant de la scène.

### 1. Pour un enjeu en biodiversité

Le nombre d'espèces microbiennes est évalué à environ dix millions. En extrapolant, si chacun de ces microbes est l'hôte d'au moins dix virus différents, alors la richesse des virus de procaryotes serait immense car l'estimation du nombre de spécimens s'élèverait alors à cent millions (Rohwer, 2003). Ce chiffre a été revu à la baisse car les virus semblent se déplacer d'un environnement à l'autre. Ainsi, la diversité virale pourrait être importante à une échelle réduite alors que d'un point de vue plus global elle serait en fait relativement limitée (Breitbart & Rohwer, 2005).

En tenant compte de ces considérations, il apparaît tout de même évident que la quantité de virus de procaryotes décrits ne représente qu'une infime partie de la population totale (Ackermann, 2001). De nouvelles approches, comme la métagénomique, semblent tout à fait indiquées pour accéder à la diversité des virus incultivés et découvrir des gènes jusqu'alors inconnus (Edwards & Rohwer, 2005).

Si la cultivabilité est un biais non négligeable, le choix très orienté des virus étudiés l'est également. En effet, la plupart des spécimens isolés à l'heure actuelle sont des prédateurs de bactéries pathogènes pour l'homme ou utilisées dans les process agro-alimentaires tels que la fermenta-

tion lactique. Ainsi, les virus de l'environnement semblent avoir été pendant longtemps peu étudiés à cause de leur apparente inutilité (Ackerman, 2001). La tendance s'inverse depuis ces dernières années et annonce l'avènement des virus de procaryotes peuplant des environnements très variés qui sont, pour certains, extrêmes.

## 2. Pour un enjeu en écologie marine

### a. Au niveau de la mortalité microbienne

Plusieurs approches ont été utilisées pour déterminer l'impact des virus vis-à-vis de la mortalité microbienne. Une technique consiste à inférer la mortalité au pourcentage de cellules visiblement infectées. Le biais de cette méthode réside dans le fait que les particules virales ne sont pas observables tout au long du cycle lytique. Ainsi, le nombre de cellules réellement infectées est largement sous-estimé. (Proctor & Fuhrmann, 1990 ; Suttle, 2005b). Une étude a montré qu'en augmentant artificiellement la quantité de virus infectant des producteurs primaires importants (tels les cyanobactéries), la production primaire était considérablement réduite. Ce résultat est donc une preuve indirecte de l'influence des virus sur la mortalité microbienne (Suttle *et al.*, 1990 ; Suttle, 2005b).

Un modèle a été proposé par K. E. Wommack et R. R. Colwell pour expliquer quelle est l'implication de l'infection virale dans le maintien de la diversité microbienne. Beaucoup d'études ont montré que, sur une période donnée, le nombre de virus et de procaryotes est relativement stable. Le modèle propose que l'espèce bactérienne, notée A, a la meilleure croissance. Elle va donc être préférentiellement détruite par le virus A. Lorsque le nombre de cellules de l'espèce A devient trop faible, le virus A s'éteint à son tour. A ce moment, un facteur de sélection va favoriser la croissance d'une autre espèce bactérienne, notée B, qui va à son tour être préférentiellement détruite par le virus B, etc. Le modèle propose donc une succession de cycles croissance/extinction qui ne sont pas observables en étudiant des prélèvements trop éloignés dans le temps (Wommack & Colwell, 2000).

Différentes études ont donc essayé de quantifier la lyse cellulaire imputable aux phages. Ainsi, M. Heldal et G. Bratbak estimaient que les virus causent la lyse de 2 à 24 % de la population bactérienne par heure (Heldal & Bratbak, 1991). Trois ans plus tard, Suttle déclarait que 10 à 20 % de la communauté microbienne sont détruits quotidiennement par les phages (Suttle, 1994). Les diverses études menées sur le sujet fournissent des valeurs extrêmement variables, certainement dues au fait que les prélèvements ont été réalisés dans diverses régions à des périodes différentes.

La tendance générale est tout de même de considérer que la mortalité microbienne imputable aux virus est loin d'être négligeable (Suttle, 2005b).

### **b. Au niveau des cycles biogéochimiques**

Le nombre estimé de procaryotes hébergés dans les océans est de  $1,1 \cdot 10^{29}$ , ce qui représente une importante proportion de la biosphère active (Whitman *et al.*, 1998). Ainsi, leur destruction, notamment par les phages, représente une source non négligeable de carbone organique, de nutriments et d'éléments trace pour les réseaux trophiques marins (Wilhelm & Suttle, 1999).

L'infection lytique aboutit à l'obtention de virions et de débris cellulaires. Ces débris sont constitués de molécules solubles additionnées de colloïdes et de fragments cellulaires dont la plupart font partie de la matière organique dissoute (Shibata *et al.*, 1997 ; Fuhrmann, 1999). Un modèle proposé en 1991 suggère que la lyse imputable aux virus libère approximativement  $1 \mu\text{g.L}^{-1}$  de carbone organique dissous par génération de bactérie (Proctor & Fuhrmann, 1991). Des études corroborant ce modèle ont été menées en 1998 dans le golfe du Mexique et montrent que la quantité de carbone relargué lors de la lyse bactérienne par les virus est de  $0,1$  à  $0,6 \mu\text{g.L}^{-1}.\text{jour}^{-1}$  dans les eaux profondes et de  $0,7$  à  $5,2 \mu\text{g.L}^{-1}.\text{jour}^{-1}$  dans les eaux côtières (Whilhelm *et al.*, 1998 ; Wilhelm & Suttle, 1999).

La plupart des produits de lyse sont à un moment ou à un autre disponibles pour les bactéries (Proctor & Fuhrmann, 1990 ; Middelboe *et al.*, 1996), bien qu'une petite proportion des virus puisse être détruite par les flagellés et que certains débris cellulaires puissent être réfractaires à la dégradation (Gonzalez & Suttle, 1993 ; Fuhrmann, 1999). La consommation rapide par les communautés microbiennes de la matière organique dissoute a deux effets majeurs : l'augmentation de la respiration de ces communautés et la diminution du transfert de carbone vers les niveaux trophiques supérieurs. De plus, la lyse cellulaire converti le carbone organique particulaire en carbone organique dissous ; ce qui a pour conséquence de réduire la vitesse de sédimentation du carbone qui est alors utilisé pour la respiration dans les eaux de surface. Ce phénomène a une grande importance dans le cycle global du carbone car la sédimentation du carbone organique particulaire correspond à un transfert net d'environ 3 Gt de carbone des eaux de surface vers les eaux profondes. Par ailleurs, le taux d'accumulation du  $\text{CO}_2$  dans l'atmosphère est diminué de moitié par rapport à une situation où le carbone organique particulaire ne serait pas du tout transformé en carbone organique dissous (Suttle, 2005b).

Les virus sont aussi très impliqués dans le cycle du phosphore. Ainsi, 1 à 12 % de l'ADN « dissous » dans l'eau de mer est en fait à l'intérieur des virions (Paul *et al.*, 1991). Ainsi, l'ADN

viral représente moins de 1 % du phosphore organique dissous. Cependant, en raison de la rapidité du turnover de l'ADN dans l'eau de mer, l'ADN viral constitue un réservoir important de phosphore organique (Bratbak *et al.*, 1994 ; Wilhelm & Suttle, 1999).

Des nutriments, autres que le carbone, sont relargués lors de la lyse cellulaire causée par les virus. Par exemple, la quantité de fer relargué correspond à la quasi-totalité des besoins des autres organismes (Poorvin *et al.*, 2004). Par ailleurs, la petite taille des virions fait d'eux d'excellents sites pour la minéralisation du fer et peut-être d'autres métaux (Daughney *et al.*, 2004 ; Suttle, 2005b).

### **c. Au niveau des flux de gènes**

Les virus peuvent agir de façon directe grâce à la transduction. C'est un processus au cours duquel l'ADN de la cellule hôte est « capturé » par le virus puis transféré à une autre cellule (Jiang & Paul, 1998). Deux types de transduction existent : celle dite généralisée et celle dite spécialisée. Les virus réalisant la transduction généralisée peuvent être tempérés ou lytiques mais ils possèdent des caractéristiques communes en ce qui concerne la réplication de leur génome et l'emballage de leur ADN. Ils sont capables d'emmener, avec une fréquence égale, n'importe quel ADN contenu dans la cellule, quelle que soit la localisation. A l'inverse, les virus réalisant la transduction spécialisée sont uniquement des virus tempérés capables d'intégration dans le chromosome de leur hôte. Ils sont capables de transférer des gènes cellulaires bien spécifiques situés près de leur site d'intégration (Wommack & Colwell, 2000). La transduction est un des principaux mécanismes de transfert horizontal de gènes dans les environnements marins (Jiang & Paul, 1998). Elle assure le maintien et/ou le contrôle de la diversité génétique des communautés de procaryotes (Wommack & Colwell, 2000).

Généralement, la transduction opère au sein d'une gamme d'hôte très restreinte. Cependant, plusieurs études récentes indiquent que certaines particules de type viral pourraient être capables de transferts horizontaux non spécifiques en milieu marin (Chiura, 1997, 2004 ; Chiura *et al.*, 1999). Il faut également considérer que l'ADN des cellules hôtes, qui est relargué lors de la lyse virale, peut être transféré à d'autres micro-organismes grâce à une transformation naturelle. Une étude a montré que la lyse virale est le phénomène majoritairement responsable de la présence d'ADN dissous dans l'eau de mer (Paul *et al.*, 1993). Le transfert génétique horizontal généralisé et la transformation permettent un brassage génétique parmi une grande variété d'espèces. Ainsi, ces processus peuvent avoir un rôle majeur sur la structure génétique et l'évolution des populations bactériennes marines. (Fuhramn, 1999).

Il existe également des transferts latéraux au cours desquels les virus récupèrent de l'ADN de leur cellule hôte pour en tirer profit. Par exemple, deux gènes impliqués dans la photosynthèse (psbA et psbD) sont retrouvés chez les phages infectant des cyanobactéries des genres *Synechococcus* et *Prochlorococcus*. L'analyse des séquences a mis en évidence qu'un transfert génétique latéral avait eu lieu plusieurs fois entre les virus et leurs hôtes. Les gènes cellulaires de la photosynthèse forment un groupe monophylétique distinct de ceux des phages qui sont devenus viraux à part entière (Mann *et al.*, 2003 ; Hambly & Suttle, 2005 ; Suttle, 2005b).

### **3. Pour un enjeu en biotechnologie moléculaire**

#### **a. En utilisant leur cycle lysogène**

Au cours de ces vingt dernières années, le clonage a pris une ampleur exceptionnelle dans divers domaines. Le besoin de cloner des fragments d'ADN de plus en plus gros, a amené les scientifiques et les industriels à développer de nouveaux outils. Un des ces derniers, appelé fosmide, utilise la propriété des phages à encapsider et à propager des acides nucléiques de grande taille.

Dans cette technique, des fragments d'ADN de 40 kpb sont clonés dans un vecteur. Le produit de ligation est encapsidé dans des capsides vides de phage  $\lambda$ . Les virus recombinants sont alors mis en contact avec une souche sensible d'*Escherichia coli* qui va recevoir le produit de ligation lors de l'infection. La cellule hôte lysogène va alors assurer la multiplication de son génome mais aussi celle de l'ADN cloné dans le vecteur. Ce système permet donc de générer des banques de clones contenant des inserts de grande taille (Yokobata *et al.*, 1991 ; Kim *et al.*, 1992).

#### **b. En utilisant leur acide nucléique**

Etant donné leur relative petite taille, les acides nucléiques phagiques sont souvent utilisés, après digestion par des enzymes de restriction adéquates, comme marqueurs de haut poids moléculaire. Les exemples les plus connus sont le phage  $\lambda$  et le phage  $\phi$ X 174.

L'acide nucléique des phages peut également être employé comme matrice pour des tests d'activité de certaines enzymes. Par exemple, le phage  $\phi$ X 174, qui possède un ADN simple brin circulaire, est utilisé pour des tests d'activité nickase (clivage d'un brin d'ADN) ou ADN polymérase (copie d'un brin d'ADN à partir d'une matrice simple brin circulaire).

Le génome de certains virus de procaryotes peut également être un outil génétique. Le virus SSV1, qui infecte le genre archéen *Sulfolobus*, a servi à l'élaboration d'un vecteur navette entre *Sulfolobus solfataricus* et *Escherichia coli*. Ce dernier, appelé pKMSD48, est inductible aux rayons

UV. De plus, il permet la production de particules virales infectieuses et morphologiquement similaires à celles de SSV1. Enfin, il est présent de façon stable dans la cellule sous deux formes : une forme double brin circulaire multicopie (20 à 40) et une forme intégrée au génome de l'hôte grâce à son intégrase. Ses caractéristiques en font le premier système efficace de transformation chez les archées (Stedman *et al.*, 1999).

### c. En utilisant leurs enzymes

A l'heure actuelle, les enzymes issues de virus archéens sont relativement nombreuses. Aucune n'est cependant exploitée par les entreprises de biotechnologie moléculaire car leur activité demeure souvent indéterminée. Elles représentent malgré tout une piste importante pour les années futures, notamment parce qu'elles sont un réservoir potentiel de nouvelles fonctions biologiques pouvant parfois se présenter sous la forme d'enzymes multifonctionnelles, comme c'est le cas chez les plasmides archéens (Lipps *et al.*, 2003, 2004).

Les premières enzymes à avoir été exploitées sont les polymérases (ARN polymérases et ADN polymérases). Une des ARN polymérase phagique les plus utilisées est l'ARN polymérase ADN dépendante du coliphage T7 (Gelfand & Hayashi, 1970 ; Davanloo *et al.*, 1984). C'est une protéine de 100 kDa qui reconnaît spécifiquement un promoteur de 17 pb. Aucun activateur ou cofacteur n'est requis pour l'initiation, la transcription processive ou la terminaison. Elle est surtout employée pour la synthèse de transcrits (ARNm et ARNi) et de sondes marquées.

En ce qui concerne les ADN polymérases, une des plus employées est celle du phage  $\phi$ 29 (Watabe *et al.*, 1984). Cette enzyme synthétise de l'ADN avec une exceptionnelle processivité doublée d'une extraordinaire activité de déplacement de brin. Ses applications sont diverses : amplification par cercle roulant pour les plasmides et génomes phagiques circulaires (Dean *et al.*, 2001), amplification par déplacement multiple pour les ADN rares (Dean *et al.*, 2002 ; Gadkar & Rillig, 2005) ou encore amplification de gros fragments par couplage avec d'autres protéines du phage  $\phi$ 29 (Blanco *et al.*, 1994).

Les autres catégories d'enzymes phagiques très fréquentes en biologie moléculaire sont les ligases. La plus connue est l'ADN ligase ATP-dépendante du bactériophage T4 (Harvey *et al.*, 1971). Elle lie de façon covalente des molécules d'ADN double brin ayant soit des extrémités 5' et 3' phosphorylées, soit des extrémités cohésives obtenues grâce à une digestion par des enzymes de restriction. La principale application de cette enzyme réside dans le clonage où elle permet l'insertion d'un fragment d'ADN dans un vecteur.

Il existe également des ARN ligases. Ainsi, très récemment, deux ARN ligases thermostables et homologues à la ligase 1 du phage T4 ont été isolées de virus infectant des bactéries des genres *Rhodothermus* et *Thermus* (Blondal *et al.*, 2003, 2005). Ces enzymes sont capables de ligaturer à haute température (optimum aux environs 60°) deux molécules d'ARN ou d'ADN simple brin. Leur thermostabilité pourrait notamment améliorer l'efficacité de la technique de RLM-RACE (« RNA Ligase Mediated - Rapid Amplification of cDNA Ends ») utilisée pour déterminer précisément les extrémités 3' et 5' des ARNs. En effet, en travaillant à haute température, il est possible d'éliminer les structures secondaires des ARNs qui empêchent la ligation lorsqu'elle se fait à 37°C (Liu & Gorovsky, 1993).

#### **4. Pour un enjeu en thérapeutique**

##### **a. En tant qu'« antibiotiques »**

Les endolysines sont des enzymes codées par les bactériophages qui détruisent le peptidoglycane des bactéries lors de la phase terminale du cycle viral. Etant donné leur spécificité et leur grande activité, elles peuvent notamment être utilisées comme agent thérapeutique. Les premières études, menées avec succès, ont été réalisées sur les streptocoques chez la souris (Loeffler *et al.*, 2001 ; Loessner, 2005).

Récemment ces enzymes phagiques ont été décrites comme étant un traitement potentiel contre certaines infections des muqueuses. C'est le cas pour la protéine PlyV12 issue du phage  $\phi 1$  qui infecte une souche d'*Enterococcus faecalis*. En effet, cette enzyme permet de lutter efficacement contre les souches d'*Enterococcus faecalis* et d'*Enterococcus faecium* extrêmement résistantes aux antibiotiques de dernière génération tels que la vancomycine (Yoong *et al.*, 2004 ; Loessner, 2005).

##### **b. Pour la synthèse d'anticorps spécifiques**

La technique du « phage display » correspond à une expression de protéines (y compris des anticorps) ou de peptides grâce à un bactériophage filamenteux. Les séquences d'ADN codant pour ces molécules sont intégrées dans le génome du virus de façon à ce que les protéines ou les peptides présentés à la surface du bactériophage soient fusionnés avec des molécules capsidiales du phage. Ceci permet de générer des banques contenant simultanément plusieurs milliards de variants.

Les applications du « phage display » sont nombreuses y compris en médecine. Par exemple, grâce à cette méthode, une équipe hollandaise a isolé la partie variable d'un anticorps humain reconnaissant un antigène associé à une tumeur. Pour accroître sa stabilité, cette molécule a été

convertie en un IgG1 après intégration dans un vecteur contenant les gènes codant pour les régions constantes des immunoglobulines. Cet IgG1 purifié entraîne la mort des cellules tumorales (Huls *et al.*, 1999).

Une autre illustration des possibilités offertes par le « phage display » concerne le traitement des maladies à prions telles que la tremblante du mouton, l'encéphalopathie spongiforme bovine et la maladie de Creutzfeld-Jakob chez l'homme (Azzazy & Highsmith, 2002). La transmission de ces maladies se fait suivant le processus suivant : lorsque la forme endogène de la protéine du prion interagit directement avec la forme infectieuse, elle devient à son tour infectieuse (Telling *et al.*, 1996). Une équipe américaine a réussi à isoler pour la première fois le fragment Fab d'un anticorps qui reconnaît un épitope présent sur les deux formes du prion (Williamson *et al.*, 1998). Lorsque ce fragment Fab est appliqué à une culture de cellules neuroblastiques infectées par le prion, il y a abolition de la transformation du prion endogène en prion infectieux. Il est ainsi possible de considérer qu'une IgG, synthétisée à partir de ce fragment Fab, peut jouer un rôle dans le traitement des maladies à prions (Peretz *et al.*, 2001).

### **c. En tant que vecteurs pour la thérapie génique**

La thérapie génique consiste principalement à injecter un gène « correcteur » dans une cellule pour remplacer le gène défectueux. Dans ce domaine, les virus sont les vecteurs les plus fréquemment utilisés. Les bactériophages présentent un intérêt tout particulier en raison de leurs intégrases qui permettent une recombinaison site-spécifique. Par exemple, celle du phage  $\phi$ C31 est fonctionnelle dans les cellules de mammifères. Elle permet à l'ADN correcteur contenant un site d'attachement bactérien (nommé attB) de s'intégrer dans le génome des cellules de mammifères au niveau d'un pseudosite d'attachement phagique (nommé attP). Ce système a été expérimenté avec succès chez de petits animaux (Held *et al.*, 2005). Etant donné que l'intégrase de  $\phi$ C31 est plus active sur une matrice circulaire que sur une matrice linéaire, l'ADN sain est sous forme plasmidique. La faible efficacité de transfert reste le désavantage principal de cette technique (Nuno-Gonzalez *et al.*, 2005).

Cependant, il apparaît dorénavant comme primordial que l'intégration n'entraîne pas d'autres troubles après la guérison (Nuno-Gonzalez *et al.*, 2005). Par exemple, des enfants immunodéficients ont été traités par thérapie génique en utilisant un rétrovirus comme vecteur (Aiuti *et al.*, 2002). Trois ans après ce traitement, deux patients ont développé une leucémie due à l'intégration du gène correcteur au niveau du promoteur d'un oncogène responsable de la maladie (Hacein-Bey-Abina *et al.*, 2003). Ainsi, les recherches s'orientent dorénavant sur le ciblage de

l'intégration. Une équipe américaine a réalisé plusieurs transplantations d'hépatocytes transfectés avec une construction contenant l'intégrase  $\phi$ C31 ainsi que le gène de l'hydrolase fumanylacétate dans des souris déficientes pour cette enzyme. Après trois transplantations successives, les hépatocytes présentaient un aspect normal. Ce résultat suggère donc que ce système représente un espoir pour l'évolution de la thérapie génique (Held *et al.*, 2005).

# Introduction au travail de thèse

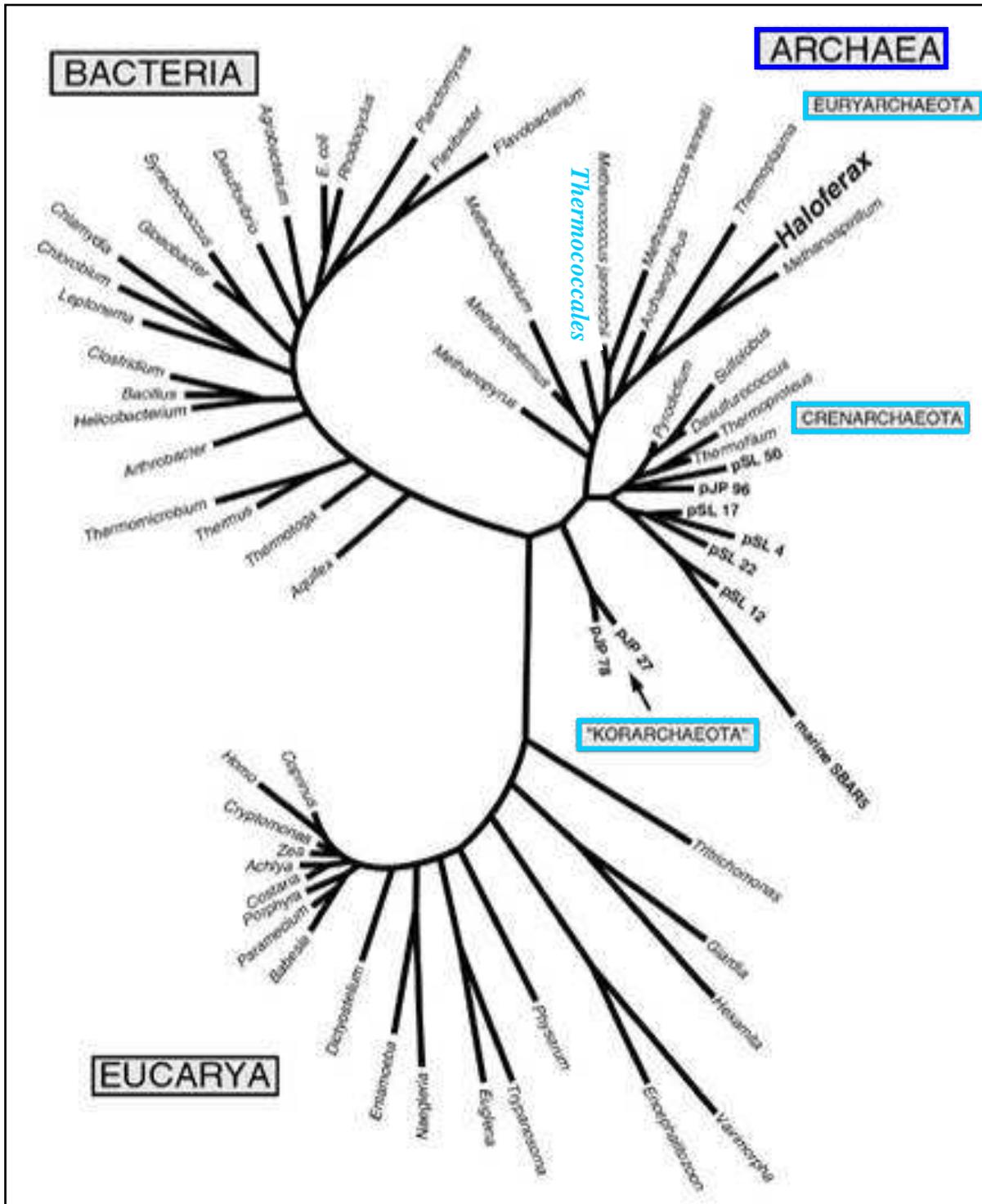
## I. Contexte de l'étude

L'avènement de la métagénomique virale de ces dernières années a permis de confirmer l'hypothèse selon laquelle les virus de procaryotes décrits ne sont que la partie émergée de la population totale (Breitbart *et al.*, 2002). Ceci est d'autant plus vrai pour les virus d'archées car ils ne constituent à l'heure actuelle que 0,8 % des spécimens caractérisés chez les procaryotes (Ackermann, 2001). Malgré leur faible nombre, les virus archéens ont réussi à conquérir de nombreux environnements et notamment des environnements dits extrêmes, comme les milieux hypersalins ou les sources chaudes terrestres (Dyall-Smith *et al.*, 2003 ; Prangishvili & Garrett, 2005). Les études portant sur les virus des archées inféodées aux sources hydrothermales marines profondes ont été initiées dans notre laboratoire, le Laboratoire de Microbiologie des Environnements Extrêmes (LM2E).

En effet, un des principaux axes de recherche du LM2E concerne l'étude de la diversité microbienne de l'environnement hydrothermal abyssal avec un volet portant sur les éléments génétiques mobiles (virus et plasmides). L'étude de ces derniers est préférentiellement menée chez les archées de l'ordre des *Thermococcales*. Notre laboratoire possède de nombreux échantillons de cheminées hydrothermales collectés au cours de diverses campagnes dans les océans Atlantique, Pacifique et Indien ainsi qu'une collection regroupant plus de 600 souches de *Thermococcales* (<http://www.ifremer.fr/souchotheque>).

Les *Thermococcales* ont été choisies comme modèle d'étude dans notre laboratoire car elles sont largement répandues au niveau des différentes sources hydrothermal marines de la planète et constituent une part très importante de la population hétérotrophe microbienne de l'écosystème (Harmsen *et al.*, 1997 ; Takai *et al.*, 2001). D'autre part, elles constituent un groupe extrêmement varié car elles regroupent 38 espèces. L'ordre des *Thermococcales* appartient au phylum des *Euryarchaeota* et regroupe trois genres : *Pyrococcus*, *Thermococcus* et *Paleococcus* (cf Figure 15). Ces sont des micro-organismes chimio-organo-hétérotrophes anaérobies stricts hyperthermophiles et soufre réducteurs. Plusieurs souches décrites appartenant à cet ordre sont devenues des micro-organismes modèles et le génome de cinq d'entre elles (*Pyrococcus abyssi*, *Pyrococcus horikoshii*, *Pyrococcus furiosus*, *Thermococcus kodaka-*

*rensis* et *Thermococcus gammatolerans*) a été entièrement séquencé (<http://www.ncbi.nlm.nih.gov>).



**Figure 15 :** Place des *Thermococcales* dans l'arbre phylogénétique du vivant non enraciné et basé sur la comparaison des séquences des gènes des ADNr 16S et 18S (d'après Barns *et al.*, 1996).

Au LM2E, l'étude des virus de *Thermococcales* est orientée pour répondre à deux objectifs distincts. Le premier consiste en une meilleure connaissance de la biodiversité et de la génomique des virus hyperthermophiles marins. Elle nécessite le criblage de nombreuses souches d'archées, la description de nouveaux virus ainsi que le séquençage et l'annotation de leurs génomes. Le second objectif, plus en aval, a trait à l'obtention de protéines thermostables (telles des polymérase, des ligases, etc) susceptibles d'appartenir à de nouvelles familles et pouvant présenter un intérêt dans le secteur des biotechnologies. Il s'inscrit le cadre de la création de l'ACI Deep Hot-Vir qui regroupe le LM2E ainsi que trois laboratoires parisiens : le Laboratoire d'Imagerie Cellulaire et Moléculaire (du CEA de Saclay), le Laboratoire de Biologie Moléculaire du Gène chez les Extrêmophiles et le Laboratoire de Génomique Structurale (de l'Université de Paris XI). Cette ACI a pour objectif l'étude structurale et fonctionnelle des protéines dites 3R (impliquées dans la recombinaison, la réplication et/ou la réparation) encodées par les virus et les plasmides portés par des archées inféodées aux sources hydrothermales marines. Cette étude consiste en une approche biochimique comprenant des étapes de surproduction, de purification, de cristallisation et de détermination de l'activité de ces protéines.

## II. Exposé du travail de thèse

PAV1, le premier virus hyperthermophile marin décrit, a été découvert dans notre laboratoire à la suite d'un criblage mené sur plus de 200 souches de *Thermococcales*. C. Geslin et ses collaborateurs ont réalisé une caractérisation de ce virus portant principalement sur sa morphologie, la nature de sa structure externe ainsi que sur la relation qu'il entretient avec son hôte, la souche GE23 de *Pyrococcus abyssi* (Geslin *et al.*, 2003). Dans le but de mieux comprendre PAV1, nous avons axé la première partie de ce travail de thèse sur l'analyse moléculaire (génomique et protéomique) du virus.

Dans un premier temps, nous avons concentré nos efforts sur l'étude du génome viral. Le but était de consolider l'analyse de la séquence réalisée au préalable et de présenter ainsi une étude globale. Certaines caractéristiques physique du génome de PAV1 ont été explorées (surenroulement, méthylations) puis nous nous sommes intéressés à la réplication ainsi qu'à la transcription. Dans un second temps, notre attention s'est portée sur les protéines codées par le génome de PAV1. Des essais de purification et des tests d'activité ont été menés sur la plus

grosse protéine de PAV1, la protéine 899. De plus, une des protéines majoritaires du virus a pu être identifiée.

En parallèle de cette étude moléculaire de PAV1, nous avons cherché un modèle plus simple que PAV1. Nous avons donc initié un nouveau criblage sur de nouvelles souches de *Thermococcales*. Pour cela, un grand nombre de cultures d'enrichissement et d'isolats ont été analysés, principalement par microscopie électronique en transmission (MET), mais aussi par polymérase chain reaction (PCR). Cette recherche a permis la découverte d'une souche nommée KAZA, proche de l'espèce *Thermococcus barophilus*, qui produit à pression atmosphérique une particule de type viral filamenteuse nommée VLP-TBFP. Cette souche étant piézophile, nous avons réalisé des cultures sous pression hydrostatique et une seconde particule de type viral de morphologie tête-queue, nommée VLP-TBPP, a été découverte. Nous avons ensuite effectué une caractérisation préliminaire de ces deux particules en recherchant la présence d'un acide nucléique dans les deux types de particules puis en étudiant la structure de la VLP-TBPP.

# Matériel et Méthode

## I. Echantillons bruts

Dans le cadre d'une approche biogéographique, 65 échantillons bruts prélevés au cours de différentes campagnes océanographiques sont sélectionnés (cf Figure 16 et Annexe 1). Ce sont principalement des fragments de cheminée hydrothermale et des modules de colonisation déposés près d'événements hydrothermaux pendant plusieurs jours. Ces échantillons bruts sont utilisés pour inoculer du milieu de culture afin de réaliser des cultures d'enrichissement d'archées de l'ordre de *Thermococcales*.



**Figure 16 :** Origine géographique des échantillons bruts utilisés.

## II. Culture des *Thermococcales*

### 1. Milieu utilisé

Le milieu utilisé pour cultiver les *Thermococcales* est le milieu de Ravot qui favorise leur type de métabolisme (Ravot *et al.*, 1995). Il contient pour 1 L d'eau distillée : 1 g de  $\text{NH}_4\text{Cl}$  ; 0,2 g de  $\text{MgCl}_2 \cdot 6\text{H}_2\text{O}$  ; 0,1 g de  $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$  ; 0,1 g de  $\text{KCl}$  ; 0,83 g de  $\text{CH}_3\text{COONa} \cdot 2\text{H}_2\text{O}$  ; 20 g de  $\text{NaCl}$  ; 5 g d'extrait de levure ; 5 g de bio-trypcase ; 3,45 g de tampon PIPES et 1 mg de rézazurine. Le milieu est ajusté à pH 7 et autoclavé 20 min à 120 °C. Puis, à froid, il est rajouté 0,03 % (v/v) de  $\text{K}_2\text{HPO}_4$  et 0,03 % (v/v) de  $\text{KH}_2\text{PO}_4$ . Le milieu est conservé à 4 °C.

## **2. Culture en milieu liquide à pression atmosphérique**

Avant utilisation, le milieu de Ravot est réparti dans des fioles pénicilline stériles à raison de 50 mL par fiole et 0,5 g de soufre élémentaire sont ensuite ajoutés dans chaque fiole. Les fioles sont bouchées hermétiquement avant d'être soumises à plusieurs cycles vide/gaz (azote) qui permettent l'obtention de l'anaérobiose. La réduction complète est assurée par un ajout au 1/100<sup>ème</sup> d'une solution de Na<sub>2</sub>S non stérile à 10 % (p/v). Selon l'objectif de l'étude, le milieu est alorsensemencé soit avec un échantillon brut (à 3 % (v/v)), soit avec une culture pure (à 1 % (v/v)). Il est ensuite incubé à 85°C sous agitation. La croissance des micro-organismes est contrôlée au microscope à contraste de phase et arrêtée soit en fin de phase exponentielle, soit en fin de phase stationnaire suivant les expériences.

## **3. Culture en milieu liquide sous pression hydrostatique**

Dans certaines expériences, les *Thermococcales* sont cultivées sous pression hydrostatique pour reproduire les conditions de pression *in situ*. Pour cela, le milieu de Ravot est préparé et inoculé dans les mêmes conditions que pour la culture à pression atmosphérique. Puis, il est réparti dans des seringues en plastique de 50 mL munies d'une aiguille bouchée. Ces dernières sont placées dans des incubateurs en inox blindés et chauffants (TopIndustrie). Lorsque la température de 85°C est atteinte, une pompe à main est utilisée pour appliquer la pression voulue (15 MPa, 25 MPa, 35 MPa ou 45 MPa). Les cultures sont arrêtées à des temps définis en sortant les seringues des incubateurs, après dépressurisation.

## **4. Isolement sur milieu solide à pression atmosphérique**

La culture en milieu solide est utilisée pour réaliser des isollements à partir des cultures d'enrichissement en milieu liquide ou pour purifier des isolats (3 passages sur boîte). Dans les deux cas, le protocole utilisé est une version modifiée de celui décrit par G. Erauso et ses collaborateurs. La manipulation s'organise en deux parties car le milieu est coulé en deux temps : d'abord la « sous-couche » (sans soufre) puis la « sur-couche » (avec soufre). La première sert plutôt de support alors que la seconde est nutritive. La plupart du temps, les colonies apparaissent à l'interface des deux couches, là où l'anaérobiose et les conditions nutritives sont optimales (Erauso *et al.*, 1995).

### **a. Réalisation de la « sous-couche »**

Pour réaliser la « sous-couche », le milieu de Ravot solide utilisé est préparé de la façon suivante. Un flacon va contenir uniquement la matière minérale (deux fois concentrée) du milieu et un autre flacon va contenir la matière organique et les phosphates (deux fois concentrés) ainsi que 1,6 % (p/v) de Gelrite® (Merk). Les deux flacons sont rentrés dans l'enceinte anaérobie et placés sur des agitateurs chauffants afin de porter à ébullition leur contenu. Chaque fraction est ensuite réduite par ajout au 1/100<sup>ème</sup> d'une solution de Na<sub>2</sub>S à 10 % (p/v) puis mélangées volume à volume dans un même flacon. Le mélange est réduit par ajout au 1/500<sup>ème</sup> d'une solution de Na<sub>2</sub>S à 10 % (p/v) puis coulé à chaud dans le fond des boîtes de Pétri en plastique. Après la prise en masse du milieu, les boîtes sont laissées sécher pendant 15 minutes à température ambiante dans l'enceinte anaérobie.

### **b. Réalisation de la « sur-couche »**

La « sur-couche » est préparée de la façon suivante : un flacon va contenir uniquement la matière minérale (deux fois concentrée) du milieu et un autre flacon va contenir la matière organique et les phosphates (deux fois concentrés) ainsi que 0,4 % (p/v) de Gelrite®. Après leur introduction dans l'enceinte anaérobie, les deux fractions du milieu sont chauffées jusqu'à ébullition. Elles sont ensuite réduites par ajout au 1/20<sup>ème</sup> d'une solution de Na<sub>2</sub>S à 10 % (p/v) et laissées sur l'agitateur chauffant. Dans des tubes à hémolyse, placés dans un bloc chauffant réglé à 85°C, sont versés 900 µL de matière minérale et 900 µL de matière organique additionnée de Gelrite®. Des dilutions au 1/10<sup>ème</sup> de la culture d'enrichissement sont ensuite réalisées dans ces tubes à hémolyse jusqu'à la dilution 10<sup>-6</sup>. Dans chaque tube à hémolyse, 300 µL de soufre colloïdal sont ajoutés puis l'ensemble du contenu est versé sur la « sous-couche » en adoptant un mouvement circulaire pour assurer l'homogénéité de la « sur-couche ».

### **c. Incubation et repiquage des colonies**

Les boîtes de Pétri sont alors placées dans une jarre anaérobie en présence de deux bandelettes de papier Whatman imbibées de 2 mL d'une solution de Na<sub>2</sub>S à 10 % (p/v). La jarre est sortie de l'enceinte anaérobie, pressurisée avec 1 bar d'azote, puis mise à incuber pendant 3 jours à 85°C. Une fois ce laps de temps écoulé, la jarre est dépressurisée puis entrée dans l'enceinte anaérobie. Dans les boîtes de Pétri, les colonies bien isolées sont prélevées à l'aide d'un cône stérile puis repiquées dans 5 mL du milieu de Ravot liquide. Les cultures

d'isolats sont incubées à leur tour à 85°C sous agitation et stoppées en fin de phase exponentielle de croissance.

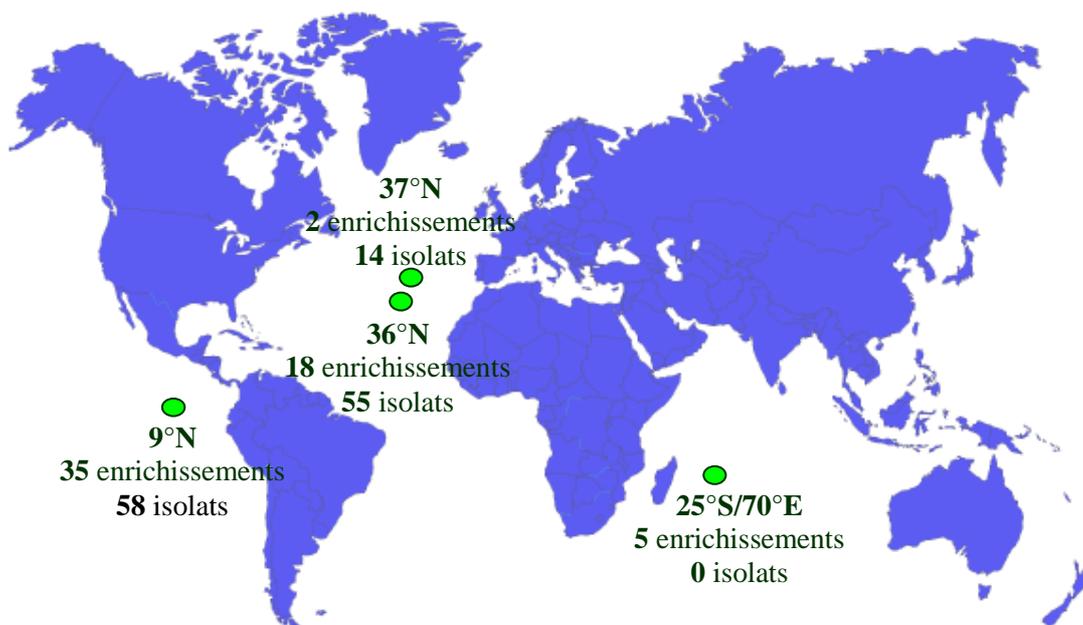
Etant donné la quantité de travail occasionnée, seul l'isolat d'intérêt, initialement nommé E2p8 puis rebaptisé KAZA, est purifié. Dans ce but, il est donc passé deux fois supplémentaires sur boîte.

### III. Recherche de nouveaux virus dans des cultures de *Thermococcales*

#### 1. La Microscopie Electronique en Transmission (MET)

##### a. Cultures d'enrichissement et isolats observés

La recherche de nouveaux virus est menée en MET dans un premier temps sur les 60 cultures d'enrichissement présentant une croissance de micro-organismes (cf Figure 17 et Annexe 2). Dans un deuxième temps, elle est menée sur 127 isolats issus des cultures d'enrichissement où des particules de type viral (VLPs) sont observées après plusieurs repiquages successifs (cf Figure 17 et Annexe 3). Les isolats, issus des cultures d'enrichissement où les VLPs ne sont observées que transitoirement, c'est-à-dire uniquement dans la culture initiale, ne sont pas analysés



**Figure 17 :** Origine géographique des différentes cultures d'enrichissement et des différents isolats observés en MET.

## b. Protocole

Une étape préalable de concentration des VLPs est nécessaire. Elle consiste à prélever 8 mL de culture qui sont clarifiés différemment suivant la nature de l'échantillon. S'il s'agit d'une culture d'enrichissement, celle-ci est filtrée sur des Acrodic PF 0,8/0,2 µm (Pall Gelman Laboratory). Dans le cas de cultures d'isolats, cette dernière est centrifugée à 4 000 g pendant 15 min à 4°C. Le filtrat ou le surnageant est ensuite ultracentrifugé à 100 000 g pendant 3 h à 4°C (rotor Beckmann à angle fixe 70.1 Ti). Le culot qui contient les VLPs est finalement repris dans du tampon TE pH 8 (10 mM Tris-HCl, 1 mM EDTA).

Cinq µL de la solution concentrée sont déposés pendant 2 min sur une grille de cuivre carbonée comportant 400 mailles (Euromedex). Puis, l'excès de liquide est absorbé à l'aide de papier Whatman. Cinq µL d'une solution d'acétate d'uranyle à 2 % (p/v) sont alors déposés sur la grille à l'abri de la lumière. Après 45 s, l'excès de contrastant est enlevé. La grille est déposée sur du papier Whatman pour sécher pendant la nuit. L'observation est effectuée sur un microscope électronique de type JEM 100 X II (Jeol) en réglant la vitesse d'accélération à 60 kV et le grossissement à 200 000.

Le mode de calcul permettant de dénombrer les VLPs est détaillé ci-dessous :

$$\text{nombre de VLPs.mL}^{-1} = \frac{\text{nombre moyen de VLPs.champ}^{-1} \times Y \times 400 \times Z}{8}$$

Dans cette opération, Y correspond au nombre de champs que contient une maille et 400 est le nombre de mailles que comporte une grille. Z est le facteur de multiplication pour tenir compte du volume total de reprise du culot par rapport au volume déposé sur la grille (5 µL). Le 8 correspond au volume de culture concentrée.

## 2. La Polymerase Chain Reaction (PCR)

### a. Isolats analysés

La recherche de nouveaux virus est menée par PCR sur 98 isolats d'origine variée (cf Figure 18 et Annexe 4). Une partie des isolats analysés est issue de cultures d'enrichissement où des particules de type viral sont observées dans les repiquages de la culture initiale. La seconde partie des isolats provient de la collection de la Souchothèque de Bretagne (<http://www.ifremer.fr/souchotheque>).



**Figure 18 :** Origine géographique des différents isolats analysés par PCR.

### b. Protocole

L'ADN total est tout d'abord extrait à partir de 2 mL d'une culture en fin de phase exponentielle réalisée en milieu de Ravot. Les cellules sont collectées par centrifugation à 6 000g pendant 15 min à 4°C. Le culot cellulaire est repris dans 80 µL de tampon TNE pH 8 (100 mM Tris-HCl, 100 mM NaCl, 50 mM EDTA). Puis, les cellules sont lysées par ajout de 10µL de SDS 10 % (p/v) et 10µL de Sarkosyl 10 % (p/v). Le lysat est alors traité avec 5 µL de protéinase K (20 mg.mL<sup>-1</sup>) puis incubé 1 h à 55°C. L'extraction d'ADN total est réalisée à l'aide d'un volume de mélange PCI (phénol/chloroforme/alcool isoamylique, 25 :24 :1) puis avec un volume de chloroforme. L'ADN est ensuite précipité en ajoutant 0,7 volume d'isopropanol, puis collecté par centrifugation à 20 800 g pendant 20 min à 4°C. Le culot d'ADN est lavé avec 400 µL d'éthanol 75 % (v/v), séché au Speed-Vac puis repris dans 20 µL de tampon TE pH 8 (10 mM Tris-HCl, 0,1 mM EDTA).

La PCR est alors réalisée sur l'ADN total. Le mélange réactionnel, réalisé dans le tampon de l'enzyme, est composé de : 2 mM de MgCl<sub>2</sub>, 50 µM de dNTP, 0,1 µM d'amorce sens REP-1S (5' GCG CGG AAT TCG GNY TNG ARG TNC CNG AG 3'), 0,1 µM d'amorce anti-sens REP-1AS (5' GCG CGC AAG CTT GGN GCN ATH GGN KTN TTY ACN TTY 3'), 1 U de Taq polymérase UptiTherm<sup>®</sup> (Interchim) et 10 ng d'ADN. Le programme utilisé est composé des phases suivantes : 3 min de dénaturation à 94°C, 30 cycles

d'amplification (30 s à 94 °C, 30 s à 48 °C et 1 min à 70 °C) et 7 min d'extension finale à 70 °C. La PCR est contrôlée en déposant 5 µL de produit d'amplification sur gel d'agarose 0,8 % (p/v) contenant 0,5 µg.mL<sup>-1</sup> de bromure d'éthidium (BET). La migration se fait en tampon TAE 1 x pendant 30 min à 8 V.cm<sup>-1</sup>.

#### **IV. Description d'un système hôte - particules de type viral chez les *Thermococcales***

##### **1. Détermination de l'identité de la souche KAZA**

Le gène codant pour l'ADNr 16S de la souche KAZA est séquencé pour déterminer de quelle espèce de *Thermococcale* la souche est la plus proche. Pour cela, une extraction de l'ADN total est pratiquée sur une culture en fin de phase exponentielle puis une PCR est réalisée sur l'ADN extrait en utilisant les mêmes conditions décrites dans le paragraphe *§§.2.5*. Seul le couple d'amorce utilisé change, il s'agit du couple 4F (5' GTA AAA CGA CGG CCA GTC CGG TTG ATC CTG CCR G 3') / 1492R (5' CAG GAA ACA GCT ATG ACC GGT TAC CTT GTT ACG ACT T 3') qui est spécifique du gène de l'ADNr 16S archéen. Le produit de PCR est ensuite cloné dans le vecteur pGME<sup>®</sup>-T en utilisant les instructions du fournisseur (Promega) puis séquencé par la société GenomExpress. La séquence obtenue est comparée avec celles présentes dans les bases de données en utilisant l'algorithme Basic Local Alignment Search Tool (Altschul *et al.*, 1990).

##### **2. Détermination de la croissance de la souche KAZA et de la production des deux particules de type viral**

Après sa purification, l'isolat de *Thermococcales* E2p8, qui porte des VLPs, est rebaptisé KAZA. La souche KAZA est cultivée à pression atmosphérique (cf paragraphe *§§.2*) et à 35 MPa (cf paragraphe *§§.3*) en présence ou non de 2,5 µg.mL<sup>-1</sup> de mitomycine C. Cette dernière possède la propriété de pouvoir induire la production de particules virales dans certains cas.

Pour chaque condition, 7 sous-cultures sont réalisées en parallèle, chacune correspondant à un point de prélèvement permettant d'établir le suivi dans le temps. Les prélèvements sur les cultures non traitées sont réalisés à t = 0, 6 h, 8 h, 12 h, 16 h, 24 h et 32 h. Pour les cultures traitées, l'ajout de mitomycine C se fait aux temps t = 0, 6 h, 8 h, 12 h, 16 h, 24 h et 32 h puis elles sont réincubées pendant 4 h dans les mêmes conditions. Les prélèvements ont

donc lieu à  $t = 4$  h, 10 h, 12 h, 16 h, 20 h, 28 h et 36 h. La croissance cellulaire est suivie en réalisant des comptages à la cellule de Thoma. La production de VLPs est suivie en réalisant des comptages en MET (cf paragraphe §§§.1).

### **3. Détermination des conditions optimales de pression hydrostatique pour la production de la VLP-TBPP**

La souche KAZA est cultivée à 4 pressions différentes (15 MPa, 25 MPa, 35 MPa et 45 MPa) pendant 32 h (cf paragraphe §§.3). Pour chaque condition de pression, une culture non traitée et une culture traitée par  $2,5 \mu\text{g.mL}^{-1}$  de mitomycine C à  $t = 8$  h sont réalisées. Toutes les cultures sont ensuite traitées pour une observation en MET et les VLPs sont comptées (cf paragraphe §§§.1).

### **4. Recherche de la VLP-TBPP dans différentes souches de *Thermococcales***

Neuf souches de *Thermococcales* issues de la collection de la Souchothèque de Bretagne sont cultivées à pression atmosphérique et à 15 MPa pendant 32 h (cf paragraphe §§§.2) en étant traitées ou non avec  $2,5 \mu\text{g.mL}^{-1}$  de mitomycine C à  $t = 8$  h. Les souches utilisées sont : *Thermococcus barophilus* souche MP, *Thermococcus barophilus* souche DV1179, *Thermococcus sp.* souche ES1, *Thermococcus fumicolans* souche ST557, *Thermococcus atlanticus* souche MA898, *Thermococcus hydrothermalis* souche AL662, *Thermococcus siculi* souche AT1243, *Pyrococcus abyssi* souche GE23 et *Pyrococcus abyssi* souche GE9. Toutes les cultures sont ensuite préparées pour la MET et les VLPs sont comptées (cf paragraphe §§§.1).

### **5. Recherche d'un acide nucléique chez les deux particules de type viral**

#### **a. Concentration des particules**

Toutes les techniques utilisées pour la recherche d'acide nucléique (sauf l'épifluorescence) sont menées sur des particules concentrées selon le protocole qui suit. La souche KAZA est cultivée pendant 32 h à pression atmosphérique et à 15 MPa avec traitement à la mitomycine C à  $t = 8$  h. Les cultures sont centrifugées à 4 000 g pendant 15 min à 4°C. Le surnageant est ensuite ultracentrifugé à 100 000 g pendant 3 h à 4°C. Les culots obtenus sont repris dans du tampon TE « low EDTA » pH 8 (10 mM Tris-HCl, 0,1 mM EDTA).

### **b. Essai d'extraction d'ADN**

Au préalable, 360  $\mu\text{L}$  de particules concentrées sont traitées avec 50  $\mu\text{g.mL}^{-1}$  de DNase I (Sigma) et 100  $\mu\text{g.mL}^{-1}$  de RNase A (Eurobio) dans un milieu réactionnel contenant 10 mM de  $\text{MgCl}_2$ . Après 1 h d'incubation à 37°C, les nucléases sont inactivées 10 min à 70°C. Les particules sont alors traitées avec 7  $\text{mg.mL}^{-1}$  de protéinase K pendant 1 h à 56°C. L'extraction d'ADN est réalisée à l'aide d'un volume de mélange PCI (phénol/chloroforme/alcool isoamylique, 25 : 24 : 1) puis avec un volume de chloroforme. L'ADN est ensuite précipité par ajout de 0,8 volume d'isopropanol puis récupéré par centrifugation pendant 20 min à 20 800 g à 4°C. Le culot d'ADN est lavé avec 500  $\mu\text{L}$  d'éthanol 75 % (v/v), séché au Speed-Vac puis repris dans 20  $\mu\text{L}$  de tampon TE « low EDTA » pH 8.

Dans l'éventualité où la quantité d'ADN extrait serait très faible, un test PCR est réalisé avec des amorces aléatoires. Les 25  $\mu\text{L}$  de mélange réactionnel sont réalisés dans le tampon de l'enzyme et contiennent : 2 mM de  $\text{MgCl}_2$ , 50  $\mu\text{M}$  de dNTP, 1 U de Taq polymérase Upti-Therm<sup>®</sup> (Interchim), 2,5  $\mu\text{L}$  de nonamères (Amersham) et 1  $\mu\text{L}$  d'ADN extrait. Le programme utilisé est composé des phases suivantes : 3 min de dénaturation à 94°C, 30 cycles d'amplification (30 s à 94°C, 30 s à 48°C et 1 min à 70°C) et 7 min d'extension finale à 70°C. La PCR est contrôlée en déposant 5  $\mu\text{L}$  de produit d'amplification sur gel d'agarose 0,8 % (p/v) contenant 0,5  $\mu\text{g.mL}^{-1}$  de BET. La migration se fait en tampon TAE 1 x (40 mM Tris-HCl pH 8, 20 mM, acétate, 1 mM EDTA pH 8) pendant 30 min à 8  $\text{V.cm}^{-1}$ .

### **c. Essai d'extraction d'ARN**

Les particules concentrées sont traitées à la DNase I, à la RNase A et à la protéinase K comme pour l'extraction d'ADN. Ensuite 0,5 volume d'éthanol absolu sont ajoutés. Après homogénéisation du mélange, l'ARN est extrait sur colonne au moyen du kit RNeasy (QIAGEN) suivant les instructions du fournisseur.

Dans l'éventualité où la quantité d'ARN extrait serait très faible, un test RT-PCR est réalisé avec des amorces aléatoires. Dans 12  $\mu\text{L}$  de volume final, 1  $\mu\text{L}$  de nonamères (Amersham) et 2  $\mu\text{L}$  d'ARN extrait sont chauffés à 70°C pendant 3 min. Le mélange réactionnel réalisé dans le tampon de l'enzyme est additionné de : 0,5 mM de dNTP, 40 U d'inhibiteur de RNase (Ambion) et 100 U de reverse transcriptase M-MLV (Ambion). La réaction de reverse transcription est réalisée 60 min à 44°C puis 10 min à 92°C. Les conditions utilisées pour la réaction de PCR et sa vérification sont les mêmes que celles de l'essai d'extraction d'ADN.

#### **d. Technique du « craking »**

Ce protocole a été mis au point par E. Marguet et N. Soler. Le tampon de « craking » (50 mM Tris-HCl pH 7,5, 2 mM EDTA pH 7,5, 1 % (p/v) SDS, 13,7 % (p/v) sucrose, 0,01 % (p/v) bleu de bromophénol) est additionné volume à volume avec les particules concentrées. Après une incubation de 10 min à 65°C, le produit de craking est déposé sur gel d'agarose 0,5 % (p/v) contenant 0,5  $\mu\text{g}\cdot\text{mL}^{-1}$  de BET. La migration se fait en tampon TAE 1 x pendant 30 min à 8  $\text{V}\cdot\text{cm}^{-1}$ .

#### **e. Test en épifluorescence**

Le protocole utilisé a été mis au point par C. Geslin et A. Lies. La souche KAZA est cultivée pendant 32 h à pression atmosphérique et à 15 MPa avec traitement à la mitomycine C à  $t = 8$  h. Un aliquot de 2 mL de chaque culture est décanté pour éliminer le soufre puis 1 mL du surnageant est centrifugé à 10 000 g pendant 20 min à 4°C. Le surnageant est collecté puis dilué au  $1/10^{\text{ème}}$  dans de l'eau milliQ. Cinq mL de cette dilution sont mélangés avec 1 mL d'une solution de pyrophosphate de sodium (10 mM) et incubés pendant 15 min. Après une agitation manuelle d'1 min, l'échantillon est filtré sur un Acrodisc PF 0,8/0,2  $\mu\text{m}$  (Pall Gelman Laboratory) puis sur un Anodisc 13 0,02  $\mu\text{m}$  (Whatman). Ce dernier est rincé deux fois avec de l'eau milliQ filtrée sur 0,02  $\mu\text{m}$ .

Cent  $\mu\text{L}$  de SYBR Green I 10 x (Molecular Probes) sont appliqués sur l'Anodisc pendant 15 min à l'obscurité puis celui-ci est séché à l'aide de papier Whatman avant d'être déposé sur une lame de verre. Cent  $\mu\text{L}$  de solution « anti-fading » à pH 7 (0,42 % (p/v) NaCl, 24,8 mM  $\text{Na}_2\text{HPO}_4$ , 49,5 % (v/v) glycérol, 10 % (p/v) N-N diméthyl-1,4-phenylene diamine sulfate) sont ensuite déposés sur l'Anodisc qui est alors recouvert d'une lamelle de verre. L'excès de liquide est enlevé en utilisant du papier Whatman puis la préparation est observée en réglant le grossissement à 100 sur un microscope à épifluorescence (Olympus BX60) grâce au filtre WG.

### **6. Purification des deux particules de type viral**

#### **a. Ultracentrifugation et gradient de densité**

La souche KAZA cultivée à pression atmosphérique (50 mL) et à 15 MPa (50 mL) est traitée comme précédemment par ajout de mitomycine C. Après 32 h d'incubation, les cultures sont arrêtées puis centrifugées à 6 000 g pendant 15 min à 4°C. Les surnageants sont ensuite ultracentrifugés à 100 000 g pendant 3 h à 4°C. Les culots obtenus sont repris dans 60

$\mu\text{L}$  de tampon « TE low » pH 8 (10 mM Tris-HCl, 0,1 mM EDTA) puis déposés sur 8 mL d'une solution de CsCl à  $0,5 \text{ mg.mL}^{-1}$  et à nouveau ultracentrifugé à 100 000 g pendant 24 h à  $4^\circ\text{C}$ . Le gradient obtenu est ensuite passé dans l'analyseur UA-6 (Isco). Les fractions présentant une absorbance à 254 nm sont collectées et diluées dans du tampon « TE low » pH 8, puis ultracentrifugées à 100 000 g pendant 3 h à  $4^\circ\text{C}$ . Les culots sont finalement repris dans 200  $\mu\text{L}$  de tampon « TE low » pH 8.

### **b. Précipitation et gradient de densité**

C'est le protocole utilisé pour purifier le virus PAV1 (Geslin *et al.*, 2003), à la différence près que la densité du gradient de chlorure de césium (CsCl) est de 1,3910 au lieu de 1,2858. Après 8 h de culture à pression atmosphérique, la souche KAZA (200 mL) est traitée par ajout de mitomycine C ( $2,5 \text{ }\mu\text{g.mL}^{-1}$ ) puis la culture est poursuivie jusqu'à la fin de la phase stationnaire ( $t = 32 \text{ h}$ ). La suite consiste en une concentration par précipitation au PEG suivie d'une purification en gradient de CsCl.

### **c. Ultrafiltration tangentielle**

Deux cultures en gros volume de la souche KAZA (1 L) sont réalisées à pression atmosphérique et à 15 MPa. A  $t = 8 \text{ h}$ , chaque culture est traitée par ajout de mitomycine C puis remise à incuber dans les mêmes conditions. Après 32 h d'incubation, les cultures sont arrêtées puis centrifugées à 6 000 g pendant 15 min à  $4^\circ\text{C}$ . Les surnageants sont ultrafiltrés en utilisant l'appareil Pellicon 2 (Millipore) muni du module P2B0 1MC 01 dont le seuil de coupure est de 1 000 kDa. Les filtrats obtenus sont séquentiellement ultrafiltrés sur le même appareil avec les modules P2B5 00C 01, P2B1 00A 01, P2B0 30A 01 qui possèdent respectivement des seuils de coupure de 500 kDa, 100 kDa et 30 kDa. Chaque concentrat est conservé à  $4^\circ\text{C}$ , centrifugé pendant 15 min à 10 000 g à  $4^\circ\text{C}$  puis préparé pour une observation en MET.

## **7. Etude préliminaire de la structure de la VLP-TBPP**

### **a. Tests de sensibilité**

La VLP-TBPP purifiée est soumise à différents agents. Les particules sont mélangées volume à volume avec du chloroforme puis incubées 10 min sous agitation à température ambiante. Parallèlement, elles sont incubées avec du Triton X-100 0,3 % (v/v) pendant 3 min. Les particules sont aussi traitées avec  $0,4 \text{ mg.mL}^{-1}$  de protéinase K pendant 1 h à  $56^\circ\text{C}$  dans un tampon réactionnel contenant 0,5 % (p/v) de SDS. Pour déterminer l'action réelle de la protéinase K deux témoins sont utilisés. Le premier correspond à des particules chauffées 1 h

à 56°C. Le second correspond à des particules traitées avec 0,5 % (p/v) de SDS pendant 1 h à 56°C.

### **b. Recherche des protéines majoritaires**

Un aliquot de 20 µL de VLP-TBPP purifiée est mélangé volume à volume avec du tampon de charge (0,2 % (p/v) bleu de bromophénol, 16 % (p/v) sucrose, 4 % (p/v) SDS, 2 % (p/v) β-mercaptoethanol, 0,124 M Tris-HCl pH 6,8) puis dénaturé pendant 5 min à 100°C. Il est ensuite déposé sur un gel d'acrylamide composé d'un gel de concentration (à 3 %) et d'un gel de séparation (gradient de 4 % à 20 %). La migration électrophorétique se fait dans un tampon dénaturant (25 mM Tris, 19 mM glycine, 1 % (p/v) SDS) pendant 1 h en voltage constant (100 V).

Après électrophorèse, le gel est coloré 1 h dans une solution de bleu de Coomassie (0,1 % (p/v) bleu de Coomassie, 50 % (v/v) éthanol, 10 % (v/v) acide acétique). Il est ensuite décoloré dans une solution contenant 30 % (v/v) d'éthanol et 10 % (v/v) d'acide acétique jusqu'à l'obtention de bandes nettes. Ces dernières sont découpées puis analysées en spectrométrie de masse MALDI-TOF par le Plateau d'Analyses Protéomique de l'INRA de Jouy-en-Josas.

## **V. Etude du génome de PAV1**

### **1. Détermination du nombre de copies du génome viral dans la cellule**

L'ADN total de la souche *Pyrococcus abyssi* GE23 est extrait comme décrit précédemment (paragraphe 3.2.5) puis digéré enzymatiquement par les endonucléases *HindIII* et *NaeI* (NEB) en suivant les instructions du fournisseur. Il est ensuite déposé sur gel d'agarose à 0,5 % (p/v). La migration se fait en tampon TAE 1 x à 4 V. cm<sup>-1</sup> pendant 6 h. Le gel est plongé pendant 15 min dans un bain de dépurination (HCl 0,25 N), puis deux fois pendant 10 min dans un bain de dénaturation (NaOH 0,5 N, NaCl 1,5 M) et un bain de neutralisation (Tris-HCl pH 7,5 0,5 M, 1,5 M de NaCl). L'ADN est ensuite transféré passivement sur une membrane de nitrocellulose comme décrit précédemment (Sambrook *et al.*, 1989).

Deux sondes, appelées « PAV1 » et « 16S GE23 », sont synthétisées. L'ADN de PAV1, extrait par lyse alcaline, est digéré enzymatiquement par *HindIII* (NEB) en suivant les instructions du fournisseur, puis déposé sur gel d'agarose 0,6 % (p/v). Après une migration en tampon TAE 1 x à 8 V.cm<sup>-1</sup> pendant 1 h, la bande de 1,7 kpb générée par la digestion du génome viral est extraite du gel à l'aide du kit Nucleo Spin Extract (Macherey-Nagel) puis mar-

quée en utilisant le kit ECL RP Labelling System (Amersham) suivant les instructions des fournisseurs. Le gène de l'ADNr 16S est amplifié par PCR comme décrit précédemment à partir de l'ADN total extrait de la souche GE23 de *Pyrococcus abyssi*. Le produit de PCR est cloné dans le vecteur pGEM<sup>®</sup>-T (Promega) puis utilisé pour transformer des cellules compétentes *E. coli* XL10-Gold (Stratagene) par choc thermique (30 s à 42°C). Après amplification, le produit de clonage est extrait par lyse alcaline comme décrit précédemment (Birnboim & Doly, 1979), digéré enzymatiquement par *NaeI* (NEB) puis déposé sur un gel d'agarose à 0,5 % (p/v) contenant 0,5 µg.mL<sup>-1</sup> de BET. La migration se fait en tampon TAE 1 x à 8 V.cm<sup>-1</sup> pendant 40 min. La bande de 0,9 kpb générée par la digestion est extraite du gel avec le kit Nucleo Spin Extract (Macherey-Nagel) puis marquée en utilisant le kit ECL RP Labelling System (Amersham) comme expliqué par les fabricants.

La membrane préparée précédemment est hybridée avec les sondes « PAV1 » et « 16S GE 23 » puis révélée en utilisant le kit ECL Detection System (Amersham) pour impressionner un film autoradiographique en suivant les instructions du fournisseur. Le film est scanné en utilisant le logiciel Adobe Photoshop puis l'intensité des bandes obtenues est ensuite mesurée avec le logiciel NIH Image. Les intensités sont corrigées de manière à tenir compte de la différence de taille des sondes « PAV1 » et « 16S GE 23 », respectivement de 1,7 kpb et 0,9 kpb. Le rapport intensité « PAV1 » / intensité « 16S GE23 » permet ensuite d'obtenir le nombre de copies du génome de PAV1 par chromosome archéen selon la formule suivante :

$$\text{nombre de copies du génome viral} = \frac{\text{intensité « PAV1 »} \times \text{taille sonde « 16S GE23 »}}{\text{taille sonde « PAV1 »} \times \text{intensité « 16S GE23 »}}$$

## 2. Détermination du surenroulement de l'ADN viral

L'ADN de PAV1 est incubé avec différentes topoisomérasés pour déterminer leur capacité à relâcher le génome viral. Le résultat de ces incubations est visualisé grâce à des migrations électrophorétiques en une et deux dimensions sur des gels d'agarose.

L'ADN de PAV1 est extrait par lyse alcaline puis mélangé avec le plasmide bactérien pTZ18R (témoin de surenroulement négatif). Les deux ADN sont incubés avec trois topoisomérasés différentes. La topoisomérase A d'*Escherichia coli* (NEB) et topoisomérase I eucaryote (Invitrogen) sont utilisées suivant les recommandations des fournisseurs. L'incubation dure 30 min à 37°C. Pour la reverse gyrase de *Sulfolobus*, les 20 µL du mélange réactionnel contiennent: 4 µL de tampon de l'enzyme (5 x) 2 µL de MgCl<sub>2</sub> (100 mM), 1,9 µL de NaCl (1

M), 0,5  $\mu\text{L}$  d'ATP (50 mM), 1  $\mu\text{L}$  de reverse gyrase purifiée, 0,2  $\mu\text{L}$  de pTZ, 5  $\mu\text{L}$  d'ADN de PAV1 extrait par lyse alcaline et 5,4  $\mu\text{L}$  d'eau pour HPLC. L'incubation dure 30 min à 75°C. Les mélanges réactionnels sont déposés, après incubation avec les différentes topoisomérases, sur un gel d'agarose à 0,5 % (p/v). La migration en une dimension se fait en tampon TAE 1 x à 2  $\text{V}\cdot\text{cm}^{-1}$  pendant 14 h à 16°C. Le gel est ensuite coloré pendant 40 min dans un bain de BET à 2  $\mu\text{g}\cdot\text{mL}^{-1}$ , puis pendant 15 min dans un bain de  $\text{MgSO}_4$  à 1 mM.

Pour la migration en deux dimensions, le protocole suivi pour les incubations avec la topoisomérase I eucaryote (Invitrogen) ainsi qu'avec la reverse gyrase de *Sulfolobus* est le même que précédemment à ceci près que pTZ18R n'est pas ajouté au mélange réactionnel et que la quantité d'ADN de PAV1 traitée est trois fois supérieure. Les mélanges réactionnels sont déposés, après incubation avec les différentes topoisomérases, sur un gel d'agarose à 0,5 % (p/v). La migration dans la première dimension se fait en tampon TAE 1 x à 2  $\text{V}\cdot\text{cm}^{-1}$  pendant 19 h à 16°C. Le gel est tourné d'un quart de tour sur la droite et plongé pendant 1 h 30 dans du tampon TAE 1 x contenant 0,013  $\mu\text{g}\cdot\text{mL}^{-1}$  de BET. Ce tampon est utilisé pour la migration dans la deuxième dimension qui se fait à 2  $\text{V}\cdot\text{cm}^{-1}$  pendant 18 h à 16°C. Le gel est ensuite coloré pendant 40 min dans un bain de BET à 2  $\mu\text{g}\cdot\text{mL}^{-1}$ , puis pendant 15 min dans un bain de  $\text{MgSO}_4$  à 1 mM.

### **3. Recherche de méthylations sur le génome viral et le chromosome archéen**

Après extraction, l'ADN de PAV1 et l'ADN total de la souche GE23 de *Pyrococcus abyssi* sont digérés séparément par quatre enzymes de restriction (*DpnI*, *MboI*, *BstNI* *ScrFI*) suivant les instructions du fournisseur (NEB). Les produits de digestion sont déposés sur gel d'agarose à 0,5 % (p/v) contenant 0,5  $\mu\text{g}\cdot\text{mL}^{-1}$  de BET. La migration se fait en tampon TAE 1 x à 3  $\text{V}\cdot\text{cm}^{-1}$  pendant 4 h.

### **4. Recherche d'un intermédiaire de réplication simple brin**

L'ADN de PAV1 est extrait et transféré sur membrane de nitrocellulose comme expliqué précédemment. Des sondes simple brin, spécifiques de chaque brin du génome viral sont synthétisées. Une PCR est réalisée sur l'ADN de PAV1 comme décrit précédemment en utilisant le couple d'amorces F1 (5' TGG TGT TCA TTA TCC TCG GC 3') / R1 (5' CTC CTC CAA TTA CCA AGT ATC 3'). Le produit d'amplification est cloné dans le vecteur pGEM<sup>®</sup>-T (Promega) puis utilisé pour transformer des cellules compétentes *E. coli* XL10-Gold (Stratagene) par choc thermique (30 s à 42°C). Le produit de clonage est extrait par lyse alcaline

puis digéré enzymatiquement soit par *Sall*, soit par *SphI* (NEB) suivant les instructions du fabricant. Les deux produits de digestion, correspondant respectivement aux sondes « brin + » et « brin - » de PAV1, sont alors marqués avec le kit DIG RNA Labelling (Roche) en suivant le protocole du fournisseur.

La membrane préparée précédemment est hybridée successivement avec la sonde « brin + » et la sonde « brin - » puis révélée avec le kit DIG Nucleic Acid Detection (Roche) comme expliqué par le fabricant. Entre les deux hybridations, la membrane est déshybridée en utilisant un bain de diméthylformamide chauffé à 55°C jusqu'à décoloration suivi de deux bains de 20 min dans la solution A (NaOH 0,2 N, SDS 0,1 % (p/v)) chauffée à 37°C.

### 5. Cartographie des transcrits du virus

Les ARNm totaux de la souche *Pyrococcus abyssi* GE23 sont extraits en utilisant du TRIzol<sup>®</sup> (Invitrogen) suivant le protocole décrit par le fournisseur. Ceci est réalisé sur 50 mL d'une culture en fin de phase exponentielle ( $2,5 \cdot 10^8$  cellules.mL<sup>-1</sup>). Les culots d'ARN sont resuspendus dans 35 µL de Tris-HCl 5 mM pH 8. La concentration et la pureté des ARN extraits sont déterminées en mesurant les absorbances à 260 et 280 nm. La qualité est estimée en déposant un aliquot sur gel d'agarose à 1,2 % (p/v). L'ADN contaminant la préparation est éliminé par un traitement à la turbo DNase (Promega) en suivant les instructions du fournisseur.

Des ADNc sont alors synthétisés. Les 12 µL de mélange réactionnel contiennent : 1 à 2 µg d'ARN et 1 µL de la solution de nonamères du kit ECL RP Labelling (Amersham), complétés par de l'eau milliQ. Une dénaturation est pratiquée à 78°C pendant 3 min. Au mélange réactionnel précédent sont ajoutés : 2 µL de tampon de la reverse transcriptase (10x), 4 µL de dNTP (2,5 mM), 1 µL d'inhibiteur de RNase (40 U/µL) (Ambion) et 1 µL de reverse transcriptase M-MLV (100 U/µL) (Ambion). Le témoin négatif est réalisé de la même façon avec de l'eau milliQ qui remplace la reverse transcriptase. La réaction de réverse transcription est réalisée avec le programme suivant : 60 min à 44°C puis 10 min à 92°C.

Le mélange réactionnel de PCR est réalisé suivant les instructions du fournisseur (Interchim) en y introduisant 2 µL de produit de RT ou 2 µL d'ADN total de la souche GE23 *Pyrococcus abyssi* (témoin positif). Le détail des couples d'amorces utilisés figure dans le tableau 4. Le programme est organisé de la façon suivante : 3 min de dénaturation à 94°C, 30 cycles d'amplification (30 s à 94 C, 30 s à 53°C et 1 min 30 à 70°C) et 7 min d'extension finale à 70°C. Pour contrôler l'amplification, les produits de PCR sont déposés sur un gel

d'agarose 0,8 % (p/v) contenant 0,5  $\mu\text{g.mL}^{-1}$  de BET. La migration se fait en tampon TAE 1 x à 8  $\text{V.cm}^{-1}$  pendant 40 min.

**Tableau 4 :** Séquences des différents couples d'amorces dessinés sur le génome de PAV1.

<b>1S / 1Ra</b> : 5' TGG CTC AAC GAG GAA GAC GC 3'	5' TAG TCG CCA CGG TTG TCC TC 3'
<b>1S / 1Rb</b> : 5' TGG CTC AAC GAG GAA GAC GC 3'	5' CGG TAG GAG CAC TCT TCA GAA TAC 3'
<b>2S / 2R</b> : 5' CCT TCT TGG GAT TGC TGA CT 3'	5' CCA AGT GCC TTC CTA ATC TTC 3'
<b>3S / 3R</b> : 5' ACA AGT ACA CCT TGG TAA TGC CTT 3'	5' CCA CAG GAT CTG GAC ATA ACC A 3'
<b>4S / 4R</b> : 5' TTT CTG GAC TTC TTT CGA GCA 3'	5' CGT TAC AAT ACC GCT CAC AGA A 3'
<b>5S / 5R</b> : 5' CAG CAG GTC TCA GGC TTA TAA TT 3'	5' CGT TCA CCA TAT GCA GAA TCA C 3'
<b>6.2S / 6.2R</b> : 5' CTT CAC GAG GTG GCT TGG TG 3'	5' AAG AGA GCC ACA AGC CAG CC 3'
<b>7S / 7R</b> : 5' TTC CTC TTG GCG AAG TTC AG 3'	5' GAG CAT TCT TTA TTT TCT CCC C 3'
<b>8S / 8R</b> : 5' CTT AGA GGG GGC TCA TTG GAT A 3'	5' CGT CTG TGG ATC CCT CGT TA 3'
<b>9S / 9R</b> : 5' GTA CTG GTT CGT GAG GTT CTT CA 3'	5' TCT CAA CAT ACT TGT TCC ATC TGG 3'
<b>10S / 10R</b> : 5' ACA TCG ATT TCA GAA CCT CGG 3'	5' CCC CAC GTT CTG ATA CAT CA 3'
<b>11S / 11R</b> : 5' CCG ACG AGA GAT TGA AGT TCT 3'	5' CGA GAG AGA AGA ACT AAA CGA ACT 3'
<b>12S / 12R</b> : 5' AGG AAA TAG GTA CAC CGA GAT TC 3'	5' TGA GAT CCC TTA GAT AAT CTG CAA 3'
<b>13S / 13R</b> : 5' CTT GGT TCT CGG GTC GGG TT 3'	5' TAG TAT CGC TAG GCG CTC GC 3'
<b>14S / 14R</b> : 5' TGG AGA AAC GCG TGT TTA TG 3'	5' AAG TTC GAG TAT TTC ATC CCA GT 3'
<b>15S / 15R</b> : 5' TCA AGC GCA TTA AGA AGA GG 3'	5' CCT TCC CCA AAT ACA CTA AT 3'
<b>16S / 16R</b> : 5' AAA GCC TAG GAC AAT GGA GG 3'	5' TTT CTC CGC GTT CTT GAA TA 3'
<b>17S / 17R</b> : 5' CCC ACT TGT CGA CTA GAT CCC 3'	5' CCT GGC GGG AAG TAT TCA AT 3'
<b>18S / 18R</b> : 5' GTG TTT GGT AAT ATA AGC CAG TTG 3'	5' GCA TTT CTC CTC AAT GGG AA 3'
<b>19S / 19R</b> : 5' TTG GGG GAA GAG TGC TGG CA 3'	5' CCG CGC GGA GAA ACT CTA CT 3'
<b>20S / 20R</b> : 5' TGT CGT TGC TGA GGC TGA AGT T 3'	5' CCT TCC CCA AAT ACA CTA AT 3'

## VI. Etude des protéines du virus PAV1

### 1. Electrophorèse de protéines en conditions dénaturantes (SDS-PAGE) et coloration au bleu de Coomassie

Tous les profils de migration de protéines de PAV1 sont obtenus par électrophorèse dite en SDS-PAGE sur gel d'acrylamide. Un aliquot de 20 µL d'échantillon est mélangé avec 5 µL de tampon de charge (0,5 % (p/v) bleu de bromophénol, 40 % (p/v) sucrose, 10 % (p/v) SDS, 5 % (p/v) β-mercaptoethanol, 0,31 M Tris-HCl pH 6,8) puis dénaturés pendant 5 min à 100°C. Il est ensuite déposé sur un gel d'acrylamide composé d'un gel de concentration (à 3 %) et d'un gel de séparation (à 7 %, 12 % ou 15 % suivant la taille des protéines analysées). La migration électrophorétique se fait dans un tampon dénaturant (25 mM Tris, 19 mM glycine, 1 % (p/v) SDS) pendant 45 min en ampérage constant (0,02 A).

Après électrophorèse, le gel est coloré 1 h dans une solution de bleu de Coomassie (0,1 % (p/v) bleu de Coomassie, 50 % (v/v) éthanol, 10 % (v/v) acide acétique). Il est ensuite décoloré dans une solution contenant 30 % (v/v) d'éthanol et 10 % (v/v) d'acide acétique jusqu'à l'obtention de bandes nettes. Puis le gel est séché pendant 2 h dans un sécheur de gel Model 583 (Bio-Rad).

### 2. Surproduction de la protéine 899 de PAV1

#### a. Couples d'amorces et vecteurs d'expression utilisés

Deux types de vecteurs d'expression sont sélectionnés pour exprimer le gène 899 codant pour la protéine du même nom : les vecteurs pET 11d, 25b+, 28a+ (Novagen) inductibles à l'IPTG et le vecteur pBAD (Invitrogen) inductible à l'arabinose. Plusieurs couples d'amorces sont dessinés pour varier les constructions.

**Tableau 5** : Détail des constructions sélectionnées.

<b>Vecteur utilisé</b>	<b>Couple d'amorces utilisés</b> (en haut : amorce sens, en bas : amorce anti-sens)	<b>Protéine attendue</b>
pET 11d Amp <sup>R</sup>	For Nco I : 5' GGAGGGAGCGCCATGGTTCGGGTAATGG 3' Rev Bam HI : 5' AATTGCAAGGATCCCAAAGACTGCG 3'	Sans étiquette
pET 25b+ Amp <sup>R</sup>	For Nco I : 5' GGAGGGAGCGCCATGGTTCGGGTAATGG 3' Rev Bam HI : 5' AATTGCAAGGATCCCAAAGACTGCG 3'	Peptide signal en N terminal
pET 28a+ Kan <sup>R</sup>	For Nco I : 5' GGAGGGAGCGCCATGGTTCGGGTAATGG 3' Rev Xho I : 5' TGAAAGGCCTTGAGGCTCGAGTATCCTGAAAAGCGGC 3'	Queue histidine en C terminal
pBAD Amp <sup>R</sup>	899-1S : 5' GGTGCCATGGTTCGGGTAATGGATAAGA 3' 899-1AS : 5' TATCCTGAAAAGCGGCTCATA 3'	Queue histidine en C terminal
pBAD Amp <sup>R</sup>	899-1S : 5' GGTGCCATGGTTCGGGTAATGGATAAGA 3' 899-2AS : 5' TTATATCCTGAAAAGCGGCTC 3'	Sans étiquette

Amp<sup>R</sup> : résistance à l'ampicilline, Kan<sup>R</sup> : résistance à la kanamycine

### **b. Clonage du gène 899 en vecteur d'expression**

#### ➤ Amplification par PCR du gène 899

L'ADN polymérase thermostable choisie est la Pfu (Promega) car elle possède une meilleure activité de correction des erreurs que la Taq polymérase (Promega). Le mélange réactionnel de PCR est réalisé suivant les instructions du fournisseur (Promega) en y introduisant 10 ng d'ADN total de la souche GE23 de *Pyrococcus abyssi* (porteuse de PAV1) extrait comme expliqué dans le paragraphe *IV.2.b*. Le programme utilisé est composé des phases suivantes : 2 min de dénaturation à 95°C, 30 cycles d'amplification (1 min à 95°C, 30 s à 56°C et 4 min à 72°C) et 5 min d'extension finale à 70°C.

Une adénine est ensuite ajoutée à l'extrémité des produits de PCR obtenus précédemment. Pour cela, ces derniers sont incubés 20 min à 70°C avec mélange réactionnel suivant : 250 µM de MgCl<sub>2</sub>, 200 µM de dATP, 1 U de Taq polymérase (Promega) et 200 ng de produit de PCR.

#### ➤ Clonage du gène 899 en vecteur d'expression pBAD et en vecteur de maintenance pCR-TOPO-XL

Les produits de PCR additionnés d'une adénine sont soit clonés directement dans le vecteur d'expression pBAD, soit clonés dans le vecteur de maintenance pCR-TOPO-XL (Invitrogen) avant leur introduction dans les différents vecteurs pET. Les deux types de constructions sont ensuite utilisés pour transformer des cellules compétentes *Escherichia coli* Sure<sup>®</sup> (Stratagene) par choc thermique (30 s à 42°C). Pour contrôler le clonage, une extraction des constructions par lyse alcaline est pratiquée comme décrit précédemment (Birnboim & Doly, 1979) sur 10 clones différents, puis déposée sur un gel d'agarose à 0,5 % (p/v) contenant 0,5 µg.mL<sup>-1</sup> de BET. La migration se fait en tampon TAE 1x à 8 V.cm<sup>-1</sup> pendant 40 min. Les clones présentant la taille attendue sont séquencés par la société GenomExpress et la séquence obtenue est comparée à celle du gène 899 de PAV1.

#### ➤ Clonage du gène 899 en vecteur s d'expression pET

Les vecteurs d'expression pET non recombinants et les vecteurs de maintenance pCR-XL-TOPO-TA contenant le gène 899 sont digérés par 12 U d'enzyme de restriction adéquate (*NcoI*, *BamHI* et *XhoI*) en suivant les instructions du fournisseur (NEB). Les vecteurs d'expression pET non recombinants sont déphosphorylés en utilisant 3 U de phosphatase alcaline de crevette (Promega) en suivant le protocole du fabricant. La totalité des produits de digestion est déposée sur un gel d'agarose à 0,8 % (p/v) contenant 0,5 µg.mL<sup>-1</sup> de BET. La

migration se fait en tampon TAE 1x à  $8 \text{ V.cm}^{-1}$  pendant 40 min. Les bandes correspondant aux vecteurs d'expression pET non recombinants linéarisés et au gène 899 digéré sont découpées puis purifiées avec le kit Nucleo Spin Extract (Macherey-Nagel) en suivant les instructions du fournisseur.

Le gène 899 digéré est alors introduit par ligation dans le vecteur d'expression pET linéarisé possédant les extrémités cohésives correspondantes, en utilisant le kit LigaFast™ Rapid DNA Ligation System (Promega) suivant le protocole du fabricant. Les produits de ligation sont ensuite utilisés pour transformer des cellules compétentes *Escherichia coli* Sure® (Stratagene) par choc thermique (30 s à  $42^\circ\text{C}$ ). La suite des opérations jusqu'au contrôle des clones par séquençage est rigoureusement la même que pour le clonage en vecteur d'expression pBAD.

### c. Surexpression du gène 899

Trois types de cellules *Escherichia coli* sont testés pour l'expression du gène 899 de PAV1. Les cellules BL<sub>21</sub> Rosetta® (Novagen) ainsi que les cellules HMS 174 (Novagen) sont utilisées pour les constructions en vecteurs pET car dans leur génome est intégré le gène codant pour l'ADN polymérase du phage T7 qui reconnaît le promoteur porté par ce type de vecteur. Par ailleurs, les BL<sub>21</sub> Rosetta® possèdent en plus le plasmide Rosetta® codant pour des ARNts rares. Celui-ci permet de faciliter la production de la protéine hétérologue en évitant les arrêts de traduction dus à la présence de codons rares dans la séquence du gène à surexprimer.. Les cellules SURE® (Stratagene) sont utilisées pour les constructions en vecteur pBAD. Les constructions vérifiées par séquençage ainsi que les vecteurs d'expression non recombinants pET et pBAD sont utilisés pour transformer les différentes cellules d'expression par choc thermique (30 s à  $42^\circ\text{C}$ ).

La surexpression du gène 899 est ensuite testée dans les différentes cellules d'expression. Pour cela, 4 mL de milieu LB (1 % (p/v) NaCl, 1 % (p/v) tryptone, 0,5 % (p/v) extrait de levure, pH 7) additionné d'antibiotique ( $50 \mu\text{g.mL}^{-1}$  d'ampicilline ou  $50 \mu\text{g.mL}^{-1}$  de kanamycine) sont inoculés à 20 % (v/v) soit avec une souche portant le vecteur non recombinant (témoin négatif), soit avec une souche transformée par le vecteur recombinant contenant l'*orf* 899. L'incubation se fait à  $37^\circ\text{C}$  sous agitation toute la nuit. Ces précultures servent à inoculer à 0,1 % (v/v) 10 mL de milieu LB additionné d'antibiotique ( $50 \mu\text{g.mL}^{-1}$  d'ampicilline ou  $50 \mu\text{g.mL}^{-1}$  de kanamycine) qui sont incubés à  $37^\circ\text{C}$  sous agitation. Lorsque l'absorbance à 600 nm des cultures atteint la valeur de 0,6, la surexpression de l'*orf* 899 est

induite par 1 mM d'IPTG (pour les vecteurs pET) ou avec 0,002 % (p/v) d'arabinose (pour le vecteur pBAD). Des prélèvements de 2 mL sont réalisés aux temps 0 et 2 h après l'induction.

#### **d Séparation des fractions contenant les protéines solubles et insolubles**

Chaque échantillon est centrifugé à 4 000 g pendant 10 min à 4°C. Le culot est repris dans 300 µL de tampon A (50 mM Tris-HCl pH 8, EDTA 0,1 mM, 2 mM β-mercaptoéthanol, 10 % (p/v) glycérol) puis soniqué deux fois (15 s et 20 s) à 40 W. Après centrifugation à 18 000 g pendant 15 min à 4°C. Le culot de sonication est repris dans 100 µL de tampon A. Il correspond à la fraction riche en membranes et en protéines insolubles. Les protéines non thermostables sont éliminées du surnageant après chauffage à 75°C pendant 15 min et centrifugation à 18 000 g pendant 15 min à 4°C. Le surnageant obtenu correspond à la fraction riche en protéines solubles. Les fractions de chaque échantillon sont alors visualisées après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie.

### **3. Essais d'amélioration de la production de la protéine 899 sous sa forme soluble**

#### **a. Abaissement de la température de culture**

Le protocole utilisé est le même que dans le paragraphe *Vogel, 2000*, à la différence près que la température de culture des cellules est modifiée. Trois températures sont testées : 30°C, 26°C et 18°C. Les différentes fractions sont visualisées après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie.

#### **b. Co-production de la protéine 899 avec des chaperons**

Le plasmide pG-KJE3 (portant la résistance au chloramphénicol) inductible à l'arabinose code pour 5 protéines chaperons d'*Escherichia coli* (Trésaugues *et al.*, 2004). Il est utilisé pour transformer par choc thermique (30 s à 42°C) des cellules HMS 174 portant soit le vecteur pET 28a+ non recombinant, soit le vecteur pET 28a+ contenant le gène 899.

Quatre mL de milieu LB additionné d'antibiotiques (50 µg.mL<sup>-1</sup> de kanamycine et 34 µg.mL<sup>-1</sup> de chloramphénicol) sont inoculés à 20 % (v/v) avec des cellules HMS 174 portant soit le plasmide pG-KJE3 ainsi que le vecteur pET 28a+ non recombinant (témoin négatif), soit le plasmide pG-KJE3 ainsi que le vecteur pET 28 a+ contenant le gène 899. L'incubation se fait à 37°C sous agitation toute la nuit. Ces précultures servent à inoculer à 0,1 % (v/v) 10 mL de milieu LB additionné d'antibiotiques (50 µg.mL<sup>-1</sup> de kanamycine et 34 µg.mL<sup>-1</sup> de chloramphénicol) qui sont incubés à 37°C sous agitation. Lorsque que l'absorbance à 600 nm

des cultures atteint la valeur de 0,5, la surexpression du gène 899 et des gènes encodant les protéines chaperons est induite par 1 mM d'IPTG et 0,2 % (p/v) d'arabinose. Les cultures sont ensuite incubées à 25°C sur la nuit. Les fractions riches en protéines solubles et insolubles sont séparées et collectées suivant le protocole décrit dans le paragraphe *Vof.z.c.* Elles sont alors visualisées après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie.

### **c. Production *in vitro* de la protéine 899**

Le système d'expression *in vitro*, Express Way Plus Expression System (Invitrogen), est utilisé pour surexprimer le gène 899 en suivant les recommandations du fournisseur. Les plasmides utilisés sont : le vecteur pET 28a+ non recombinant (témoin négatif), le vecteur pEX1-GW/*lacZ* (témoin positif du kit) et le vecteur pET 28a+ contenant le gène 899. Les protéines exprimées sont visualisées après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie.

## **4. Extraction et solubilisation des corps d'inclusion**

Trois litres de culture de cellules HMS 174 portant le vecteur pET 28a+ contenant l'*orf* 899 sont mis en culture. La surexpression du gène 899 est induite par 1 mM d'IPTG. La lyse cellulaire est réalisée en reprenant le culot cellulaire dans du tampon de lyse (100 mM Tris-HCl pH 7, 5 mM EDTA, 5 mM DTT, 5 mM hydrochloride de benzimide) puis en le passant deux fois à la presse de French (10 000 psi).

Les débris cellulaires associés aux corps d'inclusion sont éliminés par deux lavages dans le tampon 1 (100 mM Tris-HCl pH 7, 5 mM EDTA, 5 mM DTT, 2 M urée, 2 % p / v triton X-100) suivi d'un lavage dans le tampon 2 (100 mM Tris-HCl pH 7, 5 mM EDTA et 5 mM DTT). Les corps d'inclusion sont alors solubilisés par ajout du tampon d'extraction (50 mM Tris-HCl pH 7, 5 mM EDTA, 8 M hydrochloride de guanidine, 5 mM DTT) suivie d'une centrifugation à 100 000 g pendant 1 h à 4°C.. Le surnageant est récupéré puis filtré sur 0,22 µm. Le filtrat contient les protéines des corps d'inclusion solubilisés qui sont visualisées après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie.

## **5. Renaturation de la protéine 899**

### **a. Recherche d'un tampon de renaturation**

Afin de définir les conditions de renaturation de la protéine 899 solubilisée dans l'hydrochloride de guanidine, 80 tampons différents sont testés en microplaque d'après le

protocole de R. Vincentelli et ses collaborateurs (Vincentelli *et al.*, 2004). La composition de ces tampons est résumée dans le tableau 6. La protéine 899 issue de la solubilisation des corps d'inclusion est diluée au 1/20<sup>ème</sup> dans chaque tampon. Une mesure d'absorbance à 340 nm est réalisée à t = 0 et à t = 52 h sur un lecteur de plaque Multiskan MS (Labsystems).

**Tableau 6 :** Composition des tampons de renaturation testés (d'après Vincentelli *et al.*, 2004).

	1	2	3	4	5	6	7	8	9	10
A	$\beta$ -MSH Arg	KCl $\beta$ -MSH NDSB256	NaCl200	KCl Glu	KCl	-	NDSB201 Arg	KCl $\beta$ -MSH NDSB201 Arg	$\beta$ -MSH	Gly NDSB201
B	$\beta$ -MSH NDSB256	$\beta$ -MSH NDSB201 Glu	NDSB256	Cocktail	NDSB256 Glu	KCl	PEG4000 $\beta$ -MSH	Glu	NDSB201	NaCl100 Glu
C	NaCl100 $\beta$ -MSH	KCl NDSB195	EDTA $\beta$ -MSH NDSB201	PEG400 $\beta$ -MSH NDSB201	Gly $\beta$ -MSH	NaCl200 NDSB201	NDSB195 Glu	-	PEG4000 $\beta$ -MSH Glu	$\beta$ -MSH NDSB195 Arg
D	Gly	PEG4000 Glu	$\beta$ -MSH Gly	$\beta$ -MSH NDSB195 Glu	EDTA	Gly $\beta$ -MSH	EDTA $\beta$ -MSH NDSB195	PEG4000 NDSB195	KCl $\beta$ -MSH NDSB201 Glu	PEG400
E	PEG400 Glu	EDTA Arg	PEG400 $\beta$ -MSH Arg	Gly NDSB256 Arg	NaCl100 $\beta$ -MSH NDSB195	PEG4000 NDSB256 Arg	Cocktail	$\beta$ -MSH Glu	$\beta$ -MSH	-
F	EDTA	NaCl100 Arg	$\beta$ -MSH NDSB256	NaCl200 $\beta$ -MSH Glu	-	NaCl100 NDSB201	Arg	NaCl100 $\beta$ -MSH NDSB256	NaCl100 NDSB256	EDTA $\beta$ -MSH Arg
G	NDSB201	PEG4000 $\beta$ -MSH NDSB201	Gly $\beta$ -MSH NDSB195	$\beta$ -MSH NDSB201 Arg	PEG4000	$\beta$ -MSH Arg	EDTA NDSB256 Glu	NaCl200 $\beta$ -MSH Glu	$\beta$ -MSH NDSB195	NDSB195 Arg
H	NaCl200 $\beta$ -MSH NDSB195 Arg	NaCl200 NDSB195	$\beta$ -MSH	PEG400 NDSB256	PEG400 $\beta$ -MSH NDSB256 Glu	$\beta$ -MSH	Gly $\beta$ -MSH	PEG400 NDSB195	NaCl200 $\beta$ -MSH NDSB256	Cocktail

**Légende :**

pH 4 : base de 50 mM acétate de sodium    pH 5 : base de 50 mM MES    pH 6 : base de 50 mM MES  
 pH 7 : base de 50 mM Tris    pH 8 : base de 50 mM Tris    pH 9 : base de 50 mM CHES  
 NaCl100  $\Leftrightarrow$  NaCl 100 mM    NaCl200  $\Leftrightarrow$  NaCl 200 mM    KCl  $\Leftrightarrow$  KCl 100 mM    Gly  $\Leftrightarrow$  glycérol 20% (v/v)  
 PEG400  $\Leftrightarrow$  PEG 400 0,05 % (p/v)    PEG4000  $\Leftrightarrow$  PEG 4000 0,05 % (p/v)     $\beta$ -MSH  $\Leftrightarrow$   $\beta$ -mercaptoéthanol 10 mM  
 Arg  $\Leftrightarrow$  arginine 800 mM    Glu  $\Leftrightarrow$  glucose 500 mM    EDTA  $\Leftrightarrow$  EDTA 1 mM  
 NDSB 195, NDSB 201, NDSB 256  $\Leftrightarrow$  sulfobétaïnes 10 mM  
 Cocktail  $\Leftrightarrow$  NADH, thiamine.HCl, biotine, CaCl<sub>2</sub>, MgCl<sub>2</sub>, CuSO<sub>4</sub>, ZnCl<sub>2</sub>, CoSO<sub>4</sub>, ADP, NiCl<sub>2</sub> (50  $\mu$ m chacun)

### **b. Dialyse de la protéine 899**

L'hydrochloride de guanidine qui permet de solubiliser la protéine 899 est échangé par dialyse sur la nuit à 4°C contre les 4 tampons de renaturation apparaissant comme les plus efficaces : le tampon 1H (50 mM acétate de sodium pH 4, 200 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 10 mM NDSB 195, 800 mM arginine), le tampon 2H (50 mM MES pH 5, 200 mM NaCl, 10 mM NDSB 195), le tampon 7H (50 mM Tris pH 8, 20 % (v/v) glycérol, 10 mM  $\beta$ -mercaptoethanol) et le tampon 9H (50 mM CHES, 200 mM NaCl, 10 mM  $\beta$ -mercaptoethanol, 10 mM NDSB 256). Après une centrifugation de 35 min à 13 000 g, le surnageant est récupéré puis stocké à 4°C.

## **6. Essais de mise en évidence d'une implication de la protéine 899 dans le métabolisme de l'ADN viral**

### **a. Recherche d'une activité ADN polymérase chez la protéine 899**

Le test d'activité réalisé avec la protéine 899 est un test d'incorporation de thymine tritiée sur une matrice ADN simple brin polyamorcée poly(dA)<sub>265</sub>/oligo(dT)<sub>17</sub> (10 : 1). Les 20  $\mu$ L du mélange réactionnel utilisé contiennent : 0,5  $\mu$ L de dNTP (10 mM), 0,8  $\mu$ L de [<sup>3</sup>H]dTTP (5  $\mu$ M, 5 Ci.mmol<sup>-1</sup>), 0,5  $\mu$ g de matrice poly(dA)<sub>265</sub>/oligo(dT)<sub>17</sub> (10 : 1), 2  $\mu$ L de tampon de réaction de la Taq polymérase Isis (10 x) et 15,7  $\mu$ L d'eau milliQ. Le témoin négatif correspond à ce mélange sans ajout d'enzyme. Le témoin positif est le mélange additionné de 0,5  $\mu$ L de Taq polymérase Isis (1 U/ $\mu$ L). En ce qui concerne la protéine 899, ce sont 2  $\mu$ L des fractions dialysées contre les quatre tampons (1H, 2H, 7H ou 9H) qui sont ajoutés au mélange décrit ci-dessus. Le test d'activité est réalisé à 50°C pendant 20 min. L'incorporation est ensuite évaluée en mesurant la radioactivité sur un compteur à scintillation 2000 CA (Packard).

### **b. Recherche d'interactants pour la protéine 899**

Les tests d'interaction sont réalisés en résonance plasmonique de surface sur le Biacore X. Cet appareil a pour fonction de visualiser en temps réel des interactions entre biomolécules non marquées dans un flot continu de tampon. Les molécules testées, ou ligands, sont retenues de manière spécifique sur une interface appelée biocapteur. Dans cette étude, trois molécules sont testées: de l'ADN simple brin, le facteur de processivité de l'ADN polymérase de *Pyrococcus abyssi* (appelé PCNA) et une endonucléase de *Pyrococcus abyssi* (Fen I). L'analyte, dans ce cas la protéine 899 dialysée contre le tampon 1H, est injecté dans un circuit microfluidique et circule à flux constant à la surface du biocapteur. Les changements de

masse induits par l'association ou la dissociation des molécules modifient la réfringence du milieu et décalent la position de l'angle de résonance. L'enregistrement de la variation de l'angle de résonance permet de suivre en temps réel la fixation de la protéine 899 injectée sur le biocapteur. Elle est représentée ensuite sur un sensorgramme où figure le signal de résonance (en RU ou unités de résonance) en fonction du temps.

### **7. Recherche de la protéine 899 dans son hôte naturel, la souche GE23 de *Pyrococcus abyssi***

Pour tenter de mettre en évidence la protéine 899 dans l'extrait cellulaire de son hôte, 2 mL d'une culture en fin de phase exponentielle de la souche GE23 de *Pyrococcus abyssi* sont prélevés et centrifugés à 4 000 g pendant 10 min à 4°C. Le culot est repris dans 300 µL de tampon A (50 mM Tris-HCl pH 8, EDTA 0,1 mM, 2 mM β-mercaptoéthanol, 10 % (p/v) glycérol) puis soniqué deux fois (15 s et 20 s) à 40 W. Après centrifugation à 18 000 g pendant 15 min à 4°C, le surnageant est récupéré. Il correspond à la fraction riche en protéines solubles. Le culot de sonication est repris dans 100 µL de tampon A. Il correspond à la fraction riche en membranes et en protéines insolubles.

Les différentes fractions et le témoin positif (qui correspond à la fraction comportant les protéines insolubles des cellules HMS 174 portant le vecteur pET 28a contenant l'*orf* 899) sont ensuite visualisés après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie. La protéine soluble, issue de l'extrait cellulaire de la souche GE23, qui présente la même migration électrophorétique que la protéine 899 recombinante est découpée puis analysée en spectrométrie de masse MALDI-TOF par le Plateau d'Analyses Protéomique de l'INRA de Jouy-en-Josas. Le profil tryptique obtenu en spectrométrie de masse est comparé au profil tryptique virtuel de la protéine 899.

### **8. Identification d'une protéine majoritaire de PAV1**

Les virions de PAV1 sont purifiés comme décrit par C. Geslin et ses collaborateurs (Geslin *et al.*, 2003). Les protéines issues des virions purifiés sont visualisées après électrophorèse en SDS-PAGE et coloration au bleu de Coomassie. La bande d'intérêt de 15 kDa est découpée puis micro-séquencée en N-terminal par la méthode d'Edmann au Plateau d'Analyses Protéomique de l'INRA de Jouy-en-Josas. La séquence obtenue est ensuite comparée à celle des différentes protéines du virus PAV1.

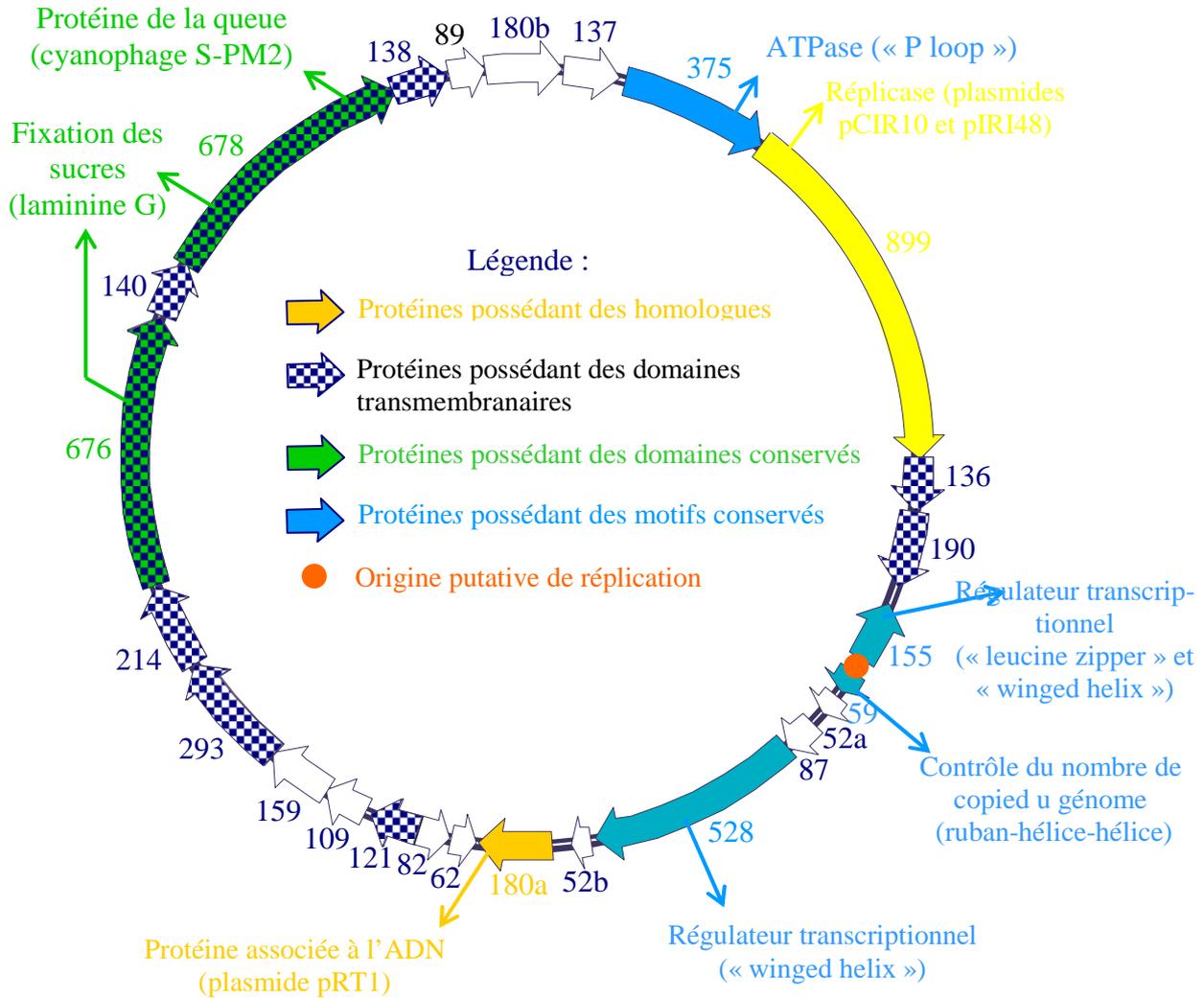
## Résultats et Discussion

### I. Préambule - Données apportées par l'analyse de la séquence du virus PAV1 (d'après Geslin C., Gaillard M., Flament D., Le Romancer M., Rouault K., Prieur D. & Erauso G., en préparation)

PAV1, premier virus hyperthermophile décrit, est porté par la souche GE23 de *Pyrococcus abyssi*. Il possède un génome à ADN double brin circulaire de 18,1 kpb qui est présent sous forme libre dans la cellule hôte (Geslin *et al.*, 2003). Après séquençage, une annotation a été réalisée et a permis de prédire la présence de 25 gènes dont 22 sont portés par le brin +. Le fait que la majorité soit orientée dans le même sens et qu'une seule origine putative de répllication ait été mise en évidence suggère que le génome viral pourrait adopter une répllication unidirectionnelle.

L'annotation de la séquence a permis de déterminer certaines fonctions putatives concernant les protéines de PAV1 (cf Figure 19). Une part importante (près de 40 %) des protéines virales ont des domaines transmembranaires prédits. Les protéines 676 et 678 possèdent en plus un ou deux domaines laminine G qui sont impliqués dans la fixation des sucres. Par ailleurs, la protéine 678 est homologue à une protéine de la queue du cyanophage marin S-PM2. Ces deux protéines sont donc probablement impliquées dans le processus de reconnaissance hôte - virus et sont peut-être constitutives des fibres caudales du virion. Ensuite, la protéine 180a présente 45 % de similarité avec la protéine 181a du plasmide pRT1 (isolé de la souche JT1 de *Pyrococcus abyssi*) qui est supposée être une protéine associée à l'ADN. La protéine 375 porte un motif « P-loop » retrouvés généralement chez les ATPases.

Une étude menée en parallèle au laboratoire sur les plasmides de *Thermococcales* a permis de révéler que la protéine 899 possède 30 % d'identité avec la protéine 3 des plasmides pCIR10 et pIRI48 portés par les souches du même nom (Erauso G., Gonnet M., Toffin A., Chevereau M. & Prieur D., en préparation). Cette protéine 3 est en opéron avec une hélicase, ce qui suggère une implication potentielle de la protéine 899 dans la répllication du génome viral. Par ailleurs, plusieurs régulateurs putatifs ont été annotés. Ainsi, la protéine 59 porte un motif ruban-hélice-hélice qui est retrouvé chez les protéines de la famille CopG. Ces dernières interviennent dans le contrôle du nombre de copies du génome chez certains plasmides. Les protéines 155 et 528 sont probablement des régulateurs transcriptionnels du fait de la présence de domaines de fixation à l'ADN de type « winged helix » et « leucine zipper ».



**Figure 19 :** Schéma du génome de PAV1 récapitulant les principales fonctions putatives assignées aux différentes protéines après l'annotation.

## II. Etude moléculaire du virus PAV1

PAV1 a été découvert dans le laboratoire en 2002 à la suite d'un criblage mené sur plus de 200 souches de *Thermococcales*. Une caractérisation morphologique, structurale et physiologique complète de ce virus a été réalisée (Geslin *et al.*, 2003). Afin de compléter la description de ce virus, nous avons poursuivi ce travail en nous intéressant aux propriétés moléculaires de PAV1. Les résultats livrés par l'analyse de la séquence du génome viral, décrits dans le paragraphe *f*, ont constitué une base solide aux travaux présentés ici.

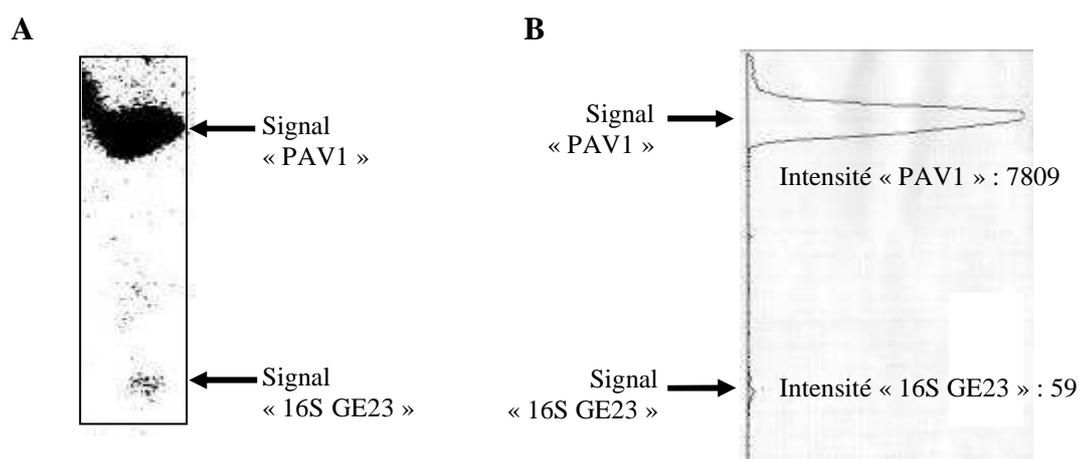
## 1. Etude du génome de PAV1

### a. Détermination du nombre de copies du génome viral dans la cellule

La technique d'extraction par lyse alcaline n'est utilisable que pour de petits ADN circulaires qui sont présents dans la cellule à fort nombre de copies. Etant donné l'intensité de la bande correspondant à l'ADN surenroulé de PAV1 qui est observable après une telle extraction, le génome viral doit sûrement exister sous forme multicopies dans son hôte. Pour déterminer un ordre de grandeur de ce nombre de copies par chromosome archéen, il est nécessaire que le gène archéen utilisé, en l'occurrence celui de l'ADNr 16S, ne soit présent qu'en une seule copie sur le chromosome de l'hôte.

Une sonde spécifique du virus (la sonde « PAV1 ») et une sonde spécifique de l'hôte (la sonde « 16S GE23 ») sont hybridées avec l'ADN total de la souche GE23 de *Pyrococcus abyssi*. Le signal obtenu est représenté sur la Figure 20A. A partir de ce signal, le logiciel NIH Image calcule les intensités correspondant aux signaux « PAV1 » et « 16S GE23 » qui sont notées sur la Figure 20B. Les intensités sont corrigées de manière à tenir compte de la différence de taille des sondes « PAV1 » et « 16S GE 23 », respectivement de 1,7 kpb et 0,9 kpb. Le rapport des intensités permet ensuite d'obtenir le nombre de copies du génome de PAV1 par chromosome archéen. Le calcul réalisé est détaillé ci-dessous :

$$\frac{7809 \times 0,9}{1,7} \div 59 = 70 \text{ copies}$$



**Figure 20 :** (A) Signal donné par l'hybridation des sondes « PAV1 » et « 16S GE23 » sur l'ADN total de la souche GE23 de *Pyrococcus abyssi*. (B) Profil généré et intensités calculées par le logiciel NIH Image à partir des signaux d'hybridation.

Cette étude nous a permis de confirmer le nombre important de copies de l'ADN viral dans la cellule tout en permettant une quantification de 70 par chromosome archéen. Aucune valeur n'est disponible dans la littérature en ce qui concerne les autres virus hyperthermophiles. En revanche, la valeur obtenue pour PAV1 est beaucoup plus élevée que ce qui a déjà été déterminé pour le plasmide hyperthermophile pGT5, porté par la souche GE5 de *Pyrococcus abyssi*. En effet, ce dernier ne comporte que 25 à 30 copies par chromosome archéen (Erauso, 1994).

La présence d'un nombre aussi important de copies du génome de PAV1 dans son hôte est peut-être due au fait que ce virus utilise son acide nucléique libre dans la cellule pour se maintenir dans la souche GE23. Ce phénomène a déjà été observé chez le coliphage P4 dans certains cas. En effet, l'ADN de ce dernier peut soit s'intégrer au chromosome d'*Escherichia coli*, soit persister sous une forme libre qui, en absence du phage « helper » P2, adopte le même fonctionnement qu'un plasmide et assure le maintien du virus (Briani *et al.*, 2001).

Ainsi, étant donné que les virions de PAV1 ne semblent pas infectieux, il est même possible d'envisager que la forme libre de son génome est en fait son principal mode de transmission. Le virus deviendrait ainsi progressivement défectif. Sa relation avec son hôte aurait petit à petit dérivé vers une situation où le virus adopte une façon rapide et simple pour se maintenir et bénéficie d'une protection vis-à-vis des rudesses du milieu extérieur. Etant donné que la souche GE23 ne semble pas du tout affectée par la présence de PAV1 (Geslin *et al.*, 2003), la relation hôte – virus tiendrait maintenant plus du commensalisme que du parasitisme.

Par ailleurs, la faible production de virions dans le milieu extérieur (au maximum 1 pour 500 cellules) (Geslin *et al.*, 2003) semble disproportionnée par rapport au nombre de copies du génome viral dans la cellule. Ceci suggère une régulation positive des fonctions précoces avec en parallèle une régulation négative très forte des fonctions d'encapsidation. Maintenir de telles régulations requiert la présence d'un certain nombre de régulateurs de la réplication et de la transcription qui pourraient être notamment les protéines encodées par les gènes 59, 155 et 528.

### b. Etude du surenroulement de l'ADN viral dans la cellule

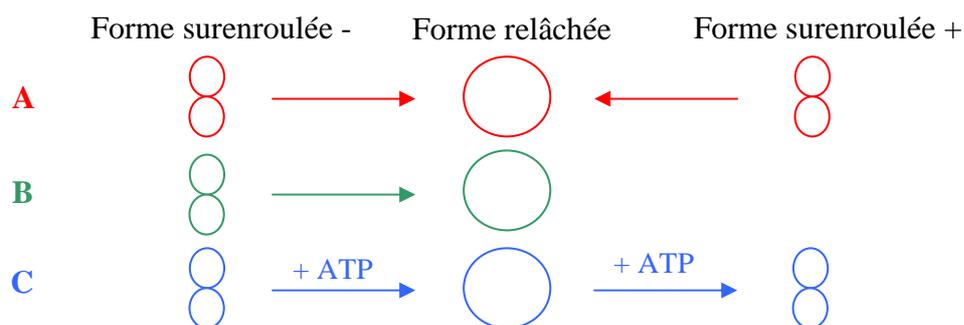
Une des caractéristiques physiques des génomes est le surenroulement. La principale relation existant en matière de topologie d'ADN double brin circulaire est la suivante :

$$L_K = \overline{T}_W + \overline{W}_R$$

Dans cette relation,  $L_K$  (« linking number ») est le nombre d'enlacements qui est constant pour une molécule donnée. Il correspond au nombre de passage minimal d'un des deux brins par l'autre nécessaire pour leur séparation.  $\overline{T}_W$  (« twist ») est le tortillement qui est égal au rapport du nombre de paires de bases sur le pas moyen de la double hélice d'ADN.  $\overline{W}_R$  (« writhe ») est le surenroulement qui correspond à un enroulement d'ordre supérieur de la double hélice (Bates & Maxwell, 1993). Il existe deux types de surenroulement. Ainsi, sous l'action de la gyrase, la gyration de la double hélice se fait vers la droite et crée un surenroulement négatif (Gellert *et al.*, 1976). A l'inverse, sous l'action de la reverse gyrase, la gyration a lieu vers la gauche et produit un surenroulement positif. La reverse gyrase n'a jusqu'alors été trouvée que chez les micro-organismes thermophiles et hyperthermophiles (Nakasu & Kikuchi, 1985).

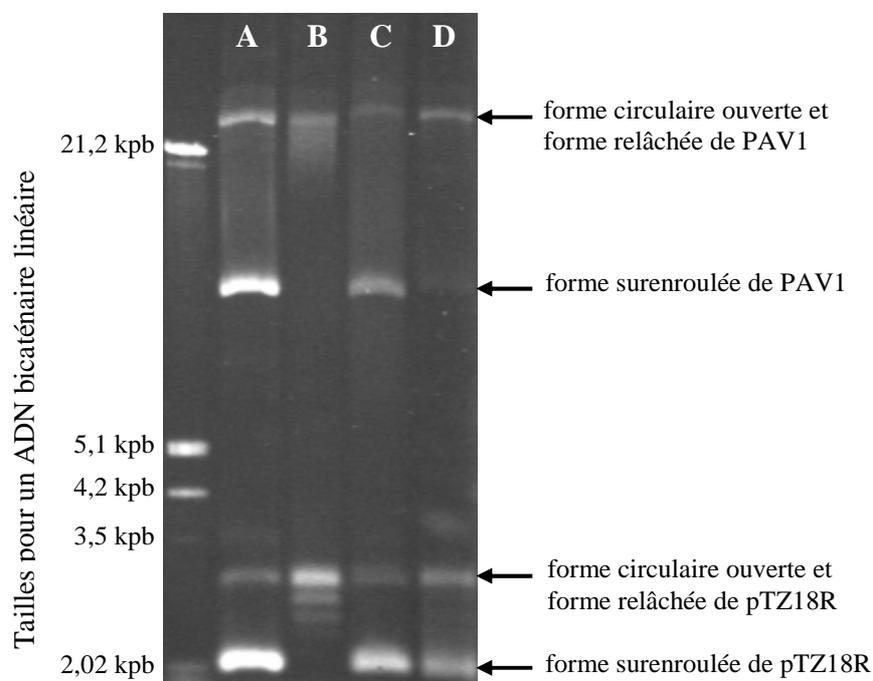
Très peu de données existent sur la topologie des génomes de virus hyperthermophiles. En réalité, seul le surenroulement du virus hyperthermophile terrestre SSV1 a été étudié. Dans le virion, le génome est fortement surenroulé positivement alors que dans la cellule il présente un spectre assez large de topologies (Nadal *et al.*, 1986). Les études menées sur les plasmides hyperthermophiles sont plus nombreuses et mettent en évidence la diversité des types de surenroulement retrouvés chez ces derniers. Ainsi, le plasmide pSL10, porté par *Desulfurolobus ambivalens*, comporte un surenroulement positif prononcé comme le virus SSV1 (Charbonnier & Forterre, 1994). pGT5, lui, possède un léger surenroulement négatif proche de l'état relâché (Charbonnier *et al.*, 1992). Enfin, le plasmide pGS5, retrouvé chez *Archaeoglobus profundus*, est fortement surenroulé négativement tout comme les plasmides de procaryotes mésophiles et les bactéries thermophiles (Lopez-Garcia *et al.*, 2000; Charbonnier & Forterre, 1994 ; Guipaud *et al.*, 1997).

Nous n'avons étudié que la topologie de la forme libre du génome de PAV1 car, comme nous l'avons montré précédemment, celle-ci correspond à la grande majorité des molécules d'ADN viral. Dans le but de déterminer la nature et le degré de surenroulement du génome de PAV1, ce dernier est incubé avec trois topoisomérases ayant des propriétés différentes qui sont résumées dans la figure 21.



**Figure 21 :** Mode d'action de la topoisomérase I eucaryote (A), de la topoisomérase A d'*Escherichia coli* (B) et de la reverse gyrase de *Sulfolobus* (C).

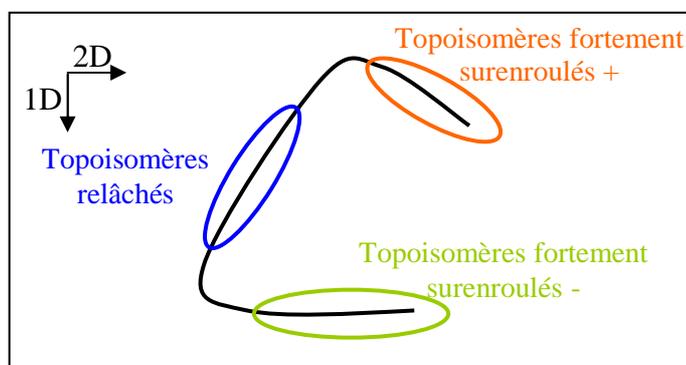
La topoisomérase I eucaryote relâche l'ADN surenroulé négativement ou positivement (Dynam *et al.*, 1981). La topoisomérase A d'*Escherichia coli* relâche uniquement l'ADN surenroulé négativement (Pruss *et al.*, 1982). La reverse gyrase de *Sulfolobus*, en présence d'ATP, relâche l'ADN surenroulé négativement et surenroule positivement l'ADN relâché (Mirambeau *et al.*, 1984). Par ailleurs, le plasmide pTZ18R surenroulé négativement, qui est utilisé comme témoin, est ajouté au génome viral lors des incubations avec les trois topoisomérases décrites ci-dessus. Le profil électrophorétique en une dimension des ADN de PAV1 et de pTZ18 de obtenu après ces incubations est représenté sur la Figure 22.



**Figure 22 :** Profils électrophorétiques des ADN de PAV1 et de pTZ18R : natifs (A), après action de la topoisomérase I eucaryote (B), après action de la topoisomérase A d'*Escherichia coli* (C) et après action de la reverse gyrase de *Sulfolobus* (D).

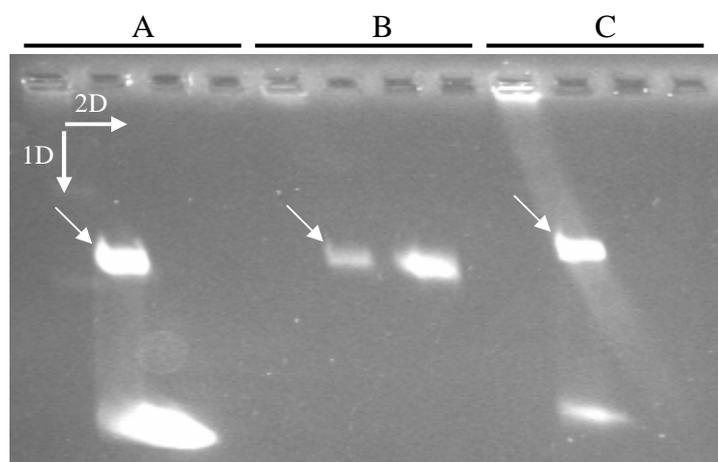
Sur le profil des ADN natifs (cf Figure 22A), les deux bandes situées dans la partie supérieure du gel correspondent au génome de PAV1 et les deux bandes observables dans la partie inférieure au génome de pTZ18R. Dans chacun des cas, la bande du haut correspond à la forme ouverte circulaire et la bande du bas à la forme surenroulée. Sur le profil des ADN incubés avec la topoisomérase I eucaryote (cf Figure 22B), la forme surenroulée de PAV1 a disparu et un smear est visible en-dessous de la forme ouverte circulaire. Ce dernier correspond aux topoisomères comportant des degrés de relaxation différents. La forme totalement relâchée migre au même endroit que la forme ouverte circulaire. La même chose est observée pour pTZ18R avec une meilleure résolution des topoisomères. Le profil des ADN incubés avec la topoisomérase A d'*Escherichia coli* (cf Figure 22C) est similaire à celui obtenu pour les ADN natifs. Le profil des ADN incubés avec la reverse gyrase de *Sulfolobus* (cf Figure 22D) est très proche de lui obtenu après incubation avec la topoisomérase I eucaryote. Cependant, une partie des formes surenroulées persiste et les topoisomères sont moins visibles. Les profils électrophorétiques obtenus montrent que la topoisomérase I eucaryote, et dans une moindre mesure, que la reverse gyrase de *Sulfolobus* relâchent l'ADN surenroulé de PAV1 et de pTZ18R. Par ailleurs, la topoisomérase A d'*Escherichia coli* est inactive sur les ADN, probablement car elle est très inhibée par la forme ouverte circulaire. C'est donc l'incubation avec la reverse gyrase de *Sulfolobus* qui suggère que l'ADN de PAV1 est surenroulé négativement.

Dans le but de mieux visualiser les topoisomères, le génome viral est soumis à une électrophorèse en deux dimensions. Le tampon utilisé au cours de la seconde dimension contient du BET qui augmente le surenroulement et permet ainsi une meilleure résolution des topoisomères. Idéalement, ce type de migration génère une arche où les topoisomères sont répartis comme expliqué dans la Figure 23 selon leur degré de surenroulement.



**Figure 23** : Schéma explicatif de la migration des topoisomères dans un gel en deux dimensions en présence de BET dans la deuxième dimension.

Les profils électrophorétiques obtenus après migration en deux dimensions du génome de PAV1 natif ou incubé avec deux topoisomérases (la topoisomérase I eucaryote ou la reverse gyrase de *Sulfolobus*) sont représentés sur la Figure 24. Dans les trois profils, la forme ouverte circulaire du génome viral est présente. Sur le profil de l'ADN natif, une quantité importante de topoisomères fortement surenroulés négativement est observable. Après incubation avec la topoisomérase I eucaryote, la forme surenroulée disparaît et une bande correspondant à des topoisomères relâchés est visible. Après incubation avec la reverse gyrase de *Sulfolobus*, les topoisomères présentent un surenroulement négatif moins prononcé que dans le cas de l'ADN natif.



**Figure 24 :** Profils électrophorétiques de l'ADN de PAV1 après migration en deux dimensions : à l'état natif (A), après incubation avec la topoisomérase I eucaryote (B) et après incubation avec la reverse gyrase de *Sulfolobus* (C). Les flèches indiquent la forme ouverte circulaire.

Les résultats obtenus avec la migration électrophorétique en deux dimensions confirment ceux obtenus en une dimension, à savoir que l'ADN de PAV1 est surenroulé négativement dans la cellule. Une nuance vis-à-vis du degré de surenroulement mérite cependant d'être apportée car l'expérience est menée à 16°C alors que la température optimale de culture pour *Pyrococcus abyssi* est de 95°C. Une précédente détaillée de M. Duguet a montré qu'il existe une variation de - 0,0105° d'angle/pb/°C (Duguet, 1993). Si cette valeur est appliquée à PAV1, la correction à apporter est donc de :

$$\frac{0,0105 \times 18\,098 \times (95-16)}{360} = 41,7 \text{ supertours}$$

En conditions physiologiques, il y a environ 42 supertours de moins, ce qui suggère que le génome viral est très légèrement surenroulé négativement, proche de l'état relâché. Cette topologie originale rappelle celle observée chez pGT5 (Charbonnier *et al.*, 1992). Mais comment expliquer une telle topologie ? L'annotation du génome de la souche GE5 a permis de mettre en évidence une reverse gyrase et une topoisomérase VI (qui relâche les supertours positifs) (Cohen *et al.*, 2003). Etant donné que la souche GE23 est relativement proche de la souche GE5, il est fort probable que ces deux enzymes y soient également présentes. Ainsi, la topologie de l'ADN de PAV1 dans la cellule serait le résultat d'une balance entre les activités antagonistes des topoisomérases de son hôte. Par ailleurs, chez les procaryotes, une majeure partie de l'ADN est surenroulé négativement grâce à son enroulement autour d'histones (Drlica, 1992 ; Forterre *et al.*, 1996). Il a récemment été prouvé que deux protéines archéennes dites « histone-like », appelées HTz et HMf, sont capables de structurer l'ADN sous forme de chromatine tout comme chez les eucaryotes et de produire un surenroulement négatif (Musgrave *et al.*, 2000 ; Tomschik *et al.*, 2001). De plus, des séquences encodant des protéines « histone-like » ont été détectés dans le génome de la souche GE5 et supportent cette hypothèse (Cohen *et al.*, 2003). Le surenroulement est donc peut-être également influencé par la présence de protéines chromatiniennes. Il peut également s'agir d'une protéine associée à l'ADN codée par le virus comme c'est le cas pour VP2 chez SSV1 (Reiter *et al.*, 1987).

Le léger surenroulement négatif de l'ADN PAV1 est peut-être nécessaire à la reconnaissance des promoteurs viraux par les enzymes impliquées dans la transcription comme c'est le cas pour le bactériophage T4 (Albright *et al.*, 1988). Par ailleurs, à haute température, les supertours positifs sont supposés prévenir l'ADN d'une dénaturation (Kikuchi & Asai, 1984). Le génome de PAV1 est probablement suffisamment protégé dans la cellule hôte pour s'affranchir d'un surenroulement positif.

Les résultats obtenus au cours de l'étude du surenroulement de l'ADN de PAV1 vont être valorisés dans une note intitulée « *DNA of a hyperthermophilic virus is slightly negatively supercoiled in its host cell* » qui sera soumise à la revue *Journal of Virology* (cf Annexe 7).

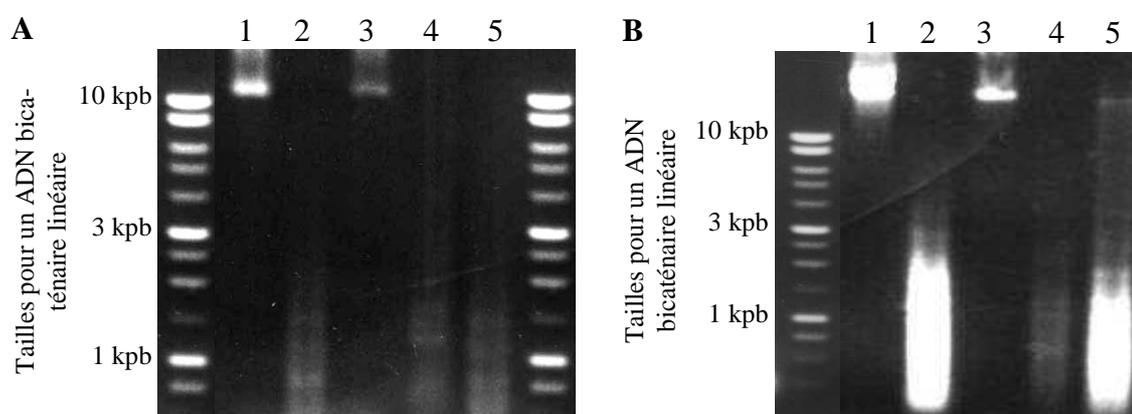
### **b. Recherche de méthylations sur le génome viral et le chromosome de l'hôte**

Un système de restriction/modification a déjà été mis en évidence chez la souche GI-H de *Pyrococcus abyssi*. Ainsi, il est envisageable que le génome de PAV1 soit modifié pour se protéger d'un tel système. Ces modifications peuvent correspondre soit à l'introduction de

nucléotides modifiés due à l'incorporation de dNTP atypiques (Mathews & Allen, 1983), soit à la méthylation de bases azotées après la réplication (Magrini *et al.*, 1997). Les méthylation les plus courantes sont celles de type *dam* et *dcm*. Une méthylation est dite de type *dam* lorsque l'adénine du site 5' GATC 3' possède un groupement méthyle en position N<sup>6</sup>. Une méthylation est de type *dcm* quand la deuxième cytosine des sites 5' CCAGG 3' ou 5' CCTGG 3' porte un groupement méthyle en C<sup>5</sup>.

Parmi les virus hyperthermophiles, deux ont été décrits comme possédant un ADN modifié. Il s'agit de SNDV qui présente des méthylation de type *dam* (Arnold *et al.*, 2000b) et de STSV1 dont le génome comporte des méthylation atypiques (Xiang *et al.*, 2005). Ces deux virus encodent leur propre méthylase. Nous avons recherché des méthylation de type *dam* et *dcm* sur le génome PAV1 et celui de son hôte. Pour ce faire, quatre enzymes de restriction (*DpnI*, *MboI*, *BstNI* et *ScrI*) ayant des propriétés différentes vis-à-vis des sites méthylés ont été utilisées.

Le résultat des digestions enzymatiques des ADN de PAV1 et de la souche GE23 est représenté sur la Figure 25.



**Figure 25** : Profils électrophorétiques l'ADN de PAV1 (A) : natif (1) et digéré par *BstNI* (2), *DpnI* (3), *MboI* (4) et *ScrI* (5) et de l'ADN total de la souche GE23 (B) : natif (1) et digéré par *BstNI* (2), *DpnI* (3), *MboI* (4) et *ScrI* (5).

*DpnI*, qui est active uniquement sur le site 5' GATC 3' *dam* méthylé, ne digère ni l'ADN viral ni le chromosome archéen (cf Figures 25A3 et 25B3). En revanche, la digestion par *MboI*, qui elle est bloquée lorsque le site est *dam* méthylé, produit un smir dans les deux cas (cf Figures 25A4 et 25B4). Les génomes de PAV1 et de son hôte ne présentent donc pas de méthylation *dam*. L'ADN viral et l'ADN de la souche GE23 semblent bien digérés par *BstNI* (cf Figures 25A2 et 25B2), qui clive les sites 5' CCAGG 3' et 5' CCTGG 3' qu'ils

soient *dcm* méthylés ou non. Le même résultat est observé avec *ScrFI* (cf Figures 25A5 et 25B5) qui est bloquée par une méthylation de ces sites. Les génomes de PAV1 et de son hôte ne présentent donc pas non plus de méthylations *dcm*.

L'ADN viral n'est donc pas modifié par les deux types de méthylations étudiés qui sont les plus courants. Il est tout à fait possible qu'il n'y ait aucune modification de l'ADN chez PAV1, comme c'est le cas pour la plupart des génomes de virus hyperthermophiles, étant donné que l'ADN de son hôte n'est pas non plus méthylé. Il est également envisageable qu'il s'agisse plutôt de méthylations atypiques, comme pour STSV1.

### **d. Etude du mode de répllication**

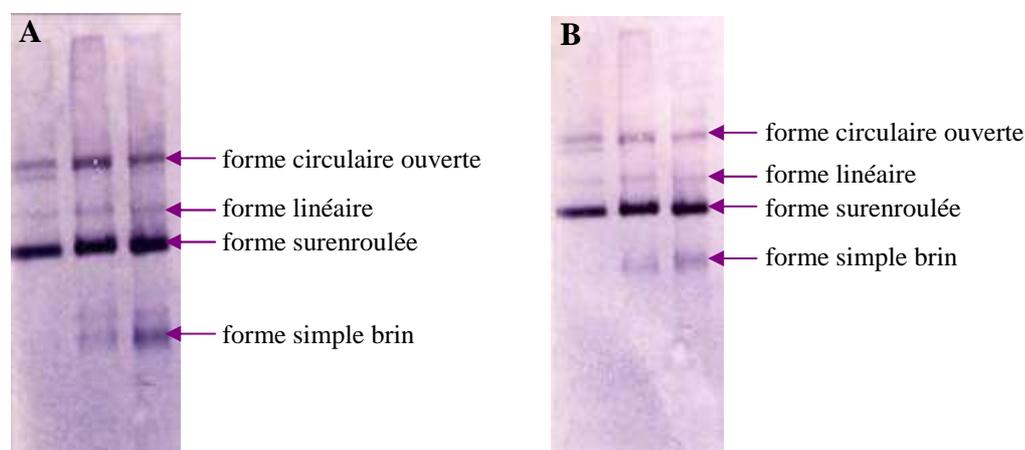
Le nombre important de copies du génome viral dans la cellule nécessite une répllication efficace qui est au cœur des processus de maintien de PAV1. D'une manière générale, les génomes circulaires de virus et de plasmides peuvent se répliquer selon trois mécanismes différents : par cercle roulant, par déplacement de brin ou selon le mode  $\theta$ . A l'heure actuelle, aucune étude n'a été menée sur le(s) mode(s) de répllication adopté(s) par les virus hyperthermophiles. Les seules données disponibles concernent le plasmide hyperthermophile pGT5. En effet, il a été montré que ce dernier se réplique par cercle roulant (Erauso *et al.*, 1996).

Le fait que la majorité des gènes de PAV1 soient orientés dans le même sens suggère un mode de répllication unidirectionnel. Or, le mécanisme de cercle roulant est le mode unidirectionnel le plus répandu. Celui-ci étant caractérisé par la présence d'un intermédiaire de répllication simple brin, nous avons recherché une telle molécule pour déterminer si PAV1 utilisait ce mécanisme.

Deux sondes simple brin ont été préparées : une est spécifique du brin + du génome alors que l'autre est spécifique du brin -. Chaque sonde a été ensuite hybridée sur l'ADN de PAV1 et dans les deux cas, quatre bandes principales sont observables sur la Figure 26. Elles correspondent aux différentes formes du génome qui sont respectivement (de haut en bas) : la forme ouverte circulaire, la forme linéaire, la forme surenroulée et la forme simple brin.

Les deux profils d'hybridation, représentés sur la Figure 26, possèdent la même intensité au niveau de la bande simple brin. Si ce signal correspondait réellement à un intermédiaire de répllication, il ne serait présent que sur le profil révélé avec la sonde spécifique du brin -. Le résultat obtenu a sûrement trait au protocole utilisé. En effet, celui-ci génère souvent un peu de forme simple brin aspécifique du fait d'une légère dénaturation de l'ADN. Celle-ci

ne correspond donc pas à un intermédiaire de réplication mais à une forme dénaturée du génome inhérente à la préparation.



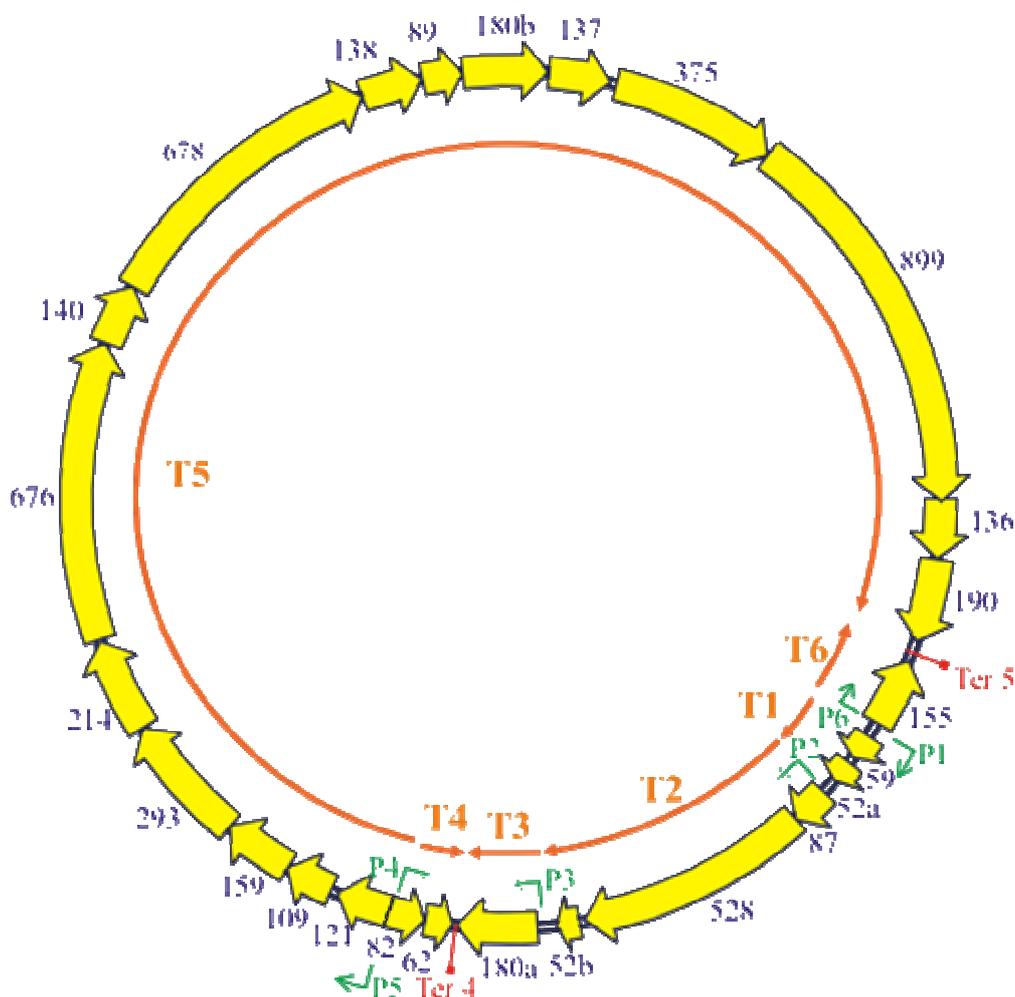
**Figure 26 :** Signal donné par l'hybridation des sondes spécifiques du brin + (A) et du brin – (B) sur l'ADN de PAV1.

PAV1 ne se réplique donc pas par cercle roulant. L'analyse *in silico* suggérant fortement une réplication unidirectionnelle, le virus se réplique donc soit par déplacement de brin soit suivant un mode  $\theta$  unidirectionnel (Del Solar *et al.*, 1998). Le mode  $\theta$  unidirectionnel est assez rare chez les virus mais il a notamment été mis en évidence chez le phage  $\lambda$  au cours de la réplication qui a lieu précocement après l'infection (Baranska *et al.*, 2001). Pour savoir si PAV1 se réplique suivant un mécanisme de déplacement de brin ou par un mode  $\theta$ , les intermédiaires de réplication du génome viral pourraient être analysés par gradient de densité couplé à des observations en MET (comme dans le cas du phage  $\phi 29$ ) et par migration électrophorétique en une ou deux dimensions (comme cela a été réalisé pour le phage  $\lambda$  (Inciarte *et al.*, 1980 ; Baranska *et al.*, 2001).

#### e. Etude des transcrits

Pour l'étude des transcrits, l'approche par RT-PCR est préférée au northern-blot car cette technique est plus sensible pour détecter les ARNm rares. De plus, elle permet de déterminer précisément quels gènes sont cotranscrits. Ainsi, des couples d'amorces sont dessinés de manière à amplifier une région chevauchant deux gènes. Dans le cas où la RT-PCR s'avère positive, il est possible de dire que les deux gènes sont cotranscrits et la même manipulation est répétée de proche en proche. Lorsque la RT-PCR est négative, il faut tout de même savoir si les deux gènes sont transcrits indépendamment. Une autre RT-PCR est alors réalisée avec

des couples d'amorces dessinés sur chaque gène. En compilant tout les résultats, la carte des transcrits du virus peut être tracée pas à pas. Elle est dessinée sur la Figure 27.



**Figure 27** : Cartographie des transcrits de PAV1 (T1 à T6) avec les promoteurs (P1 à P6) et les terminateurs de la transcription (T4 et T5) déterminés par l'analyse de séquence.

Il apparaît que les 25 gènes prédits sont tous transcrits en 6 ARNm notés T1 à T6. Quatre sont polycistroniques (T1, T2, T4 et T5) et le plus long d'entre eux, T5, couvre les deux tiers du génome avec ses 13,6 nt. Les transcrits issus du brin + sont T1, T2, T3 et T5 alors que ceux issus du brin - sont T4 et T6.

La cartographie des transcrits obtenue confirme l'analyse de la séquence qui prédisait des promoteurs putatifs comportant des motifs « TATA-like » au début de chaque transcrit. Pour ce qui est des terminateurs putatifs riches en thymine, seuls deux ont été prédits à la fin

des ARNm T4 et T5. Il est possible que les autres n'aient pas été repérés parce qu'ils sont trop atypiques.

Chez les trois virus hyperthermophiles pour lesquels la transcription a été étudiée (SSV1, SIRV1, SIRV2), les gènes sont majoritairement co-transcrits et principalement issus du brin + comme chez PAV1 (Palm *et al.*, 1991 ; Kessler *et al.*, 2004). Les ARNm polycistroniques sont très fréquemment utilisés par les procaryotes et les virus pour réguler leurs gènes. Cependant, un ARNm couvrant une portion aussi importante du génome (près de 75 %) n'a jamais été observé chez aucun virus hyperthermophile. En revanche des résultats similaires ont récemment été observés chez deux plasmides de *Thermococcales* en cours de caractérisation : pCIR10 et pIRI48. Des transcrits couvrant jusqu'à 76 % du génome ont été mis en évidence (Erauso G., Gonnet G., Toffin A., Chevereau M. & Prieur D., en préparation).

Etant donné la technique utilisée pour la cartographie, il est possible que le grand transcrit T5 « cache » la présence d'ARNm alternatifs. Cependant, l'intérêt pour PAV1 de synthétiser un transcrit d'une telle longueur reste encore à élucider. En effet, chez les virus, seuls les *Reversivirus* végétaux (les *Badnavirus* et les *Caulimovirus*) produisent un transcrit qui recouvre entièrement leur génome et qui leur permet une réplication par rétro-transcription (Medberry *et al.*, 1990).

Plus de 60 % des protéines codées par le transcrit T5 de PAV1 ont des domaines transmembranaires prédits. Il est possible d'imaginer que cet ARNm corresponde à une unité de transcription où la majorité des protéines seraient tout ou partie intégrées dans l'enveloppe du virion. Les autres protéines de cette unité pourraient avoir un rôle fonctionnel. La coexistence sur un même transcrit de gènes codant pour des protéines structurales et fonctionnelles a déjà été démontrée chez un *Caulimovirus* (Medberry *et al.*, 1990) et un *Rhabdovirus* (Wang & Walker, 1993).

### **3. Etude de deux protéines de PAV1**

#### **a. Etude de la protéine 899**

##### ➤ Pourquoi étudier la protéine 899 ?

La protéine 899, constituée de 899 résidus, a un poids moléculaire théorique de 101 kDa. Elle est codée par le gène 899 de PAV1. L'analyse bioinformatique n'a pas permis d'identifier de motif ou de domaine conservés sur cette protéine. Cependant, le gène qui

l'encode représente près de 15 % du génome viral. Etant donné la taille habituellement modérée des gènes viraux, la protéine 899 joue probablement un rôle clé dans l'existence du virus.

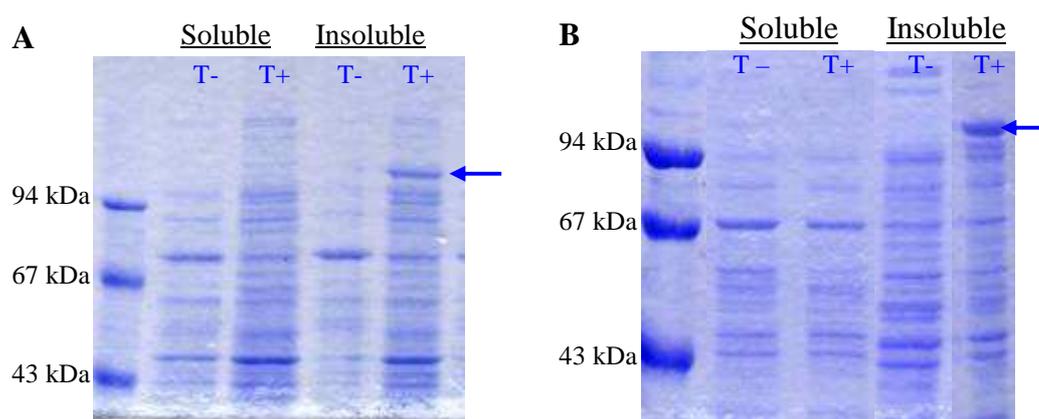
De plus, elle possède un homologue chez les plasmides pCIR10 et pIRI48 portés par les souches de *Thermococcales* du même nom isolées de sites hydrothermaux distincts et éloignés. Si ces observations confirment l'importance de la protéine 899, elles suggèrent également que cette dernière ne joue pas un rôle structural dans le virus. Enfin, sur les deux plasmides, les homologues de la protéine 899 sont en opéron avec une hélicase (Erauso G., Gonnet G., Toffin A., Chevereau M. & Prieur D., en préparation). Ceci suggère une implication possible de la protéine 899 dans la réplication du génome viral.

Les éléments génétiques mobiles se sont déjà avérés être une source d'enzymes nouvelles impliquées dans la maintenance du génome. En effet, G. Lipps et ses collaborateurs ont découvert sur le plasmide pRN1 de l'archée *Sulfolobus islandicus*, une enzyme multifonctionnelle de 904 résidus comportant des activités ATPase, primase et ADN polymérase (Lipps *et al.*, 2003). Dans ce contexte, il est décidé de focaliser le travail sur l'étude de la protéine 899 afin de mettre en évidence son rôle chez PAV1 et sa participation potentielle à la réplication de l'ADN viral. Dans ce but, la protéine 899 doit être surproduite puis purifiée pour ensuite tester son activité.

### ➤ Surexpression du gène 899

Le fait que tel ou tel vecteur va s'avérer efficace pour l'expression du gène est une donnée empirique. Il faut donc se ménager un maximum de chances de réussite et afin de pouvoir surproduire la protéine 899. Le gène 899 est amplifié par PCR puis cloné dans plusieurs vecteurs d'expression : pBAD (Invitrogen), pET 11d, pET 25b+ et pET 28a+ (Novagen). De plus, ceci permet aussi d'obtenir des variantes de la protéine native. Par exemple, une queue histidine peut être ajoutée en C terminal pour faciliter une purification ultérieure sur une colonne de nickel. La séquence du gène 899 cloné dans les différents vecteurs est comparée avec la séquence du gène 899 déterminée lors du séquençage du génome viral. Tous les clonages réalisés dans les différents vecteurs d'expression sont exempts de mutations sur le gène 899. Comme là encore, il est impossible de savoir quelle cellule va permettre une expression optimale du gène cloné, les différentes constructions obtenues sont utilisées pour transformer différentes cellules d'expression : SURE<sup>®</sup> (Stratagene), BL<sub>21</sub> Rosetta<sup>®</sup> et HMS 174 (Novagen). Les cellules BL<sub>21</sub> Rosetta<sup>®</sup> et HMS 174 sont utilisées pour les constructions en vecteur pET alors que les cellules SURE<sup>®</sup> sont utilisées pour celles en vecteur pBAD.

La protéine 899 n'est pas observable dans les extraits protéiques des cellules SURE et BL<sub>21</sub> Rosetta<sup>®</sup>. Ainsi, comme il est illustré dans la Figure 28, seules les cellules HMS 174 avec les vecteurs d'expression pET 11d et 28a+ permettent une surexpression du gène 899 et donc une surproduction de la protéine 899 (de poids moléculaire 101 kDa). La construction en vecteur pET 28a+ qui code pour une protéine avec une queue histidine en C terminal sera utilisée pour réaliser la suite des expériences. En effet, celle-ci permet de simplifier l'étape ultérieure de purification en utilisant une colonne de nickel.



**Figure 28 :** Profils électrophorétiques des protéines contenues dans les fractions solubles et insolubles obtenues 2 h après induction à l'IPTG des cellules HMS 174 contenant: (A) le vecteur pET 11d non recombinant (T-) ou le vecteur pET 11d contenant le gène 899 (T+) ; (B) le vecteur pET 28a+ non recombinant (T-) ou le vecteur pET 28a+ contenant le gène 899 (T+). Les flèches indiquent la protéine 899.

Sur les profils électrophorétiques des protéines issues des cellules HMS 174 contenant les vecteurs d'expression pET 11d et 28a+ (cf Figure 28), la totalité de la protéine 899 se trouve dans la fraction insoluble. De plus, en observant les cellules HMS 174, après induction, au microscope à contraste de phase, de petits points réfringents sont visibles au niveau du cytoplasme. Ceci suggère fortement que la protéine 899 serait entièrement produite sous forme de corps d'inclusion. Ce phénomène peut être expliqué par sa taille importante (101 kDa) et également par le fait qu'une grande partie de cette molécule est prédite hydrophobe.

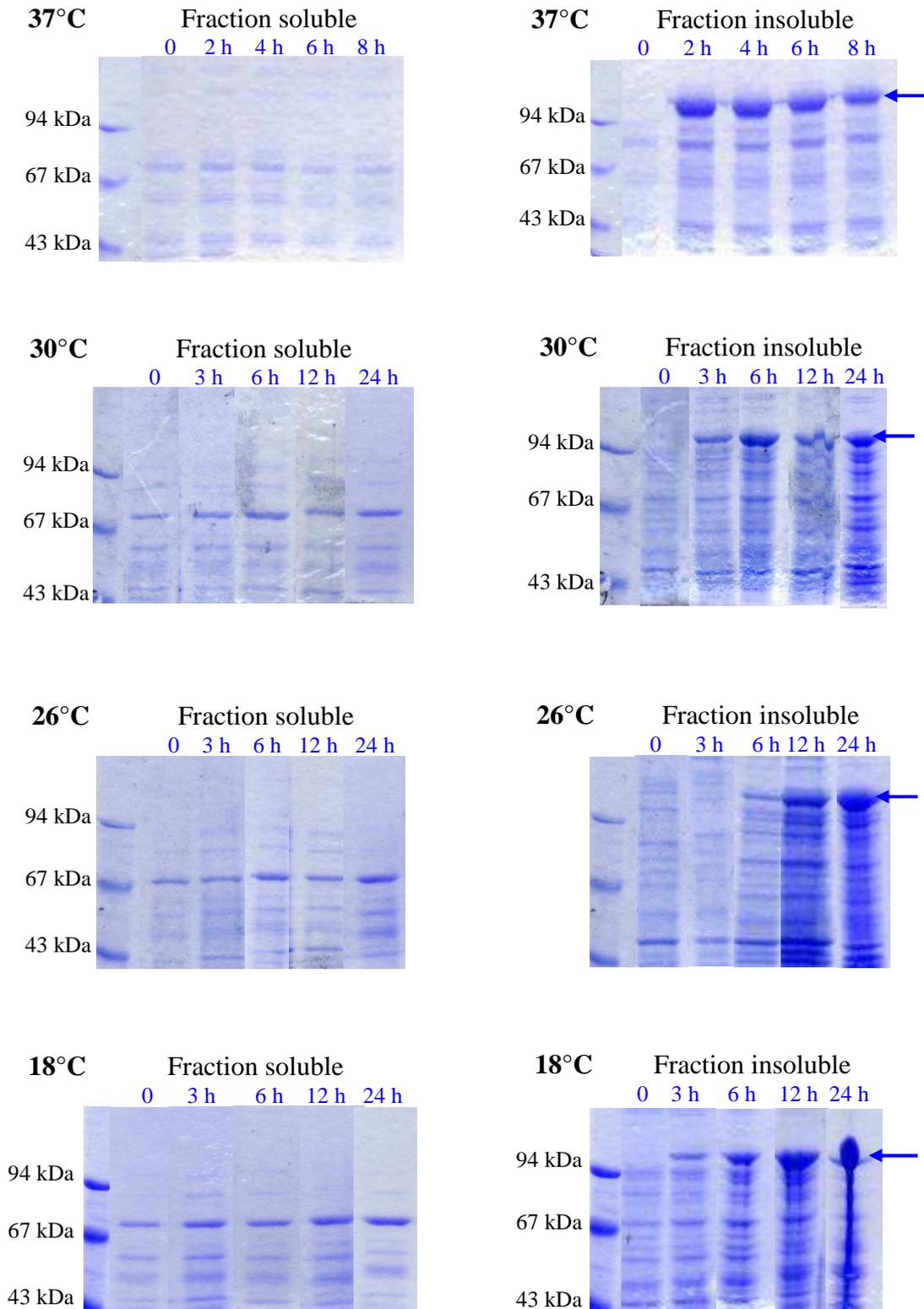
Afin de poursuivre l'étude de cette protéine, différents protocoles connus pour améliorer la production des protéines sous forme soluble, sans nécessiter de purifier les corps d'inclusion, sont testés :

- l'abaissement de la température de culture ;
- la co-production avec des chaperons ;
- la production *in vitro*.

### ➤ Essais de production de la protéine 899 sous forme soluble

La première technique consiste à abaisser la température de culture des cellules pour diminuer la vitesse de synthèse de la protéine. Souvent, cela suffit à rendre la molécule en partie soluble (Schein, 1989). Les profils électrophorétiques des protéines contenues dans les fractions solubles et insolubles des cellules HMS 174 portant le vecteur pET 28a+ contenant le gène 899 et cultivées à 37°C, 30°C, 26°C ou 18°C sont représentés sur la Figure 29. Il apparaît qu'aucune des fractions solubles ne contient la protéine 899. L'abaissement de la température de culture n'a donc eu pour effet que de retarder sa production sous forme insoluble.

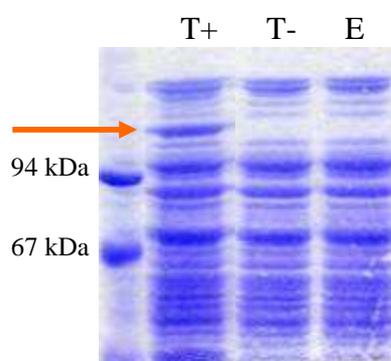
Il est aussi possible de co-produire la protéine insoluble avec des protéines chaperons. En effet, les protéines hétérologues ont tendance à être produites sous forme insoluble à cause de leur impossibilité à adopter leur structure tertiaire native. Le cytoplasme étant un milieu réduit, la formation des ponts disulfure y est difficile. Les protéines chaperons peuvent alors aider au repliement des protéines hétérologues et ainsi augmenter leur solubilité (Nishihara *et al.*, 1998 ; Trésaugues *et al.*, 2004).



**Figure 29 :** Profils électrophorétiques des protéines contenues dans les fractions solubles et insolubles obtenues après induction à l’IPTG des cellules HMS 174 portant le vecteur pET 28a+ contenant le gène 899 et cultivées à 37°C, 30°C, 26°C ou 18°C. Les flèches indiquent la protéine 899.

La protéine 899 est donc co-produite avec 5 protéines chaperons d'*Escherichia coli* (DnaK, DnaJ, GrpE, GroEL et GroES) toutes encodées par le plasmide pG-KJE3. Pour cela, ce dernier est introduit dans les cellules HMS 174 portant le vecteur pET 28a+ non recombinant ou le vecteur pET 28a contenant le gène 899, puis la surexpression est induite. Au final, la protéine 899 n'est plus visible ni dans la fraction insoluble ni dans la fraction soluble alors que les protéines chaperons sont, elles, bien exprimées. Ce problème peut-être du à la protéine chaperonne DnaK qui, parfois, débobine complètement la protéine hétérologue qui est alors rapidement dégradée (C. Lavire, communication personnelle).

Comme il est difficile d'appréhender les conditions nécessaires à une production de la protéine 899 sous forme soluble et qu'elles peuvent être atypiques, la production *in vitro*, classiquement utilisée dans le cas de protéines toxiques pour *Escherichia coli*, est également testée. Les profils électrophorétiques des protéines obtenues après expression *in vitro* sont représentés sur la Figure 30. Seule la  $\beta$ -galactosidase encodée par le plasmide pEX1-GW/*lacZ*, le témoin positif, est surproduite. La protéine 899 a disparu des profils électrophorétiques.

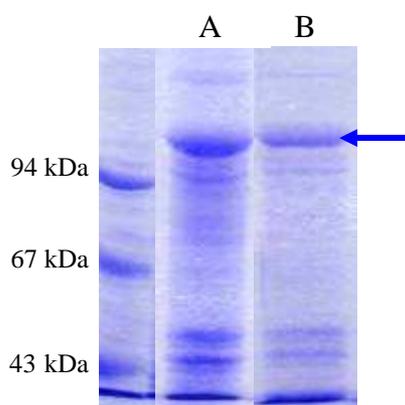


**Figure 30 :** Profils électrophorétiques des protéines obtenues après expression *in vitro* pratiquée sur le plasmide pEX1-GW/*lacZ* contenant le gène de la  $\beta$ -galactosidase (T+), le vecteur pET 28a+ non recombinant (T-) ou le vecteur pET 28a+ contenant le gène 899 (E). La flèche indique la  $\beta$ -galactosidase.

Aucun des protocoles testés n'a donc permis d'améliorer la production de la forme soluble de la protéine 899. Etant donné que cette protéine serait, *a priori*, entièrement produite sous forme de corps d'inclusion, nous réalisons une extraction et une solubilisation de ces derniers.

### ➤ Extraction et solubilisation des corps d'inclusion

Les corps d'inclusion présents dans le cytoplasme des cellules HMS 174, qui portent le vecteur pET 28a+ contenant le gène 899, sont extraits et solubilisés à l'hydrochloride de guanidine 8 M. Le profil électrophorétique des protéines collectées au cours des différentes étapes est représenté dans la Figure 31. Dans la fraction finale, une grande quantité de protéine 899 est récupérée sous forme soluble. Quatre autres protéines sont également présentes mais dans des proportions moins importantes. Ces dernières proviennent soit des cellules HMS 174 soit d'une dégradation de la protéine 899.



**Figure 31 :** Profils électrophorétiques des protéines obtenues après extraction (A) puis après solubilisation (B) des corps d'inclusion. La flèche indique la protéine 899.

L'hydrochloride de guanidine 8 M étant un agent dénaturant, la protéine virale doit être renaturée pour qu'elle retrouve sa structure native.

### ➤ Renaturation de la protéine 899

Cette étape est particulièrement délicate car chaque protéine nécessite des conditions particulières. En première approche, 80 tampons de renaturation différents sont donc testés en microplaque. Lorsque le tampon ne facilite pas le repliement de la protéine celle-ci précipite. Ce phénomène est détectable par une absorbance à 340 nm supérieure à 0,04 (Vincentelli *et al.*, 2004). La valeur des densités optiques mesurées à 340 nm à  $t = 0$  et à  $t = 52$  h pour chaque tampon additionné de la protéine 899 est notée dans la Figure 32.

## Résultats et Discussion

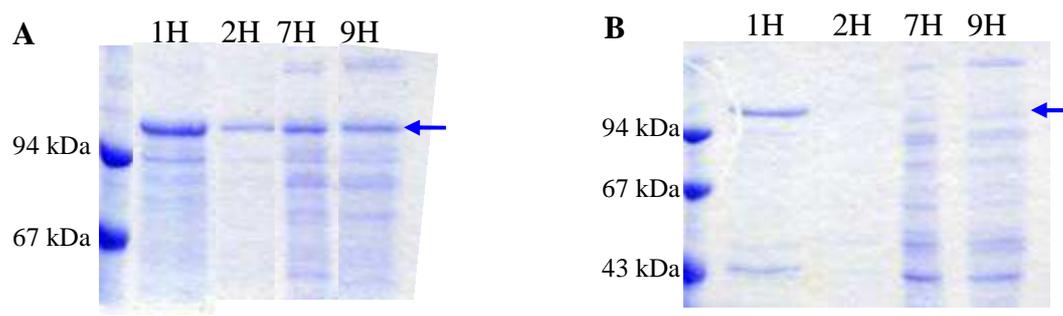
		1	2	3	4	5	6	7	8	9	10
t = 0	A	0,070	0,064	0,049	0,063	0,076	<b>0,026</b>	<b>0,024</b>	0,042	0,061	<b>0,037</b>
	B	0,085	0,096	0,080	0,119	<b>0,017</b>	0,055	0,063	0,077	0,062	<b>0,039</b>
	C	0,097	0,123	<b>0,023</b>	<b>0,021</b>	<b>0,032</b>	<b>0,029</b>	0,053	<b>0,003</b>	0,045	<b>0,038</b>
	D	0,065	<b>0,027</b>	0,073	<b>0,026</b>	<b>0,011</b>	<b>0,002</b>	<b>0,003</b>	<b>0,022</b>	0,058	<b>0,026</b>
	E	0,079	0,041	0,067	0,065	0,136	0,061	<b>0,036</b>	0,063	0,088	0,133
	F	0,051	0,050	0,095	0,132	0,135	0,090	0,113	<b>0,003</b>	0,083	0,145
	G	<b>0,043</b>	<b>0,035</b>	0,082	0,055	0,067	0,065	0,075	0,066	0,122	0,092
	H	<b>0,012</b>	<b>0,035</b>	0,044	0,051	0,067	0,072	<b>0,034</b>	<b>0,040</b>	<b>0,038</b>	0,078
			1	2	3	4	5	6	7	8	9
t = 52 h	A	0,045	0,048	0,045	0,058	0,048	0,049	<b>0,035</b>	0,035	0,054	0,044
	B	0,063	0,112	0,115	0,131	0,105	0,120	0,122	0,129	0,114	0,082
	C	0,091	0,121	0,047	0,103	0,086	0,099	0,093	0,092	0,083	0,073
	D	0,052	0,080	0,066	0,046	0,042	0,055	0,043	0,047	0,042	0,097
	E	0,053	0,065	0,060	0,058	0,044	0,054	0,045	0,075	0,076	0,102
	F	0,086	0,080	0,111	0,080	0,081	0,048	0,073	0,065	0,073	0,064
	G	0,090	0,109	0,096	0,074	0,088	0,075	0,091	0,085	0,093	0,055
	H	<b>0,026</b>	<b>0,034</b>	0,057	0,039	0,039	0,032	<b>0,023</b>	<b>0,029</b>	<b>0,030</b>	0,054

**Figure 32** : Densité optiques mesurées à 340 nm pour la protéine 899 mélangée à chacun des 80 tampons aux temps t = 0 et t = 52 h.

Sur les 80 tampons testés, 27 ne font pas précipiter la protéine 899 immédiatement. Après 52 h d'incubation, seulement 6 tampons semblent convenir. Leurs noms correspondent aux cases dans lesquelles ils sont décrits au niveau du tableau 6 du paragraphe *Vf.5.a* de la partie *Matériel et Méthode*. Ils sont dénommés : 1H, 2H, 7A, 7H, 8H et 9H. La protéine 899 va devoir être dialysée contre ces différents tampons pour éliminer l'hydrochloride de guanidine. Cependant, cette étape nécessitant une certaine quantité de protéine, nous ne pouvons tester que quatre tampons : 1H, 2H, 7H et 9H.

La protéine 899 est donc dialysée contre ces quatre tampons. Les fractions récupérées après dialyse sont centrifugées pour éliminer la part de protéine qui a pu précipiter. Le surnageant est alors collecté. Les profils électrophorétiques des protéines contenues dans les fractions après dialyse contre les 4 tampons et dans le surnageant après centrifugation des fractions dialysées contre les 4 tampons sont représentés dans la Figure 33.

Il apparaît que seul le tampon 1H permet de conserver en grande quantité la protéine 899 sous forme soluble. Une réserve peut tout de même être émise car ce tampon contient 800 mM d'arginine. Or, à de telles concentrations cette molécule peut agir exactement comme l'hydrochloride de guanidine (S. Quevillon-Cheruel, communication personnelle). Il est donc possible que la protéine soit encore dénaturée. Nous décidons malgré tout d'explorer l'implication possible de la protéine 899 dans la réplication de l'ADN viral.



**Figure 33 :** Profils électrophorétiques des protéines contenues : dans les fractions après dialyse contre les tampons 1H, 2H, 7H ou 9H (A) puis dans le surnageant après centrifugation des fractions dialysées contre les tampons 1H, 2H, 7H ou 9H (B). Les flèches indiquent la protéine 899.

- Essais de mise en évidence d’une implication de la protéine 899 dans la réplication de l’ADN viral

Le test d’activité réalisé sur la protéine 899 dialysée contre les quatre tampons est un test d’incorporation de thymine tritiée sur une matrice ADN simple brin linéaire polyamorcée. La radioactivité incorporée est mesurée sur un compteur à scintillation et les mesures obtenues figurent dans le Tableau 7. La radioactivité présente dans les mélanges réactionnels contenant la protéine 899 est inférieure à celle du témoin négatif. Nous n’avons donc pas pu mettre en évidence une activité d’incorporation de thymine tritiée par la protéine 899 dialysée.

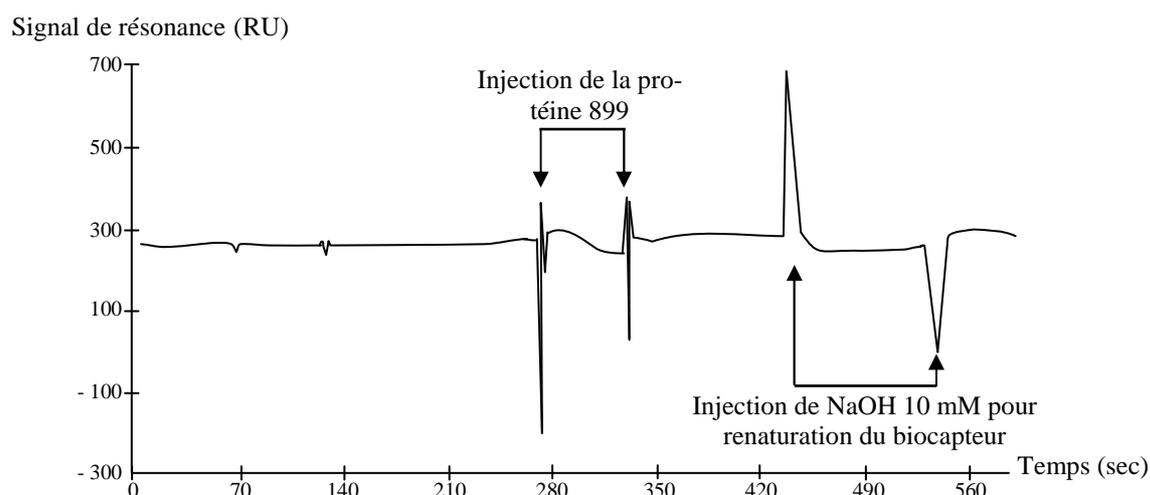
**Tableau 7 :** Mesures de la radioactivité (en coups par minute) incorporée après incubation de la matrice avec la protéine 899 dialysée contre les tampons 1H, 2H, 7H et 9H.

	Témoin -	Témoin +	Protéine 899 dans tp 1H	Protéine 899 dans tp 2H	Protéine 899 dans tp 7H	Protéine 899 dans tp 8H
Mesure 1	30	3285	10	14	35	30
Mesure 2	45	2624	7	9	23	16
<b>Mesure moyenne</b>	<b>38</b>	<b>2955</b>	<b>9</b>	<b>12</b>	<b>29</b>	<b>15</b>

Les raisons de cette absence d’activité peuvent se situer à trois niveaux. Il peut tout d’abord s’agir du mélange réactionnel utilisé lors du test d’incorporation qui n’est pas optimal pour la protéine. En effet, d’infimes variations de pH ou de concentration en sels suffisent parfois à rendre la molécule inactive. De plus, chaque enzyme nécessite des conditions très particulières. Ainsi, l’ADN polymérase B et l’ADN polymérase D de l’archée *Pyrococcus abyssi* sont actives respectivement à pH 8 et à pH 6,6. De plus, seule la polymérase B requiert

la présence de KCl. Le problème peut également provenir des conditions utilisées pour la renaturation. Comme il a été expliqué précédemment, la quantité d'arginine est telle dans le tampon de renaturation qu'elle peut gêner ou même complètement empêcher le repliement de la protéine dans sa forme native. Enfin, la protéine 899 peut ne pas avoir d'activité ADN polymérase. Cependant, il est possible d'envisager qu'elle joue tout de même un rôle dans la réplication du génome viral en tant que qu'hélicase, ligase ou primase.

Pour mieux définir son rôle biologique, nous avons mené une recherche d'interactants potentiels. Les tests d'interaction sont réalisés en résonance plasmonique de surface qui permet de visualiser en temps réel des interactions entre biomolécules non marquées dans un flot continu de tampon. Nous avons testés les interactions de la protéine 899 dialysée contre le tampon 1H avec trois molécules : de l'ADN simple brin, le facteur de processivité de l'ADN polymérase de *Pyrococcus abyssi* (appelé PCNA) et une endonucléase de *Pyrococcus abyssi* (Fen I). Le sensorgramme obtenu dans le cas du PCNA est reproduit sur la Figure 34.

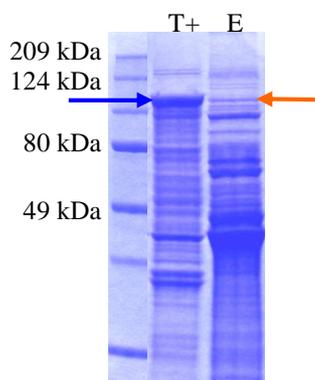


**Figure 34** : Profil d'interaction protéine 899 / PCNA obtenu avec le Biacore X.

Tous les tracés obtenus ont la même allure que celui observé pour le test de l'interaction protéine 899 / PCNA. Or, ces tracés correspondent à une absence d'interaction. Ces résultats ne permettent pas de discriminer entre une mauvaise renaturation de la protéine 899 et une absence d'implication de cette dernière dans le réplication de l'ADN viral. Dans le but de purifier sa forme native, nous menons une recherche préalable de cette protéine dans son hôte naturel, la souche GE23 de *Pyrococcus abyssi*.

### ➤ Recherche de la protéine 899 native dans son hôte naturel, la souche GE23 de *Pyrococcus abyssi*

Les profils électrophorétiques obtenus pour la protéine 899 recombinante et pour les protéines présentes dans la fraction soluble de l'extrait brut de la souche GE23 de *Pyrococcus abyssi* sont illustrés sur la Figure 35.



**Figure 35 :** Profils électrophorétiques des protéines contenues dans : la fraction insoluble des cellules HMS 174 portant le vecteur pET 28a+ contenant le gène 899 (T +) ainsi que dans la fraction soluble de la souche GE23 de *Pyrococcus abyssi* (A). La flèche bleue indique la protéine 899 recombinante et la flèche orange la protéine analysée en spectrométrie de masse.

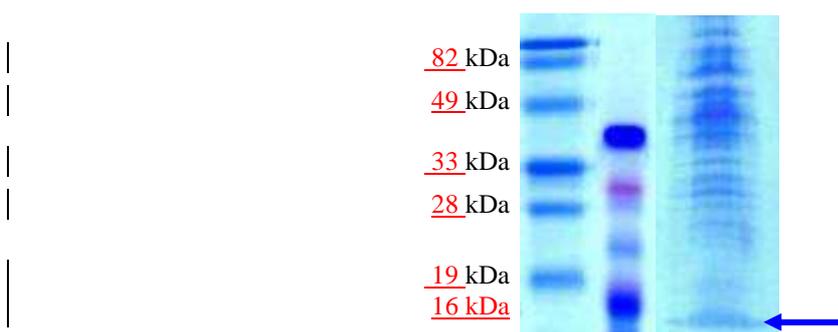
Il apparaît qu'une bande issue de la fraction soluble de l'archée migre de la même façon que la protéine 899 recombinante. Après analyse en spectrométrie de masse MalDI-tof, il s'avère qu'il ne s'agit pas d'une protéine virale mais d'une protéine de l'hôte. Par ailleurs, l'analyse indique également que la bande ne résulte pas d'un mélange de plusieurs protéines. Il est tout à fait possible que la protéine native migre différemment de la protéine recombinante. Si c'est le cas, elle serait un peu au-dessus ou au-dessous de la protéine analysée. Or, dans cette zone les bandes sont assez peu intenses et très resserrées. Une analyse en spectrométrie de masse LC-MS/MS pourrait être réalisée sur toute cette zone pour déterminer si la protéine 899 s'y trouve effectivement. Il est cependant possible qu'elle soit trop peu exprimée dans son hôte pour être détectée.

Pour découvrir la fonction de la protéine 899, le chemin risque d'être encore long et semé de problèmes techniques. Cependant, en raison notamment de la découverte très récente d'un autre d'homologue de cette protéine chez un nouveau plasmide de *Thermococcales*, pTN2 (N. Soler, communication personnelle), il apparaît qu'il s'agit d'une molécule importante pour les éléments génétiques qui la possèdent. Ainsi, comprendre quel est le rôle biologique de la protéine 899 chez PAV1 permettra de mieux appréhender le fonctionnement intime de ce virus encore très énigmatique.

### b. Etude d'une protéine majoritaire du virus

Des travaux antérieurs ont montré la présence de trois protéines majoritaires (6 kDa, 13 kDa et 36 kDa) chez le virion de PAV1 (Geslin et *al.*, 2003). Cependant, la très faible quantité de matériel n'avait pas permis d'accéder à la séquence des protéines contenues dans les bandes. Pour établir la correspondance entre les protéines encodées par le virus et les celles observables sur le gel, la même analyse que celle menée par C. Geslin et ses collaborateurs est réalisée en utilisant une quantité non limitante de matériel biologique.

Le profil électrophorétique des protéines obtenues à partir des virions purifiés est représenté sur la Figure 36. Il est assez complexe car certaines protéines de l'hôte semblent persister dans la fraction purifiée de PAV1. Cependant une bande majoritaire d'un poids moléculaire proche de 15 kDa est bien visible. Etant donné l'approximation de la taille des molécules, cette protéine correspond certainement à la bande de 13 kDa mise en évidence au cours des travaux précédents.



**Figure 36 :** Profils électrophorétiques des protéines issues des virions de PAV1 purifiés. La flèche indique la protéine majoritaire.

La protéine de 15 kDa séquencée par la méthode d'Edmann se révèle être la protéine 121 du virus PAV1. Cependant, un alignement des deux séquences représenté sur la Figure 37 met en évidence que la protéine contenue dans la bande de 15 kDa possède 26 résidus de moins que la protéine prédite.

Protéine prédite :	M-T-V-V-V-Y-A-P-T-S-G-W-N-F-R-G-V-W-Q-W-L-N-E-E-D-A-A-M-M-D-A-L-E-D-V---
Protéine séquencée :	M-M-D-A-L-E-D-V---

**Figure 37 :** Alignement des séquences de la protéine prédite et de la protéine séquencée expérimentalement.

L'hypothèse selon laquelle le gène a été mal annoté ne peut être écartée. Cependant, une analyse statistique des peptides signaux retrouvés chez les archées a permis de déterminer que leur taille moyenne est de 25 acides aminés avec un clivage après une alanine (Bardy *et al.*, 2003). Or, le dernier résidu de la partie manquante de la protéine 121 est une alanine. Ceci suggère fortement qu'elle pourrait posséder un peptide signal qui est clivé après adressage.

Par ailleurs, le clivage N-terminal a déjà été observé chez plusieurs protéines virales. C'est le cas, par exemple, des protéines constituant la tête des phages lambdaïdes (Black, 1989) et de certaines protéines du virus archéen  $\psi$ M2 (Pfister *et al.*, 1998). En considérant ces différents éléments, il est possible de poser l'hypothèse selon laquelle la protéine 121 est en fait un des constituants protéiques majoritaires de l'enveloppe de PAV1, adressé à cette dernière grâce à un peptide signal.

Des anticorps dirigés contre la protéine 121 pourraient être synthétisés de façon à réaliser une décoration directe des particules de PAV1 sur des coupes ultrafines. Ceci permettrait de préciser la localisation de cette protéine au niveau de la structure du virus, comme cela se pratique couramment pour les virus végétaux (Nass *et al.*, 1998).

L'analyse de la séquence du génome de PAV1 et l'étude moléculaire du virus qui vient d'être présentée vont être valorisées dans un article intitulé « *Genome organization and molecular analysis of PAV1, the first marine hyperthermophilic virus-like particle* » qui va être soumis à la revue *Journal of Bacteriology* (cf Annexe 8).

### **III Recherche de nouveaux virus chez les *Thermococcales* des sources hydrothermales océaniques profondes**

#### **1. Préambule**

La quantité d'informations concernant les virus hyperthermophiles terrestres s'est considérablement accrue ces dernières années. En revanche, l'étude des virus hyperthermophiles marins, beaucoup plus récente, a jusqu'à présent livré moins de données. De plus, le nombre d'équipes les étudiant est restreint (seulement trois dans le monde). A l'heure actuelle, le seul virus hyperthermophile marin décrit est PAV1. Or, ce dernier n'est pas infectieux (Geslin *et al.*, 2003). Nous avons donc mis en œuvre une recherche de virus pouvant infecter des *Thermococcales* hydrothermales marines.

## 2. Recherche de particules de type viral dans des cultures d'enrichissement de *Thermococcales*

Les 65 échantillons bruts sélectionnés servent d'inoculum pour réaliser des cultures d'enrichissement de *Thermococcales*. La durée de l'incubation à 85°C nécessaire pour obtenir une densité cellulaire acceptable varie de 14 h à 10 j (cf Annexe 5). Ceci peut s'expliquer par la quantité très variable de cellules apportées lors de la mise en culture. En effet, l'échantillon brut n'est pas homogène et il arrive de prélever une partie pauvre en cellules ou, au contraire, une partie très riche. Au final, 60 cultures présentent une croissance avérée de micro-organismes, pour la plupart très certainement apparentés aux *Thermococcales* étant donné les conditions de culture utilisées (hétérotrophie, chimio-organotrophie, température de 85°C, anaérobiose et présence de soufre).

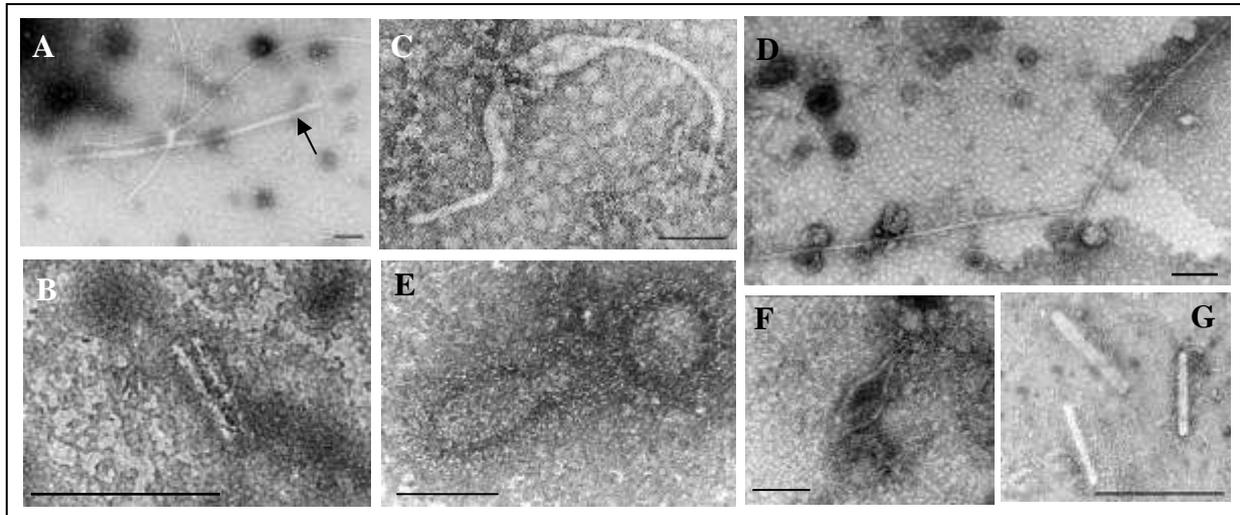
Obtenir 92 % de cultures positives est en accord avec plusieurs études antérieures qui montrent que les *Thermococcales* sont très répandues au niveau de l'écosystème hydrothermal (Harmsen *et al.*, 1997 ; Takai *et al.*, 2001). Il est possible que les échantillons bruts qui n'ont pas permis le développement de *Thermococcales*, soient en réalité beaucoup trop pauvres en matière organique et/ou trop exposés au fluide hydrothermal. Ce type d'échantillon est, en général, beaucoup plus propice à la croissance de micro-organismes autotrophes hyperthermophiles comme les méthanogènes.

Les 60 cultures d'enrichissement positives sont regardées en microscopie électronique en transmission pour détecter la présence de particules de type viral (VLPs). Finalement, 15 présentent des VLPs et seulement 9 conservent leurs particules après plusieurs repiquages de la culture initiale. Les caractéristiques concernant les VLPs observées sont récapitulées dans les Tableaux 8 et 9 ainsi que dans les Figures 38 et 39.

Le fait que les VLPs ne se maintiennent pas dans toutes dans les cultures peut s'expliquer de deux façons. D'abord, les particules observables ne sont pas nécessairement produites par les micro-organismes présents dans l'enrichissement. En effet, il est probable que ces VLPs, relarguées dans le milieu extérieur puis transportées par les courants, soient prélevées sans leurs cellules hôtes. Il est aussi possible d'envisager que l'espèce microbienne qui produit ces particules ait été supplantée par d'autres espèces au cours du repiquage de la culture initiale.

**Tableau 8 :** Caractéristiques des VLPs ne se maintenant pas dans la culture d'enrichissement après plusieurs repiquages.

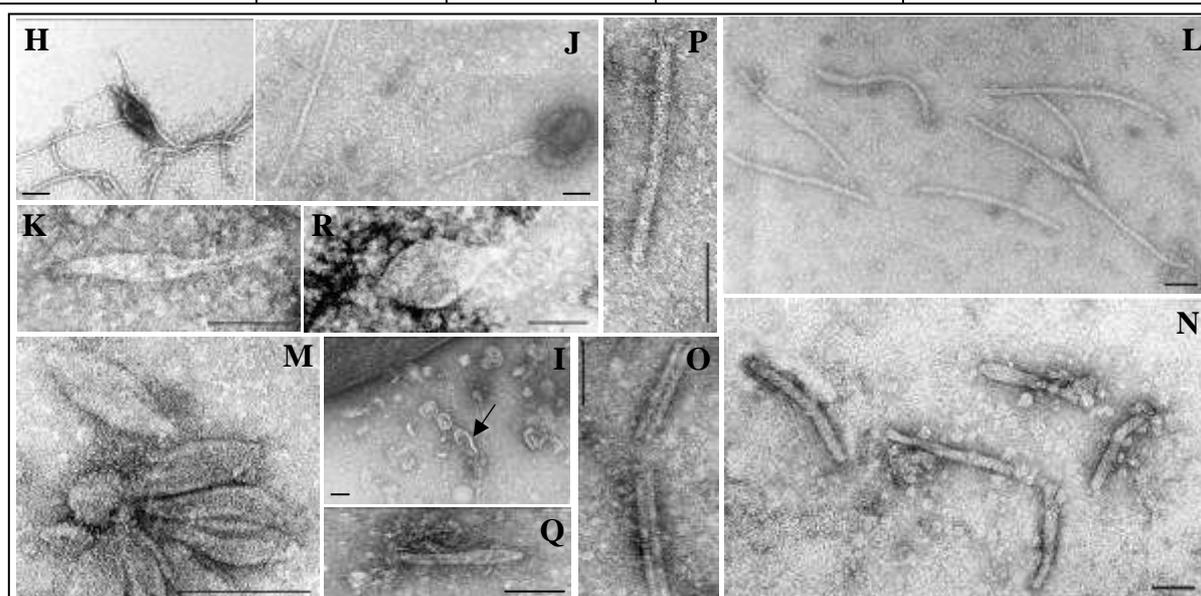
Nom de la culture d'enrichissement	Origine de l'échantillon	Morphologie des VLPs	Taille moyenne des VLPs	Référence des photos des VLPs
IR-96-01 int	MAR 36°N	Filament flexueux	925 nm x 25 nm	A
IR-96-CO3-03 bas	MAR 36°N	Bâtonnet à cavité centrale	80 nm x 20 nm	B
IR-98-VC	MAR 36°N	Cuillère	400 nm x 60 nm	C
		Bâtonnet simple	1300 nm x 20 nm	D
EX-08	EPR 9°N	Citron	130 nm x 60 nm	E
EX-12	EPR 9°N	Citron	160 nm x 50 nm	F
EX-37 NE	EPR 9°N	Bâtonnet simple	66 nm x 8 nm	G



**Figure 38 :** Photographies des VLPS ne se maintenant pas dans la culture d'enrichissement après plusieurs repiquages observées en MET après coloration négative. La barre correspond à 100 nm.

**Tableau 9 :** Caractéristiques des VLPs se maintenant dans la culture d'enrichissement après plusieurs repiquages.

Nom de l'échantillon brut	Origine de l'échantillon	Morphologie des VLPs	Taille des VLPs	Référence des photos des VLPs
IR-95-02	MAR 36°N	Citron	100 nm x 40 nm	H
IR-95-02 NE	MAR 36°N	Citron	200 nm x 60 nm	I
IR-96-01 ext	MAR 36°N	Bâtonnet Simple	575 nm x 15 nm	J
		Cuillère	325 nm x 50 nm	
IR-96-03 int	MAR 36°N	Cuillères	275 nm x 25 nm	K
IR-96-04 int	MAR 36°N	Cuillères	430 nm x 25 nm	L
		Citron à extensions	430 nm x 25 nm	
		Bâtonnet simple	400 nm x 25 nm	
IR-96-05 int	MAR 36°N	Citron	120 nm x 45 nm	M
IR-102-04	MAR 37°N	Massues	260 nm x 20 nm	N
		Bâtonnet à cavité centrale	185 nm x 20 nm	O
		Bâtonnets à arborisation	275 nm x 20 nm	P
EX-38	EPR 9°N	Massue	225 nm x 30 nm	Q
EX-39	EPR 9°N	Citron	225 nm x 100 nm	R



**Figure 39 :** Photographies des VLPS se maintenant dans la culture d'enrichissement après plusieurs repiquages observées en MET après coloration négative. La barre correspond à 100 nm.

Les particules présentent neuf morphologies distinctes et 4 cultures d'enrichissement présentent plusieurs morphotypes. Les VLPs peuvent être séparées en deux groupes suivant le critère morphologique. Le premier groupe rassemble les particules filamenteuses rigides et flexueuses. Les bâtonnets fins et rigides de 66 nm par 8 nm (cf Figure 38G) et les longs filaments flexueux d'environ 925 nm de long et 25 nm de large (cf Figure 38A) sont observées dans des cultures distinctes. Par ailleurs, trois morphologies différentes sont présentes dans la même culture d'enrichissement : (1) des bâtonnets rigides présentant une arborisation terminale de 275 nm x 20 nm (cf Figure 39P) qui ressemble à celle des *Rudivirus* SIRV1, SIRV2 et ARV1 (Prangishvili *et al.*, 1999 ; Vesterggard *et al.*, 2005), (2) des particules d'environ 185 nm x 25 nm caractérisées par une canalisation centrale qui rappelle la structure interne de SIRV1 et SIRV2 (Prangishvili *et al.*, 1999) (cf Figure 39O), (3) des VLPs avec une morphologie intermédiaire ressemblant à des massues. Il s'agit des bâtonnets plus ou moins rigides avec une extrémité plus large que l'autre dont la taille est approximativement de 260 nm x 20 nm (cf Figure 39N). Nous ne savons pas si la dernière morphologie observée correspond réellement à un nouveau morphotype. En effet, la structure des particules semble très flexible car tous les spécimens n'ont pas exactement la même allure. Etant donné l'extrême plasticité apparente de ces VLPs, il est possible qu'elles soient en réalité très déformées. Ce phénomène a déjà été observé chez le virus PAV1 lorsqu'il cultivé dans un milieu minimum (Geslin *et al.*, 2003). Ainsi, il apparaît que la différence de pression osmotique peut influencer la morphologie des VLPs. Dans le cas présent, il est possible d'imaginer que les particules sont très sensibles à la différence de pression. Ainsi, étant donné qu'elles sont issues d'échantillons prélevés à environ 900 m de profondeur, le fait de les cultiver à pression atmosphérique pourrait entraîner leur déformation.

Le second groupe de VLPs inclus exclusivement des particules en forme de citron. Deux sous-groupes peuvent être définis sur la base de la présence d'une queue courte ou plus longue. Les particules du sous-groupe I présentent une morphologie de citron à courte queue observée notamment chez PAV1, SSV1 et His1 (Geslin *et al.*, 2003 ; Martin *et al.*, 1984 ; Bath *et al.*, 1998). La longueur et la largeur de ces VLPs sont variables : de 100 nm x 40 nm pour les plus petites (cf Figure 39H) jusqu'à 225 nm x 125 nm pour les plus grandes (cf Figure 39P). Ces particules sont parfois organisées en rosette, attachées par leur queue à un fragment de membrane cellulaire (cf Figure 39M). Plusieurs de ces VLPs semblent posséder une certaine plasticité car elles des formes exagérément allongées ont pu être observées. Ces deux phénomènes ont pu être également observés chez PAV1 (Geslin *et al.*, 2003). La mor-

phologie de citron semble largement répandue aux quatre coins du globe étant donné qu'elle a pu être mise en évidence au niveau de la ride Médio-Atlantique (36°N) et de la ride Est-Pacifique (9°N) (cf Tableaux 8 et 9).

Le sous-groupe II est caractérisé des particules ressemblant à des cuillères qui possèdent une queue plus longue que leur tête en forme de citron. Ce type de VLPs n'a été observé que sur les cultures d'enrichissement issues d'échantillon de la ride Médio-Atlantique (36°N). La queue peut-être rigide (cf Figures 39J et 39K) ou flexible (cf Figures 38C et 39L). Les particules à queue rigide ont une tête d'environ 100 nm x 40 nm prolongée par une queue d'une longueur moyenne de 200 nm. Les VLPs à queue flexible possède une tête de taille comparable mais la queue est beaucoup plus longue car elle mesure en moyenne 300 nm. La plupart des particules en forme de cuillère sont retrouvées en mélange avec des VLPs ayant une morphologie de bâtonnets (cf Tableaux 8 et 9). Dans deux de ces cas, les bâtonnets présentent le même aspect et la même largeur que les queues des cuillères. Il est donc possible qu'ils correspondent à des queues séparées de la tête (cf Figures 39J et 39L). Le second morphotype rencontré au sein des particules du sous-groupe II correspond à des VLPs en forme de citron présentant des extensions bipolaires (cf Figure 39L). Le corps central fait environ 110 nm x 25 nm et les extensions mesurent approximativement 160 nm de long. Au total, la particule est longue de 430 nm. Cette morphologie rappelle celle du virus archéen récemment décrit nommé ATV. Ce dernier est capable de synthétiser ses deux extensions, de manière autonome, à l'extérieur de son hôte (Häring *et al.*, 2005c).

Cette étude a permis d'explorer pour la première fois la diversité morphologique des particules de type virale hyperthermophiles marines. Le résultat des observations montre une diversité insoupçonnée des VLPs produites par les *Thermococcales* des sources hydrothermales océaniques profondes de l'océan Atlantique et de l'océan Pacifique. A l'exception des formes filamenteuses connues également chez les bactéries (Ackermann, 2001), les morphologies observées ici semblent être caractéristiques des virus d'archées (Rice *et al.*, 2001 ; Rachel *et al.*, 2002). Des particules en forme de citron ont déjà été mises en évidence dans divers environnements, parfois extrêmes comme les sources chaudes acides terrestres (pour SSV1), les milieux hypersalins (pour His1) et les sédiments de subsurface (pour la VLPA3) (Martin *et al.*, 1984 ; Bath *et al.*, 1998 ; Wood *et al.* 1989). Cette étude agrandit leur répartition à l'environnement hydrothermal profond. Par ailleurs, les particules en forme de cuillère et celles ayant une morphologie de citron à extensions bipolaires semblent être inféodée à des environnements chauds. De plus, étant donné que ce sont majoritairement des archées qui peu-

plent ces milieux extrêmes, aucun des bactériophages connus à ce jour ne possède une telle morphologie.

Les résultats exposés ci-dessus ont été valorisés dans un article publié en 2003 dans la revue *Research in Microbiology* intitulé « *Occurrence of virus-like particles in high temperature enrichment cultures from deep-sea hydrothermal vents* » (cf Annexe 9).

### **3. Recherche de particules de type viral dans des isolats de *Thermococcales***

#### **a. Par microscopie électronique en transmission**

La première approche utilisée, pour trouver des particules de type viral infectant des souches de *Thermococcales*, consiste à réaliser des isolements sur milieu solide à partir des 9 cultures d'enrichissement où les VLPs se maintiennent après repiquage. Chaque isolat est ensuite observé en MET pour détecter une éventuelle production de particules de type viral.

L'effort d'isolement permet ainsi d'obtenir 127 isolats (cf Annexe 3). Après leur observation en MET, il s'avère qu'un seul, nommé E2p8, est porteur d'une VLP filamenteuse. Cet isolat est purifié puis rebaptisé KAZA. Le résultat obtenu est assez inattendu. En effet, la particule présente dans l'isolat n'a pas du tout la morphologie des VLPs initialement observées dans la culture d'enrichissement dont elle est issue (cf Figure 39D). Ces dernières sont des citrons de 200 nm de long au maximum alors que la particule filamenteuse atteint quasiment la longueur d'1µm. Il est donc assez peu probable qu'elle soit la résultante d'une déformation de la morphologie de citron. Une des explications à ce phénomène peut être que l'isolat porteur de la particule filamenteuse n'est pas majoritaire dans la culture d'enrichissement. Ainsi, les VLPs qu'il produit ne sont pas détectables en MET car pas assez nombreuses.

Le nombre de particules découvertes au cours de ce travail peut paraître dérisoire. En effet, le pourcentage d'isolats virosés issus des cultures d'enrichissement IR-95-02 et IR-95-02 NEensemencées avec l'échantillon brut *IR-95-02* n'est que de 3,7. Cependant, une étude récente, menée sur la diversité virale d'un échantillon de source chaude terrestre en MET, a montré que le pourcentage de souches virosées est de 2,5 (Håring et al., 2005b). Il apparaît nécessaire que le nombre d'isolats analysés soit considérable pour en découvrir un virosé. Par ailleurs, la MET permet de ne détecter que des virions produits en grande quantité par les cellules. En effet, la limite de détection est de  $10^4$  particules.mL<sup>-1</sup> (Ewert & Paynter, 1980). De plus, les virus lysogènes, qui sont présents uniquement par leur génome dans la cellule hôte,

sont indétectables. Nous avons donc décidé d'utiliser une seconde approche pour pallier les inconvénients de la MET.

### **b. Par Polymerase Chain Reaction**

Nous avons choisi la PCR car elle est simple à mettre en œuvre tout en étant très sensible. Le gène cible utilisé est le gène 899 de PAV1 qui, comme nous l'avons déjà évoqué, possède un homologue chez plusieurs plasmides de *Thermococcales* (pCIR10, pIRI48 et pTN2). Ce gène semble très conservé car si PAV1 est issu du bassin Nord Fidjien, les plasmides proviennent de la ride Centro-Indienne et de la ride Médio-Atlantique. Ce gène apparaît donc, à priori, comme un bon marqueur. Le couple d'amorces dégénérées REP-1S / REP-1AS est dessiné à partir de l'alignement des différentes séquences codant pour les gènes homologues. Ces amorces sont ensuite utilisées pour réaliser une PCR sur l'ADN total de 98 isolats de *Thermococcales*.

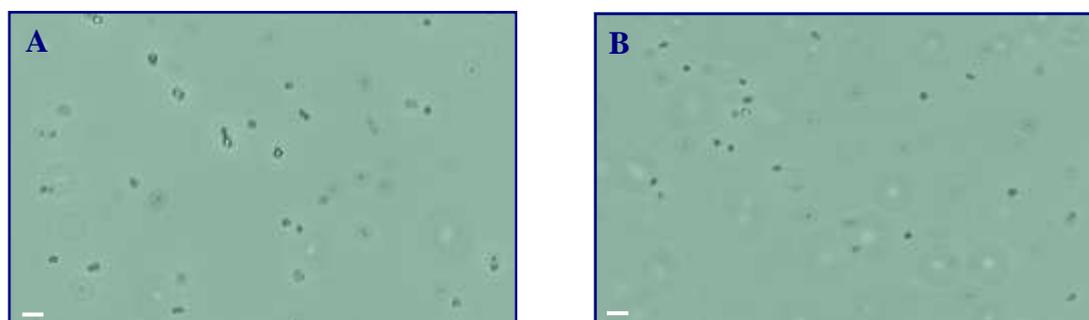
Seules trois souches donnent un résultat positif, à savoir : *Pyrococcus abyssi* GE23 (porteuse de PAV1), CIR10 (porteuse de pCIR10) et IRI48 (porteuse de pIRI48). Ces dernières correspondent en fait à des témoins positifs utilisés pour valider le couple d'amorces dessiné. Etant donné qu'aucune autre souche ne donne de signal, deux explications peuvent être envisagées. Premièrement, le gène conservé est peut-être moins répandu qu'il n'est apparu au début. Cependant, il est également tout à fait possible que le couple d'amorce utilisé reste trop spécifique. En effet, disposer seulement de trois séquences à aligner n'est peut être pas suffisant pour assurer la dégénérescence des amorces dessinées. A ce stade, il apparaît nécessaire qu'un plus grand nombre d'éléments génétiques mobiles de *Thermococcales* soient séquencés. S'il s'avère que ces derniers portent également un gène homologue au gène 899 de PAV1, il sera alors possible de concevoir d'autres amorces moins spécifiques qui permettront un criblage rapide et sensible pour la recherche de nouveaux virus producteurs ou non de particules.

## IV. Caractérisation préliminaire deux particules de type viral portées par une souche de *Thermococcales*

### 1. Etude de la souche hôte KAZA

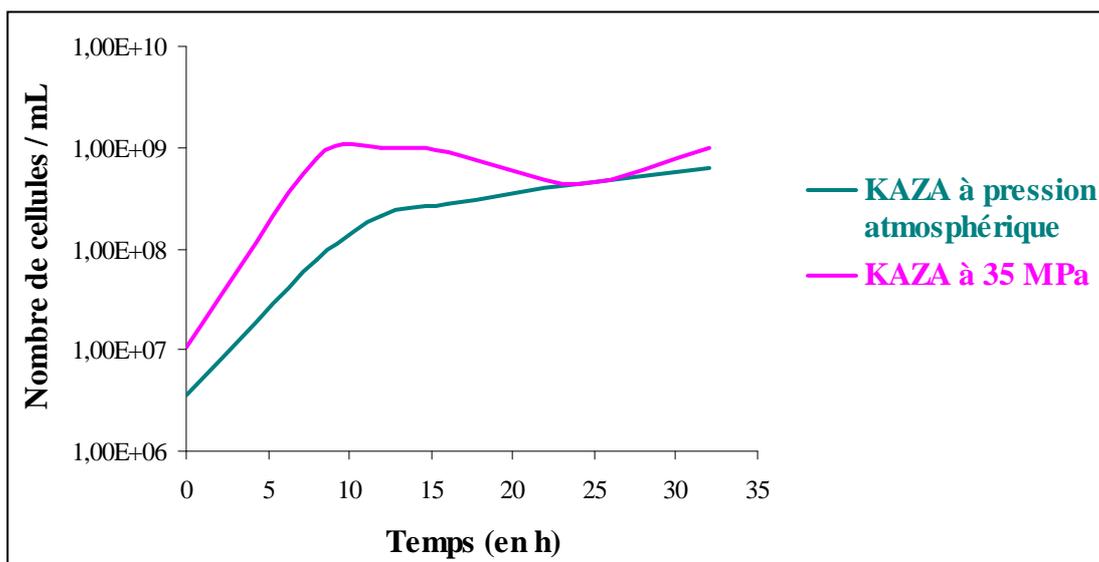
Pour déterminer à quelle espèce de *Thermococcales* appartient la souche KAZA, sur laquelle une VLP filamenteuse a été observée, le gène de son ADNr 16S est partiellement séquencé. La séquence obtenue (477 pb) présente 99 % d'homologie avec le gène de l'ADNr 16S de la souche MP qui est la souche de référence de l'espèce *Thermococcus barophilus* (Marteinsson *et al.*, 1999). Cette dernière présente la particularité d'avoir un taux de croissance deux fois supérieur sous une forte pression hydrostatique (40 MPa) par rapport à une faible pression (0,3 MPa). Nous avons donc décidé d'étudier l'influence de la pression hydrostatique sur la morphologie et la croissance de la souche KAZA. J. Deming et J. Baross ont montré que les micro-organismes thermophiles présentent un optimum de croissance à une pression supérieure à la pression existant sur le site de prélèvement (Deming & Baross, 1993). Ainsi, la souche KAZA provenant d'un échantillon prélevé à – 2300 m de profondeur (ce qui équivaut à 23 MPa), elle est cultivée à 35 MPa.

Dans un premier temps, nous avons réalisé des observations au microscope à contraste de phase des cultures de KAZA. A pression atmosphérique, les cellules sont très irrégulières et ont parfois la forme de faucilles. De plus, elles sont comme gonflées (jusqu'à 4  $\mu\text{m}$  de diamètre) et très réfringentes (cf Figure 40A). A l'inverse, à 35 MPa, les cellules sont rondes et très régulières. Par ailleurs, leur diamètre moyen est classique pour des *Thermococcales* (1  $\mu\text{m}$ ) et le cytoplasme apparaît plus dense (cf Figure 40B). Des résultats similaires ont déjà été obtenus pour la souche de référence MP (Marteinsson *et al.*, 1999).



**Figure 40** : Photographie en microscopie optique à contraste de phase de la souche KAZA à pression atmosphérique (A) et à 35 MPa (B). La barre représente 5  $\mu\text{m}$ .

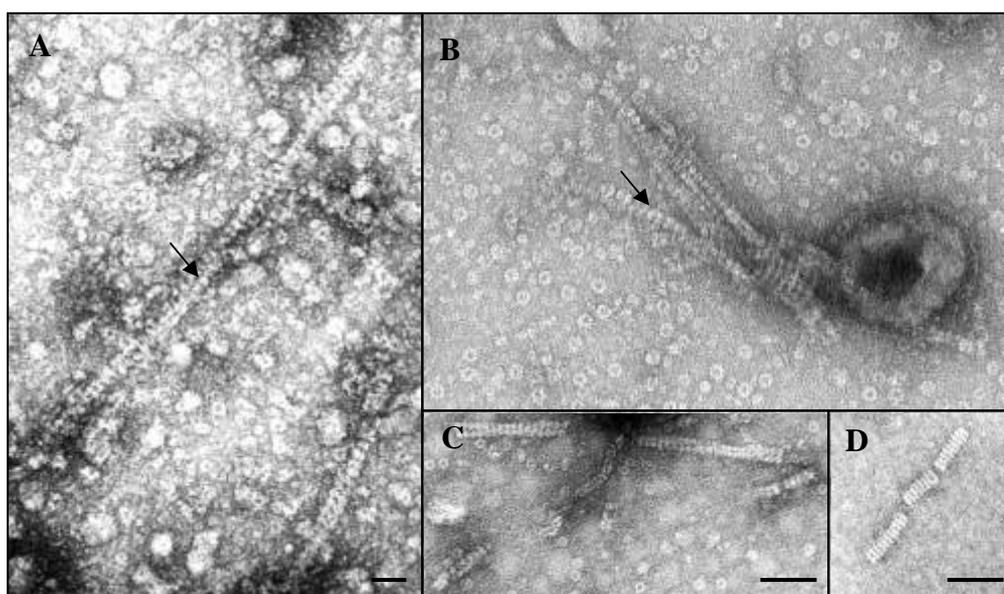
Nous avons ensuite réalisé en parallèle des cinétiques de la croissance de la souche KAZA à pression atmosphérique et à 35MPa. Les courbes obtenues, représentées sur la Figure 41, montrent que le taux de croissance de KAZA est plus élevé à 35 MPa qu'à la pression atmosphérique. Cette souche apparaît adaptée pour croître sous une pression hydrostatique importante, celle-ci étant caractéristique de son environnement naturel.



**Figure 41** : Cinétiques de croissance de la souche KAZA à pression atmosphérique et à 350 MPa.

## 2. Description morphologique de la particule de type viral filamenteuse

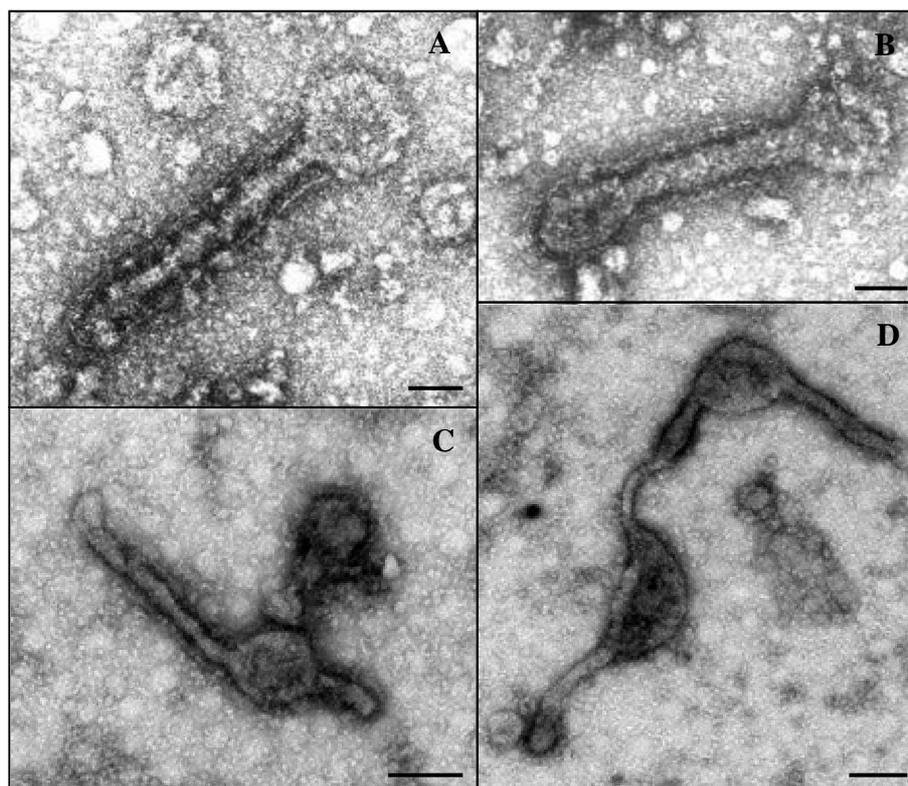
La particule de type virale observée sur la souche KAZA présente une morphologie filamenteuse particulière. En effet, le filament semble constitué d'une succession de petites sous-unités sphériques appariées par deux (cf Figure 42A, 42B et 42C). Etant donné sa forme, elle a été nommée VLP-TBFP pour « *Thermococcus barophilus* filamentous particle ». Elle mesure 900 nm x 20 nm (cf Figures 42A et 42B) mais il est très fréquent d'observer des fragments de plus faible longueur dont la taille varie de 100 nm à 300 nm (cf Figures 42C et 42D). Ces derniers semblent résulter de cassures, peut-être engendrées par la préparation de l'échantillon pour la MET. La structure de la VLP-TBFP apparaît donc relativement fragile. Une structure similaire à celle de la VLP-TBFP a déjà observé chez une VLP issue d'une source chaude acide terrestre. Cependant, les sous-unités qui la constituent ont une forme triangulaire et cette particule est six fois plus courte que la VLP-TBFP. Ainsi, il apparaît que la VLP-TBFP présente des caractéristiques tout à fait originales, jamais observées auparavant dans le monde des virus de procaryotes.



**Figure 42 :** Photographies de la VLP-TBFP entière (A, B) et de fragments (C, D) observés en MET après coloration négative. La barre représente 100 nm et les flèches pointent des sous-unités.

### 3. Etude de l'influence de la pression hydrostatique sur la morphologie de la VLP-TBFP

La souche KAZA est cultivée à 35 MPa puis le surnageant de culture est observé au MET après concentration par ultracentrifugation. Contre toute attente, aucune particule filamenteuse précédemment observée à pression atmosphérique n'est détectée ici sous pression hydrostatique. Par contre, des VLPs de morphologie tête-queue sont observables dans les préparations. La structure des VLPs observées semble être flexible car elle présente des variantes de trois grands types : le type « haltère », le type « à extensions bipolaires » et le type « bourgeonnant ». La forme tête-queue consiste en une tête sphérique d'un diamètre moyen de 120 nm et une queue de longueur variable (250 nm à 1 300 nm) et d'une largeur de 20 nm (cf Figure 43A). Le type haltère correspond à un bâtonnet plus ou moins long (en moyenne 300 nm) avec à chaque extrémité des têtes de 120 nm de diamètre (cf Figure 43B). Le type à extensions bipolaires consiste en une tête d'environ 100 nm possédant deux queues d'une longueur pouvant varier de 50 à 400 nm (cf Figure 43C). Le type bourgeonnant correspond à un enchaînement de bâtonnets d'environ 250 nm de long et de têtes parfois irrégulières d'un diamètre moyen de 120 nm (cf Figure 43D). L'enchaînement est plus ou moins long et peut atteindre 1 000 nm de long. Cette nouvelle VLP a été nommée VLP-TBFP pour « *Thermococcus barophilus* polymorphous particle ».



**Figure 43 :** Photographies des morphologies de la VLP-TBPP : le type tête-queue (A), le type « haltère » (B), le type « à extensions bipolaires » (C) et le type « bourgeonnant » (D) observées en MET après coloration négative. La barre représente 100 nm.

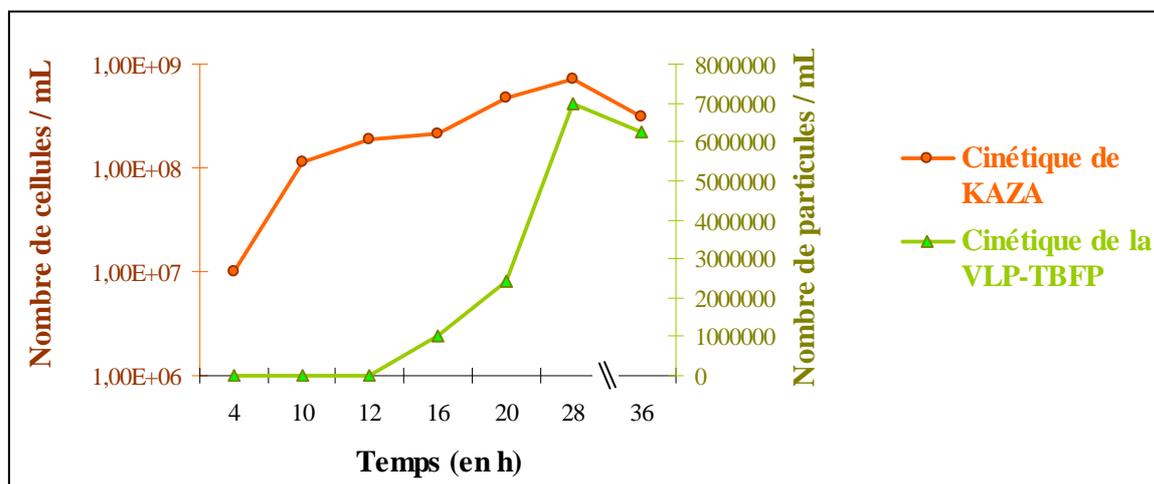
Deux explications peuvent être formulées pour expliquer ce ne soit pas le même type de particules observables à pression atmosphérique et sous pression hydrostatique. Premièrement, la VLP-TBFP n'est peut-être qu'une particule non achevée qui ne peut adopter sa forme définitive que sous pression hydrostatique. Le résultat correspondrait alors à la VLP-TBPP observées à 35 MPa. Cependant, il peut aussi s'agir de deux particules distinctes se développant dans des conditions de pression différentes. Pour répondre à cette question, il sera nécessaire d'avoir accès à leurs acides nucléiques et à leurs protéines pour les comparer.

#### **4. Etude de la production de la VLP-TBFP et de la VLP-TBPP**

##### **a. Etude de la production de la VLP-TBFP**

La souche KAZA semble être curée de la VLP-TBFP lors de repiquages successifs. Il arrive cependant que la particule réapparaisse sporadiquement dans les cultures sans raison particulière apparente. En traitant les cultures avec de la mitomycine C, la production de VLPs s'avère stable. Il apparaît donc qu'il n'y a pas réellement de curage de la souche mais plutôt une diminution du nombre de particules qui finit par ne plus être détectable en MET.

Par ailleurs, nous ne pouvons pas à proprement parler d'induction de la production de la VLP-TBFP par la mitomycine C, cependant cette dernière nous assure une quantité suffisante de particules pour des observations en MET. Nous avons alors réalisé des cinétiques de la croissance cellulaire et de la production de VLPs en présence de mitomycine C à pression atmosphérique et les courbes obtenues sont représentées sur la Figure 44.



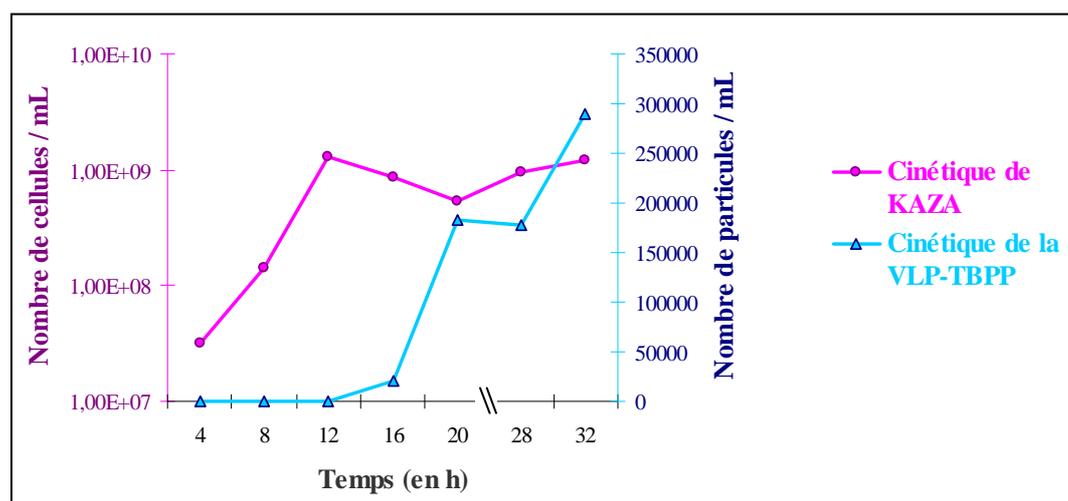
**Figure 44** : Cinétiques de croissance de la souche KAZA et de production de la VLP-TBFP avec traitement à la mitomycine C à pression atmosphérique.

La souche KAZA ne présente pas de lyse brutale car la lyse observée après 28 h de culture est tout à fait normale. Ainsi le fait pour l'archée de produire des particules, ne semble pas perturber sa croissance. En ce qui concerne la production de la VLP-TBFP, elle atteint son maximum après 28 h de culture, c'est-à-dire au moment où le nombre de cellules est également le plus important. Ce résultat suggère que la production de la VLP-TBFP est totalement dépendante de la croissance de la souche KAZA. Néanmoins, la quantité de particules reste faible car elle atteint au maximum  $7.10^6 \cdot \text{mL}^{-1}$ . Cependant, le même phénomène a déjà été observé pour PAV1 dont la production de virions ne dépasse pas  $1,6.10^6 \text{ particules} \cdot \text{mL}^{-1}$  (Geslin *et al.*, 2003). La relation entre la souche KAZA et la VLP-TBFP semble donc être une infection permanente sans mort de la cellule infectée. Etant donné que la production de particules n'est pas, à proprement parler, inductible, cette relation correspond à un état porteur, très fréquent chez les virus archéens.

#### **b. Etude de la production de la VLP-TBFP**

Dans un premier temps, nous avons réalisé des cultures de la souche KAZA à 35 MPa avec et sans traitement à la mitomycine C. Nous avons compté les particules dans les deux types de cultures, en phase de déclin pour s'assurer qu'un maximum de VLPs sont relarguées

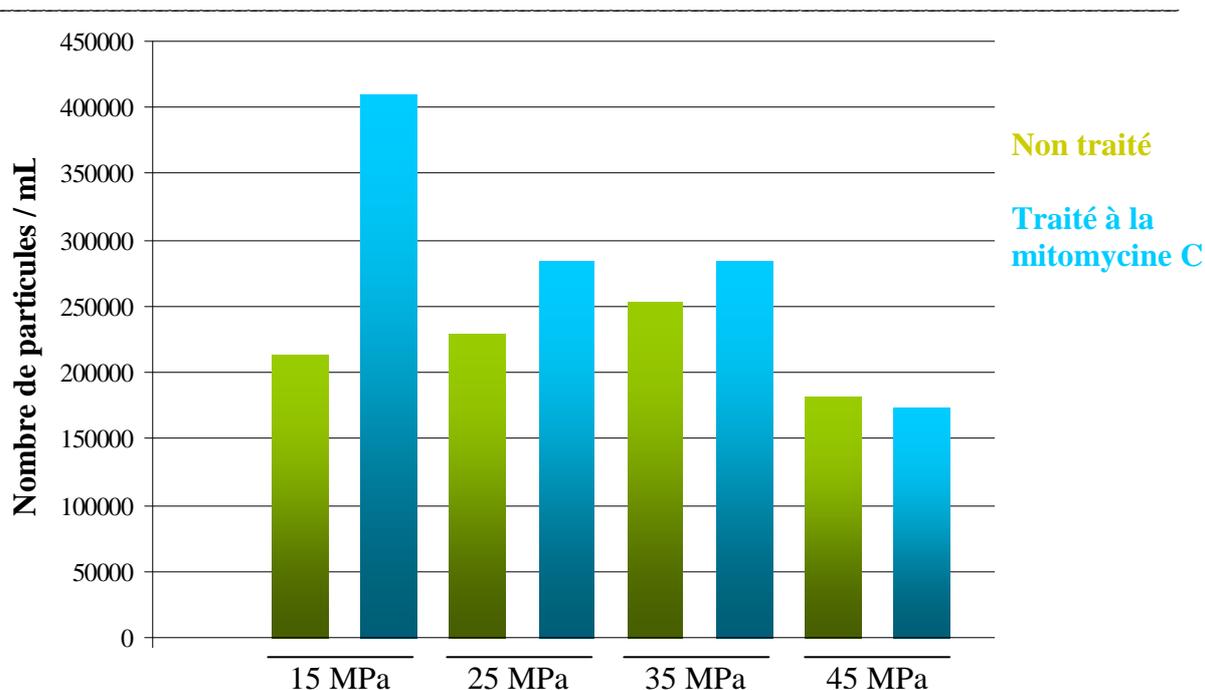
dans le milieu extérieur. Dans la culture témoin (non traitée), le nombre de particules est de  $2,53.10^5.mL^{-1}$  alors que dans la culture traitée il atteint  $2,84.10^5.mL^{-1}$ . Il existe donc une légère augmentation de la production de VLPs après traitement à la mitomycine C, ce qui suggère une faible induction de la VLP-TBPP. Nous avons ensuite réalisé des cinétiques de la croissance cellulaire et de la production de VLPs en présence de mitomycine C à 35 MPa et les courbes obtenues sont représentées sur la Figure 45.



**Figure 45 :** Cinétiques de la souche KAZA et de production de la VLP-TBPP avec traitement à la mitomycine C à 35 MPa.

La souche KAZA ne subit pas de lyse brutale car la lyse observée après 12 h de culture est normale étant donné la rapidité de sa croissance à 35 MPa. La production de la VLP-TBPP atteint un plateau entre 20 h et 28 h de culture, puis elle devient maximale à  $t = 32$  h. Elle est donc très importante lorsque la souche est en plein dans sa phase de déclin. Ces résultats mettent en évidence une corrélation entre la croissance de l'archée KAZA et la production de la VLP-TBPP. Ces dernières sont donc bien produites par la cellule sous pression hydrostatique. Cependant, la quantité de particules produites est extrêmement faible et plus de dix fois inférieure à ce qui est observé pour la VLP-TBFP.

Nous avons donc cherché des solutions pour augmenter cette production. Dans un premier temps, nous avons fait varier la pression hydrostatique en présence et en absence de mitomycine C. Les particules sont comptées sur des cultures de 32 h réalisées à 15 MPa, 25 MPa, 35 MPa et 45 MPa avec ou sans traitement. Les diagrammes obtenus à partir des comptages sont représentés sur la Figure 46.

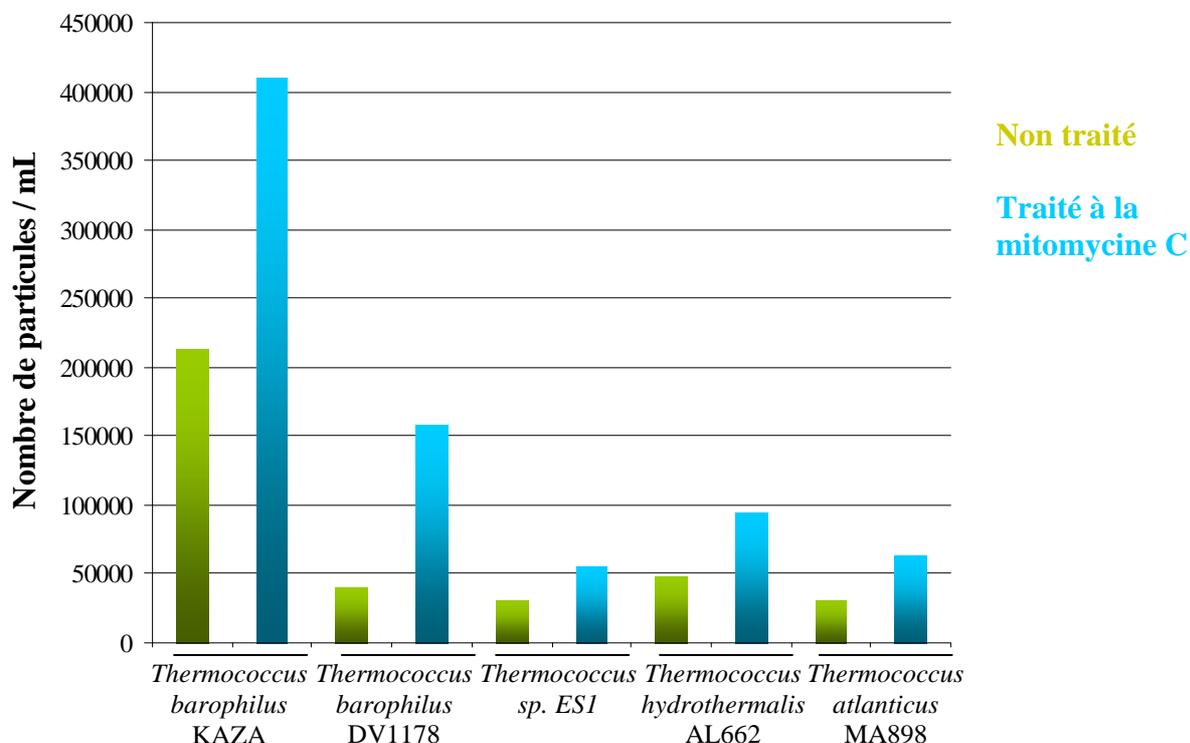


**Figure 46 :** Production des la VLP-TBPP à différentes pressions hydrostatiques avec et sans traitement à la mitomycine C.

Sur la culture témoin, la production de la VLP-TBPP est relativement homogène quelque soit la pression appliquée avec tout de même un maximum à 35 MPa. En effet, la pression optimale pour la croissance de la souche KAZA se situe aux environs de 35 MPa (Deming & Baross, 1993) et nous avons montré que la production de la VLP-TBPP est directement dépendante de la croissance archéenne. En revanche, avec un traitement de la culture à la mitomycine C, la production de particules atteint un maximum à 15 MPa pour diminuer graduellement jusqu'à 45 MPa. De plus, la mitomycine C semble augmenter la production de la VLP-TBPP à 15 MPa, 25 MPa et 35 MPa. Ce phénomène est particulièrement prononcé à 15 MPa car la quantité de particules est quasiment doublée par rapport à la culture témoin. En appliquant une pression hydrostatique de 15 MPa sur une culture traitée à la mitomycine C, nous avons réussi à augmenter la production de la VLP-TBPP d'un facteur 1,5, ce qui reste modeste.

Nous avons alors recherché des souches de *Thermococcales* capables d'une production de VLPs supérieure à celle observée sur la souche KAZA. Neuf souches ont été sélectionnées. et parmi elles, 7 appartiennent au genre *Thermococcus* (*Thermococcus barophilus* souche MP, *Thermococcus barophilus* souche DV1178, *Thermococcus* sp. souche ES1, *Thermococcus fumicolans* souche ST557, *Thermococcus atlanticus* souche MA898, *Thermococcus hydrothermalis* souche AL662, *Thermococcus siculi* souche AT1243) et 2 au genre *Pyrococcus*

(*Pyrococcus abyssi* souche GE23 et *Pyrococcus abyssi* souche GE9). Ces souches ont été cultivées à pression atmosphérique à 15 MPa avec ou sans traitement à la mitomycine C puis les VLPs ont été comptées. Le diagramme résultant de ces comptages correspond à la Figure 47.



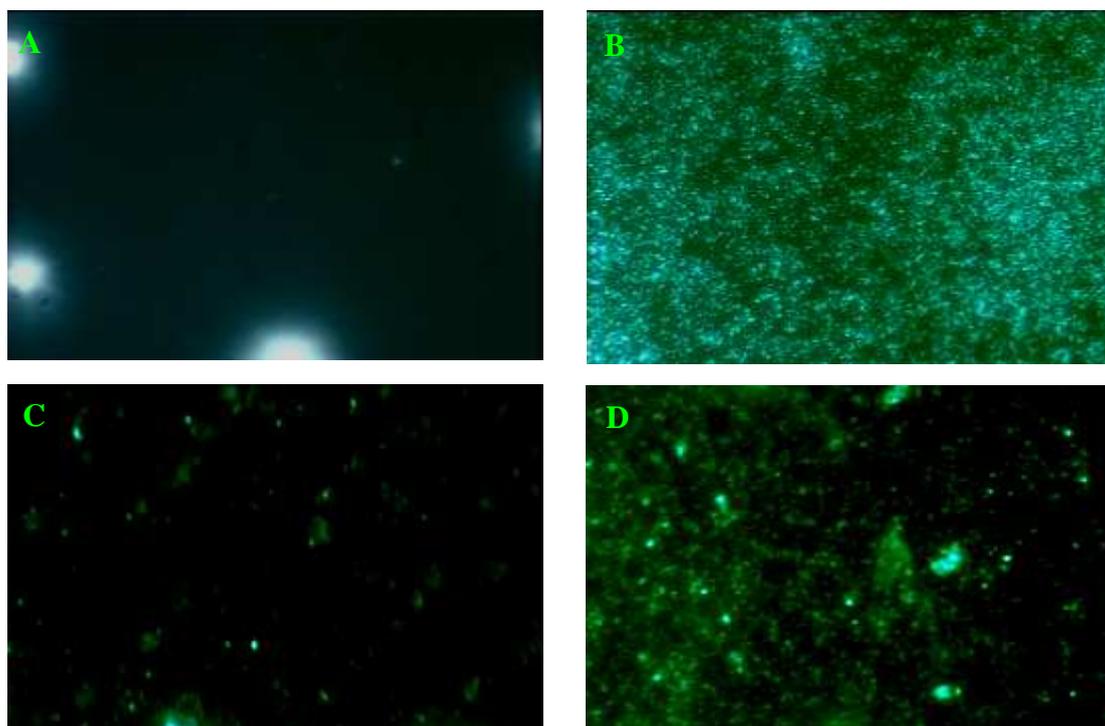
**Figure 47 :** Production de la VLP-TBPP par différentes souches de *Thermococcales* cultivées à 15 MPa avec et sans traitement à la mitomycine C.

A pression atmosphérique, aucune des souches, même traitées à la mitomycine C, ne porte de VLP-TBPP. A 15 MPa, quatre souches de *Thermococcus* produisent des particules de morphologie similaire à la VLP-TBPP : *Thermococcus barophilus* DV1178, *Thermococcus sp. ES1*, *Thermococcus hydrothermalis* AL662 et *Thermococcus atlanticus* MA898. Cette production est accrue par un ajout de mitomycine C. Cependant, la quantité de particules observées reste très faible par rapport à celle observée chez la souche KAZA.

## 5. Recherche d'un acide nucléique chez la VLP-TBFP et la VLP-TBPP

Les techniques classiques d'extraction réalisées sur les particules, concentrées par ultracentrifugation, n'ont pas permis de mettre en évidence de l'ADN ou de l'ARN dans la VLP-TBFP et la VLP-TBPP. La quantité d'acide nucléique étant peut-être très faible, nous

avons décidé de réaliser un test en microscopie à épifluorescence sur les deux types de particules. Les photographies obtenues sont regroupées dans la Figure 48.



**Figure 48** : Photographies après marquage au SYBR Green I du témoin négatif (eau milliQ) (A), du témoin positif (PAV1) (B), de la VLP-TBFP (C) et de la VLP-TBPP (D) observés en microscopie à épifluorescence.

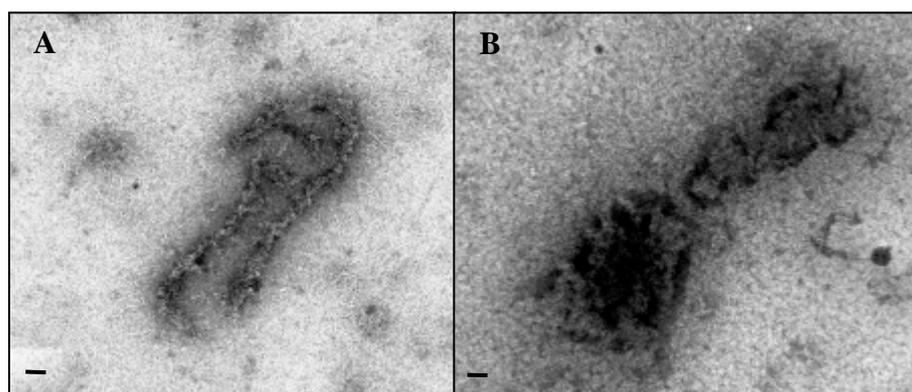
Etant donné la quasi-absence de signal observée sur la Figure 48C, aucun acide nucléique ne semble être présent dans la VLP-TBFP. Une absence de matériel génétique a déjà été constatée chez une particule de type viral filamenteuse possédant des motifs triangulaires issue d'un échantillon de source chaude terrestre. D'après D. Prangishvili, cette particule serait un assemblage protéique sécrété par la souche productrice en condition de stress (D. Prangishvili, communication personnelle). La VLP-TBFP peut également correspondre à une particule non achevée. En revanche, le signal enregistré sur la Figure 48D, bien que d'une intensité inférieure à celle observée pour PAV1, suggère que la VLP-TBPP renfermerait de l'acide nucléique. La faiblesse de l'intensité se justifie par la faible quantité de particules produites. Dans le but d'obtenir une preuve directe de l'absence ou de la présence d'acides nucléiques dans les deux types de VLPs, des expériences d'hybridation *in situ* sur des coupes ultra-fines de suspension de particules sont actuellement en cours. Il pourra ensuite être envisagé d'utiliser des kits commerciaux spécialisés pour l'extraction acides nucléiques viraux présents en très faible quantité.

## 6. Purification de la VLP-TBFP et de la VLP-TBPP

Les trois protocoles de purification présentés dans la partie *Matériel et Méthode* sont testés sur la VLP-TBFP sans aucun succès. En effet, très peu de petits fragments de 100 nm de long sont récoltés dans les fractions purifiées alors que la taille attendue des particules est de 900 nm. Des contrôles intermédiaires réalisés au cours des différents protocoles montrent que la particule est très sensible au PEG, à l'ultracentrifugation et au CsCl. L'ultrafiltration tangentielle, plus douce, ne permet cependant pas de purifier la VLP-TBFP car celle-ci ne semble pas passer la première cartouche possédant des pores de 1 000 kDa. D'une manière très regrettable, aucune manipulation complémentaire n'a pu être faite sur cette particule du fait de sa trop grande fragilité. En revanche, le protocole comportant une ultracentrifugation suivie d'un gradient de densité nous a permis de purifier la VLP-TBPP. Les 4 types de morphologies de la VLP-TBPP sont en mélange avec quelques débris de membranes archéennes. Il est donc possible que la VLP-TBPP soit constituée en partie par la membrane cellulaire.

## 7. Etude préliminaire de la structure de la VLP-TBPP

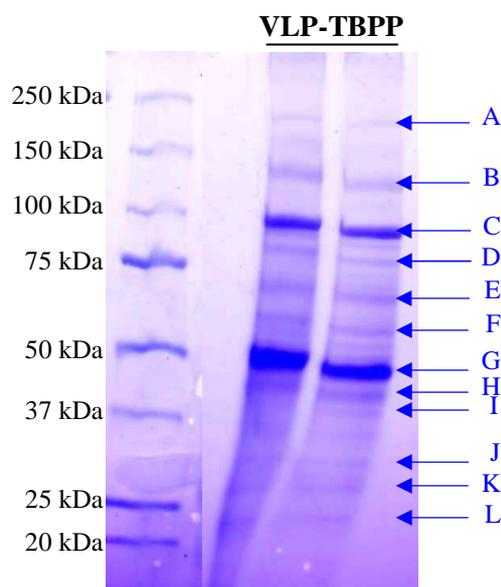
Les traitements au Triton X-100, au SDS et à la protéinase K, qui ont la propriété de désorganiser les structures protéiques, entraînent une disparition des VLPs. De petits amas ressemblant à des pelotes de laine emmêlée sont majoritairement observables. Des structures en cours de désagrégation sont aussi présentes comme, par exemple, des particules surdimensionnées aux contours fantomatiques (cf Figure 49). La VLP-TBPP est donc majoritairement composée de protéines.



**Figure 49** : Photographies de la VLP-TBPP de type « haltère » (A) et de type tête-queue (B) traitées à la protéinase K et observées en MET après coloration négative. La barre représente 100 nm.

Par ailleurs, le traitement au chloroforme, qui est un solvant des lipides, est sans effet notable sur les particules. La VLP-TBPP, tout comme les bactériophages tête-queue, n'est pas enveloppée. La membrane observable au niveau de la queue résulterait donc plutôt d'un artefact de coloration.

Nous avons ensuite réalisé une électrophorèse en SDS-PAGE des particules purifiées. Le profil électrophorétique des protéines contenues dans la fraction obtenue après purification de la VLP-TBPP présente 12 bandes majoritaires de poids moléculaires compris entre 23 kDa et 175 kDa (cf Figure 50).



**Figure 50 :** Profil électrophorétique des protéines contenues dans la fraction obtenue après purification de la VLP-TBPP. Les flèches bleues indiquent les bandes envoyées pour analyse en spectrométrie de masse Maldi-tof.

Etant donné qu'il est très probable que certaines aient une origine cellulaire, elles ont toutes été en cours analysées par spectrométrie de masse Maldi-tof. Les résultats obtenus sont résumés dans le tableau 10. Il apparaît que la majorité des protéines sont issues de l'hôte. Cependant, il est intéressant de noter que deux protéines ont un profil tryptique similaire à celui d'une protéine hypothétique d'un cyanophage marin (S-PM2) et à celui d'une protéine de nucléocapside d'un virus eucaryote (IBV). Il va donc être nécessaire de micro-séquencer ces molécules d'intérêt pour compléter ces résultats préliminaires très encourageants.

**Tableau 10 :** Protéines assignées aux différentes bandes protéiques obtenues après électrophorèse en SDS-PAGE de la VLP-TBPP purifiée.

Bande	Protéine assignée	Base de données utilisée	Score	Taux de couverture
A	Phosphoenolpyruvate synthase de <i>Pyrococcus abyssi</i>	PAbyssi30mai06	6,40e+004 (MOWSE)	17 %
B	Pas d'identification possible	-	-	-
C	Phosphoenolpyruvate synthase de <i>Thermococcus kodakarensis</i>	NCBI	91 (Mascott)	31 %
D	Phosphoenolpyruvate synthase de <i>Pyrococcus abyssi</i>	PAbyssi30mai06	8,24e+003 (MOWSE)	15 %
E	Protéine hypothétique du cyanophage S-PM2	TrEMBL	8,50e-006 (pValue)	27 %
F	Protéine de nucléocapside du virus de la bronchite infectieuse	TrEMBL	9,80e-008 (pValue)	21 %
G	Glutamate déshydrogénase de <i>Thermococcus littoralis</i>	Swissprot	1,37e+009 (MOWSE)	31 %
H	Glutamate déshydrogénase de <i>Pyrococcus abyssi</i>	PAbyssi30mai06	3,76e+003 (MOWSE)	18 %
I	Protéine hypothétique de <i>Pyrococcus abyssi</i>	PAbyssi30mai06	1,13e+003 (MOWSE)	26 %
J	Protéine hypothétique de <i>Pyrococcus abyssi</i>	PAbyssi30mai06	1,13e+003 (MOWSE)	26 %
K	Enolase de <i>Pyrococcus abyssi</i>	PAbyssi30mai06	2,33e+003 (MOWSE)	17 %
L	Protéine hypothétique de <i>Methanococcus jannaschii</i>	Swissprot	1,83e+004 (MOWSE)	35 %



## Conclusion générale et Perspectives

Au cours de ces dernières années, l'essor de la métagénomique virale et les comptages directs de particules ont permis de propulser les virus au rang des entités biologiques les plus abondantes et les plus diversifiées, à tel point que D. Bamford parle d'une vie biologique dans un « océan viral ». Les virus représentent, après les procaryotes, le deuxième constituant le plus important de la biomasse océanique. En effet, la quantité de carbone qui constitue les virus marins est estimée à près de 200 Mt, ce qui en fait un des éléments majeurs de stockage du carbone organique sur la planète. La viriosphère influence fortement les grands cycles biogéochimiques et même le climat. De plus, en régulant les populations microbiennes par son pouvoir lytique, elle occupe une place prépondérante dans la dynamique et l'équilibre des écosystèmes aquatiques. Par ailleurs, les virus, de part leur capacité de transfert des gènes, participent à l'évolution et à l'adaptation des communautés microbiennes.

Concernant la diversité, il a été montré que l'eau de mer est extrêmement riche en virus car 200 L peuvent contenir entre 3 000 et 7 000 génotypes viraux différents. Et pourtant, seuls 5 génomes de virus procaryotes marins ont été entièrement séquencés à ce jour et 4 sont des cyanophages. Ceci est dicté par le fait que les cyanobactéries sont les plus importants producteurs primaires au niveau des écosystèmes aquatiques de surface qui dépendent de la photosynthèse. Par opposition, une telle démarche n'avait pas encore été menée dans un écosystème basé sur la chiotrophie comme l'environnement hydrothermal abyssal qui est pourtant aujourd'hui reconnu comme étant une oasis de vie. Dans ce contexte, l'étude du génome de PAV1 prend tout son sens et apparaît comme pionnière. Les approches de métagénomique virale génèrent des séquences de virus incultivées qui ne permettent pas de pouvoir aborder l'écologie de l'environnement dans sa globalité. Ce genre d'étude n'est possible qu'en étudiant des systèmes hôte – virus caractérisés. Les archées de l'ordre des *Thermococcales* constituent, de part leur abondance et leur diversité au niveau des événements hydrothermaux abyssaux, une réserve importante de nutriments carbonés, azotés et phosphatés. C'est pourquoi l'étude des virus de *Thermococcales* revêt un intérêt tout particulier.

La forme très originale de citron que présentent SSV1 (virus hyperthermophile terrestre de *Crenarchaeota*), His1 (virus mésophile halophile d'*Euryarchaeota*) et PAV1 (virus hyperthermophile marin d'*Euryarchaeota*) semble spécifique des virus du domaine *Archaea*.

En effet, ces trois virus présentent une morphologie similaire et une aptitude à vivre dans un environnement dit extrême. Cependant, leurs génomes sont totalement différents. Ceci montre bien que ces virus fuselliformes ont certainement des origines phylogénétiques différentes. La forme citron n'est donc probablement qu'une convergence fonctionnelle due à la pression exercée par l'environnement (pression osmotique, pression hydrostatique, haute température). Il serait donc plus judicieux de parler d'« analogie » de forme plutôt que d'« homologie ». Dans le même ordre d'idée, PAV1 étant totalement différent de ses comparses morphologiques, il est possible d'envisager de proposer une nouvelle famille virale dont il serait le premier représentant.

D'une manière générale, les virus archéens du phylum des *Crenarchaeota* possèdent un pool de gènes communs qui leur sont propres. Pour D. Prangishvili, ces virus ont donc une origine différente de ceux du phylum des *Euryarchaeota* et des bactériophages. Par ailleurs, au sein même des virus d'*Euryarchaeota*, il semble qu'il existe deux origines distinctes : une commune avec les bactériophages et une totalement différente de tous les autres virus. En effet, la plupart des spécimens décrits sont des virus ayant une morphologie et une parenté très proche des phages. La découverte de la particule de type viral VLP-TBPP nous montre que la structure tête-queue n'est pas réservée aux spécimens mésophiles. En revanche, SH1, qui possède une capsid polyédrique, est génétiquement très différent des bactériophages. Les virus en forme de citron, eux, semblent également posséder des génomes très originaux avec un seul gène qui s'avère être d'origine phagique.

Une autre grande découverte issue des analyses métagénomiques est le fait que plus de 65 % des gènes portés par les virus n'ont pas de fonction connue. Mais l'ensemble des travaux de génomique portant sur les virus hyperthermophiles, y compris ceux menés sur PAV1, montre que ces derniers présentent un taux de gènes orphelins bien supérieur (supérieur à 90 %). Ces virus constituent donc un énorme réservoir de nouvelles structures protéiques et fonctions biologiques. Un coin du voile a été levé avec la découverte d'une nouvelle enzyme multifonctionnelle (aux activités polymérase, primase et ATPase) chez un plasmide hyperthermophile de *Crenarchaeota*. D'après l'hypothèse de P. Forterre, les virus seraient à l'origine de la création des protéines intervenant dans les mécanismes de réplication, de recombinaison et de réparation dans le monde cellulaire. Il apparaît donc qu'il est du plus grand intérêt d'explorer ces nouveaux réservoirs de gènes que sont les virus hyperthermophiles. Ainsi, malgré les dif-

difficultés rencontrées, les travaux réalisés sur la protéine 899 de PAV1, il apparaît nécessaire de persévérer dans cette voie.

Le résultat de la « chasse » au virus de *Thermococcales* des sources hydrothermales abyssales engagée lors de cette thèse est assez inattendu. En effet, si la faible quantité de particules de type viral isolées était prévisible, leur « piézo-dépendance » est tout à fait originale dans le monde des virus de procaryotes. Ceci constitue une étude miroir à ce qui a été réalisée, il y a de cela 40 ans sur les bactériophages et leur inactivation sous pression hydrostatique. Il semble donc que les *Thermococcales* soient un réservoir de virus émergents (du fond des océans) aux propriétés insoupçonnées qui confortent les espoirs mis dans ce projet de longue haleine.

Les perspectives à envisager pour ce travail de thèse concernent tout d'abord le cycle viral de PAV1. En effet, les différents travaux réalisés jusqu'à présent ont permis de répondre à de nombreuses questions cependant, plusieurs points restent à élucider :

- la reconnaissance hôte – virus se fait-elle réellement grâce à une interaction entre les glycoprotéines (de la S-layer ou de la membrane) de l'hôte et les protéines virales 676 et 678 ?
- le génome viral pénètre-t-il tout seul dans la cellule ou est-ce le virion dans sa totalité ? Les deux possibilités peuvent également être envisagées de concert.
- le mode de réplication du génome viral adopté est-il réellement unidirectionnel, quel est-il et quelles sont les protéines impliquées dans ce processus ?
- la transcription comporte-t-elle des fonctions précoces et tardives ?
- la sortie des virions se fait-elle par bourgeonnement, par excrétion ou par un autre mécanisme ?

La découverte de la VLP-TBPP nécessite une étude approfondie pour démontrer qu'il s'agit effectivement d'un virus. Il est primordial de mettre en évidence de manière directe la présence d'un acide nucléique dans la particule. Il serait également très intéressant de tester son caractère infectieux. Par ailleurs, les observations en microscopie électronique en transmission suggèrent que la VLP-TBPP pourrait être à assemblage piézo-dépendant. Ceci sous-entend que la VLP-TBFP correspondrait en fait à la queue de la VLP-TBPP. Pour prouver cette hypothèse, des protéines communes aux deux types de particules pourraient être recherchées et une reconstruction tridimensionnelle des structures des VLPs pourrait être réalisée.

Dans le cadre d'une approche plus globale, il semble intéressant de se diriger vers une approche de génomique structurale et fonctionnelle pour rechercher de nouvelles structures protéiques et de nouvelles fonctions biologiques codées par les éléments génétiques des hyperthermophiles marins. Cet axe de recherche est déjà en place avec la création récente de l'ACI Deep Hot Vir et de l'ANR *Genoarchaea*.

Changer d'échelle d'étude et avoir une vue générale de l'environnement hydrothermal abyssal est également du plus grand intérêt. Ainsi, deux projets vont être proposés très prochainement pour étudier la diversité (ANR Deep Oases) et la dynamique (ANR COD) des populations virales.

## Références Bibliographiques

### A

- Ackermann, H.W. (2001) Frequency of morphological phage descriptions in the year 2000. Brief review. *Arch Virol* **146**: 843-857.
- Aertsen, A., Faster, D., and Michiels, C.W. (2005) Induction of Shiga toxin-converting prophage in *Escherichia coli* by high hydrostatic pressure. *Appl Environ Microbiol* **71**: 1155-1162.
- Aiuti, A., Slavin, S., Aker, M., Ficara, F., Deola, S., Mortellaro, A. *et al.* (2002) Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* **296**: 2410-2413.
- Albright, L.M., Kassavetis, G.A., and Geiduschek, E.P. (1988) Bacteriophage T4 late transcription from plasmid templates is enhanced by negative supercoiling. *J Bacteriol* **170**: 1279-1289.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.
- Arnold, H.P., Ziese, U., and Zillig, W. (2000b) SNDV, a novel virus of the extremely thermophilic and acidophilic archaeon *Sulfolobus*. *Virology* **272**: 409-416.
- Arnold, H.P., She, Q., Phan, H., Stedman, K., Prangishvili, D., Holz, I. *et al.* (1999) The genetic element pSSVx of the extremely thermophilic crenarchaeon *Sulfolobus* is a hybrid between a plasmid and a virus. *Mol Microbiol* **34**: 217-226.
- Arnold, H.P., Zillig, W., Ziese, U., Holz, I., Crosby, M., Utterback, T. *et al.* (2000a) A novel lipothrixvirus, SIFV, of the extremely thermophilic crenarchaeon *Sulfolobus*. *Virology* **267**: 252-266.
- Azzazy, H.M.E., and Highsmith, W.E.J. (2002) Phage display technology: clinical applications and recent innovations. *Clinical biochemistry* **35**: 425-445.

### B

- Baltimore, D. (1971) Expression of animal virus genomes. *Bacteriological Reviews* **35**: 235-241.

- Bamford, D.H., Burnett, R.M., and Stuart, D.I. (2002) Evolution of viral structure. In *Theor Popul Biol*, pp. 461-470.
- Bamford, D.H., Ravantti, J.J., Ronnholm, G., Laurinavicius, S., Kukkaro, P., Dyall-Smith, M. *et al.* (2005) Constituents of SH1, a novel lipid-containing virus infecting the halophilic euryarchaeon *Haloarcula hispanica*. *J Virol* **79**: 9097-9107.
- Baranska, S., Gabig, M., Wegrzyn, A., Konopa, G., Herman-Antosiewicz, A., Hernandez, P. *et al.* (2001) Regulation of the switch from early to late bacteriophage lambda DNA replication. *Microbiology* **147**: 535-547.
- Baranyi, U., Klein, R., Lubitz, W., Kruger, D.H., and Witte, A. (2000) The archaeal halophilic virus-encoded Dam-like methyltransferase M. phiCh1-I methylates adenine residues and complements dam mutants in the low salt environment of *Escherichia coli*. *Mol Microbiol* **35**: 1168-1179.
- Bardy, S.L., Eichler, J., and Jarrell, K.F. (2003) Archaeal signal peptides--a comparative survey at the genome level. *Protein Sci* **12**: 1833-1843.
- Barns, S.M., Delwiche, C.F., Palmer, J.D., and Pace, N.R. (1996) Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *PNAS* **93**: 9188-9193.
- Bates, A.D., and Maxwell, A. (1993) *DNA topology*. Oxford: IRL Press.
- Bath, C., and Dyall-Smith, M.L. (1998) His1, an archaeal virus of the Fuselloviridae family that infects *Haloarcula hispanica*. *J. Virol.* **72**: 9392-9395.
- Bath, C., Cukalac, T., Porter, K., and Dyall-Smith, M.L. (2006) His1 and His2 are distantly related, spindle-shaped haloviruses belonging to the novel virus group, Salterprovirus. *Virology* **349**: In press.
- Benbouzid-Rollet, N., Lopez-Garcia, P., Watrin, L., Erauso, G., Prieur, D., and Forterre, P. (1997) Isolation of new plasmids from hyperthermophilic Archaea of the order Thermococcales. *Res Microbiol* **148**: 767-775.
- Bettarel, Y., Sime-Ngando, T., Amblard, C., and Laveran, H. (2000) A comparison of methods for counting viruses in aquatic systems. *Appl Environ Microbiol* **66**: 2283-2289.
- Bettstetter, M., Peng, X., Garrett, R.A., and Prangishvili, D. (2003) AFV1, a novel virus infecting hyperthermophilic archaea of the genus acidianus. *Virology* **315**: 68-79.
- Bird, D.F., Juniper, K., Ricciardi-Rigault, M., Martineu, P., Prairie, Y.T., and Calvert, S.E. (2001) Subsurface viruses and bacteria in Holocene/Late Pleistocene sediments of Saanich Inlet, BC: ODP Holes 1033B and 1034B, Leg 169S. *Marine Geology* **174**: 227-239.

- Birkenbihl, R.P., Neef, K., Prangishvili, D., and Kemper, B. (2001) Holliday junction resolving enzymes of archaeal viruses SIRV1 and SIRV2. *J Mol Biol* **309**: 1067-1076.
- Birnboim, H.C., and Doly, J. (1979) A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**: 1513-1523.
- Blanco, L., Lazaro, J.M., de Vega, M., Bonnin, A., and Salas, M. (1994) Terminal protein-primed DNA amplification. *PNAS* **91**: 12198-12202.
- Blondal, T., Hjorleifsdottir, S.H., Fridjonsson, O.F., Aevarsson, A., Skirnisdottir, S., Hermannsdottir, A.G. *et al.* (2003) Discovery and characterization of a thermostable bacteriophage RNA ligase homologous to T4 RNA ligase 1. *Nucleic Acids Res* **31**: 7247-7254.
- Blondal, T., Thorisdottir, A., Unnsteinsdottir, U., Hjorleifsdottir, S., Aevarsson, A., Ernstsson, S. *et al.* (2005) Isolation and characterization of a thermostable RNA ligase 1 from a *Thermus scotoductus* bacteriophage TS2126 with good single-stranded DNA ligation properties. *Nucleic Acids Res* **33**: 135-142.
- Blum, H., Zillig, W., Mallok, S., Domdey, H., and Prangishvili, D. (2001) The genome of the archaeal virus SIRV1 has features in common with genomes of eukaryal viruses. *Virology* **281**: 6-9.
- Borriss, M., Helmke, E., Hanschke, R., and Schweder, T. (2003) Isolation and characterization of marine psychrophilic phage-host systems from Arctic sea ice. *Extremophiles* **19**: 19.
- Bratbak, G., Thingstad, F., and Heldal, M. (1994) Viruses and the microbial loop. *Microb Ecol* **28**: 209-221.
- Breitbart, M., and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends Microbiol* **13**: 278-284.
- Breitbart, M., Hewson, I., Felts, B., Mahaffy, J.M., Nulton, J., Salamon, P., and Rohwer, F. (2003) Metagenomic analyses of an uncultured viral community from human feces. *J Bacteriol* **185**: 6220-6223.
- Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M., Mead, D. *et al.* (2002) Genomic analysis of uncultured marine viral communities. *PNAS* **99**: 14250-14255.
- Briani, F., Deho, G., Forti, F., and Ghisotti, D. (2001) The plasmid status of satellite bacteriophage P4. *Plasmid* **45**: 1-17.
- Brussaard, C.P., Marie, D., and Bratbak, G. (2000) Flow cytometric detection of viruses. *J Virol Methods* **85**: 175-182.
- Brüssow, H., and Hendrix, R.W. (2002) Phage genomics: small is beautiful. *Cell* **108**: 13-16.

Brüssow, H., Fremont, M., Bruttin, A., Sidoti, J., Constable, A., and Fryder, V. (1994) Detection and classification of *Streptococcus thermophilus* bacteriophages isolated from industrial milk fermentation. *Appl Environ Microbiol* **60**: 4537-4543.

Büchen-Osmond, C. (2003a) Taxonomy and classification of viruses. In *Manual of clinical microbiology, 8th ed.* Washington, pp. 1217-1225.

Büchen-Osmond, C. (2003b) The universal virus database ICTVdb. *Computing in Science & Engineering* **5**: 2-11.

## C

Casjens, S.R. (2005) Comparative genomics and evolution of the tailed-bacteriophages. *Curr Opin Microbiol* **8**: 451-458.

Casjens, S.R., and Hendrix, R.W. (2001) Bacteriophage lambda and its relatives. In *Encyclopedia of life science*. www.els.net (ed): Nature publishing group.

Casjens, S.R., Gilcrease, E.B., Huang, W.M., Bunny, K.L., Pedulla, M.L., Ford, M.E. *et al.* (2004) The pKO2 linear plasmid prophage of *Klebsiella oxytoca*. *J Bacteriol* **186**: 1818-1832.

Charbonnier, F., and Forterre, P. (1994) Comparison of plasmid DNA topology among mesophilic and thermophilic eubacteria and archaeobacteria. *J Bacteriol* **176**: 1251-1259.

Charbonnier, F., Erauso, G., Barbeyron, T., Prieur, D., and Forterre, P. (1992) Evidence that a plasmid from a hyperthermophilic archaeobacterium is relaxed at physiological temperatures. *J Bacteriol* **174**: 6103-6108.

Chen, F., Lu, J.R., Binder, B.J., Liu, Y.C., and Hodson, R.E. (2001) Application of digital image analysis and flow cytometry to enumerate marine viruses stained with SYBR gold. *Appl Environ Microbiol* **67**: 539-545.

Chen, H., Joerger, R.D., Kingsley, D.H., and Hoover, D.G. (2004) Pressure inactivation kinetics of phage lambda cI 857. *J Food Prot* **67**: 505-511.

Chiura, H.X. (1997) Generalized gene transfer by virus-like particules from marine bacteria. *Aquatic Microbial Ecology* **13**: 75-83.

Chiura, H.X. (2004) Novel broad-host range gene transfer particles in nature. *Microbes and environment* **19**: 249-264.

Chiura, H.X., Velimirov, B., and Kogure, K. (1999) Virus-like particles in microbial population control and horizontal gene transfer in the aquatic environment. In *Microbial Biosystems : New Frontiers*

*8th Symposium on Microbial Ecology*. Bell, C.R., Brylinsky, M., and Johnson-Green, P. (eds). Halifax (Canada).

Clark, A.J., Inwood, W., Cloutier, T., and Dhillon, T.S. (2001) Nucleotide sequence of coliphage HK620 and the evolution of lambdoid phages. *J Mol Biol* **311**: 657-679.

Cohen, G.N., Barbe, V., Flament, D., Galperin, M., Heilig, R., Lecompte, O. *et al.* (2003) An integrated analysis of the genome of the hyperthermophilic archaeon *Pyrococcus abyssi*. *Mol Microbiol* **47**: 1495-1512.

## D

Daniels, L.L., and Wais, A.C. (1984) Restriction and modification of halophage S45 in Halobacterium. *Curr Biol* **10**: 133-136.

Daniels, L.L., and Wais, A.C. (1990) Ecophysiology of Bacteriophage S5100 Infecting Halobacterium cutirubrum. *Appl Environ Microbiol* **56**: 3605-3608.

Danovaro, R., Manini, E., and Dell'Anno, A. (2002) Higher abundance of bacteria than of viruses in deep Mediterranean sediments. *Appl Environ Microbiol* **68**: 1468-1472.

Danovaro, R., Dell'Anno, A., Trucco, A., Serresi, M., and Vanucci, S. (2001) Determination of virus abundance in marine sediments. *Appl Environ Microbiol* **67**: 1384-1387.

Daughney, C.J., Chatellier, X., Chan, A., Kenward, P., Fortin, D., Suttle, C., and Fowle, D.A. (2004) Adsorption and precipitation of iron from seawater on a marine bacteriophage (PWH3A-P1). *Marine Chemistry* **91**: 101-115.

Davanloo, P., Rosenberg, A.H., Dunn, J.J., and Studier, F.W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. *PNAS* **81**: 2035-2039.

Dean, F.B., Nelson, J.R., Giesler, T.L., and Lasken, R.S. (2001) Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Res* **11**: 1095-1099.

Dean, F.B., Hosono, S., Fang, L., Wu, X., Faruqi, A.F., Bray-Ward, P. *et al.* (2002) Comprehensive human genome amplification using multiple displacement amplification. *PNAS* **99**: 5261-5266.

Del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., and Diaz-Orejas, R. (1998) Replication and control of circular bacterial plasmids. *Microbiol Mol Biol Rev* **62**: 434-464.

Deming, J.W., and Baross, J.A. (1993) Deep-sea smokers: windows to a subsurface biosphere? *Geochim Cosmochim Acta* **57**: 3219-3230.

D'Hérelle, F. (1917) Sur un microbe invisible antagoniste des bacilles dysentériques. *C R Acad Sci II* **165**: 373-375.

Drlica, K. (1992) Control of bacterial DNA supercoiling. *Mol Microbiol* **6**: 425-433.

Duguet, M. (1993) The helical repeat of DNA at high temperature. *Nucleic Acids Res* **21**: 463-468.

Dyall-Smith, M., Tang, S.L., and Bath, C. (2003) Haloarchaeal viruses: how diverse are they? *Res Microbiol* **154**: 309-313.

Dynan, W.S., Jendrisak, J.J., Hager, D.A., and Burgess, R.R. (1981) Purification and characterization of wheat germ DNA topoisomerase I (nicking-closing enzyme). *J Biol Chem* **256**: 5860-5865.

## E

Edelman, D.C., and Barletta, J. (2003) Real-time PCR provides improved detection and titer determination of bacteriophage. *Biotechniques* **35**: 368-375.

Edwards, R.A., and Rohwer, F. (2005) Viral metagenomics. *Nat Rev Microbiol* **3**: 504-510.

Erauso, G. (1994) Les Thermococcales du Bassin Nord-Fidjien, description d'une nouvelle espèce : *Pyrococcus abyssi* et caractérisation de son plasmide. In *Sciences de la Mer*. Brest: Université de Bretagne Occidentale.

Erauso, G., Prieur, D., Godfroy, A., and Raguénes, G. (1995) Plate cultivation technique for strictly anaerobic, thermophilic, sulfur-metabolizing Archaea. In *Archaea*. Robb, F.T., and Place, A.R. (eds), pp. 25-29.

Erauso, G., Marsin, S., Benbouzid-Rollet, N., Baucher, M.F., Barbeyron, T., Zivanovic, Y. *et al.* (1996) Sequence of plasmid pGT5 from the archaeon *Pyrococcus abyssi*: evidence for rolling-circle replication in a hyperthermophile. *J Bacteriol* **178**: 3232-3237.

Erauso G., Gonnet M., Toffin A., Chevereau M. & Prieur D., en préparation

Ewert, D.L., and Paynter, M.J. (1980) Technique for determining total bacterial virus counts in complex aqueous systems. *Appl Environ Microbiol* **39**: 253-260.

## F

Fauquet, C.M. (2002) Viral classification and nomenclature. In *Encyclopedia of life science*. www.els.net (ed).

Filée, J., Tetart, F., Suttle, C.A., and Krisch, H.M. (2005) Marine T4-type bacteriophages, a ubiquitous component of the dark matter of the biosphere. *PNAS* **102**: 12471-12476.

Forterre, P. (2005) The two ages of the RNA world, and the transition to the DNA world: a story of viruses and cells. *Biochimie* **87**: 793-803.

Forterre, P., Bergerat, A., and Lopez-Garcia, P. (1996) The unique DNA topology and DNA topoisomerases of hyperthermophilic archaea. *FEMS Microbiol Rev* **18**: 237-248.

Foster, R.A.C., Johnson, F.H., and Miller, V.K. (1949) The influence of hydrostatic pressure and urethane on the thermal inactivation of bacteriophage. *The Journal of General Physiology* **33**: 1-16.

Fuhrman, J.A. (1999) Marine viruses and their biogeochemical and ecological effects. *Nature* **399**: 541-548.

## G

Gadelle, D., Filee, J., Buhler, C., and Forterre, P. (2003) Phylogenomics of type II DNA topoisomerases. *Bioessays* **25**: 232-242.

Gadkar, V., and Rillig, M.C. (2005) Application of Phi29 DNA polymerase mediated whole genome amplification on single spores of arbuscular mycorrhizal (AM) fungi. *FEMS Microbiol Lett* **242**: 65-71.

Gelfand, D.H., and Hayashi, M. (1970) In vitro synthesis of a DNA dependent RNA polymerase coded on coliphage T7 genome. *Nature* **228**: 1162-1165.

Gellert, M., Mizuuchi, K., O'Dea, M.H., and Nash, H.A. (1976) DNA gyrase: an enzyme that introduces superhelical turns into DNA. *PNAS* **73**: 3872-3876.

Geslin, C., Le Romancer, M., Erauso, G., Gaillard, M., Perrot, G., and Prieur, D. (2003) PAV1, the first virus-like particle isolated from a hyperthermophilic euryarchaeote, "Pyrococcus abyssi". *J Bacteriol* **185**: 3888-3894.

Gibson, U.E.M., Heid, C.A., and Williams, P.M. (1996) A novel method for real time quantitative RT-PCR. *Genome research* **6**: 995-1001.

Gonzalez, J.M., and Suttle, C. (1993) Grazing by marine nanoflagellates on viruses and virus-sized particles : ingestion and digestion. *Marine Ecology Progress Series* **94**: 1-10.

Griffin, D.W., Gibson, C.J., 3rd, Lipp, E.K., Riley, K., Paul, J.H., 3rd, and Rose, J.B. (1999) Detection of viral pathogens by reverse transcriptase PCR and of microbial indicators by standard methods in the canals of the Florida Keys. *Appl Environ Microbiol* **65**: 4118-4125.

Guipaud, O., Marguet, E., Noll, K.M., de la Tour, C.B., and Forterre, P. (1997) Both DNA gyrase and reverse gyrase are present in the hyperthermophilic bacterium *Thermotoga maritima*. *PNAS* **94**: 10606-10611.

Guixa-Boixareu, N., Calderon-Paz, J.I., Heldal, M., Bratbak, G., and Pedros-Alio, C. (1996) Viral lysis and bacterivory as prokaryotic loss factors along a salinity gradient. *Marine Ecology Progress Series* **11**: 215-227.

## H

Hacein-Bey-Abina, S., Von Kalle, C., Schmidt, M., McCormack, M.P., Wulffraat, N., Le-boulch, P. *et al.* (2003) LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**: 415-419.

Hambly, E., and Suttle, C.A. (2005) The virosphere, diversity, and genetic exchange within phage communities. *Curr Opin Microbiol* **8**: 444-450.

Hara, S., Terauchi, K., and Koike, H. (1991) Abundance of viruses in marine waters: assessment by epifluorescence and transmission electron microscopy. *Appl Environ Microbiol* **57**: 2731-2734.

Häring, M., Rachel, R., Peng, X., Garrett, R.A., and Prangishvili, D. (2005b) Viral diversity in hot springs of Pozzuoli, Italy, and characterization of a unique archaeal virus, Acidianus bottle-shaped virus, from a new family, the Ampullaviridae. *J Virol* **79**: 9904-9911.

Häring, M., Vestergaard, G., Brugger, K., Rachel, R., Garrett, R.A., and Prangishvili, D. (2005a) Structure and genome organization of AFV2, a novel archaeal lipothrixvirus with unusual terminal and core structures. *J Bacteriol* **187**: 3855-3858.

Häring, M., Vestergaard, G., Rachel, R., Chen, L., Garrett, R.A., and Prangishvili, D. (2005c) Virology: independent virus development outside a host. *Nature* **436**: 1101-1102.

Häring, M., Peng, X., Brugger, K., Rachel, R., Stetter, K.O., Garrett, R.A., and Prangishvili, D. (2004) Morphology and genome organization of the virus PSV of the hyperthermophilic archaeal genera *Pyrobaculum* and *Thermoproteus*: a novel virus family, the Globuloviridae. *Virology* **323**: 233-242.

Harmsen, H., Prieur, D., and Jeanthon, C. (1997) Distribution of microorganisms in deep-sea hydrothermal vent chimneys investigated by whole-cell hybridization and enrichment culture of thermophilic subpopulations. *Appl Environ Microbiol* **63**: 2876-2883.

Harvey, C.L., Gabriel, T.F., Wilt, E.M., and Richardson, C.C. (1971) Enzymatic breakage and joining of deoxyribonucleic acid. IX. Synthesis and properties of the deoxyribonucleic acid adenylate in the phage T4 ligase reaction. *J Biol Chem* **246**: 4523-4530.

Heid, C.A., Stevens, J., Livak, K.J., and Williams, P.M. (1996) Real time quantitative PCR. *Genome Research* **6**: 986-994.

- Held, P.K., Olivares, E.C., Aguilar, C.P., Finegold, M., Calos, M.P., and Grompe, M. (2005) In vivo correction of murine hereditary tyrosinemia type I by phiC31 integrase-mediated gene delivery. *Mol Ther* **11**: 399-408.
- Heldal, M., and Bratbak, G. (1991) Production and decay of viruses in aquatic environments. *Marine Ecology Progress Series* **72**: 205-212.
- Hendrix, R.W. (1999) Evolution: the long evolutionary reach of viruses. *Curr Biol* **9**: R914-917.
- Hendrix, R.W. (2002) Bacteriophages: evolution of the majority. *Theor Popul Biol* **61**: 471-480.
- Hendrix, R.W. (2003) Bacteriophage genomics. *Curr Opin Microbiol* **6**: 506-511.
- Hennes, K.P., and Suttle, C.A. (1995) Direct counts of viruses in natural waters and laboratory cultures by epifluorescence microscopy. *Limnol. Oceanogr.* **40**: 1050-1055.
- Hercher, M., Mueller, W., and Shapiro, H.M. (1979) Detection and discrimination of individual viruses by flow cytometry. *J Histochem Cytochem* **27**: 350-352.
- Huls, G.A., Heijnen, I.A.F.M., Cuomo, M.E., Koningsberger, J.C., Wiegman, L., Boel, E. *et al.* (1999) A recombinant, fully human monoclonal antibody with antitumor activity constructed from phage-displayed antibody fragments. *Nature Biotechnology* **17**: 276-281.

I

- Inciarte, M.R., Salas, M., and Sogo, J.M. (1980) Structure of replicating DNA molecules of Bacillus subtilis bacteriophage phi29. *J Virol* **34**: 187-199.
- Inman, R.B., Schnos, M., and Howe, M. (1976) Location of the "variable end" of Mu DNA within the bacteriophage particle. *Virology* **72**: 393-401.

J

- Janekovic, D., Wunderl, S., Holz, I., Zillig, W., Gierl, A., and Neumann, H. (1983) TTV1, TTV2 and TTV3, a family of viruses of the extremely thermophilic, anaerobic, sulfur reducing archaeobacterium Thermoproteus tenax. *Mol Gen Genet* **192**: 39-45.
- Jiang, S., Steward, G., Jellison, R., Chu, W., and Choi, S. (2004) Abundance, distribution, and diversity of viruses in alkaline, hypersaline Mono Lake, California. *Microb Ecol* **47**: 9-17.
- Jiang, S.C., and Paul, J.H. (1998) Gene transfer by transduction in the marine environment. *Appl Environ Microbiol* **64**: 2780-2787.

Jiang, W., Li, Z., Zhang, Z., Baker, M.L., Prevelige, P.E., Jr., and Chiu, W. (2003) Coat protein fold and maturation transition of bacteriophage P22 seen at subnanometer resolutions. *Nat Struct Biol* **10**: 131-135.

Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F., and Hendrix, R.W. (2000) Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. *J Mol Biol* **299**: 27-51.

## K

Kamimura, T., Yoshikawa, A., Ichida, F., and Sasaki, H. (1981) Electron microscopic studies of Dane particles in hepatocytes with special reference to intracellular development of Dane particles and their relation with HBeAg in serum. *Hepatology* **1**: 392-397.

Keeling, P.J., Klenk, H.P., Singh, R.K., Schenk, M.E., Sensen, C.W., Zillig, W., and Doolittle, W.F. (1998) *Sulfolobus islandicus* plasmids pRN1 and pRN2 share distant but common evolutionary ancestry. *Extremophiles* **2**: 391-393.

Kessler, A., Brinkman, A.B., van der Oost, J., and Prangishvili, D. (2004) Transcription of the rod-shaped viruses SIRV1 and SIRV2 of the hyperthermophilic archaeon *Sulfolobus*. *J Bacteriol* **186**: 7745-7753.

Khayat, R., Tang, L., Larson, E.T., Lawrence, C.M., Young, M., and Johnson, J.E. (2005) From the Cover: Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *PNAS* **102**: 18944-18949.

Kikuchi, A., and Asai, K. (1984) Reverse gyrase--a topoisomerase which introduces positive superhelical turns into DNA. *Nature* **309**: 677-681.

Kim, U.J., Shizuya, H., de Jong, P.J., Birren, B., and Simon, M.I. (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. *Nucleic Acids Res* **20**: 1083-1085.

Klein, R., Baranyi, U., Rössler, N., Greineder, B., Scholz, H., and Witte, A. (2002) Natrialba magadii virus phiCh1: first complete nucleotide sequence and functional organization of a virus infecting a haloalkaliphilic archaeon. *Mol Microbiol* **45**: 851-863.

## L

Lawrence, J.G., Hatfull, G.F., and Hendrix, R.W. (2002) Imbroglions of viral taxonomy: genetic exchange and failings of phenetic approaches. *J Bacteriol* **184**: 4891-4905.

- Lee, M.S., LeMaistre, A., Kantarjian, H.M., Talpaz, M., Freireich, E.J., Trujillo, J.M., and Stass, S.A. (1989) Detection of two alternative bcr/abl mRNA junctions and minimal residual disease in Philadelphia chromosome positive chronic myelogenous leukemia by polymerase chain reaction. *Blood* **73**: 2165-2170.
- Lipps, G., Rother, S., Hart, C., and Krauss, G. (2003) A novel type of replicative enzyme harbouring ATPase, primase and DNA polymerase activity. *Embo J* **22**: 2516-2525.
- Lipps, G., Weinzierl, A.O., von Scheven, G., Buchen, C., and Cramer, P. (2004) Structure of a bifunctional DNA primase-polymerase. *Nat Struct Mol Biol* **11**: 157-162.
- Liu, X., and Gorovsky, M.A. (1993) Mapping the 5' and 3' ends of Tetrahymena thermophila mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic Acids Res* **21**: 4954-4960.
- Loeffler, J.M., Nelson, D., and Fischetti, V.A. (2001) Rapid killing of Streptococcus pneumoniae with a bacteriophage cell wall hydrolase. *Science* **294**: 2170-2172.
- Loessner, M.J. (2005) Bacteriophage endolysins - current state of research and applications. *Curr Opin Microbiol* **8**: 480-487.
- Lopez-Garcia, P., Forterre, P., van der Oost, J., and Erauso, G. (2000) Plasmid pGS5 from the hyperthermophilic archaeon Archaeoglobus profundus is negatively supercoiled. *J Bacteriol* **182**: 4998-5000.
- Lwoff, A., Horne, R., and Tournier, P. (1962) A system of viruses. *Cold Spring Harbor Symposium on Quantitative Biology* **27**: 51-55.

## M

- Magrini, V., Salmi, D., Thomas, D., Herbert, S.K., Hartzell, P.L., and Youderian, P. (1997) Temperate Myxococcus xanthus phage Mx8 encodes a DNA adenine methylase, Mox. *J Bacteriol* **179**: 4254-4263.
- Maniloff, J. (2001) Bacteriophages. In *Encyclopedia of life science*. www.els.net (ed): Nature publishing group.
- Mann, N.H., Cook, A., Millard, A., Bailey, S., and Clokie, M. (2003) Bacterial photosynthesis genes in a virus. *Nature* **424**: 741.
- Marie, D., Vaultot, D., and Partensky, F. (1996) Application of the novel nucleic acid dyes YOYO-1, YO-PRO-1, and PicoGreen for flow cytometric analysis of marine prokaryotes. *Appl Environ Microbiol* **62**: 1649-1655.

- Marie, D., Partensky, F., Jacquet, S., and Vaulot, D. (1997) Enumeration and cell cycle analysis of natural populations of marine picoplankton by flow cytometry using the nucleic acid stain SYBR Green I. *Appl Environ Microbiol* **63**: 186-193.
- Marie, D., Brussaard, C.P.D., Thyrrhaug, R., Bratbak, G., and Vaulot, D. (1999) Enumeration of marine viruses in culture and natural samples by flow cytometry. *Appl Environ Microbiol* **65**: 45-52.
- Marteinsson, V.T., Birrien, J.L., Reysenbach, A.L., Vernet, M., Marie, D., Gambacorta, A. *et al.* (1999) *Thermococcus barophilus* sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent. *Int J Syst Bacteriol* **49**: 351-359.
- Martin, A., Yeats, S., Janekovic, D., Reiter, W.-D., Aicher, W., and Zillig, W. (1984) SAV1, a temperature u. v.-inducible DNA virus-like particle from the archaeobacterium *Sulfolobus acidocaldarius*. *EMBO J* **3**: 2165-2168.
- Mathews, C.K., and Allen, J.R. (1983) DNA precursor synthesis. In *Bacteriophage T4*. C. K. Mathews et al., E. (ed). Washington, DC: Am. Soc. Microbiol.
- Matthews, C.K. (2001) Bacteriophage T4. In *Encyclopedia of life science*. www.els.net (ed): Nature publishing group.
- Medberry, S.L., Lockhart, B.E., and Olszewski, N.E. (1990) Properties of Commelina yellow mottle virus's complete DNA sequence, genomic discontinuities and transcript suggest that it is a pararetrovirus. *Nucleic Acids Res* **18**: 5505-5513.
- Middelboe, M., Jorgensen, N.O.G., and Kroer, N. (1996) Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. *Appl Environ Microbiol* **62**: 1991-1997.
- Mirambeau, G., Duguet, M., and Forterre, P. (1984) ATP-dependent DNA topoisomerase from the archaeobacterium *Sulfolobus acidocaldarius*. Relaxation of supercoiled DNA at high temperature. *J Mol Biol* **179**: 559-563.
- Musgrave, D., Forterre, P., and Slesarev, A. (2000) Negative constrained DNA supercoiling in archaeal nucleosomes. *Mol Microbiol* **35**: 341-349.
- Muskhelishvili, G. (1994) The archaeal SSV integrase protease intermolecular excisive recombination in vitro. *Systematic and Applied Microbiology* **16**: 605-608.
- Muskhelishvili, G., Palm, P., and Zillig, W. (1993) SSV1-encoded site-specific recombination system in *Sulfolobus shibatae*. *Mol Gen Genet* **237**: 334-342.
- Myllykallio, H., Lipowski, G., Leduc, D., Filee, J., Forterre, P., and Liebl, U. (2002) An alternative flavin-dependent mechanism for thymidylate synthesis. *Science* **297**: 105-107.

**N**

- Nadal, M., Mirambeau, G., Forterre, P., Reiter, W.-D., and Duguet, M. (1986) Positively supercoiled DNA in a virus-like particle of an archaebacterium. *Nature* **321**: 256-258.
- Nakasu, S., and Kikuchi, A. (1985) Reverse gyrase; ATP-dependent type I topoisomerase from *Sulfolobus*. *Embo J* **4**: 2705-2710.
- Nass, P.H., Domier, L.L., Jakstys, B.P., and D'Arcy, C.J. (1998) In situ localization of barley yellow dwarf virus-PAV 17-kDa protein and nucleic acids in oats. *Virology* **88**: 1031-1039.
- Nelson, D. (2004) Phage taxonomy: we agree to disagree. *J Bacteriol* **186**: 7029-7031.
- Newcomb, W.W., Juhas, R.M., Thomsen, D.R., Homa, F.L., Burch, A.D., Weller, S.K., and Brown, J.C. (2001) The UL6 gene product forms the portal for entry of DNA into the Herpes Simplex Virus capsid. *J Virol* **75**: 10923-10932.
- Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998) Chaperone co-expression plasmids: differential and synergistic roles of DnaK-DnaJ-GrpE and GroEL-GroES in assisting folding of an allergen of Japanese cedar pollen, Cryj2, in *Escherichia coli*. *Appl Environ Microbiol* **64**: 1694-1699.
- Noble, R.T., and Fuhrman, J.A. (1998) Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquatic Microbial Ecology* **14**: 113-118.
- Nuno-Gonzalez, P., Chao, H., and Oka, K. (2005) Targeting site-specific chromosome integration. *Acta Biochim Pol* **52**: 285-291.
- Nuttall, S.D., and Dyal Smith, M.L. (1993) HF1 and HF2: Novel bacteriophages of halophilic archaea. *Virology* **197**: 678-684.

**O**

- Oren, A., Bratbak, G., and Heldal, M. (1997) Occurrence of virus-like particles in the Dead Sea. *Extremophiles* **1**: 143-149.
- Ortmann, A.C., and Suttle, C.A. (2005) High abundance of viruses in a deep-sea hydrothermal vent system indicates viral mediated microbial mortality. *Deep Sea Research I* **52**: 1515-1527.

## P

- Palm, P., Schleper, C., Grampp, B., Yeats, S., McWilliam, P., Reiter, W.D., and Zillig, W. (1991) Complete nucleotide sequence of the virus SSV1 of the archaeobacterium *Sulfolobus shibatae*. *Virology* **185**: 242-250.
- Paul, J.H., Jiang, S.C., and Rose, J.B. (1991) Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. *Appl Environ Microbiol* **57**: 2197-2204.
- Paul, J.H., Rose, J.B., Jiang, S.C., Kellogg, C.A., and Dickson, L. (1993) Distribution of viral abundance in the reef environment of Key Largo, Florida. *Appl Environ Microbiol* **59**: 718-724.
- Pauling, C. (1982) Bacteriophages of *Halobacterium halobium*: isolation from fermented fish sauce and primary characterization. *Can J Microbiol* **28**: 916-921.
- Pedulla, M.L., Ford, M.E., Houtz, J.M., Karthikeyan, T., Wadsworth, C., Lewis, J.A. *et al.* (2003) Origins of highly mosaic mycobacteriophage genomes. *Cell* **113**: 171-182.
- Peng, X., Kessler, A., Phan, H., Garrett, R.A., and Prangishvili, D. (2004) Multiple variants of the archaeal DNA rudivirus SIRV1 in a single host and a novel mechanism of genomic variation. *Mol Microbiol* **54**: 366-375.
- Peng, X., Blum, H., She, Q., Mallok, S., Brugger, K., Garrett, R.A. *et al.* (2001) Sequences and replication of genomes of the archaeal rudiviruses SIRV1 and SIRV2: relationships to the archaeal lipothrixvirus SIFV and some eukaryal viruses. *Virology* **291**: 226-234.
- Peretz, D., Williamson, R.A., Kameko, K., Vergara, J., Leclerc, E., Schmitt-Ulms, G. *et al.* (2001) Antibodies inhibit prion propagation and clear cell cultures of prion infectivity. *Nature* **412**: 739-743.
- Pfankuch, E., and Kausche, G.A. (1940) Isolierung un übermikroskopishce Abbildung eines Bacteriophagen. *Die Naturwissenschaften* **28**: 46.
- Pfister, P., Wasserfallen, A., Stettler, R., and Leisinger, T. (1998) Molecular analysis of *Methanobacterium* phage psiM2. *Mol Microbiol* **30**: 233-244.
- Porter, K., Kukkaro, P., Bamford, J.K., Bath, C., Kivela, H.M., Dyall-Smith, M.L., and Bamford, D.H. (2005) SH1: A novel, spherical halovirus isolated from an Australian hypersaline lake. *Virology* **335**: 22-33.
- Porter, K.G., and Feig, Y.S. (1980) The use of DAPI for identifying and counting aquatic microflora. *Limnol Oceanogr* **25**: 943-948.
- Prangishvili, D. (2003) Evolutionary insights from studies on viruses of hyperthermophilic archaea. *Res Microbiol* **154**: 289-294.

- Prangishvili, D., and Garrett, R.A. (2005) Viruses of hyperthermophilic Crenarchaea. *Trends Microbiol* **13**: 535-542.
- Prangishvili, D., Arnold, H.P., Gotz, D., Ziese, U., Holz, I., Kristjansson, J.K., and Zillig, W. (1999) A novel virus family, the Rudiviridae: structure, virus-host interactions and genome variability of the sulfolobus viruses SIRV1 and SIRV2. *Genetics* **152**: 1387-1396.
- Prigent, M., Leroy, M., Confalonieri, F., Dutertre, M., and DuBow, M.S. (2005) A diversity of bacteriophage forms and genomes can be isolated from the surface sands of the Sahara Desert. *Extremophiles* **9**: 289-296.
- Proctor, L.M., and Fuhrman, J.A. (1990) Viral mortality of marine bacteria and cyanobacteria. *Nature* **343**: 60-62.
- Proctor, L.M., and Fuhrman, J.A. (1991) Roles of viral infection in organic particle flux. *Marine Ecology Progress Series* **69**: 133-142.
- Pruss, G.J., Manes, S.H., and Drlica, K. (1982) Escherichia coli DNA topoisomerase I mutants: increased supercoiling is corrected by mutations near gyrase genes. *Cell* **31**: 35-42.
- Puig, M., Pina, S., Lucena, F., Jofre, J., and Girones, R. (2000) Description of a DNA amplification procedure for the detection of bacteriophages of Bacteroides fragilis HSP40 in environmental samples. *J Virol Methods* **89**: 159-166.

## R

- Rachel, R., Bettstetter, M., Hedlund, B.P., Haring, M., Kessler, A., Stetter, K.O., and Prangishvili, D. (2002) Remarkable morphological diversity of viruses and virus-like particles in hot terrestrial environments. *Arch Virol* **147**: 2419-2429.
- Ravot, G., Magot, M., Fardeau, M.L., Patel, B.K., Prensier, G., Egan, A. *et al.* (1995) Thermotoga elfii sp. nov., a novel thermophilic bacterium from an African oil-producing well. *Int J Syst Bacteriol* **45**: 308-314.
- Reiter, W.D., Zillig, W., and Palm, P. (1988) Archaeobacterial viruses. *Adv Virus Res* **34**: 143-188.
- Reiter, W.D., Palm, P., Voos, W., Kaniecki, J., Grampp, B., Schulz, W., and Zillig, W. (1987) Putative promoter elements for the ribosomal RNA genes of the thermoacidophilic archaeobacterium Sulfolobus sp. strain B12. *Nucleic Acids Res* **15**: 5581-5595.
- Rice, G., Stedman, K., Snyder, J., Wiedenheft, B., Willits, D., Brumfield, S. *et al.* (2001) Viruses from extreme thermal environments. *PNAS* **98**: 13341-13345.

- Rice, G., Tang, L., Stedman, K., Roberto, F., Spuhler, J., Gillitzer, E. *et al.* (2004) The structure of a thermophilic archaeal virus shows a double-stranded DNA viral capsid type that spans all domains of life. *PNAS* **101**: 7716-7720.
- Rohrmann, G.F., Cheney, R., and Pauling, C. (1983) Bacteriophages of *Halobacterium halobium*: virion DNAs and proteins. *Can J Microbiol* **29**: 627-629.
- Rohwer, F. (2003) Global phage diversity. *Cell* **113**: 141.
- Rohwer, F., and Edwards, R. (2002) The Phage Proteomic Tree: a genome-based taxonomy for phage. *J Bacteriol* **184**: 4529-4535.
- Rohwer, F., Seguritan, V., Choi, D.H., Segall, A.M., and Azam, F. (2001) Production of shotgun libraries using random amplification. *Biotechniques* **31**: 108-112, 114-106, 118.
- Ruska, H. (1940) Die Sichtbarmachung der bacteriophagen Lyse im Übermikroskop. *Die Naturwissenschaften* **28**: 45-46.
- Rutberg, L. (1965) On the Induction of *Escherichia Coli* K (Lambda) with High Hydrostatic Pressure. *Virology* **26**: 190-198.

## S

- Saiki, R.K., Scharf, S., Faloona, F., Mullis, K.B., Horn, G.T., Erlich, H.A., and Arnheim, N. (1985) Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Biotechnology* **24**: 476-480.
- Salas, M. (1991) Protein-priming of DNA replication. *Annu Rev Biochem* **60**: 39-71.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Analysis of genomic DNA by southern hybridization. In *Molecular cloning A laboratory manual*. Nolan, C. (ed), pp. 9.31-39.62.
- Sanger, F., Coulson, A.R., Friedmann, T., Air, G.M., Barrell, B.G., Brown, N.L., Fiddes, J.C., Hutchison, C.A. III, Slocombe, P.M., and Smith, M. (1978) The nucleotide sequence of bacteriophage phiX174. *J. Mol. Biol.* **125**: 225-246.
- Sato, T., Fukui, T., Atomi, H., and Imanaka, T. (2003) Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. *J Bacteriol* **185**: 210-220.
- Schein, C.H. (1989) Production of soluble recombinant proteins in bacteria. *Biotechnology* **7**: 1141-1149.
- Schnabel, H. (1984a) An immune strain of *Halobacterium halobium* carries the invertible L segment of phage phiH as a plasmid. *PNAS* **81**: 1017-1020.

- Schnabel, H. (1984b) Integration of plasmid pphiHL into phage genome during infection of *Halobacterium halobium* R1-L with phage phiHL1. *Mol Gen Genet* **197**: 19-23.
- Schnabel, H., and Zillig, W. (1984) Circular structure of the genome of phage phiH in a lyso-genic *Halobacterium halobium*. *Mol Gen Genet* **193**: 422-426.
- Schnabel, H., Schramm, E., Schnabel, R., and Zillig, W. (1982b) Structural variability in the genome of phage phiH of *Halobacterium halobium*. *Mol Gen Genet* **188**: 370-377.
- Schnabel, H., Palm, P., Dick, K., and Grampp, B. (1984) Sequence analysis of the insertion element ISH1.8 and of associated structural changes in the genome of phage phiH of the archaeobacterium *Halobacterium halobium*. *Embo J* **3**: 1717-1722.
- Schnabel, H., Zillig, W., Pfäffle, M., Schnabel, R., Michel, H., and Delius, H. (1982a) *Halobacterium halobium* phage phiH. *Embo J* **1**: 87-92.
- Shibata, A., Kogure, K., Koike, H., and Ohwada, K. (1997) Formation of submicron colloidal particles from marine bacteria by viral infection. *Marine Ecology Progress Series* **155**: 303-307.
- Shin, G.A., and Sobsey, M.D. (2003) Reduction of Norwalk virus, poliovirus 1, and bacteriophage MS2 by ozone disinfection of water. *Appl Environ Microbiol* **69**: 3975-3978.
- Signer, E.R. (1967) A plaque assay for lambda attachment. *Virology* **33**: 352-354.
- Solomon, L., Zeegen, P., and Eiserling, F.A. (1966) The effects of high hydrostatic pressure on coliphage T-4. *Biochim Biophys Acta* **112**: 102-109.
- Stedman, K.M., Schleper, C., Rumpf, E., and Zillig, W. (1999) Genetic requirements for the function of the archaeal virus SSV1 in *Sulfolobus solfataricus*: construction and testing of viral shuttle vectors. *Genetics* **152**: 1397-1405.
- Stedman, K.M., She, Q., Phan, H., Arnold, H.P., Holz, I., Garrett, R.A., and Zillig, W. (2003) Relationships between fuselloviruses infecting the extremely thermophilic archaeon *Sulfolobus*: SSV1 and SSV2. *Res Microbiol* **154**: 295-302.
- Stolt, P., and Zillig, W. (1992) In vivo studies on the effects of immunity genes on early lytic transcription in the *Halobacterium salinarium* phage phi H. *Mol Gen Genet* **235**: 197-204.
- Stolt, P., and Zillig, W. (1993) In vivo and in vitro analysis of transcription of the L region from the *Halobacterium salinarium* phage phi H: definition of a repressor-enhancing gene. *Virology* **195**: 649-658.
- Stolt, P., and Zillig, W. (1994) Gene regulation in halophage phiH ; more than promoters. *Systematic and Applied Microbiology* **16**: 591-596.
- Suttle, C. (2005a) The virosphere: the greatest biological diversity on Earth and driver of global processes. *Environ Microbiol* **7**: 481-482.

Suttle, C.A. (1994) The significance of viruses to mortality in aquatic microbial communities. *Microb Ecol* **28**: 237-243.

Suttle, C.A. (2005b) Viruses in the sea. *Nature* **437**: 356-361.

Suttle, C.A., Chan, A.M., and Cottrell, M.T. (1990) Infection of phytoplankton by viruses and reduction of primary productivity. *Nature* **347**: 467-471.

## T

Takai, K., Komatsu, T., Inagaki, F., and Horikoshi, K. (2001) Distribution of archaea in a black smoker chimney structure. *Appl Environ Microbiol* **67**: 3618-3629.

Telling, G.C., Parchi, P., DeArmond, S.J., Cortelli, P., Montagna, P., Gabizon, R. *et al.* (1996) Evidence for the conformation of the pathologic isoform of the prion protein enciphering and propagating prion diversity. *Science* **274**: 2079-2082.

Tomschik, M., Karymov, M.A., Zlatanova, J., and Leuba, S.H. (2001) The archaeal histone-fold protein HMf organizes DNA into bona fide chromatin fibers. *Structure* **9**: 1201-1211.

Torsvik, T., and Dundas, I.D. (1974) Bacteriophage of *Halobacterium salinarium*. *Nature* **248**: 680-681.

Torsvik, T., and Dundas, I.D. (1980) Persisting phage infection in *Halobacterium salinarium* str. 1. *J Gen Virol* **47**: 29-36.

Tresaugues, L., Collinet, B., Minard, P., Henckes, G., Aufrere, R., Blondeau, K. *et al.* (2004) Refolding strategies from inclusion bodies in a structural genomics project. *J Struct Funct Genomics* **5**: 195-204.

Twort, F.W., Lond, L.R.C.P., and Lond, M.R.C.S. (1915) An investigation on the nature of ultra-microscopic viruses. *The Lancet* **189**: 1241-1243.

## V

Van Regenmortel, M.H.V. (1999) How to write the names of virus species. *Arch Virol* **144**: 1041-1042.

Vestergaard, G., Haring, M., Peng, X., Rachel, R., Garrett, R.A., and Prangishvili, D. (2005) A novel rudivirus, ARV1, of the hyperthermophilic archaeal genus *Acidianus*. *Virology* **336**: 83-92.

Vincent, W.F., Gibson, J.A., Pienitz, R., Villeneuve, V., Broady, P.A., Hamilton, P.B., and Howard-Williams, C. (2000) Ice shelf microbial ecosystems in the high arctic and implications for life on snowball earth. *Naturwissenschaften* **87**: 137-141.

Vincentelli, R., Canaan, S., Campanacci, V., Valencia, C., Maurin, D., Frassinetti, F. *et al.* (2004) High-throughput automated refolding screening of inclusion bodies. *Protein Sci* **13**: 2782-2792.

Vogelsang-Wenke, H., and Oesterhelt, D. (1988) Isolation of a haobacterial phage with a fully cytosine-methylated genome. *Mol Gen Genet* **211**: 407-414.

## W

Wais, A.C., Kon, M., MacDonald, R.E., and Stollar, B.D. (1975) Salt-dependent bacteriophage infecting *Halobacterium cutirubrum* and *H. halobium*. *Nature* **256**: 314-315.

Wang, Y., and Walker, P.J. (1993) Adelaide river rhabdovirus expresses consecutive glycoprotein genes as a polycistronic mRNA: new evidence of gene duplication as an evolutionary process. *Virology* **195**: 719-731.

Watabe, K., Leusch, M., and Ito, J. (1984) Replication of bacteriophage phi 29 DNA in vitro: the roles of terminal protein and DNA polymerase. *PNAS* **81**: 5374-5378.

Weinbauer, M.G., and Suttle, C.A. (1997) Comparison of epifluorescence and transmission electron microscopy for counting viruses in natural marine waters. *Aquatic Microbial Ecology* **13**: 225-232.

Wen, K., Ortmann, A.C., and Suttle, C.A. (2004) Accurate estimation of viral abundance by epifluorescence microscopy. *Appl Environ Microbiol* **70**: 3862-3867.

Whitman, W.B., Coleman, D.C., and Wiebe, W.J. (1998) Prokaryotes: the unseen majority. *PNAS* **95**: 6578-6583.

Wiedenheft, B., Stedman, K., Roberto, F., Willits, D., Gleske, A.K., Zoeller, L. *et al.* (2004) Comparative genomic analysis of hyperthermophilic archaeal Fuselloviridae viruses. *J Virol* **78**: 1954-1961.

Wilhelm, S.W., and Suttle, C.A. (1999) Viruses and nutrient cycles in the sea. *BioScience* **49**: 781-788.

Wilhelm, S.W., Weinbauer, M.G., Suttle, C., and Jeffrey, W.H. (1998) The role of sunlight in the removal and repair of viruses in the sea. *Limnol Oceanogr* **43**: 586-592.

Williamson, R.A., Peretz, D., Pinilla, C., Ball, H., Bastidas, R.B., Rozenshteyn, R. *et al.* (1998) Mapping the prion protein using recombinant antibodies. *J Virol* **72**: 9413-9418.

Witte, A., Baranyi, U., Klein, R., Sulzner, M., Luo, C., Wanner, G. *et al.* (1997) Characterization of *Natronobacterium magadii* phage Phi Ch1, a unique archaeal phage containing DNA and RNA. *Mol Microbiol* **23** (3): 603-616.

Wommack, K.E., and Colwell, R.R. (2000) Virioplankton: viruses in aquatic ecosystems. *Microbiol Mol Biol Rev* **64**: 69-114.

Wood, A.G., Whitman, W.B., and Konisky, J. (1989) Isolation and characterization of an archaeobacterial viruslike particle from *Methanococcus voltae* A3. *J Bacteriol* **171**: 93-98.

## X

Xenopoulos, M.A., and Bird, D.F. (1997) Virus à la sauce Yo-Pro : microwave-enhanced staining for counting viruses by epifluorescence microscopy. *Limnol Oceanogr* **42**: 1648-1650.

Xiang, X., Chen, L., Huang, X., Luo, Y., She, Q., and Huang, L. (2005) *Sulfolobus tengchongensis* spindle-shaped virus STSV1: virus-host interactions and genomic features. *J Virol* **79**: 8677-8686.

## Y

Yokobata, K., Trenchak, B., and de Jong, P.J. (1991) Rescue of unstable cosmids by in vitro packaging. *Nucleic Acids Res* **19**: 403-404.

Yoong, P., Schuch, R., Nelson, D., and Fischetti, V.A. (2004) Identification of a broadly active phage lytic enzyme with lethal activity against antibiotic-resistant *Enterococcus faecalis* and *Enterococcus faecium*. *J Bacteriol* **186**: 4808-4812.

## Z

Zillig, W., Holz, I., Klenk, H.P., Trent, J., Wunderl, S., Janecovik, D. *et al.* (1987) *Pyrococcus woesei*, *sp. nov.*, an ultra-thermophilic marine archaeobacterium, representing a novel order, *Thermococcales*. *Systematic and Applied Microbiology* **9**: 62-69.

Zillig, W., Kletzin, A., Schleper, C., Holz, I., Janecovik, D., Hain, J. *et al.* (1994) Screening for *Sulfolobales*, their plasmids and viruses in Icelandic solfataras. *Systematic and Applied Microbiology* **16**: 609-628.

Zillig, W., Prangishvilli, D., Schleper, C., Elferink, M., Holz, I., Albers, S. *et al.* (1996) Viruses, plasmids and other genetic elements of thermophilic and hyperthermophilic Archaea. *FEMS Microbiol Rev* **18**: 225-236.

Zillig, W., Arnold, H.P., Holz, I., Prangishvili, D., Schweier, A., Stedman, K. *et al.* (1998) Genetic elements in the extremely thermophilic archaeon Sulfolobus. *Extremophiles* **2**: 131-140.