

2010

THÈSE

Présentée à la

FACULTÉ DES SCIENCES DE LA VIE

en vue de l'obtention du titre de

DOCTEUR DE L'UNIVERSITÉ DE STRASBOURG

Domaine : Sciences du Vivant

Spécialité : Aspects moléculaires et cellulaires de la biologie

par

Pascale Schellenberger

**Caractérisation structurale et fonctionnelle
de déterminants de la capsidie du *Grapevine fanleaf virus*
impliqués dans la transmission par *Xiphinema index***

Soutenue publiquement le 7 juin 2010 devant la Commission d'Examen :

Pr. Catherine FLORENTZ

Dr. Guy SCHOEHN

Dr. Olivier LE GALL

Dr. Stéphane BLANC

Dr. Christophe RITZENTHALER

Dr. Olivier LEMAIRE

Examineur Interne

Rapporteur Externe

Rapporteur Externe

Examineur Externe

Directeur de Thèse

Co-encadrant de Thèse

Institut de Biologie Moléculaire des Plantes (CNRS, UPR 2357) - Strasbourg
Institut National de Recherche Agronomique (INRA, UMR 1131) - Colmar

à mes Parents

à Caroline et Alexia

à mes Grands Parents

à Swen

J'adresse mes plus vifs remerciements aux Pr. Catherine Florentz, Dr. Guy Schoehn, Dr. Olivier Legall et au Dr. Stéphane Blanc, pour l'honneur qu'ils me font de juger mon travail de thèse. Je remercie également le Département Santé des Plantes et Environnement de l'INRA et le Conseil Interprofessionnel des Vins d'Alsace pour avoir co-financé ce travail.

Je remercie mes directeurs de thèse Christophe Ritzenthaler et Olivier Lemaire pour m'avoir permis de vivre cette expérience si enrichissante, ce projet aux facettes si variées. Merci Christophe, Corinne Keichinger et Gérard Demangeat pour votre disponibilité infaillible et vos conseils, entre autres lors des questionnements qui ont jalonné mes travaux.

Merci à tous les partenaires du projet, cette thèse c'est aussi grâce à vous : Bernard Lorber, Claude Sauter, Marc Bergdoll, Patrick Bron et Gérard Pehau-Arnaudet. Chaque collaboration m'a énormément apportée, grâce à votre enthousiasme et vos qualités scientifiques. Merci de m'avoir accueilli dans vos équipes de recherche. Je pense particulièrement à Claude et Bernard, qui m'ont fait découvrir le monde « atomique ». L'IBMC est avec le temps devenu un troisième laboratoire d'accueil pour moi !

Au début de la thèse, je ne pensais pas que la recherche reposait autant sur le travail d'équipe. Je souhaite remercier l'équipe Virologie et Vection de Colmar, particulièrement Molly, Manue, Valérie, Aurélie et Sophie, pour leur amabilité, leur disponibilité, et leur aide pendant ces plusieurs années passées ensemble. Je n'oublie pas de remercier Jacky qui s'est démené pour produire les tonnes de plantes herbacées que je lui demandais pour les purifications de virus. Petit clin d'œil à la qualité de l'environnement de travail du site de l'INRA : travailler avec une vue sur les Vosges, près du vignoble alsacien, c'est sympa !! J'ai aussi passé de bons moments avec Claire, Elodie, Kamal, Khalid, Fanny et Anette, de l'équipe de virologie de l'IBMP de Strasbourg. Entre mes passages épisodiques pour essayer de produire des protoplastes et ma demi-paillasse partagée avec Corinne, c'est surtout pendant ces derniers mois que j'ai vraiment appris à vous connaître, lors de la rédaction de ma thèse.

Mes derniers mots vont à Gérard, qui m'a aidé du tout début de la thèse jusqu'à ces dernières heures. Merci pour ton soutien.... Tu peux arrêter de m'envoyer des corrections maintenant !

Sommaire

SOMMAIRE	1
ABREVIATIONS	3
CHAPITRE 1 : CONTEXTE SCIENTIFIQUE	5
1. Concepts fondamentaux de la vection des phytovirus	5
1. 1 Définition de la vection.....	5
1. 2 Vecteur et spécificité de vection des phytovirus	7
1. 3 Différents modes de transmission	8
1. 4 Déterminants de la transmission	9
2. Architecture des particules virales icosaédriques	11
2. 1 Les capsides icosaédriques.....	11
2. 3 Variation de la symétrie icosaédrique pseudo $T = 3$ chez les Picornavirales.....	14
3. Les modèles biologiques	17
3. 1 Structures de particules virales résolues	17
3. 2 Les virus infectant les bactéries : les phages.....	18
3. 3 Les virus infectant les animaux.....	18
3. 4 Les virus infectant les végétaux ou phytovirus	23
3. 4. 1 Déterminants de la vection du Cucumber necrosis virus (CNV).....	28
3. 4. 2 Déterminants de la vection du Cucumber mosaic virus (CMV)	30
3. 4. 3 Déterminants de la vection du Cauliflower mosaic virus (CaMV).....	31
3. 5 La vection des Nepovirus.....	33
3. 5. 1 Caractéristiques générales de la vection.....	33
3. 5. 2 Déterminants viraux de la transmission	34
3. 5. 3 Caractéristiques de rétention des virus	36
CHAPITRE 2 : LE GRAPEVINE FANLEAF VIRUS ET SON VECTEUR	41
1. La maladie du court - noué de la vigne	41
1. 1 Virus responsables de la maladie	41
1. 2 Description des symptômes sur vigne.....	42
1.3 Épidémiologie de la maladie du court-noué.....	44
1.4 Biologie de <i>Xiphinema index</i> , le vecteur du GFLV.....	44
1. 5 Quelles stratégies de lutte contre la maladie ?.....	45
2. Cycle viral du Grapevine fanleaf virus	46
OBJECTIFS DU TRAVAIL DE THÈSE	49
CHAPITRE 3 : LA CP DÉTERMINE ÉGALEMENT LA SPÉCIFICITE DE TRANSMISSION DU VIRUS PAR XIPHINEMA DIVERSICAUDATUM	51
1. Caractéristiques de la population de <i>Xiphinema diversicaudatum</i>	51
1. 1 Un peu d'écologie concernant <i>X. diversicaudatum</i>	51
1. 2 Multiplication de la population de <i>X. diversicaudatum</i>	52
2. Choix des conditions expérimentales de transmission de virus par <i>X. diversicaudatum</i>	53
3. Adaptation du test de transmission par <i>X. index</i>	55
4. Article 1 : The coat protein determines the specificity of virus transmission by <i>X. diversicaudatum</i>.	56

CHAPITRE 4 : REGIONS DE LA PROTÉINE DE CAPSIDE DU GFLV IMPLIQUÉES DANS LA TRANSMISSION PAR SON NEMATODE VECTEUR X. INDEX 67

- 1. Présentation 67**
- 2. Article 2 : A stretch of 11 amino acids in the β B - β C loop of the coat protein of *Grapevine fanleaf virus* is essential for transmission by the nematode *Xiphinema index* 68**

CHAPITRE 5 : ÉTUDE STRUCTURALE DE LA CAPSIDE DU GFLV ET D'UN VARIANT FAIBLEMENT TRANSMIS PAR X. INDEX 95

- 1. Introduction - Stratégies de détermination des structures de capsid de virus à haute résolution 95**
 - 1. 1 Les stratégies 95*
 - 1. 2 Les résolutions obtenues 96*
- 2. Démarche expérimentale 97**
- 3. Article 3 : Strategies for the crystallization of Viruses : The case of *Grapevine fanleaf virus* 98**
- 4. Article 4 : Structural insights into the molecular mechanisms governing *Grapevine fanleaf nepovirus* transmission by nematodes 115**
- 5. Compléments d'article 147**
 - 5. 1 Structure quaternaire de la capsid 147*
 - 5. 2 Les régions 1 à 5 délimitées d'après la structure du TRSV 148*
 - 5. 3 Résidus de surface du GFLV divergeant entre le GFLV, l'ArMV et le TRSV 151*
 - 5. 4 La cavité : site potentiel de liaison à X. index 151*

CHAPITRE 6 : RECHERCHE DES DÉTERMINANTS VIRAUX DE LA TRANSMISSION PAR COMPARAISON STRUCTURALE ENTRE LE GFLV ET L'ARMV 155

- 1. Démarche expérimentale 155**
- 2. Qualité de la purification de l'ArMV 156**
- 3. Obtention de la structure de l'ArMV par cryo- MET 157**
- 3. Comparaison entre la capsid de GFLV et de l'ArMV 159**
- 4. Conclusion 161**

CONCLUSION GÉNÉRALE ET PERSPECTIVES 162

MATERIEL ET MÉTHODES 168

- 1. Analyse d'images de cryo-MET 168**
 - 1. 1 Principe général 168*
 - 1. 1. 1 Détermination de l'orientation des particules virales d'après la méthode PFT 169*
 - 1. 1. 2 Correction de la CTF 170*
 - 1. 2 Préparation des échantillons et conditions d'acquisition des images 171*
 - 1. 3 Reconstruction 3D de la structure de l'ArMV-C 172*
- 2. Calcul des surfaces accessibles au solvant 173**

ANNEXES 174

- 1. Annexe 1. Sources internet utilisées pour illustrer la figure 1.1 174**
- 2. Annexe 2. Sources internet et références bibliographiques utilisées dans la table 1. 3 175**
- 3. Annexe 3. Crystal growth of proteins, nucleic acids, and viruses in gels 177**
- 4. Annexe 4. Surface accessible des résidus de la capsid de GFLV-TD. 191**
- 5. Annexe 5. Alignement des séquences protéiques des CP codées par l'ARN2-U (X81814) et l'ARN2-L (X81815) de l'ArMV-S. 198**

BIBLIOGRAPHIE 200

Abréviations

°C	: degrés celsius
2D	: bidimensionnel
3D	: tridimensionnel
Å	: angstrom
µm	: micrometre
aa	: acide aminé
ADN	: acide désoxyribonucléique
ADNc	: ADN complémentaire
ArMV	: <i>Arabis mosaic virus</i>
ARN	: acide ribonucléique
ARNm	: ARN messenger
ATP	: adénosine triphosphate
BET	: bromure d'éthidium
BSA	: bovine serum albumin
CaMV	: <i>Cauliflower mosaic virus</i>
CMV	: <i>Cucumber mosaic virus</i>
CNV	: <i>Cucumber necrosis virus</i>
CP	: protéine de capsid
CPMV	: <i>Cowpea mosaic virus</i>
Cryo-MET	: cryomicroscopie électronique à transmission
EDTA	: ethylene diamine tetraacetic acid
ELISA	: enzyme linked immunosorbent assay
Fig.	: figure
FEG	: Field emission gun (canon à émission de champ)
FMDV	: Foot and mouth disease virus
GDP	: guanosine di-phosphate
GFLV	: <i>Grapevine fanleaf virus</i>
GTP	: guanosine triphosphate
h	: heure
HEPES	: acide N-2, hydroxyl-ethylpipérazine, N'2 ethane sulfonique
IC	: immuno capture
kb	: kilobase
kDa	: kilodalton
l	: litre
M	: molaire
m/v	: masse / volume
ME	: microscope électronique
MET	: ME à transmission
min	: minute
MOPS	: acide morpholinopropanesulfonique
MP	: protéine de mouvement
NDP	: nucléosides di-phosphates
nm	: nanomètre
nt	: nucléotide
ORF	: cadre ouvert de lecture
p	: pico
pb	: paire de base
PAGE	: électrophorèse sur gel de polyacrylamide
PBS	: tampon phosphate salin
PCR	: polymerase chain reaction
PFT	: polar fourier transform (transformée de Fourier polaire)
pH	: potentiel hydrogène
PIPES	: acide pipérazine-1,4-diéthanesulfonique
RE	: réticulum endoplasmique
RT	: reverse transcription

SDS	: dodécyl sulfate de sodium
SDS PAGE	: sodium deodecyl sulfate polyacrylamide gel electrophoresis
TBE	: tris borate EDTA
T-DNA	: ADN de transfert
TEMED	: NNN'N'tetraméthylène diamine
TF	: transformé de Fourier
TMV	: <i>Tobacco mosaic virus</i>
UA	: unité asymétrique
UTP	: uridine 5' triphosphate
UV	: ultraviolet
v / v	: volume / volume

Codes à une lettre et à trois lettres des acides aminés :

A	Ala	Alanine
C	Cys	Cystéine
D	Asp	acide Apsartique
E	Glu	acide Glutamique
F	Phe	Phénylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Méthionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Trp	Thréonine
V	Val	Valine
W	Trp	Tryptophane
Y	Tyr	Tyrosine

Chapitre 1

Contexte scientifique

Afin de comprendre un mécanisme moléculaire correspondant à une fonction biologique, l'étude de la structure des macromolécules est un outil de choix. Elle permet de connaître l'aspect des molécules et de replacer les connaissances biologiques, biochimiques et biophysiques dans leur contexte structural. C'est une des nombreuses facettes de la biologie intégrative (Wake, 2003). Le cadre de mes recherches de thèse se place dans ce contexte. J'ai en effet cherché à caractériser les déterminants viraux de la transmission d'un virus icosaédrique de plante, le *Grapevine fanleaf virus* (GFLV), en combinant des approches de biologie moléculaire et structurale.

Dans ce premier chapitre introductif, j'expliquerai tout d'abord les concepts fondamentaux de la vection des virus et les concepts d'architecture des particules virales icosaédriques qui sont deux thématiques fondamentales de mon travail de thèse. Puis, je présenterai les déterminants de leur transmission, au travers d'exemples de virus dont la vection est largement étudiée. J'accorderai un intérêt particulier aux virus pour lesquels des données structurales sont disponibles afin de montrer comment l'obtention d'informations structurales a permis d'avancer dans la compréhension des bases moléculaires de l'interaction virus-vecteur. Le chapitre 2 présentera les caractéristiques biologiques du modèle étudié et l'objectif de l'étude.

1. Concepts fondamentaux de la vection des phytovirus

1. 1 Définition de la vection

Les virus sont des parasites cellulaires obligatoires à organisation simple. Parmi les phases fondamentales du cycle viral, la transmission à un nouvel hôte est primordiale pour la survie du virus. Afin d'atteindre et infecter un nouvel hôte, diverses stratégies ont été développées par les virus. La transmission est dite « interne » ou « verticale » lorsque le virus infecte la descendance de l'hôte infecté. Ainsi chez certains phytovirus elle se fait par la semence, le pollen ou par multiplication végétative. La transmission est appelée « externe » ou « horizontale » lorsque le virus infecte un nouvel hôte qui ne fait pas partie de sa descendance. Dans le cas des phytovirus, la transmission horizontale s'effectue soit par voie mécanique (inoculation par contact au travers de blessures) soit par

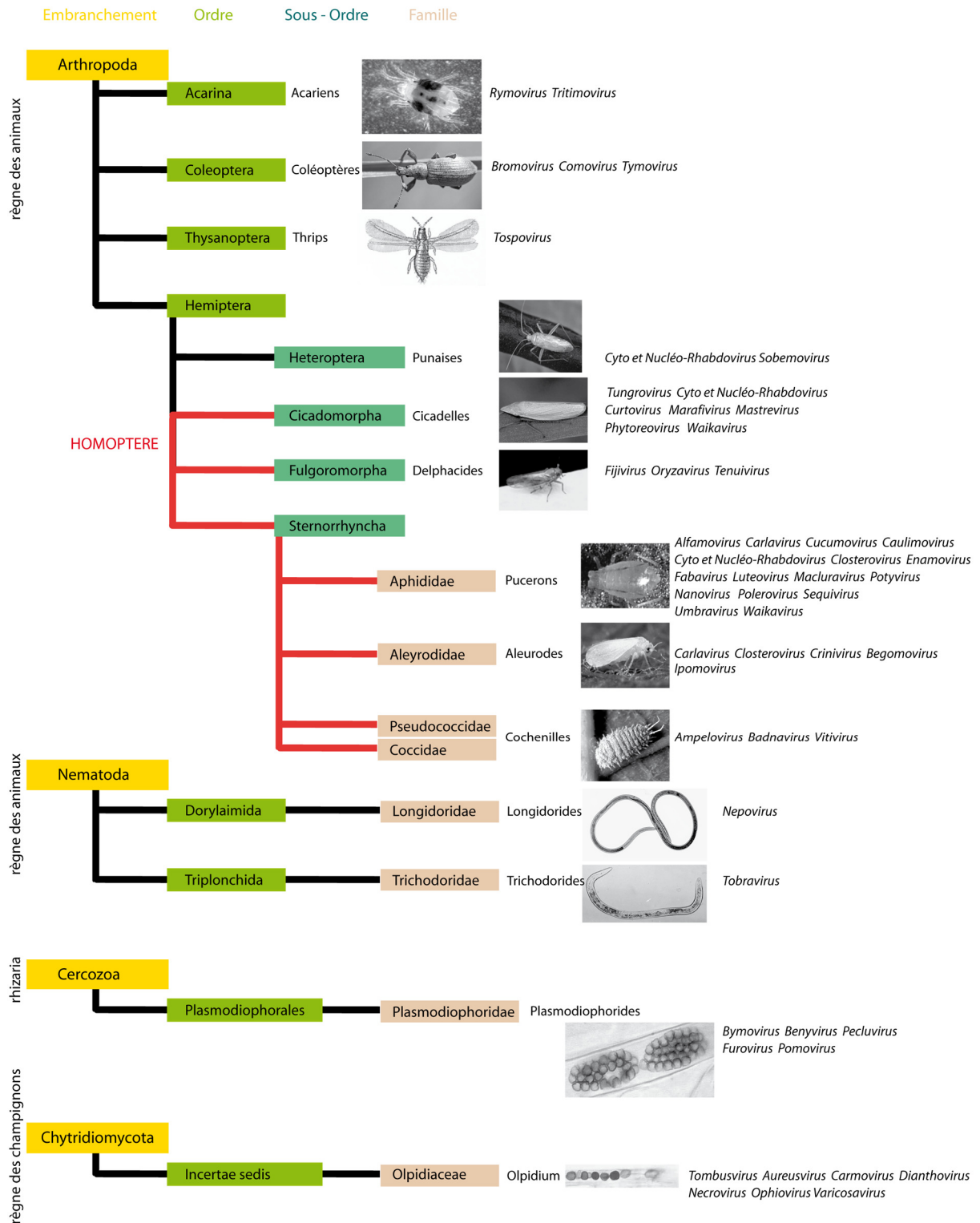


Fig. 1.1. Les principaux vecteurs de virus de plantes. Le nom usuel des vecteurs est indiqué à gauche de chacune des vignettes illustrant une espèce vectrice. Les genres viraux transmis par chaque famille de vecteur sont indiqués à droite des vignettes. Les familles faisant partie des homoptères (ligne magenta) sont vecteurs de la majorité des virus de plantes. L'origine de chaque image est donnée dans l'annexe 1.

l'intermédiaire d'organismes vivants dénommés vecteurs. Un agent pathogène peut utiliser plusieurs stratégies de transmission. Ainsi, le *Cucumber mosaic virus* (CMV) peut être transmis à la fois par la graine et par ses vecteurs (Ali and Kobayashi, 2010).

La vection est la transmission d'un agent pathogène d'un hôte infecté à un nouvel hôte par l'intermédiaire d'un organisme vivant, le vecteur. C'est un processus actif, puisque les vecteurs inoculent les particules virales directement dans les cellules, généralement lors de prises alimentaires. Ils permettent ainsi au virus de franchir la cuticule et la paroi cellulaire des plantes, ou encore l'épiderme des animaux dans le cas de virus de vertébrés. La grande majorité des phytovirus recensés par l'ICTV (International Committee on Taxonomy of Viruses) est transmise par arthropodes (environ 75%), mais cette stratégie est minoritaire dans le cas des virus d'animaux (environ 40 %) (Hogenhout *et al.*, 2008; van den Heuvel *et al.*, 1999).

1. 2 Vecteur et spécificité de vection des phytovirus

Les espèces vectrices de phytovirus appartiennent à des taxons très divers parmi les arthropodes (insectes et arachnides), les nématodes, les champignons et plasmidiophorales et même certaines plantes parasites (exemple de la cuscute et l'orobanche). La transmission des virus par arthropodes a été la première à être mise en évidence par l'identification, en 1895, d'une cicadelle vectrice du virus nanisant du riz (*Rice dwarf virus*, RDV) et en 1893 du moustique vecteur de la fièvre bovine du Texas, maladie infectant les vertébrés (Beaty and Marquardt, 1996; Campbell and Bruehl, 1986). Ce n'est qu'en 1958, que la transmission par nématodes a été démontrée, par l'identification de *Xiphinema index* comme vecteur de la maladie du court noué de la vigne (*Grapevine fanleaf virus*, GFLV) (Hewitt *et al.*, 1958). A la même période, la transmission de la maladie des grosses nervures de la laitue (*Lettuce-big vein virus*, LBVV) a été associée à la présence d'un champignon, *Olpidium brassicae* (Fry, 1958; Grogan *et al.*, 1958). La transmission par nématode et par champignon a été appelée «soil-borne transmission», faisant référence au fait que ces vecteurs vivent dans le sol. Il est important de préciser que le type de vecteur utilisé par le virus, tout comme son mode de vection, aura des conséquences majeures sur l'épidémiologie de la maladie virale et les méthodes de luttés choisies contre sa propagation (Spence, 2001).

Les arthropodes faisant partie du groupe des Homoptères, qui inclut les sous-ordres des *Cicadomorpha*, des *Fulgoromorpha* et des *Sternorrhyncha*, assurent la transmission de 80% des phytovirus transmis par insectes (Fereser and Moreno, 2009) et 55% des virus de plantes dont le vecteur a été caractérisé (Nault, 1997). Parmi les arthropodes, les deux principaux vecteurs sont les pucerons, avec 275 virus transmis recensés dans 19 genres différents (Nault, 1997) et les aleurodes, avec 114 virus transmis recensés dans 5 genres différents (Jones, 2003)(fig. 1.1). La transmission par nématodes, plasmidiophorales et champignons concerne une minorité de genres viraux. Seules deux

familles de nématodes, les *Longidoridae* et les *Trichodoridae* comprennent des vecteurs de virus (MacFarlane, 2003). Les champignons et les plasmidiophorales transmettent quant à eux, respectivement 16 et 21 virus, appartenant aux genres viraux détaillés dans la figure 1.1 (Rochon *et al.*, 2004).

Les relations entre vecteurs et virus montrent différents niveaux de spécificité selon la gamme d'espèces vectrices que possède un virus donné. Cela peut aller d'un groupe d'espèces appartenant soit à des taxons différents, tels les *Rhabdovirus* qui sont transmis à la fois par des cicadelles, les pucerons et les punaises, soit à la même famille, tels les *Potyvirus* qui sont transmis par de multiples espèces de pucerons, soit à une seule espèce vectrice, tel l'*Arabid mosaic virus* qui est transmis par une seule espèce de nématode (*Xiphinema diversicaudatum*). Dans le cas de la plus haute spécificité, la relation entre vecteur et virus est dite « exclusive » (Brown and Weischer, 1998): le GFLV est exclusivement transmis par *Xiphinema index* et ce nématode est vecteur du seul GFLV.

1. 3 Différents modes de transmission

Le mode de transmission est établi en fonction du type d'interaction virus-vecteur. Si on considère la transmission en terme de phases, le vecteur prélève tout d'abord le virus sur une plante infectée. C'est pendant cette étape d'ingestion que le virus est retenu au niveau de l'appareil alimentaire du vecteur. Cette première phase, nommée « acquisition », va impliquer une spécificité de reconnaissance entre des protéines virales et des sites spécifiques du vecteur. Puis, différents scénarios se présentent. Dans un mode de transmission **non circulant**, le virus reste adsorbé à son site de rétention, jusqu'à la prochaine prise alimentaire du vecteur où il sera injecté dans une nouvelle plante ou dans la même plante si le vecteur est non ailé par exemple. Le virus est ainsi inoculé - phase d'« inoculation » - à partir du même site de rétention. Dans le mode de transmission **circulant** - encore appelé **persistant non multipliant** - le virus est internalisé dans le vecteur. Dans ce cas, concernant la transmission par arthropodes, les particules virales migrent vers les glandes salivaires avant d'être relarguées avec la salive dans un nouvel hôte. Dans le mode **circulant multiplicatif**, encore appelé **persistant multipliant**, le virus est internalisé et se réplique dans le vecteur.

Tous ces termes ont été établis par des études effectuées sur la transmission par arthropodes. La notion de persistance du virus dans le vecteur a été définie d'après le temps pendant lequel le vecteur reste infectieux, c'est-à-dire capable d'inoculer l'agent pathogène à une plante saine (Sylvester, 1956; Watson and Robers, 1939). Ces caractéristiques sont résumées dans le tableau 1.1.

La majorité des phytovirus sont transmis selon le mode non-circulant, alors que dans le cas des virus d'animaux le mode multipliant est majoritaire (van den Heuvel *et al.*, 1999). Il convient de noter que dans le cas de la vection des virus infectant les vertébrés, une terminologie différente est souvent appliquée. La transmission non persistante est dite « mécanique », car elle est considérée comme une

association quasi non spécifique entre virus et vecteur (Carn, 1996). Par opposition, la transmission persistante dite « biologique » est considérée comme étant spécifique et circulante. La vection de virus d'animaux est assistée par des arthropodes appartenant majoritairement à l'ordre des Diptères (moustiques, mouches). Par conséquent, les virus d'animaux transmis par vecteurs sont référencés comme étant des « *arthropode-borne viruses* » duquel découle le terme d' « arbovirus ».

Tableau 1.1. Caractéristiques des différents modes de vection de virus par insectes
(d'après Ng & Falk, 2006)

Caractéristiques	Modes de transmission			
	non-circulant		circulant	
	non-persistant	semi-persistant	persistant-non-multipliant	persistant-multipliant
Durée d'acquisition et d'inoculation	secondes à minutes	minutes à heures	heures	heures
Durée de latence	aucune	aucune	heures à jours	jours à semaines
Durée de rétention	minutes à heures	heures à jours	plusieurs jours	toute la vie
Passage trans-stadial	non	non	oui	oui
Réplication	non	non	non	oui
Passage trans-ovarien	non	non	non	souvent
Spécificité vectrice	large	étroite	étroite	étroite

1. 4 Déterminants de la transmission

La transmission des virus de plantes repose sur une association spécifique entre virus et vecteurs. Cette interaction implique donc une reconnaissance spécifique entre une ou des protéines virales et un interactant chez le vecteur. Les sites d'attachement du virus dans le vecteur sont en général identifiés et limités à des structures bien déterminées comme la membrane externe des zoospores des champignons, la cuticule de la partie antérieure de l'appareil alimentaire des arthropodes et nématodes, ou encore la membrane intestinale dans le cas de la transmission circulante par arthropodes. Cependant, aucun récepteur n'est encore totalement caractérisé à ce jour. Les données les plus avancées au niveau du vecteur concernent la vection du CaMV, pour lequel la nature protéique du site récepteur du virus a été démontrée (Uzest *et al.*, 2007).

Toutes les données établies sur les protéines virales et leurs motifs protéiques déterminant la vection des phytovirus ont été synthétisées et publiées sous forme de revues, auxquelles je renvoie le lecteur pour des informations complémentaires à la thèse : sur la transmission par tous types de vecteurs (Andret-Link and Fuchs, 2005; Ng and Falk, 2006; Ziegler-Graff and Brault, 2008), par nématodes (MacFarlane, 2003), par *Olpidium* et plasmodiophorides (Rochon *et al.*, 2004), par arthropodes (Herrbach, 2004; Hogenhout *et al.*, 2008). Concernant les déterminants viraux de la

transmission, il est établi que la capsid joue un rôle primordial dans la rétention par le vecteur, que se soit dans le cas des insectes, des nématodes, des champignons ou des plasmodiophorides (Andret-Link and Fuchs, 2005; MacFarlane, 2003; Ng and Falk, 2006; Rochon *et al.*, 2004; Ziegler-Graff and Brault, 2008). La transmission non-circulante peut dans certains cas impliquer une protéine virale non structurale, parfois deux, qui assiste la rétention du virus par le vecteur, en faisant le lien entre le virion et le site de rétention (fig. 1.2). Ces protéines virales sont globalement nommées « protéines assistantes » ou « facteurs assistants », en anglais « HC » pour « *helper component* » (Syller, 2006). Les domaines de la capsid ou des protéines HC impliqués dans la transmission non circulante ont été pour la plupart identifiés sur la base de comparaison de séquences d'isolats transmissibles et non-transmissibles. Le rôle de ces domaines a ensuite été confirmé par des études de génétique inverse par mutagenèse effectuée dans le génome viral ou dans des clones infectieux (Atreya *et al.*, 1990; Huet *et al.*, 1994; Lopez-Moya *et al.*, 1999; Peng *et al.*, 1998; Perry *et al.*, 1998; Perry *et al.*, 1994; Robbins *et al.*, 1997; Vellios *et al.*, 2002).

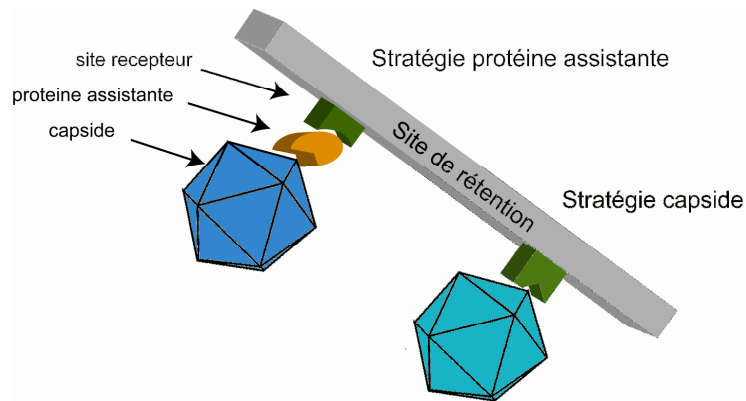
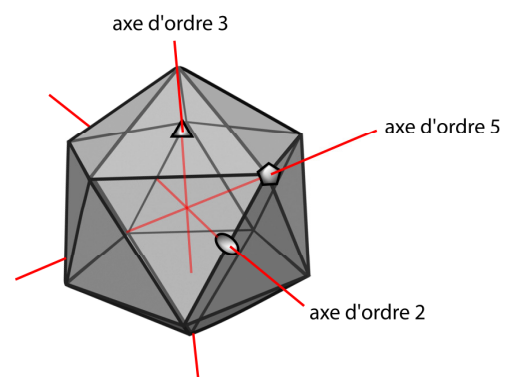


Fig. 1.2. Les stratégies d'interaction entre la particule virale (particule icosaédrique en bleu) et le site de rétention du vecteur (symbolisé par un rectangle gris) dans le cas du mode non-persistant. A gauche, une protéine virale non structurale (orange) assiste l'interaction entre la particule et le récepteur potentiel. A droite, la capsid interagit directement avec le site récepteur (vert). Dans le cas de la transmission par insectes et par nématodes, le site de rétention correspond à la cuticule interne de l'appareil alimentaire, dans le cas de la transmission par champignons, il correspond à la membrane externe des zoospores.

2. Architecture des particules virales icosaédriques

Qu'est ce que la capsid ? La forme libre d'un virus s'appelle le virion, encore dénommée particule virale. Il est composé du génome viral, responsable de l'infectivité du virion, et de la capsid qui protège le génome. Les protéines constituant la capsid sont les protéines structurales du virus. Les protéines non structurales ne forment pas la capsid, mais dans certains cas elles peuvent se retrouver encapsidées avec le génome : on parle alors de nucléoprotéines. Les capsides peuvent adopter des morphologies totalement différentes suivant la famille de virus, en fonction de l'organisation des sous-unités au sein de la capsid. Deux types de symétries permettent de décrire une très grande partie des virus de plante : les virus à symétrie hélicoïdale se présentant sous forme de bâtonnet plus ou moins flexueux et les virus à symétrie icosaédrique sous forme parasphérique. Les études de diffractions aux rayons X sur les premiers cristaux de virus de plantes (*Tomato bushy stunt virus* (TBSV), *Tobacco mosaic virus* (TMV) et *Turnip yellow mosaic virus* (TYMV)) ont permis de définir cette organisation (Crick and Watson, 1956; Franklin, 1955; Klug *et al.*, 1961; Watson, 1954). Ainsi, la problématique de savoir comment plusieurs copies d'une seule protéine asymétrique peuvent être disposées selon un réseau régulier dans la capsid s'est posée très tôt en virologie. Cette partie sera consacrée à l'étude de l'architecture des capsides icosaédriques et plus particulièrement à celle des *Picornavirales*, puisqu'il s'agit de l'ordre viral auquel appartient le GFLV (Le Gall *et al.*, 2008).

Fig. 1.3. Schéma de l'icosaèdre, polyèdre régulier comprenant 20 faces et 12 sommets



2. 1 Les capsides icosaédriques

Le paradigme de la construction des capsides icosaédriques repose sur le fait qu'elles sont formées de l'assemblage de 60 copies identiques d'un bloc nommé unité asymétrique (UA) de l'icosaèdre (Caspar and Klug, 1962). Ces 60 copies sont assemblées selon la symétrie de l'icosaèdre. Ce polyèdre convexe régulier est composé de 20 triangles équilatéraux se rejoignant en 12 sommets (fig. 1.3). Chacun des axes de rotation de l'icosaèdre se croise en son centre, il y a donc 6 axes d'ordre 5 passant par les sommets, 10 axes d'ordre 3 passant par le milieu des faces, et 15 axes d'ordre 2 passant par le milieu des arêtes. Même si ces virus sont dits isométriques, il faut bien comprendre que leur capsid possède des orientations différentes. En effet, l'image d'un icosaèdre orienté selon un axe d'ordre 2 n'est pas la même que si ce dernier est orienté selon un axe d'ordre 5, ou d'ordre 3.

La capsid virale la plus basique possède 60 sous-unités identiques, avec 1 sous-unité par UA. Dans ce cas, les sous-unités sont strictement équivalentes les unes par rapport aux autres. Caspar et Klug (1962) ont démontré qu'en introduisant de manière régulière plus de 60 sous-unités par capsid, on obtenait une variante de l'icosaèdre, les icosadeltaèdres, ressemblant au dôme géodésique. Dans ce modèle, certains contacts entre sous-unités sont de type pentamérique et d'autre de type hexamérique, ce qui nécessite des réarrangements structuraux réguliers de certaines sous-unités. Les protéines ne sont donc plus équivalentes, comme dans l'icosaèdre de base, mais quasi-équivalentes. (Caspar and Klug, 1962)

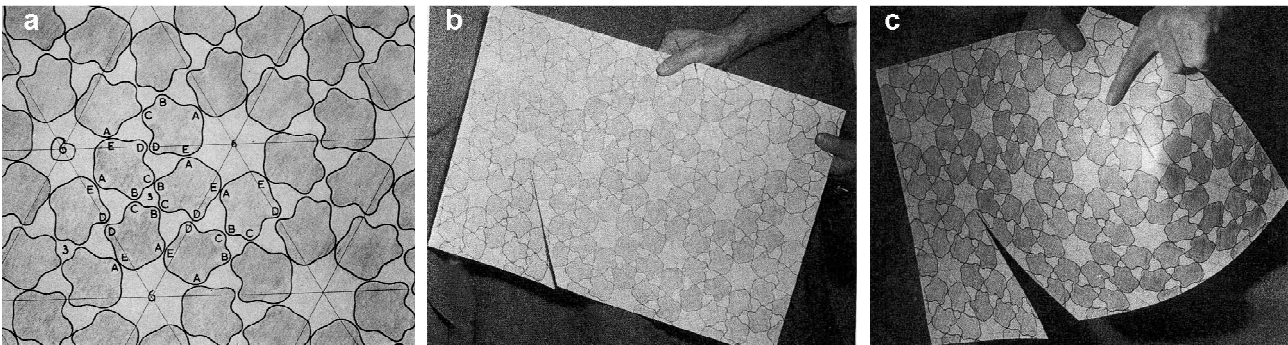


Fig. 1.4. Illustration d'un réseau équivalent d'unités asymétriques (UA) (a) au sein d'un réseau 2D de triangles équilatéraux. Le motif possède des axes de rotation d'ordre 6. Même si les UA sont positionnées dans 6 orientations différentes dans l'espace, elles sont reliées de manière équivalente : les contacts (A-E ; C-B ; D-D) sont conservés dans le motif. Il est possible de rendre le réseau hexagonal qui par définition est plan (b) en une surface convexe (c) par l'introduction d'un pentamère à la place d'un hexamère. Figure adaptée d'après Caspar et Klug, 1962).

Comment calculer le nombre de triangulation ?

L'organisation des faces triangulaires de l'icosaèdre forme sur un plan 2D un réseau hexagonal qu'il est possible de rendre convexe par l'introduction de pentamères à la place d'hexamères (fig. 1.4). C'est l'opération qui est faite quand nous plions le modèle de l'icosaèdre. Les pentamères forment les sommets de l'icosaèdre et définissent le triangle équilatéral qui fait la face de l'icosaèdre (Johnson and Speir, 1997). Suivant la position des pentamères insérés dans le réseau hexagonal, la face de l'icosaèdre contiendra plus ou moins d'hexamères, donc plus ou moins de protéines de capsides (fig. 1.5). Il est ainsi possible d'introduire plus de sous-unités dans la capsid tout en conservant la symétrie icosaédrique. Le nombre de triangulation est calculé d'après les positions des pentamères insérés dans le réseau hexagonal selon la formule $T = h^2 + hk + k^2$ où h et k sont les nombres entiers positifs correspondant à la position d'insertion (fig. 1.5). La méthode de construction reprise dans la figure 1.5, décrite par Johnson et Speir en 1997, ne correspond pas exactement à celle de Caspar et Klug qui, dans leur concept d'origine, divisent le triangle équilatéral

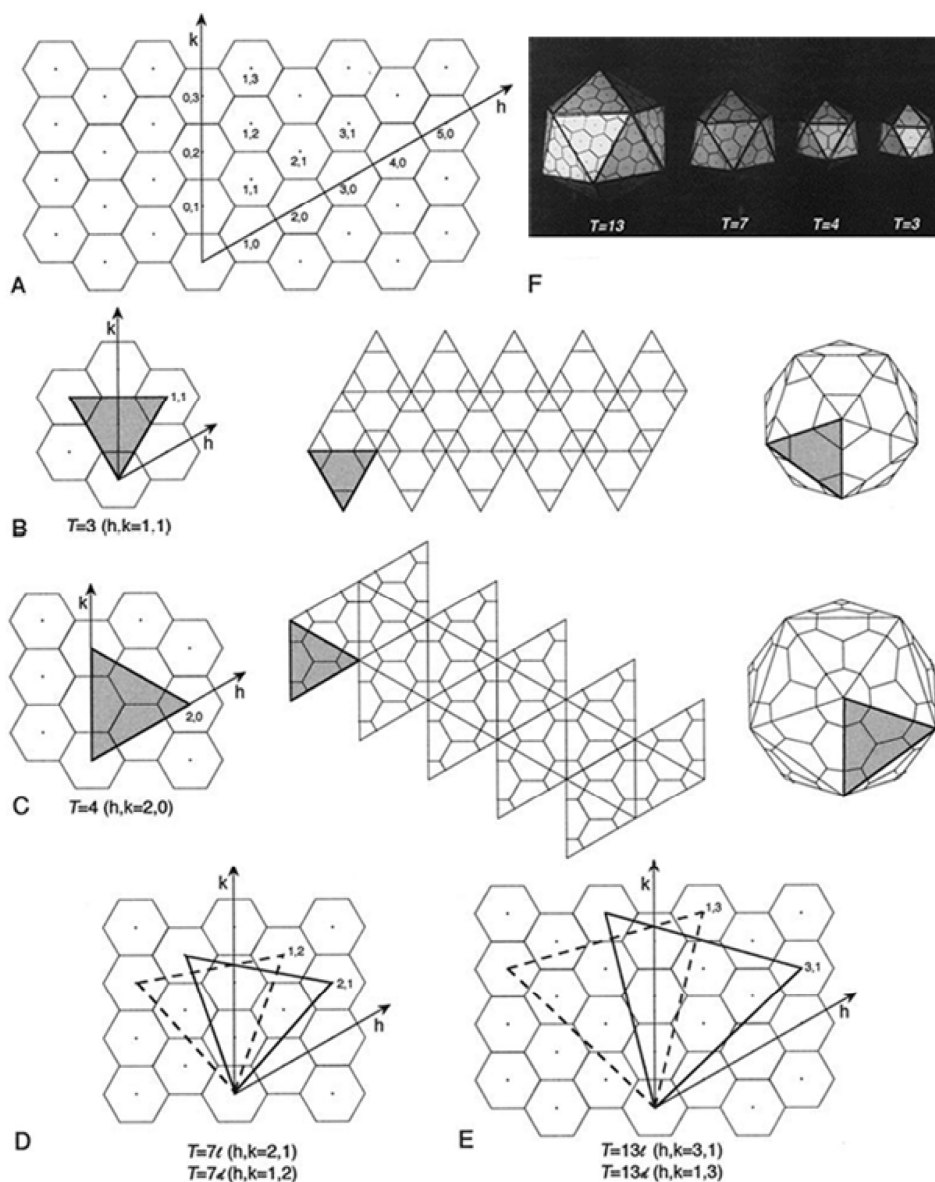


Fig. 1.5. Principe géométrique de génération d'un réseau de surface quasi-équivalent icosaédrique. (A) La répétition régulière d'hexagones est la base géométrique de génération de l'icosaèdre. La position des hexamères est assignée selon sa position sur l'axe h et k . D'après le principe de quasi-équivalence, l'icosaèdre contiendra $60T$ sous-unités organisées en hexamères et en pentamères. Comme illustré par la figure 1.4, un motif hexagonal est considéré comme plan et les pentagones sont considérés comme étant convexes, introduisant une courbure dans le réseau hexagonal à la position où il est inséré. Un icosaèdre est construit en introduisant 12 pentamères, chacun à une position (h, k) définies par le centre de l'hexamère qu'il remplace. Pour construire le modèle d'un motif quasi-équivalent particulier, une face de l'icosaèdre est générée dans le réseau hexagonal. Pour cela, l'hexagone d'origine $(h, k = 0)$ et le (h, k) hexagone sont remplacés par un pentamère. Le nombre de triangulation T est défini par la formule $T = h^2 + hk + k^2$. Le premier coté du triangle est formé par le segment rejoignant le point d'origine $(h, k = 0, 0)$ et le point (h, k) . Les deux autres côtés permettent de former un triangle équilatéral incluant le premier segment. (A) La face de l'icosaèdre pour un motif $T = 3$ est définie par le triangle grisé $(h, k = 1, 1)$. Le motif créé est répété 20 fois, puis replié pour former l'icosaèdre. (C) Même opération pour un motif $T = 4$ $(h, k = 2, 0)$. Deux motifs $T = 7$ (D) et $T = 13$ (E) sont possibles, les deux étant des images miroirs. (F) Icosaèdres possédant des unités de triangulation différentes construit avec la méthode décrite au dessus. Figure adaptée d'après Baker *et al.*, 1999 et Johnson *et al.*, 1997.

de l'icosaèdre primitif en sous-triangulations successives. Malgré cela, les deux méthodes aboutissent aux mêmes modèles (Johnson and Speir, 1997).

Le nombre de triangulation T , définit le nombre de sous-unités protéiques retrouvées dans une UA icosaédrique. Si la capsid est formée à partir de la même sous-unité, le nombre total de sous-unités sera de $60T$ et le nombre d'hexamères sera de $10(T-1)$. La formule nous confirme bien qu'un virus, dont les protéines de capsid s'organisent selon un réseau icosaédrique $T = 1$, possède 60 sous-unités et aucun hexamère. C'est le cas de l'icosaèdre dont toutes les sous-unités sont équivalentes. L'illustration F de la figure 1.5 donne plusieurs exemples d'icosaèdres à surface quasi-équivalente possédant des nombres de triangulations différents. Si la taille de la sous-unité est maintenue constante, les capsides à nombre de triangulations plus élevés seront plus volumineuses et pourront ainsi encapsider plus de matériel viral. Par définition, si l'UA contient plus d'un type de protéines, le nombre de triangulation est dénommé pseudo T .

2. 3 Variation de la symétrie icosaédrique pseudo $T = 3$ chez les *Picornavirales*

Le GFLV (genre, *Nepovirus*) appartient à la famille des *Secoviridae* qui, de part ses caractéristiques génétiques et morphologiques communes avec les virus animaux appartenant à la famille des *Picornaviridae*, est inclu dans l'ordre de virus nommé *Picornavirales* (Le Gall *et al.*, 2008; Sanfaçon *et al.*, 2009). La conservation de l'architecture de capsid de plusieurs représentants du genre *Secoviridae* met également en lumière ce lien dans l'évolution (Chandrasekar and Johnson, 1998; Lin and Johnson, 2003). Parmi les *Picornavirales*, la structure de 4 virus infectant les végétaux (genres *Comovirus* et *Nepovirus*) et de 21 virus infectant les animaux (genres *Aphthovirus*, *Cardiovirus*, *Enterovirus* et *Senecavirus*) a été résolue par cristallographie aux rayons X (Tableau 1.2). Leur capsid est non enveloppée, de l'ordre de 30 nm de diamètre et adopte une symétrie icosaédrique pseudo $T = 3$. Ce sont donc des capsides très simples, par rapport aux capsides à symétrie bien plus élevée des *Adenovirus* (pseudo $T = 25$), ou encore par rapport aux particules virales composées de plusieurs coques protéiques, comme celles des *Reovirus* (capsid intermédiaire $T = 13$). Néanmoins, l'originalité des capsides de *Picornavirales* réside dans le fait que plusieurs genres viraux ont adapté la symétrie de la particule en $T = 3$, mais chacun à sa manière.

Dans l'organisation $T = 3$, trois sous-unités possédant la même chaîne polypeptidique sont présentes dans l'UA icosaédrique (fig. 1.6). Puisque leur structure n'est pas strictement équivalente au sein de la capsid (voir partie 2.1), elles sont nommées différemment d'après leur position dans la capsid : sous-unité A si elles se placent au niveau des axes d'ordre 5 ; sous-unité C au niveau des axes d'ordre 2 et 3, sous-unité B au niveau des axes d'ordre 3 (fig. 1.6) (Lucas *et al.*, 2002; Smith *et al.*, 2000; Speir *et al.*, 1995). De plus, d'autres axes de symétrie - les « pseudo » axes se rajoutent aux axes principaux de symétries de l'icosaèdre. Premièrement, un pseudo-axe 3 est présent entre les 3

sous-unités de l'UA. Deuxièmement, les sous-unités A et B provenant de 2 UA différentes se rejoignent en un pseudo-axe 2 (fig. 1.6). Troisièmement, l'axe de rotation icosaédrique d'ordre 3 correspond, dans une capsidite $T = 3$, à un pseudo-axe d'ordre 6, puisqu'il se place au centre d'un hexamère de sous-unités de capsidite.

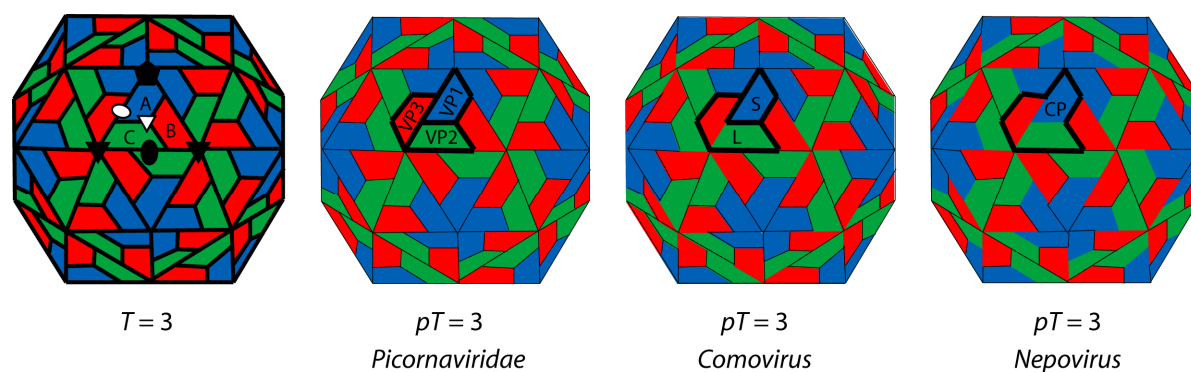
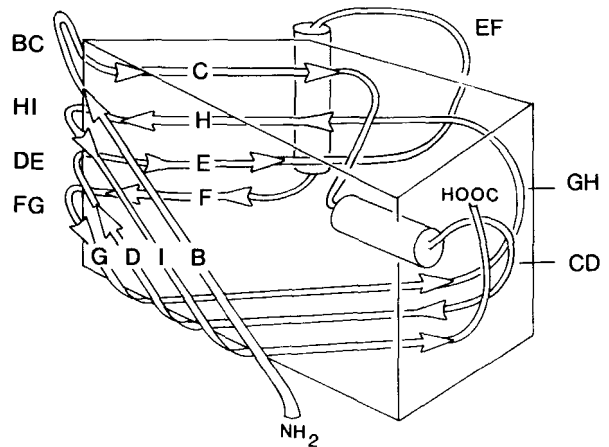


Fig. 1.6. Modèles d'organisation de surface de la capsidite $T = 3$ et ses dérivés pseudo $T (pT) = 3$ chez les virus de la famille *Picornaviridae*, et genres *Comovirus* et *Nepovirus*. Les 180 sous-unités de la capsidite en $T = 3$ (chacune symbolisée par un trapèze) se subdivisent en 3 classes, A, B et C, suivant leurs positions dans l'unité asymétrique (UA) et coloriées d'après le code couleur conventionnel. L'UA de l'icosadeltaèdre de gauche est surlignée par un trait noir épais. Le pentagone, le triangle et l'ovale noir représentent respectivement les axes d'ordre 5, 3, 2, d'une UA de l'icosaèdre. Le triangle et l'ovale blanc, représentent les pseudo-axes d'ordre 3 et 2. Chez les *Picornaviridae*, 60 copies des protéines virales VP1, VP2 et VP3 forment la capsidite, avec VP4 qui n'est pas visible sur ce schéma. Chez les *Comovirus*, 60 copies de la sous unité *small* (S) et *large* (L) forment la capsidite. Chez les *Nepovirus*, 60 copies de la protéine de capsidite (CP) forment la particule virale. Puisque dans les deux derniers icosaédres, les trois domaines de l'UA sont dans une configuration similaire au virus $T = 3$, à l'exception du fait que les trois domaines proviennent d'une chaîne polypeptidique différente, les virus sont décrits comme des $pT = 3$.

Chez les *Picornaviridae*, comme par exemple le *Poliiovirus* et le *Rhinovirus*, les trois sous-unités formant l'UA sont trois protéines virales (VP) différentes nommées VP1, VP2 et VP3. Une protéine supplémentaire VP4 est présente dans l'UA, mais n'intervient pas sur la symétrie car elle est interne à la particule virale (Acharya *et al.*, 1989; Hogle *et al.*, 1985; Luo *et al.*, 1987; Rossmann *et al.*, 1985). Chez les *Comovirus*, 2 protéines de capsidite sont présentes : La protéine S (*Small*) et la protéine L (*Large*) qui correspondent respectivement à la sous-unité A, et la sous-unité B+C (Lin *et al.*, 1999; Lin *et al.*, 2000). Chez les *Nepovirus*, 1 seule protéine forme l'UA, mais elle contient 3 domaines structuraux correspondant aux positions des sous-unités A, B et C du réseau icosaédrique $T = 3$. Les domaines de la CP sont donc nommés C, B et A, en partant de la partie N-terminale. Ils sont reliés par une boucle de liaison d'une quinzaine de résidus interne à la particule virale (Chandrasekar and Johnson, 1998). La figure 1.6 montre que la position d'UA est décalée par rapport à sa position dans un réseau icosaédrique $T = 3$, mais cela ne change pas l'organisation des domaines A, B et C au sein du virion.

La structure tertiaire des CP des virus appartenant à la super-famille des *Picornavirales*, est bien conservée (Harrison, 1991; Le Gall *et al.*, 2008). Les chaînes polypeptidiques adoptent un repliement en *sandwich* β , constitué de brins β antiparallèles, dont certaines boucles se replient en hélices α . La structure est fortement vrillée et adopte une forme trapézoïdale, communément appelée *jelly-roll* (ou *β barrel*). La nomenclature du *jelly roll* modèle a été établie sur la structure du *Poliovirus* (figure 1.7) (Hogle *et al.*, 1985) qui contient 8 brins β antiparallèles. Cette structure tertiaire diffère selon le genre viral par son nombre de brins β , d'hélices α , la longueur des boucles, ou encore l'orientation de la sous-unité dans la capsid. Dans le cas du CPMV, qui est le membre type des *Comovirus*, les 3 domaines C, B et A, possèdent tous de légers réarrangements structuraux (Lin *et al.*, 1999). La différence la plus marquante étant certainement l'existence d'un feuillet β supplémentaire dans le domaine A qui entraîne le repliement du domaine en un sandwich à 10 brins β .

Fig. 1.7. Schéma du repliement en *jelly roll* de la chaîne polypeptidique des sous-unités protéique de capsides de plusieurs virus icosaédriques. Les flèches et les cylindres représentent respectivement les brins β et les hélices α . Figure d'après Hogle *et al.*, 1985.



3. Les modèles biologiques

Pour plusieurs modèles viraux, les analyses structurales entreprises pour les protéines virales impliquées dans la transmission ont permis d'avancer dans la compréhension du mécanisme d'interaction virus-vecteur. Cette partie du manuscrit se consacre aux virus pour lesquels des avancées scientifiques majeures sur la structure des déterminants viraux de la vection ont été obtenues. Pour avoir une vision plus globale et puisque beaucoup de données structurales sont disponibles dans le cas des virus animaux, j'ai également inclus dans cette partie les connaissances établies sur les déterminants de la vection des ces virus.

3. 1 Structures de particules virales résolues

La capside de virus est une structure multifonctionnelle et dynamique. Le rôle primordial de la capside est de protéger le génome viral. La capside doit donc à la fois être résistante et refermée, afin de protéger les acides nucléiques lors du passage d'une cellule à une autre, mais elle doit également pouvoir se relâcher, se désassembler, afin de libérer les acides nucléiques dans une nouvelle cellule. Elle est souvent essentielle au mouvement du virus d'un hôte à un autre, mais également pour le mouvement de cellule à cellule. En effet, dans le cas de la majorité des phytovirus ce sont des particules qui traversent la paroi cellulaire au niveau des plasmodesmes modifiés par la protéine de mouvement d'origine virale (Ritzenthaler and Hoffmann, 2007). Dans le cas des virus infectant les vertébrés, la reconnaissance des virions au niveau de récepteurs cellulaires entraîne l'internalisation du virus à travers la membrane plasmique (Marsh and Helenius, 1989). Dans certains cas, la capside est également impliquée dans d'autres mécanismes tels que la réplication du génome viral, la symptomatologie (Bol, 2008), ou encore la défense du virus contre son hôte via son intervention dans la suppression de silencing (Liu *et al.*, 2004; Shi *et al.*, 2009). La capside est un énorme complexe macromoléculaire impliquant des protéines, des glycoprotéines, des acides nucléiques, et dans le cas des virus enveloppés des lipides. L'obtention de la structure tridimensionnelle à haute résolution de particules virales représente par conséquent un vif intérêt, tant au niveau fondamental, afin d'étudier les interactions entre macromolécules, qu'au niveau appliqué, afin de développer des agents anti-viraux.

En 2005, Natarajan et collaborateurs recensaient dans la banque de donnée VIPER (Virus Particle Explorer) spécialisée dans les structures de capsides complètes, 62 structures cristallographiques uniques de virus et 25 obtenues par reconstruction 3D à partir d'images de cryomicroscopie électronique (cryo-ME), en excluant les structures redondantes telles que les sérotypes, isolats, mutants, et virus en complexe avec un ligand (Natarajan *et al.*, 2005). Par exemple,

la structure du bactériophage MS2 ne sera comptée qu'une seule fois, alors qu'on trouve plus d'une dizaine d'entrées en rapport à ce phage. En 2009, on dénombre 106 espèces virales dont la structure de capsid est résolue et déposée sur VIPER et/ou sur Protein Data Bank (PDB) (Tableau 1.2). La majorité des structures sont recensées à la fois sur VIPER et la PDB, mais certaines d'entre elles ne sont déposées que sur VIPER (numéro d'identification avec une étoile dans le tableau 1.2) et certaines ne sont disponibles que sur la PDB (virus filamenteux, numéro d'identification : 2COW; 1IFP; 1EI7; 1CGM ; et icosaédrique : 1YXN; 3CNF; 3KK5). A cela se rajoute 39 virus non répertoriés dans les banques de données, dont la structure a été reconstruite à partir d'image de Cryo-EM (références bibliographique du tableau 1.2).

Comme mon travail de thèse utilise des approches de biologie structurale, j'ai recensé parmi les différents virus dont la structure est résolue, ceux dont on dispose des informations sur la vection. Ce travail est rassemblé dans le tableau 1.2 et fait l'objet d'une description détaillée ci-après.

3. 2 Les virus infectant les bactéries : les phages

Les phages présentent une diversité de structure de capsides très intéressante. Pour beaucoup de familles virales, tel que les *Myoviridae*, *Podoviridae*, ou encore les *Siphoviridae*, les capsides sont composées d'éléments supplémentaires non icosaédriques. Plusieurs de ces éléments permettent au phage de s'accrocher directement à la paroi cellulaire de leur hôte et ainsi de pouvoir infecter une bactérie par contact direct (Casjens, 2008). Ils se reproduisent en suivant une alternance entre cycle lytique et lysogène et ne rentrent donc pas dans le cadre de l'étude de la vection des virus.

3. 3 Les virus infectant les animaux

La majorité des virus « animaux » (infectant les vertébrés et les invertébrés) sont transmis sans vecteurs. Ceci inclut la transmission par contact sexuel (*Papillomavirus*, virus de l'immunodéficience humaine HIV), via des aérosols (*Influenza*, *Adenovirus*, *Reovirus*), par transmission féco-orale (*Poliovirus*, *Rotavirus*, *Reovirus*) et par voie parentérale (*Hepatitis*, HIV). Les structures de ces virus ont été résolues soit dans leur intégralité (Tableau 1.2), soit en cristallisant les protéines de capsid de manière isolée (cas du HIV), et elles sont étudiées dans un autre contexte que celui de la transmission par vecteur.

Tableau 1.2. Structures de particules virales résolues par cristallographie aux rayons X ou par reconstruction 3D à partir d'images de cryo-microscopie électronique à transmission

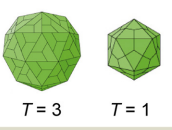
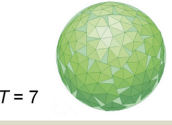

Hôte	Virus : famille, (genre), espèce		références ^a	
Virus	Virophage	Spuntnik	3KK5	
Protozoaires	Mimiviridae (Mimivirus)	Acanthamoeba polyphaga mimivirus	Xiao et al., 2009	
Archea		Sulfolobus turreted icosahedral virus	Kayat et al., 2005	
Champignons	Totiviridae (Totivirus)	Saccharomyces cerevisiae virus L - A 1	1M1C	
		P4 virus	Cheng et al., 1994	
	non assigné (Partitivirus)	Penicillium stoloniferum virus F	3ES5	
Algues	Phycodnaviridae (Chlorovirus)	Paramecium bursaria chlorella virus 1	1M4X	
	(non assigné)	Phaeocystis pouchetii virus	Yan et al., 2005	
Bactéries	Corticoviridae (Corticovirus)	Pseudoalteromonas phage PM2	383*	
	Cystoviridae (Cystovirus)	Pseudomonas phage phi6	Butcher et al., 1997	
	Inoviridae (Inovirus)	Enterobacteria phage fd	2C0W	
		Pseudomonas phage PF3	1IFP	
	Leviviridae (Levivirus)	Enterobacteria phage MS2	2BU1	
		Enterobacteria phage fr	1FR5	
		Enterobacteria phage GA	1GAV	
		Pseudomonas phage PP7	1DWN	
		(Allolevivirus)	Enterobacteria phage Qbeta	1QBE
		(non assigné)	Pseudomonas phage PRR1	2VF9
	Microviridae (Microvirus)	Caulobacter phage cphiCb5	2W4Z	
		Enterobacteria phage alpha3	1MO6	
		Enterobacteria phage G4	1GFF	
		Enterobacteria phage phiX174	1CD3	
		Spiroplasma phage SpV4	1KVP	
	Myoviridae (P2-like viruses)	Enterobacteria phage P2	Dokland et al., 1992	
		Enterobacteria phage P4	Dokland et al., 1992	
		(SPO1-like viruses)	Bacillus phage SPO	Robert et al., 2006
	Podoviridae (Phi29-like viruses)	Bacillus phage phi29	1YXN	
		(P22-like viruses)	Enterobacteria phage P22	Thuman -Comminke et al., 1996
		(Epsilon15 -like viruses)	Salmonella phage epsilon15	3C5B
	Siphoviridae (Lambda-like viruses)	Enterobacteria phage HK97	1OHG	
		Enterobacteria phage lambda	Dockland et al., 1993	
Tectiviridae (Tectivirus)	Bacteriophage PRD1	1W8X		
Plante	 T = 3 T = 1	Bromoviridae (Alfavirus)	Alfalfa mosaic virus	201*
		(Bromovirus)	Cowpea chlorotic mottle virus	1ZA7
			Brome mosaic virus	1JS9
	 T = 7	(Cucumovirus)	Tomato aspermy virus	1LAJ
		Caulimoviridae (Caulimovirus)	Cucumber mosaic virus	1F15
		non assigné (Potexvirus)	Cauliflower mosaic virus	Plisson et al., 2005
	 proteine de capside	Potyviridae (Potyvirus)	Papaya mosaic potexvirus	Kendall et al., 2008
			Soybean mosaic virus	Kendall et al., 2008

Tableau 1.2. Suite

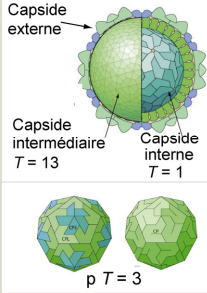
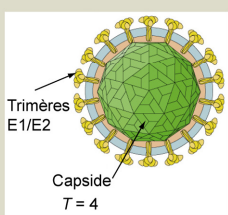
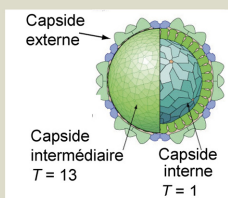
Hôte	Virus : famille, (genre), espèce		références ^a	
	Reoviridae	(Phytoreovirus)	Rice dwarf virus 1UF2	
	Secoviridae	(Comovirus)	Bean-pod mottle virus 1PGL Cowpea mosaic virus 1NY7 Cowpea severe mosaic virus Olson et al., 1995 Red clover mottle virus 202*	
		(Nepovirus)	Tobacco ringspot virus 1A6C Blackcurrant reversion nepovirus Seitsonen et al., 2009	
	Virgaviridae	(Tobamovirus)	Tobacco mosaic virus 1EI7 Cucumber green mottle mosaic virus 1CGM Carnation mottle virus 10PO Melon necrotic spot virus 2ZAH	
		(Tombusvirus)	Tomato bushy stunt virus 2TBV Cucumber necrosis virus Katpally et al., 2007 Tobacco necrosis virus 1C8N	
		(Necrovirus)	Physalis mottle Virus 1E57 Turnip yellow mosaic virus 2FZ2 Desmodium yellow mottle virus 1DDL	
	Tymoviridae	(Tymovirus)	Rice yellow mottle virus 1F2N Sesbania mosaic virus 1X36 Cocksfoot Mottle Virus 1NG0 Ryegrass mottle virus 2IZW Southern Bean Mosaic Virus 4SBV	
		non assigné	(Sobemovirus)	
		Satellite		Tobacco mosaic satellite virus 1A34 Panicum mosaic satellite virus 1STM
	Animal vertébrés et invertébrés	Adenoviridae	(Mastadenovirus)	Human Adenovirus 2 1X9T Human Adenovirus 3 2C9F Human Adenovirus 5 2BVI Human Adenovirus 12 Chiu et al., 1999
		Birnaviridae	(Avibirnavirus)	Avian infectious bursal disease virus 1WCE
		Caliciviridae	(Lagovirus)	Rabbit hemorrhagic disease Thouvenin et al., 1997
			(Norovirus)	Norwalk virus 1IHM Calicivirus Prasad et al., 1994
			(Vesivirus)	San Miguel sea lion virus 4 2GH8
		Dicistroviridae	(Cripavirus)	Cricket Paralysis Virus 1B35
Flaviviridae		(Flavivirus)	West Nile virus 2OF6 Dengue virus 2 1P58 Tick-borne encephalitis virus Furlong et al., 2001 Yellow Fever Virus 1NA4	
		Hepadnaviridae	(Orthohepadnavirus)	Human Hepatitis B 1QGT
		(Avihepadnavirus)	Duck Hepatitis B Kenney et al., 1995	
Hepeviridae		(Hepevirus)	Hepatitis E 3HAG	
Herpesviridae		(Simplexvirus)	Herpes simplex virus Zhou et al., 1999	
		(Varicellovirus)	Equid herpesvirus 1 Baker et al., 1990	
		(Cytomegalovirus)	Human cytomegalovirus Butcher et al., 1998	
		(non assigné)	Channel catfish herpesvirus Booy et al., 1996	
Iridoviridae		(Iridovirus)	Chilo iridescent virus Yan et al., 2009	
Nodaviridae	(Alphanodavirus)	Black beetle virus 2BBV Flock house virus 2Z2Q Nodamura virus 1NOV Pariacoto virus 1F8V		
	(Betanodavirus)	Malabaricus grouper nervous necrosis virus Tang et al., 2002		

Tableau 1.2. Suite

Hôte	Virus : famille, (genre), espèce		références ^a
	Papillomaviridae	(Deltapapillomavirus) Bovine papillomavirus type 1	Baker et al., 1991
		(Mupapillomavirus) Human papillomavirus type 1	Baker et al., 1991
		(Kappapapillomavirus) Cottontail rabbit papillomavirus	Belnap et al., 1996
	Parvoviridae	(Amdovirus) Aleutian mink disease parvovirus	McKenna et al., 1999
		(Densovirus) Galleria mellonella par vovirus	1DNV
		(Dependovirus) Adeno-associated virus 2	1LP3
		Adeno-associated virus 4	2G8G
		Adeno-associated virus 5	Walter et al., 2004
		Adeno-associated virus 8	2QA0
	(Erythrovirus)	Human parvovirus B19	1S58
	(Parvovirus)	Murine minute virus (MVM)	1Z14
		Porcine parvovirus	1K3V
		Feline panleukopenia virus	1C8E
		Canine parvovirus	4DPV
	Picobirnaviridae	(Picobirnavirus) Rabbit picobirnavirus	2VF1
	Picornaviridae	(Aphthovirus) Foot and Mouth Disease Virus	1QQP
		(Cardiovirus) Theiler's Murine Encephalomyelitis Virus	1TME
		Mengo encephalomyocarditis virus	2MEV
	(Enterovirus)	Coxsackievirus B3	1JEW
		Coxsackievirus A21	1Z7S
		Coxsackievirus A9	1D4M
		Human Poliovirus type 1	1PO2
		Human Poliovirus type 2	1EAH
		Human Poliovirus type 3	1PVC
		Bovine enterovirus	1BEV
		Echovirus 1	1EV1
		Echovirus 7	1M11
		Echovirus 11	1H8T
		Echovirus 12	2C8I
		Human Rhinovirus 1A	2HWD
		Human Rhinovirus 2	3DPR
		Human Rhinovirus 3	1RHI
		Human Rhinovirus type 14	1R09
		Human Rhinovirus type 16	1AYM
		Swine Vesicular disease virus	1OOP
	(Senecavirus)	Seneca valley virus	3CJI
	Polyomaviridae	(Polyomavirus) Murine Polyoma Virus	1SID
		Simian Virus 40	1SVA
	Reoviridae	(Cypovirus) Bombyx mori cytoplasmic polyhedrosis virus	3CNF
		(Orbivirus) Bluetongue virus	2BTV
		Broadhaven virus	Schoehn et al., 1997
		(Rotavirus) Rhesus rotavirus	3GZT
		Simian rotavirus	Yeager et al., 1994
		(Orthoreovirus) Human reovirus	Zang et al., 2002b
		Rhesus reovirus	1EJ6
		Avian orthoreovirus	Zang et al., 2005
		(Aquareovirus) Aquareovirus	Shaw et al., 1996
	Tetraviridae	(Omegatetravirus) Nudaurelia capensis omega virus	1OHF
		(Betatetravirus) Providence virus	2QQP
	Togaviridae	(Alphavirus) Semliki forest virus	1DYL
		Sindbis virus	1LD4
		Aura virus	Zhang et al., 2002a
		Ross river virus	Smith et al., 1995
	non assigné	Pyrrococcus furiosus virus	2EOZ



Légende tableau 1.2. Les virus dont la transmission est assistée par un vecteur arthropode, nématode ou champignon sont respectivement encadrés en rouge, bleu, et vert. Une représentation schématique de la particule virale est indiquée pour chaque famille de virus transmis par vecteurs (d'après le site internet ViralZone, référence citée en annexe). Elle n'est pas proportionnelle à la taille réelle des particules. La symétrie icosaédrique de chaque capsid est indiquée par le nombre de triangulation ($T = n$). (a) et (*) numéros d'accès à la Protein Data Bank (PDB, référence citée en annexe) ou à la Viral particle Explorer data base (VIPERdb) respectivement. Les références bibliographiques sont indiquées pour les structures non répertoriées dans les 2 banques de données. Les références sont citées en annexe 2.

Dans le Tableau 1.2, les virus animaux transmis par arthropodes selon le mode persistant propagatif sont au nombre 10, comprenant des virus icosaédriques enveloppés - dans la famille des *Togaviridae* et des *Flaviviridae* - et des virus icosaédriques non enveloppés - dans la famille des *Reoviridae*.

La famille des *Flaviviridae* comprend des pathogènes très répandus chez l'homme, qui sont transmis soit par moustiques, soit par tiques (Vasilakis and Weaver, 2008). Ce sont les virus du genre *Flavivirus* qui sont transmis par vecteur. Parmi eux, le virus de la dengue et le virus de la fièvre jaune (*Yellow fever virus*) sont transmis par les moustiques du genre *Aedes*, le *West nile virus* est transmis par les moustiques du genre *Culex*, et le virus de l'encéphalite est transmis par tiques. La capsid des *Flavivirus* comporte 3 protéines structurales, la protéine d'enveloppe (E), la protéine de membrane (M) et la protéine de capsid organisée dans un réseau icosaédrique $T = 3$. Des protéines d'enveloppes sont responsables de la reconnaissance et de la fusion cellulaire chez l'hôte (Heinz *et al.*, 2003). Il a également été démontré que la protéine E est essentielle à la propagation du *Yellow fever virus* dans le moustique (McElroy *et al.*, 2006).

Les *Alfavirus* (famille *Togaviridae*) comprennent 26 virus transmis par arthropodes (Strauss *et al.*, 1994) dont le *Semliki forest virus* (SFV), le *Sindbis virus* (SINV), le *Ross river virus* (RRV) et l'*Aura virus* transmis par moustiques. Ils possèdent une capsid icosaédrique enveloppée de type $T = 4$. A la surface du virus sont ancrées des spicules, ou en anglais « spikes », formées par des trimères d'hétérodimères de deux glycoprotéines E1 et E2. Ces protéines sont connues pour leur rôle essentiel à l'adhésion cellulaire et à l'entrée des *Alfavirus* dans les cellules de mammifères (Strauss *et al.*, 1994). Elles exposent les domaines de reconnaissance du SINV, SFV et RRV à des récepteurs de type heparan sulfate (HS, une glycosaminoglycan) (Zhang *et al.*, 2005). E2 et E1 influencent également la réplication et le mouvement du SINV dans les cellules de moustiques (Pierro *et al.*, 2007, 2008) (Knight *et al.*, 2009). Les protéines de capsides d'autres *Alfavirus* répandus, tel que le Venezuelan equine encephalitis virus (VEEV), ont également été résolues par cristallographie aux rayons X, mais hors du contexte de la capsid entière (numéro d'accès PDB, 1EP5). Les récepteurs du VEEV sont présents dans l'intestin moyen des moustiques et correspondrait à une protéine de 32kDa de type récepteur à lamines (van den Heuvel *et al.*, 1999).

La famille des *Reoviridae* contient 7 genres viraux infectant les plantes et les vertébrés, mais tous ne sont pas transmis par arthropodes (Hogehout *et al.*, 2008). Parmi les *Reoviridae* infectant les vertébrés, seuls les *Orbivirus* (*Orbivirus*) et les *Coltivirus* sont transmis par arthropodes. Le *Bluetongue virus* est principalement transmis par moucheron de manière persistante (Weaver and Reisen, 2010), mais une transmission non circulante (mécanique) par un autre diptère, *Melophagus ovinus* a également été rapportée (Carn, 1996). Le *Broadhaven virus* (BDV) est transmis par tiques. Les particules virales des deux virus possèdent une architecture similaire en trois couches protéique, dont deux sont organisées selon la symétrie icosaédrique : $T = 2$ pour la capsid interne, $T = 13$ pour la capsid intermédiaire (Grimes *et al.*, 1998; Grimes *et al.*, 1997). La couche externe est constituée de deux protéines virales, chacune organisée en trimère, dont l'organisation est différente entre les deux virus (Schoehn *et al.*, 1997). Il est également notable que dans le cas des réovirus végétaux, comme le *Rice dwarf virus* (RVD, famille *Phytoreovirus*), la protéine de surface nommée P2 est impliquée dans l'invasion du virus dans le vecteur (Hogehout *et al.*, 2008).

Les informations sur les déterminants de la transmission des arbovirus sont donc limitées, même pour les virus dont la structure est déterminée. Comme nous venons de le voir, il a pu être démontré que les protéines d'enveloppe de plusieurs virus influencent le cycle infectieux du virus dans l'insecte, certainement via leur rôle dans l'interaction avec des récepteurs cellulaires. Des études récentes, portent sur la recherche de facteurs immuns anti-viraux produits en réponse à l'infection du virus dans le vecteur (voir pour revue Fragkoudis *et al.*, 2009). L'épidémiologie des arbovirus (diversité génétique, dynamique des populations, risques d'épidémies) est quant à elle bien documentée (Vasilakis and Weaver, 2008; Weaver and Vasilakis, 2009) (Weaver and Barrett, 2004; Weaver and Reisen, 2010).

La transmission dite « mécanique », qui s'apparente plus au mode de transmission du GFLV, est faiblement étudiée, car ce mode de transmission est considéré comme ayant peu d'impact épidémiologique (Carn, 1996; Gray and Banerjee, 1999). Pourtant certains cas pourraient être intéressants, comme par exemple certains *Lentivirus* transmis occasionnellement par des diptères, ou encore le cas du *Foot and mouth disease virus* dont la transmission par des mouches tsé-tsé a été rapportée (Carn, 1996).

3. 4 Les virus infectant les végétaux ou phytovirus

On dispose de beaucoup d'information sur la transmission des phytovirus. Elles proviennent majoritairement de virus dont la structure tridimensionnelle n'est pas encore résolue (pour revues : Andret-Link and Fuchs, 2005; Herrbach, 2004; Hogehout *et al.*, 2008; MacFarlane, 2003; Ng and Falk, 2006; Ziegler-Graff and Brault, 2008). Parmi ces virus, les *Potyvirus* sont certainement les

exemples les plus intéressants. Ils représentent les virus dont l'impact économique est le plus sévère pour les productions végétales (Lopez-Moya and A., 2008), et dont la biologie de la vécition est bien étudiée. Il est donc important de faire une description des déterminants de leur transmission, même si pour la grande majorité d'entre eux les données structurales sont inexistantes. Le genre *Potyvirus*, composé de virus flexueux à ARN simple brin, est transmis par pucerons selon un mode non persistant. L'implication dans la transmission d'une protéine virale assistant la transmission a été mise en évidence par l'étude d'un isolat non transmissible du *Tobacco etch virus* (TEV) en 1975 (Simons, 1976). Depuis, il a été établi que cette protéine est le facteur HC-Pro, une protéine multifonctionnelle impliquée à la fois dans le mouvement, le clivage des polyprotéines virales et les mécanismes de défense (suppression de silencing) (Syller, 2006). L'implication de la protéine de capsid et de la protéine HC-Pro dans la transmission a été démontrée pour de nombreux *Potyvirus*, incluant le *Potato virus Y* (PVY), *Tobacco etch virus* (TEV), *Tobacco vein mottling virus* (TVMV), *Zucchini yellow mosaic virus* (ZYMV), *Turnip mosaic virus* (TuMV), ou encore le *Kalanchoe mosaic potyvirus* (KMV) (Atreya *et al.*, 1991; Canto *et al.*, 1995; Husted, 1995; Lopez-Moya *et al.*, 1999; Peng *et al.*, 1998; Wang *et al.*, 1998). Les motifs protéiques essentiels à la transmission ont été identifiés grâce à l'analyse de séquences conservées parmi des isolats transmissibles et par génétique inverse. Trois motifs protéiques sont déterminants:

- les résidus **DAG** de la partie N-terminale de la CP sont impliqués dans l'interaction avec HC-Pro (Blanc *et al.*, 1997, interaction démontrée pour le TVMV).

- les résidus **KITC** proche de la partie N-terminale de HC-pro sont impliqués dans l'interaction avec le site de rétention du vecteur (Blanc *et al.*, 1998, interaction démontrée pour le TEV).

- les résidus **PTK** de la partie centrale de HC-pro sont impliqués dans l'interaction avec la CP (Peng *et al.*, 1998, interaction démontrée pour le ZYMV).

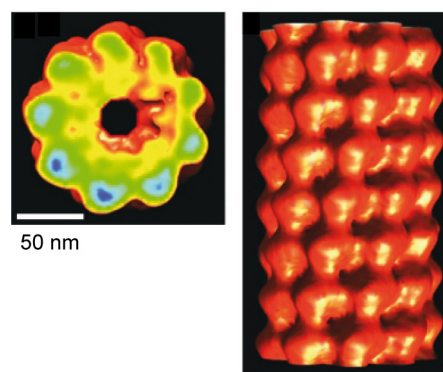
Les mutations ponctuelles effectuées dans chacun de ces motifs, ainsi que sur les résidus situés en aval du motif DAG, affectent l'efficacité de transmission par pucerons (Atreya *et al.*, 1990; Atreya *et al.*, 1991; Atreya *et al.*, 1995; Huet *et al.*, 1994; Peng *et al.*, 1998). Parmi les nombreuses mutations effectuées, on ne retrouve pas de corrélation entre la nature biochimique des acides aminés substitués et l'effet sur la transmission. Ainsi, même si, dans la majorité des cas le changement de charge dans le motif DAG provoque une réduction drastique de l'efficacité de transmission, les contre-exemples existent : par exemple le remplacement du motif DAG, soit en EAG, soit en DGG (même charge latérale) entraîne la perte de transmission par puceron du TVMV (Atreya *et al.*, 1995).

Pour ce genre viral, seule la structure du *Soybean mosaic virus* (SMV) a pu être obtenue à basse résolution (22 Å), permettant ainsi de déterminer son architecture hélicoïdale (Kendall *et al.*, 2008) (fig. 1.10A). Les analyses d'image de cryo-EM, effectuées sur HC-Pro du LMV (Plisson *et al.*, 2003) et du TEV (Ruiz-Ferrer *et al.*, 2005) montrent que la protéine possède une forme allongée et se dimérise selon une orientation parallèle pour le LMV et antiparallèle dans le cas du TEV (fig. 1.10 B,

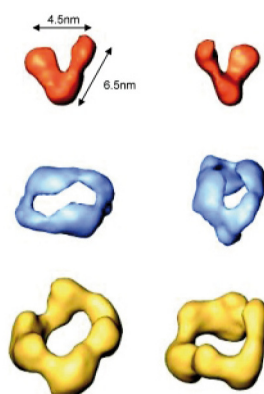
C). L'étude structurale de la protéine HC-Pro du LMV suggère que la partie N-terminale et la partie centrale ne sont pas responsables de la dimérisation. Dans le cas du TEV, des formes tétramériques et hexamériques ont également pu être mises en évidence et leurs structures reconstruites respectivement à 30 et 36 Å.

Fig. 1.10. Potyvirus : Structure de capsid et protéine assistante HC-Pro. (A) Reconstruction 3D du SMV à 22 Å soit en coupe transversale (gauche) soit en vue de surface (droite). (B) : deux vues des reconstructions 3D de la protéine HC-Pro du TEV en état dimérique, tétramérique ou hexamérique (du haut vers le bas). (C) : projection de la carte de densité de la protéine HC-Pro du LMV organisée selon un réseau de symétrie $p4$. La densité a été calculée à 9 Å. Le carré noir représente la maille élémentaire du réseau ($a=103$ Å, $b=103$ Å). Le monomère HC-Pro est entouré par une ligne noire. Figures adaptées d'après Kendall et al., 2008, Plisson et al., 2003 et Ruiz-Ferrer et al., 2005.

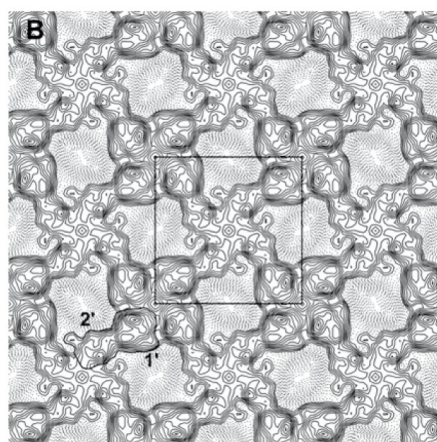
(A) Reconstruction 3D de la capsid du SMV



(B) Reconstruction 3D de HC-Pro du TEV



(C) Projection 2D de HC-Pro du LMV



De manière globale, les structures 3D des protéines virales participant à la transmission des *Potyvirus* n'atteignent pas une résolution suffisante afin de localiser précisément les déterminants de la transmission, mais ces études structurales prouvent l'existence de différents états oligomériques des protéines HC-Pro, qui pourrait expliquer leur rôle multifonctionnel.

A l'opposé de la vaste majorité des phytovirus, il existe des espèces virales dont la structure à haute résolution de la capsid e a été obtenue, mais pour lesquelles on ne connaît pas de vecteurs (Tableau 1.2). Il s'agit du *Tobacco mosaic virus* (TMV) et *Cucumber green mottle mosaic virus*, (deux *Tobamovirus*), le *Tomato bushy stunt virus* (TBSV) (*Tombusvirus*), le *Carnation mottle virus* (*Carmovirus*), le *Red clover mottle virus* (*Comovirus*), le *Brome mosaic virus* (*Bromovirus*), le *Potato virus X* (*Potexvirus*), le *Physalis mottle virus*, le *Desmodium yellow mottle virus* (deux *Tymovirus*) et le *Ryegrass mottle virus* (*Sobemovirus*). Le cas du TMV est remarquable puisque, en plus de son statut de modèle dans l'histoire de la virologie – première découverte par Beijerinck en 1892 d'un agent infectieux capable de passer à travers des filtres de porcelaine à bactéries, puis dans les années 30 par la démonstration par Bawden et Pirie que le virus est un complexe nucléo-protéique - il est l'un des rares phytovirus dont la transmission horizontale est exclusivement mécanique et transmis verticalement occasionnellement par la semence. Le TBSV a également une importance historique, du fait qu'il soit le premier virus parasphérique à avoir été cristallisé dans un réseau 3D (Bawden and Pirie, 1938) et le premier dont la structure atomique a été résolue après diffraction au rayon X à une résolution de 2,9 Å (Harrison, 1972; Harrison and Jack, 1975; Harrison *et al.*, 1978). Aucun vecteur ne lui est assigné, malgré plusieurs tentatives de transmission par champignon puisqu'un membre du genre *Tombusvirus*, le *Cucumber necrosis virus* (CNV) est transmis par *Olpidium*.

Concernant les 23 autres phytovirus dont la structure de capsid e est connue, pour beaucoup d'entre eux le vecteur a été identifié, mais il y a peu d'informations sur les déterminants de la v ection. C'est le cas des virus du genre *Comovirus*, avec le *Cowpea mosaic virus* (CPMV), *Cowpea severe mosaic virus* et le *Bean pod mottle virus* dont seule l'espèce vectrice a été identifiée (transmis par coccinelles). Dans le cas du *Cowpea chlorotic virus* (*Bromovirus*), il a récemment été démontré que la CP détermine sa spécificité de transmission par les coccinelles (Mello *et al.*, 2009). Étant donné que les déterminants de la transmission du RDV ont été mentionnés dans la partie précédente, nous pouvons encore décrire le cas du *Melon necrotic spot virus* (MNSV, genre *Carmovirus*) et du *Tobacco necrosis virus* (TNV, genre *Necrovirus*) qui sont transmis respectivement par *Olpidium bornovanus* et *O. brassicae* (Rochon *et al.*, 2004). Ils possèdent une capsid e icosaédrique $T = 3$ de 35 nm et 32 nm de diamètre, respectivement. Leur v ection est très spécifique, ce qui est en accord avec les observations faites pour d'autres virus transmis par des espèces appartenant au genre *Olpidium* (Rochon *et al.*,

2004). La rétention du TNV à la surface des zoospores est spécifique de son espèce vectrice (Temmink *et al.*, 1970).

Pour le *Cauliflower mosaic virus* (CaMV), le *Cucumber necrosis virus* (CNV) et le *Cucumber mosaic virus* (CMV), les déterminants viraux de la vection sont bien documentés et des données structurales ont été acquises. Je propose de décrire les déterminants de la vection de ces trois virus dans les paragraphes suivants. Pour éviter toute redondance dans le manuscrit, les données sur le *Tobacco ringspot virus* (TRSV) seront également détaillées plus loin, dans la partie consacrée aux déterminants de la vection des népovirus.

3. 4. 1 Déterminants de la vexion du *Cucumber necrosis virus* (CNV)

Le CNV (famille *Tombusviridae*, genre *Tombusvirus*) est un virus icosaédrique non enveloppé, à ARN simple brin monopartite de polarité positive. Il est transmis par *Olpidium bornovanus* selon un mode non circulant (Dias, 1970a, b). Les virions sont adsorbés à la surface de la zoospore du champignon, et sont inoculés dans la cellule végétale lors de l'internalisation de la zoospore dans la racine de la plante. Des expériences de génétique inverse échangeant la protéine de capsid du CNV avec celle du TBSV, qui n'est pas transmis par *Olpidium*, ont démontré que seule la CP détermine la spécificité de transmission du CNV par son vecteur (McLean *et al.*, 1994).

Le virus possède une structure icosaédrique $T = 3$, un diamètre d'environ 30 nm, et est composé de 180 sous-unités. Sa structure a été obtenue par analyse d'image de cryo-ME et reconstruction 3D à une résolution de 10 Å (fig. 1.11). (Katpally *et al.*, 2007). Un modèle de la CP a été construit par homologie avec la structure cristallographique du TBSV, le membre type du genre *Tombusvirus* (Harrison *et al.*, 1978; Kakani *et al.*, 2001). Chaque protéine de capsid se replie en 3 domaines distincts : un domaine flexible *R* (RNA) qui interagit avec l'ARN, un domaine *S* (*Shell*) et le domaine *P* (*Protruding*) qui interagit en dimère pour former les protubérances de surface localisées sur les axes d'ordre 2 et les quasi-axes d'ordre 2 de la particule virale. Les domaines *R* et *S* sont reliés par un domaine *a* (*arm*) et les domaines *S* et *P* par un domaine charnière *h* (*hinge*) (Fig. 1.11).

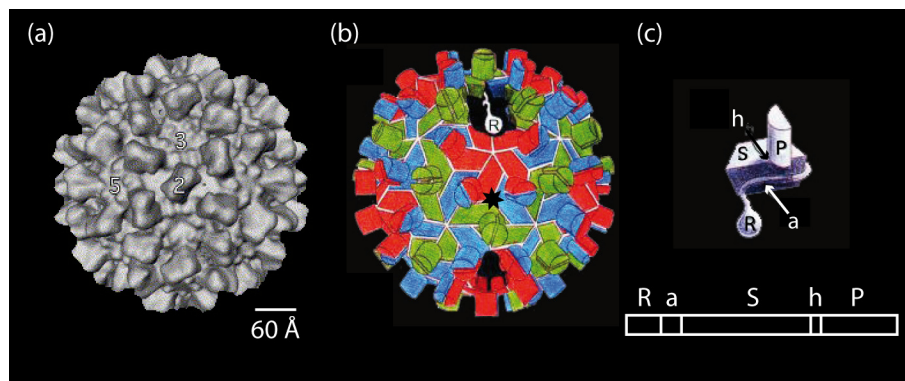


Fig. 1.11. Structure du CNV. (a) Structure cryo-ME à 10 Å vue depuis un axe 2 du virion d'après Katpally *et al.* 2007. (b) Schématisation de la capsid, vue selon un axe différent de (a). Les sous-unités A, B et C sont coloriées respectivement en rouge, bleu et vert. La section en haut de la capsid montre que le domaine R de la CP est à l'intérieur de la particule. L'étoile noire indique le quasi-axe 3. (c) Représentation de la CP en volume (en haut), ou linéaire (en bas). La région R correspond au RNA binding domain, a, arm ; S, Shell ; h, hinge ; P, Protruding domain. (b) et (c) d'après Kakani *et al.*, 2001.

Sur le modèle créé par homologie avec la structure du TRSV, on retrouve à l'intérieur de la capsid une structure appelée anneau β (Kakani *et al.*, 2001) - dont des équivalents structuraux sont

retrouvés pour d'autres virus (Willits *et al.*, 2003) - dans lequel, les domaines *arm* de 3 sous-unités C interagissent en forme d'anneau autour de l'axe d'ordre 3 de la particule virale.

Cette structure quaternaire en anneau est importante pour l'infectivité du virus, sa stabilité et sa transmission. En effet, une seule mutation dans ce domaine (substitution de la Pro⁸⁵ par Gly⁸⁵), influence le changement de conformation du virus en $T = 1$ (Katpally *et al.*, 2007). La délétion des 16 derniers résidus du domaine *arm* qui, d'après le modèle 3D, sont les résidus directement impliqués dans l'anneau β , entraîne la perte d'infectivité du virus. Si les 18 premiers résidus du domaine sont supprimés, l'infectivité du virus est diminuée, et la transmission par champignon est impossible (Hui and Rochon, 2006).

Ce qui est particulièrement intéressant dans les études menées sur la transmission du CNV, c'est qu'un changement de conformation de la particule provoqué par l'accrochage au vecteur est nécessaire à la transmission (Kakani *et al.*, 2004). En effet, il a été remarqué que les particules de CNV ont une conformation plus large, dite gonflée (*swollen form*) quand elles sont liées aux zoospores. Ce relâchement s'effectue au niveau de l'axe d'ordre 3 et rend 2 sites tryptiques du domaine *arm* accessibles à la protéolyse une fois la capsidie attachée au vecteur. Or, un mutant non transmis possédant une mutation ponctuelle dans le domaine *arm* est capable de se lier au vecteur efficacement, mais ne peut plus opérer le changement de conformation (résistant à la digestion tryptique), démontrant que le changement de conformation est requis pour assurer la transmission.

Plusieurs variants naturels du CNV, faiblement transmis par *O. bornovanus*, ont été identifiés (Kakani *et al.*, 2001; Robbins *et al.*, 1997) et étudiés. Dans le cas de ces variants, la réduction de la transmission n'implique pas forcément des mutations de résidus du domaine *arm* mais majoritairement des mutations ponctuelles localisées dans la cavité proche du pseudo-axe 3 du virion (étoile noire sur la figure. 1.11) (Kakani *et al.*, 2001).

Des tests d'interaction *in vitro* entre les particules virales purifiées et les zoospores démontrent que les mutations empêchent la rétention du CNV sur son vecteur (Kakani *et al.*, 2001). Ces expériences confirment que la rétention du virus est une étape cruciale dans la transmission. Une étude de rétention *in vitro* entre le CNV et les zoospores montre qu'un nombre défini de sites de rétention du virus existent ($\sim 2,7 \times 10^4$ sites) (Kakani *et al.*, 2003). La rétention des virus sur le vecteur est altérée en faisant un traitement des zoospores au périodate et à la trypsine et en incubant des zoospores avec un compétiteur de type sucre avant l'interaction *in vitro*. (Kakani *et al.*, 2003). Ces résultats soulignent l'existence de récepteurs spécifique du virus, et suggèrent que ces récepteurs putatifs contiennent des oligosaccharides.

3. 4. 2 Déterminants de la vection du *Cucumber mosaic virus* (CMV)

Le CMV (famille *Bromoviridae*, genre *Cucumovirus*) est un virus icosaédrique non enveloppé, à ARN simple brin tripartite de polarité positive. Le CMV possède une gamme d'hôtes et de vecteurs parmi les plus larges pour les phytovirus (Ali and Kobayashi, 2010). Il est transmis selon un mode non persistant, par plus de 86 espèces de pucerons (Edwardson and Christie, 1991). La recherche des déterminants viraux de la transmission du CMV bénéficie de 2 atouts majeurs. D'une part, la capsidie du CMV peut être assemblée *in vitro* (Gera *et al.*, 1979). D'autre part, comme pour tous les virus transmis par pucerons, l'alimentation du vecteur peut se faire à travers une membrane artificielle, ce qui permet de faire des tests utilisant des protéines purifiées et quantifiées (Sylvester, 1988). Une étude comparant la transmissibilité de virions produits soit par assemblage *in vitro* de CMV sauvage, soit par assemblage *in vitro* d'un isolat faiblement transmis par puceron, a initialement montré l'importance de la CP dans la transmission (Gera *et al.*, 1979). Ce résultat a été confirmé quelques années plus tard par une approche similaire à partir d'autres isolats de *Cucumovirus* (Chen and Francki, 1990).

Le virus possède une structure icosaédrique $T = 3$ d'environ 30 nm de diamètre, composée de 180 sous-unités. Sa structure cristalline a tout d'abord été obtenue à une résolution de 8 Å, en combinant cristallographie aux rayons X et une reconstruction 3D à partir d'images de cryo-ME à 28 Å (Wikoff *et al.*, 1997), puis a été améliorée jusqu'à une résolution de 3,2 Å par un autre groupe de recherche (Smith *et al.*, 2000). La morphologie de la capsidie est différente de celle du CNV. Ici, des unités pentamériques et hexamériques sont protubérantes et les dépressions sont présentes au niveau des axes d'ordre 2 et des pseudo-axes 2 (fig. 1.12). De plus, un espace béant est présent au niveau de l'axe 3. La CP a une structure en tonneau β dont la conformation est différente suivant sa position dans la capsidie (Smith *et al.*, 2000).

L'identification de plusieurs variants naturels faiblement transmis a mis en évidence l'implication des acides aminés 129, 162 et 168 dans l'efficacité de vection. De manière intéressante, les mutations ponctuelles effectuées à ces positions de la séquence codante de la CP réduisent la transmission du CMV par *Myzus persicae* sans affecter celle par un autre vecteur *Aphis gossypii* (Perry *et al.*, 1998; Perry *et al.*, 1994). Ces résidus ne sont pas impliqués directement dans l'interaction avec le vecteur puisqu'ils ne sont pas en surface de la particule virale, mais réduisent la stabilité de la capsidie (Ng *et al.*, 2005; Ng *et al.*, 2000).

L'analyse de la structure cristalline du CMV a révélé l'existence d'une boucle de surface chargée négativement, fortement conservée parmi différents isolats de CMV et d'autres *Cucumovirus* (Smith *et al.*, 2000). Cette boucle β H- β I de 8 résidus est localisée au niveau de l'axe 3 (quasi-axe 6) du virus, dans ces protubérances hexamériques si particulières (fig. 1.12). Le repliement de la boucle et les distances entre les chaînes latérales suggèrent la présence d'un cation dans la boucle. Des

mutations ponctuelles effectuées sur les résidus chargés de la boucle n'affectent pas l'infectivité, ni la stabilité des virions, mais réduisent considérablement la transmission par puceron (Liu *et al.*, 2002).

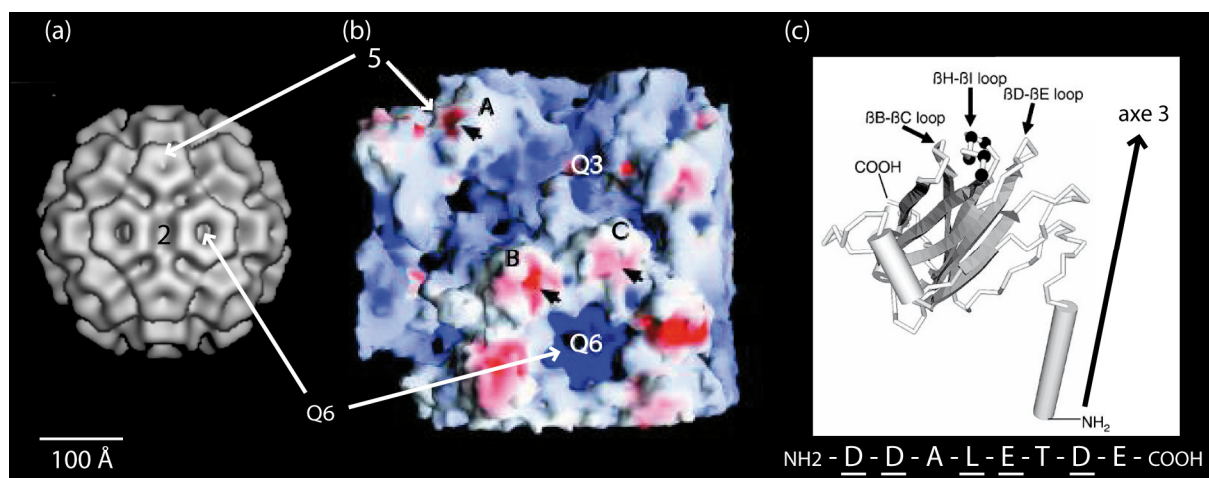


Fig. 1.12. Structure du CMV et de ses déterminants viraux de la transmission. (a) Structure cryo-ME à 28 Å vue depuis un axe 2 du virion d'après Wikoff et al., 1997. (b) Charges électrostatiques présentes à la surface d'un pentamère avec en bleu les charges positives et en rouges les charges négatives. La position de la boucle βH-βI déterminant la transmission du virus, est indiquée par une flèche noire pour une sous-unité A, B et C (figure adaptée d'après Smith et al., 2000). (c) Représentation « cartoon » de la structure cristallographique de la CP du CMV, orientée approximativement d'après l'axe 3 (la flèche pointe vers l'extérieur de la particule virale). La position des résidus chargés de la boucle est représentée par des sphères noires et est soulignée dans la séquence primaire de la boucle du CMV sauvage (figure adaptée d'après Liu et al., 2002).

3. 4. 3 Déterminants de la vection du *Cauliflower mosaic virus* (CaMV)

Le CaMV (famille *Caulimoviridae*, genre *Caulimovirus*) est un virus à ADN double brin circulaire. Sa capsid est icosaédrique de symétrie $T = 7$, non enveloppée, composée d'une seule protéine structurale et mesure environ 50 nm de diamètre. Le CaMV est transmis par plus d'une vingtaine d'espèce de pucerons, selon un mode semi persistant, non circulant. La particule virale n'est pas suffisante pour la transmission. Ici, deux autres protéines virales non structurales assistent la vection : la P3 qui interagit avec la capsid (Leh *et al.*, 2001) et la P2 qui interagit à la fois avec le complexe P3-virions (Leh *et al.*, 2001) et avec le vecteur (Uzest *et al.*, 2007). Plusieurs études complémentaires menées sur ces déterminants montrent que l'acquisition des protéines virales se fait de manière séquentielle et implique une dynamique moléculaire complexe. Selon le modèle de formation du complexe transmissible du CaMV, la P2 va d'abord s'accrocher sur le site récepteur du vecteur, puis le complexe virion-P3 est recruté sur le site récepteur (Drucker *et al.*, 2002).

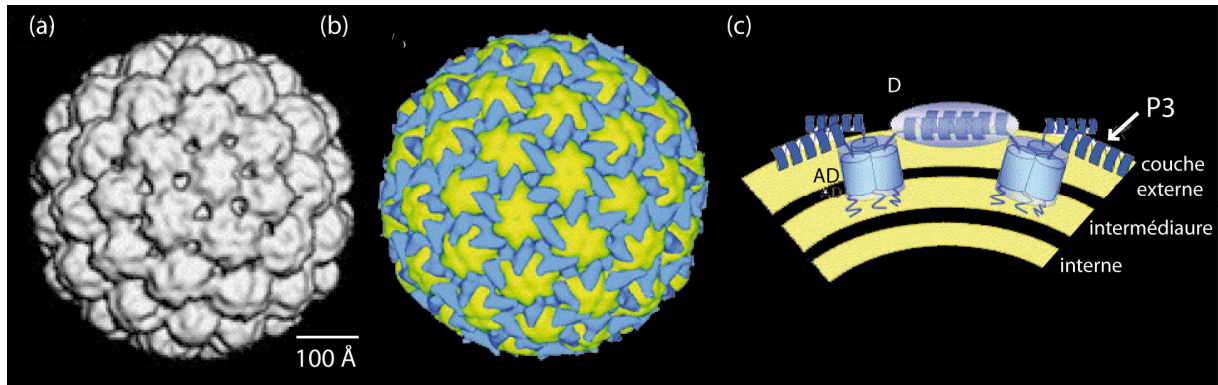


Fig. 1.13. Structure du CaMV et son interaction avec P3. (a) Structure cryo-ME à 18 Å d'après un axe d'ordre 2 du virion. (b) Structure cryo-ME du complexe CaMV-P3 avec la densité électronique correspondant au virion en jaune, et celle de la protéine P3 en bleu. (c) Modèle d'interaction entre le virus et P3, dans lequel le domaine N-terminal de P3 serait sous forme trimérique dans la particule virale (AD= *anchoring domain*), et le domaine C-terminale formerait des dimères à l'extérieur de la particule virale (D= *digitation*). La partie C-terminale de P3 pénètre dans la couche intermédiaire de la capsid. Figures adaptées d'après Plisson *et al.*, 1997.

La structure du complexe CaMV-P3 a été résolue à 26 Å par analyse d'image de cryo-ME et reconstruction 3D, et comparée à celle du CaMV, obtenue à 18 Å (Plisson *et al.*, 2005). D'après ces analyses, les protéines P3 se fixeraient sous une forme trimérique dans les canyons présents à la surface de la capsid (fig. 1.13). La partie C-terminale de la P3, à l'origine de la trimerisation, s'ancrerait dans la coque protéique, alors que la partie N-terminale serait présente du côté externe, interagissant en dimère avec d'autres P3. Des tests de far-western ont montré que la partie N-terminale de la P3 est responsable de l'interaction avec la P2 (Leh *et al.*, 1999). D'après des prédictions de structures secondaires, la région N-terminale de la P3 et la région C-terminale de la P2 adopteraient une conformation secondaire en hélice α , ce qui permettrait leur interaction via une structure *coiled-coil* (Hebrard *et al.*, 2001; Leh *et al.*, 1999). L'obtention récente de la structure cristallographique de la P3 à 2,5 Å confirme le repliement de la protéine en plusieurs hélices α (Hoh *et al.*, 2010). La partie N-terminale de la P2 est, quant à elle, impliquée dans l'interaction avec le vecteur. En effet, le changement d'un seul résidu, l'aa en position 6 de la protéine, entraîne une réduction de la transmission par différentes espèces de pucerons, et est capable de modifier la spécificité de vection (Moreno *et al.*, 2005). De plus, la réduction de la transmission est corrélée avec une non-rétention de la P2 mutée sur le stylet du vecteur (Uzest *et al.*, 2007).

C'est par une étude d'interaction entre des stylets disséqués de plusieurs espèces de pucerons vecteurs du CaMV et des protéines P2 fusionnées à la *Green Fluorescent Protein* (GFP), qu'il a été démontré qu'un récepteur de type protéique est responsable de l'adsorption du CaMV dans son vecteur (Uzest *et al.*, 2007). Ce résultat est la première démonstration de l'existence d'un récepteur de type protéique pour la vection des phytovirus. Ces récepteurs sont localisés au bout des stylets

maxillaires de pucerons vecteurs comme *Aphis pisum*, *Myzus persicae*, et *Brevicoryne brassicae*, dans un organe spécialisé dénommé l'acrostyle (Uzest *et al.*, 2010).

3. 5 La vection des *Nepovirus*

3. 5. 1 Caractéristiques générales de la vection

Les népovirus sont des virus non enveloppés icosaédriques, d'un diamètre d'environ 30 nm. Le virion est composé d'une seule protéine de capsid (CP) contenant 3 domaines structuraux, C, B, et A dont la symétrie pseudo $T = 3$ a été décrite dans la partie 2.3 précédente. Ils possèdent un génome bipartite à ARN simple brin de polarité positive et sont transmis selon un mode semi-persistant. Le terme *Nepovirus* désigne les virus icosaédriques transmis par nématode, puisque **nepovirus** provient de l'abréviation de « *nematode-transmitted virus with polyhedral particles* » (Cadman, 1963). Cependant, seul un tiers des virus regroupés dans ce genre sont effectivement transmis par nématode (12 espèces sur 32) (Tableau 1.3). Les autres espèces étant soit transmises uniquement par pollen (3 espèces) soit n'ont pas de vecteurs identifiés. Le *Blackcurrant reversion virus* (BRV) est le seul népovirus transmis par un arthropode : *Cecidophyopsis ribis*. La découverte initiale de la transmission du GFLV par *X. index* a engendré une recherche intensive pour la mise en évidence d'autres associations nématodes/népovirus. Trois genres de nématodes—*Xiphinema*, *Longidorus* et *Paralongidorus* (famille *Longidoridae*, Ordre des *Dorylaimidae*) sont responsables de la transmission naturelle de népovirus. Ces trois genres de nématodes regroupent plus de 400 espèces différentes mais, à ce jour, seulement 18 espèces (huit espèces de *Longidorus*, une espèce de *Paralongidorus* et neuf espèces de *Xiphinema*) ont été démontrées comme étant vecteurs de népovirus (Tableau 1.3). Ce tableau illustre bien que la gamme d'espèce vectrice des népovirus est restreinte (Brown *et al.*, 1995; Harrison *et al.*, 1974b).

En effet, beaucoup de nématodes ne transmettent que 2 ou 3 espèces virales différentes. C'est le cas pour *X. diversicaudatum* qui transmet l'ArMV et le SLRV (virus anciennement classé dans le genre *Nepovirus*, mais réassigné au genre *Sadwavirus*, famille *Secoviridae* (Sanfaçon *et al.*, 2009). Le niveau de spécificité peut même aller jusqu'au niveau de l'isolat : les isolats écossais et anglais du *Raspberry ringspot virus* (RpRSV) sont transmis spécifiquement par *Longidorus elongatus* et *L. macrosoma*, respectivement (Murant *et al.*, 1996).

Ainsi, la transmission des *Nepovirus* par nématodes selon le mode semi persistant non circulant est très spécifique. Cette particularité reflète une interaction de haute affinité entre les virus et les sites de rétention au niveau de l'appareil alimentaire du vecteur. La deuxième singularité de cette association nepovirus/nématodes est le temps de rétention extrêmement long du virus chez les

Xiphinema. Ces derniers sont capables de rester virulifères pendant de longues périodes : neuf mois pour *X. americanum* (McGuire, 1973), deux ans pour *X. rivesi* (Bitterlin and D., 1987), et probablement au moins quatre ans pour *X. index* (Demangeat *et al.*, 2005b). Malgré la grande période de rétention observée, qui pourrait suggérer un mode de transmission circulant, il n'y a jamais eu aucune évidence de multiplication des virus dans les nématodes. De plus, le caractère virulifère du nématode est perdu après la mue et il n'y a pas de passage transovarien (Taylor and Brown, 1997).

3. 5. 2 Déterminants viraux de la transmission

Les données concernant les déterminants de la transmission sont encore limitées. L'analyse de la transmissibilité de pseudo-recombinants produit à partir d'échanges génomique de deux souches du RpRSV transmis par deux nématodes vecteurs différents, a démontré que les déterminants de la transmission sont codés par l'ARN2, celui contenant le gène codant pour la CP (Harrison *et al.*, 1974a). Le rôle essentiel des particules virales dans la transmission des népovirus fut mise en évidence par une étude avec le TBRV, associant une étude de transmissibilité de pseudo recombinants et une étude sérologique (Harrison and Murant, 1977). La protéine de mouvement (la protéine 2B) codée par le gène situé en amont du gène codant pour la CP, a également été suspectée pour jouer un rôle dans la transmission, en tant que protéine assistante (Mayo and Robinson, 1996), mais cette hypothèse a été réfutée par la suite. Tout d'abord, l'analyse de séquences codant pour la protéine 2B et la CP de 6 isolats de RpRSV - isolats écossais et anglais sérologiquement différents et transmis par des nématodes différents - révèle que les différences significatives ne se localisent pas sur la protéine 2B, mais sont retrouvées au niveau de la CP (Scott *et al.*, 2000). Certaines de ces différences correspondent à des positions potentiellement à la surface de la particule virale.

Par la suite, la preuve moléculaire que seule la CP est responsable de la spécificité de transmission a été produite par des études sur le GFLV. Ici, des expériences échangeant séquentiellement les gènes de l'ARN2 du GFLV par leurs équivalents ArMV, virus très proche du GFLV qui n'est pas transmis par *X. index* mais par *X. diversicaudatum*, montrent que seuls les recombinants possédant un gène codant pour une CP de nature GFLV sont transmis par *X. index* (Andret-Link *et al.*, 2004b; Belin *et al.*, 2001). Cette étude a également permis d'exclure les deux autres protéines codées par l'ARN2 de la spécificité de transmission

Tableau 1.3. Association spécifique entre népovirus et nématodes vecteurs de la famille des *Longidoridae*
(d'après Demangeat, 2007)

Genre	Espèce	Nepovirus	Acronyme
<i>Longidorus</i>	<i>apulus</i>	Artichoke italian latent virus (isolat italien)	AILV
	<i>arthensis</i>	Cherry rosette virus	CRV
	<i>attenuatus</i>	Tomato black ring virus (isolat allemand/anglais)	TBRV
	<i>diadecturus</i>	Peach rosette mosaic virus	PRMV
	<i>elongatus</i>	Raspberry ringspot virus (isolat écossais)	RpRSV
		Tomato black ring virus (isolat écossais)	TBRV
	<i>fasciatus</i>	Artichoke italian latent virus (isolat grecque)	AILV
	<i>macrosoma</i>	Raspberry ringspot virus (isolat anglais)	RpRSV
	<i>martini</i>	Mulberry ringspot virus	MLRV
	<i>Paralongidorus</i>	<i>maximus</i>	Raspberry ringspot virus (isolat allemand)
<i>americanum sensu lato</i>		Cherry rasp leaf virus	CRLV
		Peach rosette mosaic virus	PRMV
		Tobacco ringspot virus	TRSV
		Tomato ringspot virus	ToRSV
<i>Xiphinema</i>	<i>americanum sensu stricto</i>	Cherry rasp leaf virus	CRLV
		Tobacco ringspot virus	TRSV
		Tomato ringspot virus	ToRSV
	<i>bricolensis</i>	Tomato ringspot virus	ToRSV
		Cherry rasp leaf virus	CRLV
	<i>californicum</i>	Tobacco ringspot virus	TRSV
		Tomato ringspot virus	ToRSV
	<i>diversicaudatum</i>	Arabidopsis mosaic virus	ArMV
	<i>index</i>	Strawberry latent ringspot virus *	SLRV
		Grapevine fanleaf virus	GFLV
	<i>intermedium</i>	Tobacco ringspot virus	TRSV
		Tomato ringspot virus	ToRSV
	<i>rivesi</i>	Cherry rasp leaf virus	CRLV
		Tobacco ringspot virus	TRSV
		Tomato ringspot virus	ToRSV
	<i>tarjanense</i>	Tobacco ringspot virus	TRSV
		Tomato ringspot virus	ToRSV

* SLRV est assigné au genre *Sadwavirus*

Une seule structure cristalline de virus transmis par nématode était disponible au début du travail de thèse, celle du *Tobacco ringspot virus* (TRSV) (Chandrasekar and Johnson, 1998). C'est le membre type du genre *Nepovirus*. Il est transmis par des nématodes du genre *X. americanum* et des espèces étroitement liées (voir Tableau 1.3). L'observation de coupes de *X. americanum* après marquages fluorescents des particules virales et l'observation de coupes ultrafines en microscopie électronique, ont montré que la présence de virions au niveau des sites de rétention du virus (lumen de l'odontophore et œsophage des nématodes) est corrélée à une transmission efficace par le vecteur (Wang and Gererich, 1998; Wang *et al.*, 2002).

L'obtention de sa structure cristalline à 3,5 Å a été utilisée pour identifier les acides aminés en surface de la particule virale pour 11 népovirus. L'alignement de la séquence primaire de la CP des 11 népovirus juxtaposé à la position des acides aminés de la CP du TRSV indique que les résidus en surface sont très variables. Seulement deux régions de surface (sites 1 et 2) sont conservées parmi au moins 8 *Nepovirus* (fig. 1.14) (Chandrasekar and Johnson, 1998). Le site 1 se localise entre le pseudo-axe 2 et le pseudo-axe 3 de la capsid. Il inclut les résidus de la boucle $\beta E-\beta F$ du domaine C ; la boucle $\beta C-\beta D$ du domaine A et la boucle $\beta G-\beta H$ des trois domaines. Le site 2 inclut les résidus de la

boucle $\beta\text{B}-\beta\text{C}$ et $\beta\text{H}-\beta\text{I}$ du domaine C et B, se joignant au niveau de l'axe d'ordre 3 (fig. 1.14). Cependant, en l'absence de conservation de ces séquences entre les népovirus transmis par la même espèce vectrice, la signification biologique de ces deux sites conservés n'est pas du tout établie.

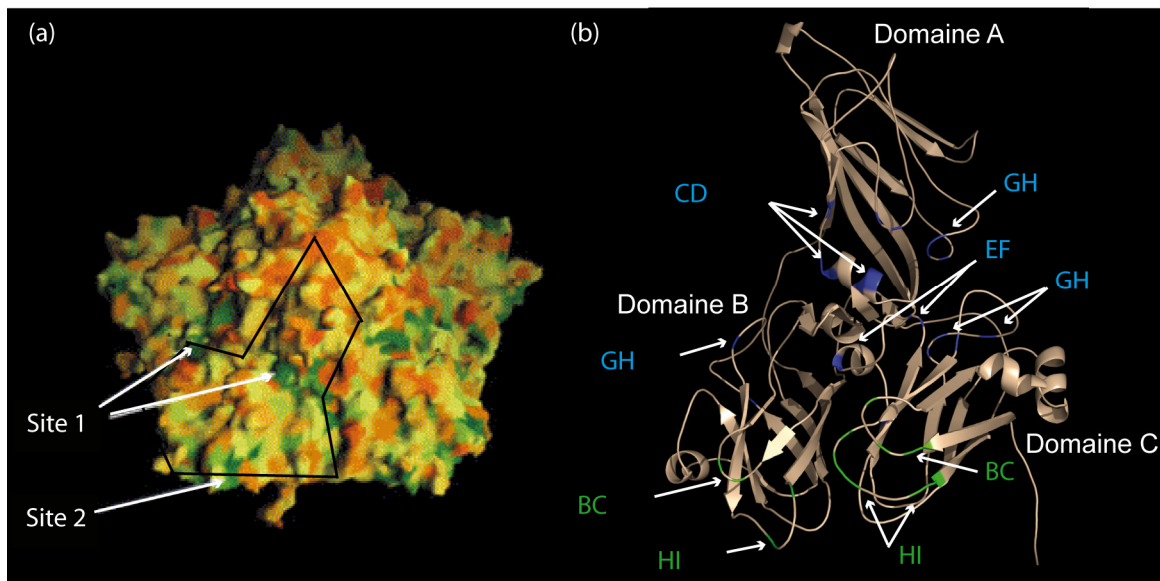


Fig. 1.14. Localisation sur la structure 3D du TRSV des régions de surface (sites 1 et 2) conservés chez les népovirus. (a) Vue externe de la surface d'un pentamère de sous-unités de protéine de capsid (CP). La variabilité des résidus de surface présents chez les népovirus sont colorés en fonction de leur conservation en allant du rouge (pas de conservation) au jaune (conservation entre 6 népovirus), jusqu'au vert (conservation entre 11 népovirus). Les deux régions site 1 et site 2 sont des régions conservées parmi au moins 8 népovirus. La position d'une CP est indiquée par un tracé noir. (Figure adaptée d'après Chandrasekhar *et al.*, 1998). (b) Représentation « cartoon » de la CP orientée selon le côté externe de la particule virale. Elle est composée de 3 domaines *jelly roll*, C, B et A. Les sites 1 et 2 sont coloriés respectivement en vert et bleu.

Enfin, même si cette partie est consacrée à la transmission des *Nepovirus*, il est important de noter que d'autres nématodes appartenant à l'ordre des *Dorylaimidae*, mais dans la famille des *Trichodoridae* (genres *Trichodorus* et *Paratrachodorus*) transmettent des virus à particules hélicoïdales, les *Tobraviridae*. Dans leur cas, la protéine 2B, dont la fonction putative est d'assurer le mouvement du virus, codée par le gène en amont du gène codant pour la CP assure la fonction de protéine assistant la transmission (Hernandez *et al.*, 1997; MacFarlane *et al.*, 1996; Vassilakos *et al.*, 2001). Il a également été démontré que les 15 acides aminés positionnés en C-terminal de la CP sont requis pour la transmission (MacFarlane *et al.*, 1996)

3. 5. 3 Caractéristiques de rétention des virus

Les nématodes du genre *Xiphinema* et *Londidorus* sont des vers filiformes ectoparasites, au corps non segmenté de 2 à 12 millimètres de long (Brown and MacFarlane, 2001), dont la morphologie de l'appareil alimentaire est semblable (fig. 1.15). Il comporte un odontostyle, aiguille

creuse aux bords non soudés de 60 à 250 μm de long qui permet au nématode de se nourrir en pénétrant la paroi végétale pour atteindre le cytoplasme des cellules végétales. Suit l'odontophore, deuxième pièce creuse connectant l'odontostyle à l'œsophage. Des muscles rattachés à la base et aux cotés de l'odontophore, permettent de rétracter ou protracter l'odontostyle du corps du nématode. L'odontophore est prolongé par l'œsophage qui est un tube flexible et étroit. L'œsophage se termine par le bulbe œsophagien. Ce bulbe assure le rôle de pompe refulante et aspirante. Il est associé à des glandes salivaires qui déversent leurs productions dans la lumière du bulbe œsophagien. Il contrôle par sa rétraction ou dilatation, respectivement l'injection de la salive dans les cellules végétales ou la ponction du cytoplasme digéré de la cellule. Enfin, l'intestin est la partie la plus longue de l'appareil alimentaire, il s'étend jusqu'au scrotum.

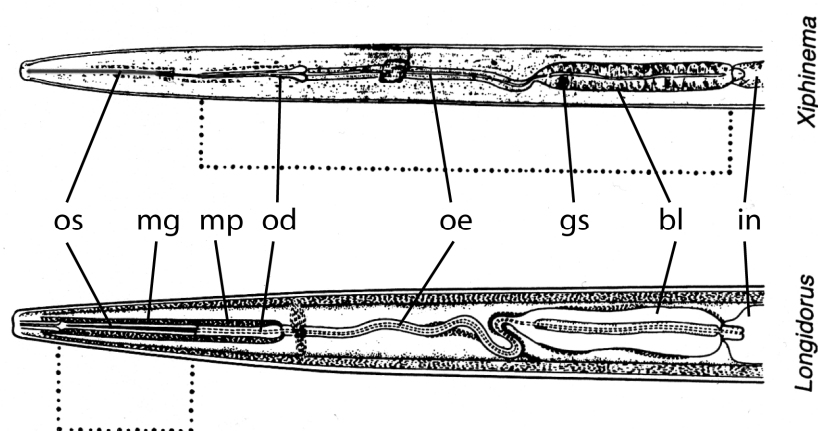


Fig. 1.15. Représentation schématique de la partie antérieure de l'appareil alimentaire des *Longidorus* (en bas) et des *Xiphinema* (en haut). Le site de rétention des népovirus dans le tractus de l'appareil alimentaire est indiqué par des pointillés. Os: odontostyle, mg: membrane guide, mp: muscle protracteur, od: odontophore, oe: œsophage, gs : glande salivaire, bl : bulbe musculo-glandulaire (bulbe œsophagien), in : intestin. mg et mp ne sont pas visibles sur le croquis du *Longidorus* et inversement les glandes salivaires représentées pour le croquis du *Xiphinema*. (Figure adaptée d'après Taylor et Brown, 1997).

Les particules virales sont adsorbées en une monocouche à la surface du tractus alimentaire du nématode (Taylor and Robertson, 1970), mais le site de rétention diffère suivant le genre de nématode vecteur (Taylor and Brown, 1997). Pour les vecteurs du genre *Longidorus*, les virus sont retenus dans le lumen de l'odontostyle, mais également entre l'odontostyle et la membrane guide. Chez les *Xiphinema*, les virus sont adsorbés depuis le début de l'odontophore jusqu'à la fin du bulbe œsophagien (fig. 1.15).

Cette adsorption des *Nepovirus* est un phénomène sélectif, qui implique certainement des molécules spécifiques du nématode. Dans le cas de *X. diversicaudatum*, il a pu être démontré que la capacité du nématode à retenir le virus peut être hérité génétiquement. En 1986, une étude a montré que le croisement entre une population de *X. diversicaudatum* transmettant efficacement l'ArMV et

une autre ne la transmettant qu'occasionnellement, générerait des nématodes « hybrides » transmettant l'ArMV avec une efficacité intermédiaire (Brown, 1986b).

La nature biochimique des sites de rétention pourrait être des molécules polysaccharidiques. Chez *X. diversicaudatum*, le site de rétention des particules virales répond positivement au marquage pour les carbohydrates - composé polysaccharidique -, alors que le même site chez *L. elongatus*, vecteur du RpRSV et du TBRV, n'est pas marqué, ni d'ailleurs l'odontostyle de *X. diversicaudatum* et de *X. index* (Robertson and Henry, 1986a; Robertson and Henry, 1986b). D'autre part, le marquage à la ferritine du lumen de l'odontostyle de *Longidorus* indique la présence de fortes charges négatives à ce niveau de l'appareil alimentaire, ce qui pourrait influencer l'interaction avec des charges positives de surfaces des virions (Robertson, 1987).

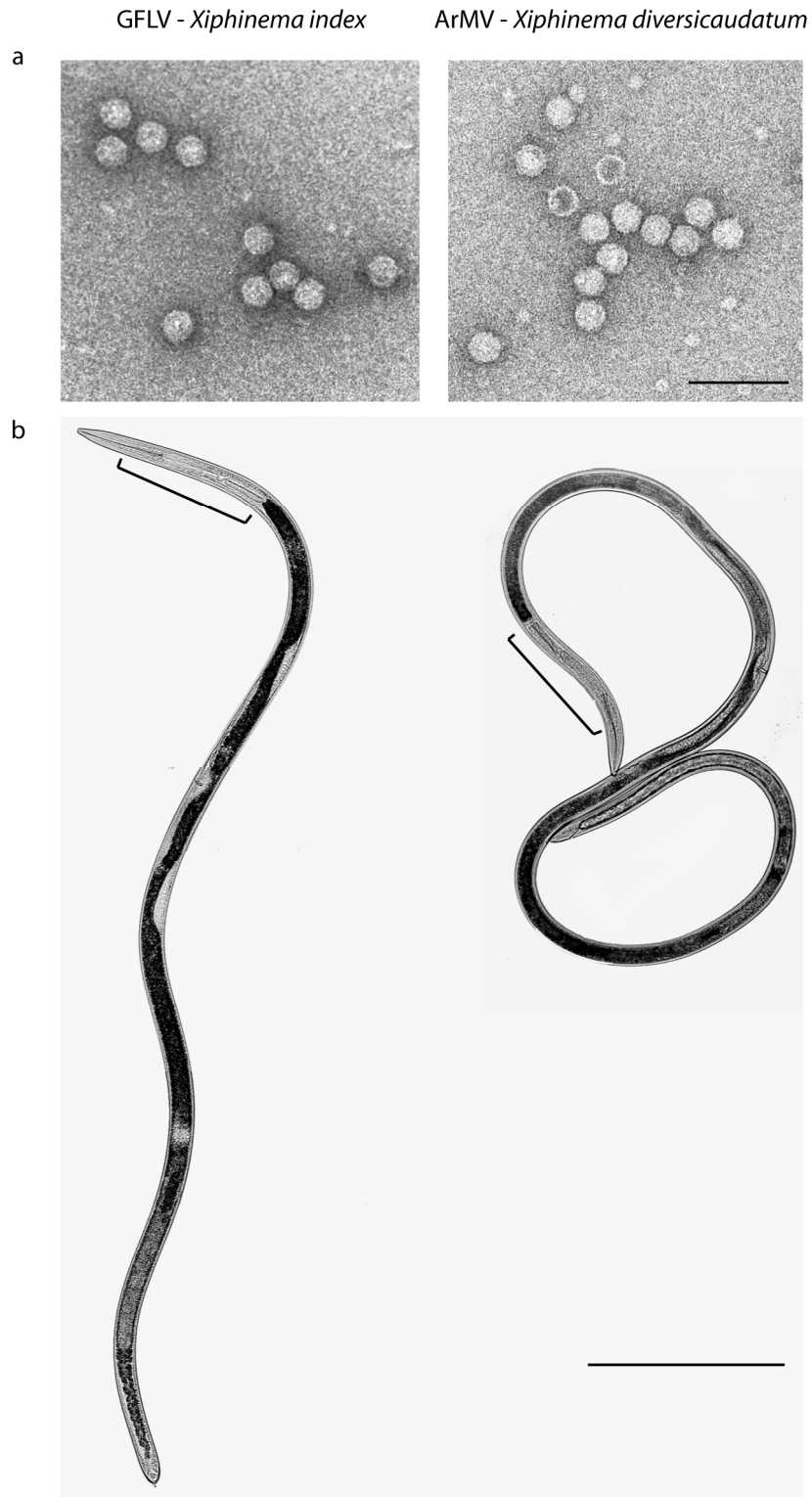


Fig. 2.1. Le GFLV et l'ArMV agent viraux principalement responsables de la maladie du court-noué de la vigne et leurs nématodes vecteurs en France. (a) Clichés de microscopie électronique à transmission après une coloration négative des particules virales purifiées de GFLV et ArMV. La barre indique 100 nm. (b) Photographies de microscopie optique en fond clair et lumière transmise d'un individu adulte femelle de *Xiphinema index* à gauche et d'un individu femelle de *X. diversicaudatum* à droite. La partie antérieure du nématode correspondant au site de rétention du virus, est indiquée par une accolade. La barre indique 1 mm. (Photos P. Schellenberger, INRA Colmar).

Chapitre 2

Le Grapevine fanleaf virus et son vecteur

1. La maladie du court - noué de la vigne

1. 1 Virus responsables de la maladie

La maladie du court-noué connaît beaucoup de noms différents : arrciamento, urticado, Reisigkrankheit, fanleaf degeneration. En Europe, elle fut rapportée pour la première fois par Casalis-Allut en 1865, puis par Rathay en 1883, qui avait déjà remarqué que la transmission de la maladie était d'origine tellurique. Aux Etats-Unis, la première mise en évidence de la maladie du court-noué est rapportée en 1950 dans un vignoble de Californie (Hewitt, 1950). Cette maladie est répandue dans la quasi-totalité des vignobles du monde et est l'une des maladies à virus les plus dommageables pour la viticulture (Andret-Link *et al.*, 2004a; Martelli and Taylor, 1990). En France, 60 % des parcelles de vignes sont touchées par cette maladie et on considère qu'un tiers du vignoble français est très atteint (Demangeat *et al.*, 2005a). Seize espèces virales, dont quinze *Nepovirus* et un *Sadwavirus* induisent la maladie du court-noué chez la vigne (Martelli and Boudon-Padieu, 2006). Parmi les 16 espèces virales impliquées dans la dégénérescence infectieuse, le *Grapevine fanleaf virus* (GFLV) est le principal agent responsable de cette maladie. Le GFLV, répandu mondialement, est transmis de vigne à vigne exclusivement par le nématode ectoparasite *Xiphinema index* (Thorne et Allen, 1950) (Fig 2.1). L'association entre GFLV et *X. index* a été prouvée en 1958 (Hewitt *et al.*, 1958). Les particules virales de GFLV ont été isolées de la vigne et purifiées quelques années après par Dias and Harrison (1963). Le nématode *X. italiae* a également été proposé comme vecteur naturel du GFLV par Cohn et collaborateurs en 1970. Cependant cette association n'a jamais pu être confirmée par d'autres travaux (Martelli and Taylor, 1990; Taylor and Brown, 1997). Par conséquent, *X. index* est considéré comme le vecteur naturel exclusif du GFLV.

Tout comme le GFLV, l'*Arabis mosaic virus* (ArMV) est un népovirus également responsable de la maladie du court-noué chez la vigne (Fig 2.1). Il est sérologiquement relié au GFLV et induit chez la vigne une symptomatologie similaire à celle du GFLV (Dias and Harrison, 1963). Il est répandu sur tout le territoire européen mais également au Moyen-Orient, en Asie, au Japon, au Canada et aux USA (Californie). L'ArMV possède une organisation génétique similaire à celle du GFLV. En revanche, il est transmis spécifiquement par *X. diversicaudatum* mais par *X. index* (Tableau 1.3). La découverte de cette association a été publiée juste après celle du GFLV-*X. index* (Harrison and

Cadman, 1959; Jha and Posnette, 1961). En France, l'ArMV et le GFLV sont les deux seuls agents de la maladie du court noué de la vigne (Demangeat *et al.*, 2005a).

1. 2 Description des symptômes sur vigne.

Les *Nepovirus* infectent toutes les variétés de vignes. Des symptômes variables sont observés sur les feuilles, les rameaux et les grappes des vignes infectées (fig. 2.2). L'expression des symptômes varie en fonction de l'isolat viral, du cépage et des conditions pédoclimatiques. Par exemple, ils seront plus importants sur un cep de Chardonnay ou de Gewurztraminer, que sur du Riesling ou du Pinot noir. De plus, ils ne touchent pas forcément la plante entière et peuvent s'exprimer uniquement sur un ou deux sarments (tige), bien que la plante entière soit infectée.

Sur les feuilles, les symptômes primaires correspondent à des décolorations (jaunissement, panachures réticulées, tâches annulaires) ou des déformations du limbe (fig. 2.2b). Les sinus pétiolaires sont élargis et les nervures primaires sont rapprochées. On peut également observer une déformation des bords de la feuille leur donnant une forme en éventail (fig. 2.2b) d'où la dénomination de la maladie en anglais « *fanleaf* ».

Sur les sarments, on remarque fréquemment un raccourcissement des entre-nœuds (d'où le nom de la maladie du court-noué), des doubles nœuds et des symptômes de fasciations (fig 2.2d). Les vignes infectées peuvent subir après plusieurs années une réduction significative de la végétation entraînant des difficultés lors de la taille, surtout lorsqu'elle est mécanisée.

Au niveau des grappes, une réduction de leur nombre et de leur taille, des symptômes de coulure (avortement des ovules) et de millerandage (retards de la maturation des baies) sont couramment observés (fig 2.2c). L'infection virale se traduit encore par des altérations du métabolisme de la plante. Les plus connues concernent la photosynthèse, les processus respiratoires, les activités enzymatiques, les transports à longue distance par le phloème, les équilibres hormonaux, la nutrition minérale (Walter, 1998).

Enfin, la maladie provoque un affaiblissement progressif de la souche qui réduit ainsi la longévité de la plante et peut conduire à sa mort. L'ensemble de ces phénomènes conduisent à des pertes de récolte pouvant dépasser les 80% et altérer la qualité du vin (Andret-Link *et al.*, 2004a; Bovey *et al.*, 1990; Komar *et al.*, 2008).

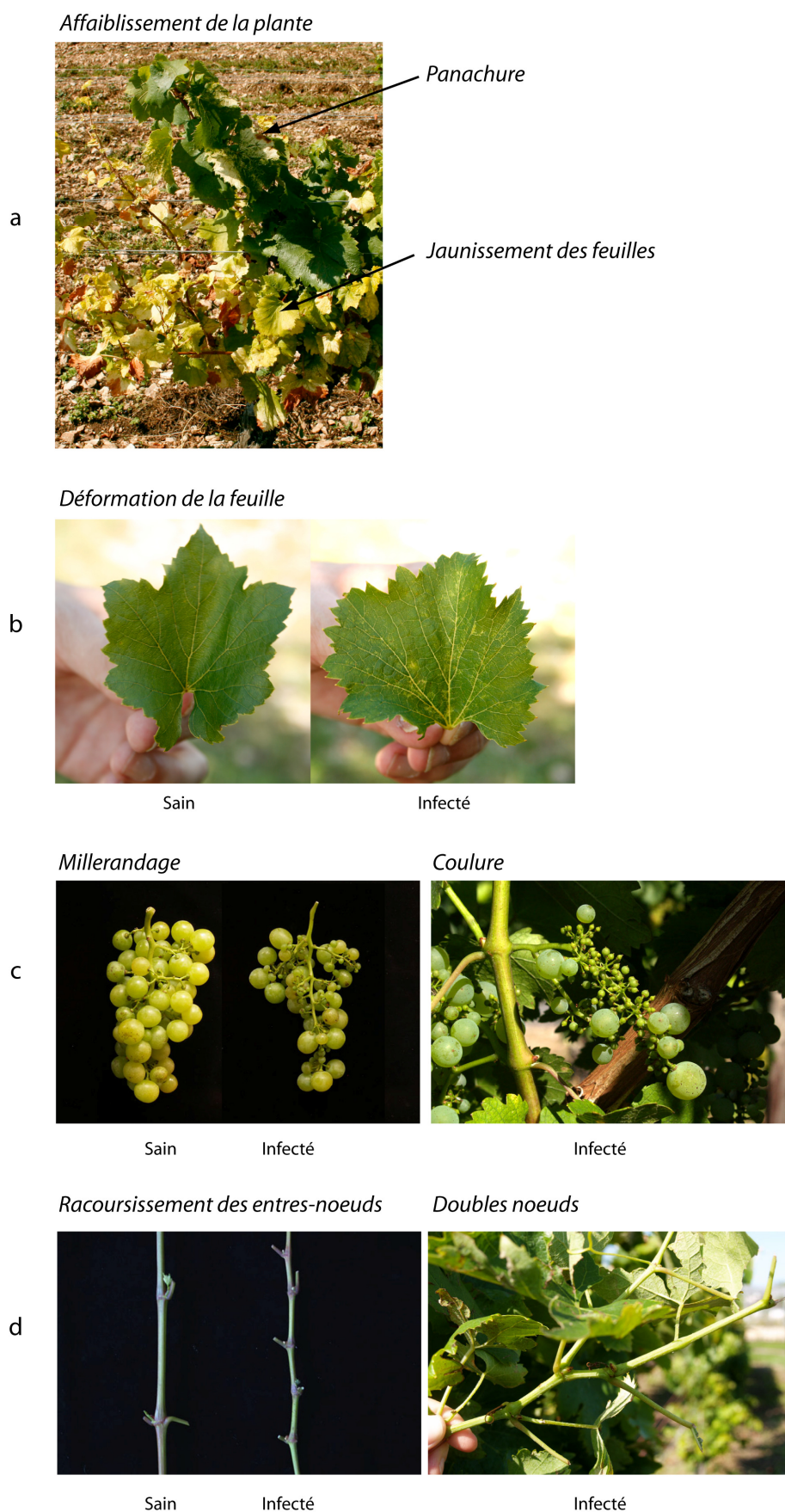


Fig. 2.2. Symptômes de la maladie du court-noué de la vigne au niveau de la plante entière (a), des feuilles (b), des fruits (c) et des tiges (d). Photographies P. Schellenberger, INRA Colmar.

1.3 Épidémiologie de la maladie du court-noué

La maladie du court-noué est distribuée au vignoble sous forme de foyers de pieds de vigne malades (Fig 2.3). Cette répartition des plantes malades est caractéristique d'une transmission naturelle par un vecteur souterrain. L'extension de cette maladie est souvent lente mais inexorable. Cette extension progressive lente est liée au déplacement très faible des nématodes dans le sol qui est estimé à quelques centimètres par an (Lamberti *et al.*, 1975). La dissémination du nématode se fait de proche en proche par le transport de terre ou avec celle qui adhère aux instruments de travail du sol (Fig 2.3). La contamination des parcelles situées en aval d'un secteur contaminé est souvent due à l'érosion du sol et au ruissellement des eaux (Villate, 2008).



Fig. 2.3. Foyers de court-noué dans un vignoble en bourgogne. Un foyer majeur est présent au premier plan, et deux foyers mineurs sont de l'autre côté de la route. Tous sont indiqués par une flèche. Photo P. Schellenberger, INRA Colmar.

La vitesse de recontamination d'une nouvelle parcelle dépend de la densité en nématodes virulifères dans le sol. Celle-ci est liée au nombre de générations qui se sont succédées et qui ont permis l'installation du ou des nématodes infectieux dans la parcelle. Il est cependant important de noter que les effectifs communément retrouvés dans les parcelles atteintes sont faibles (en moyenne une dizaine d'individus par kg de sol). En sol fortement contaminé, l'installation de jeunes plants est généralement compromise car la probabilité de rencontre entre nématodes et jeunes racines est très élevée. Dans ces conditions, des ceps de vigne peuvent être réinfectés dans l'année qui suit la plantation (Demangeat, communication personnelle).

1.4 Biologie de *Xiphinema index*, le vecteur du GFLV

L'aire d'origine de *X. index* est vraisemblablement le Moyen-Orient d'où l'espèce semble avoir été disséminée avec la culture de la vigne (Villate, 2008). Le nématode *X. index* a une gamme d'hôte très étroite. Il est inféodé à la vigne qui est aussi l'hôte naturel du GFLV. Le figuier est également une plante hôte mais cette espèce végétale n'est pas un hôte du virus. *X. index* affectionne

les sols lourds, car il y subirait moins la dessiccation à laquelle il est sensible. Après éclosion des œufs pondus isolément dans le sol, le nématode passe par quatre stades larvaires. Les adultes et tous les stades larvaires sont capables d'acquérir et de transmettre le GFLV (Taylor and Brown, 1997). Chez *X. index*, les mâles sont très rares et ne participent que très occasionnellement à la reproduction qui est majoritairement de type parthénogénétique (Villate *et al.* 2010). Ce mode de reproduction permet à un seul individu virulifère femelle d'engendrer un foyer de propagation du virus.

Le cycle complet (œuf à adulte) dure 2 à 3 mois dans les conditions les plus favorables mais peut s'étendre à plusieurs années au vignoble. Ce cycle de vie peut être prolongé par des périodes de quiescence à n'importe quel stade de développement lorsque *X. index* rencontre des conditions très défavorables (Esmenjaud, 2000; Taylor and Brown, 1997).

L'analyse de la distribution verticale de *X. index* a été réalisée dans plusieurs parcelles de vignes naturellement infectées. Ces analyses ont permis de montrer que *X. index* est souvent absent des horizons supérieurs et que les effectifs les plus importants sont retrouvés dans la rhizosphère des racines en croissance, c'est-à-dire entre 50 et 100 cm de profondeur (Villate, 2008).

1. 5 Quelles stratégies de lutte contre la maladie ?

Les caractéristiques biologiques de *X. index* (faibles effectifs, survie dans les horizons profonds du sol, et période de quiescence) ainsi que la rétention du GFLV par *X. index* pendant au moins de 4 ans en l'absence de plante hôte (Demangeat *et al.*, 2005b), expliquent les difficultés rencontrées pour lutter efficacement contre cette maladie et la recontamination quasi systématique des nouvelles plantations de vignes. La stratégie de lutte contre cette maladie est différente selon que l'on considère un vignoble sain ou un vignoble infecté.

Dans le cas d'un vignoble sain, la stratégie est prophylactique et consiste à éviter d'introduire une contamination dans la parcelle. L'objectif est de replanter du matériel végétal certifié exempt de virose.

Dans le cas d'un vignoble infecté, la stratégie vise à éradiquer le virus et/ ou le nématode. Les pieds malades sont arrachés après avoir été dévitalisés pour détruire les racines et priver les nématodes de ressources alimentaires. Cette pratique culturale peut être accompagnée d'un repos du sol pour une période de 7 ans minimum. Cependant cette période est souvent incompatible avec les enjeux économiques de la viticulture. Jusqu'en 2009, les populations de nématodes pouvaient également être contrôlées plus ou moins efficacement par l'utilisation de nématicides. Mais la toxicité aigüe, la rémanance et l'inefficacité de ces nématicides (Abawi and Widmer, 2000) a conduit à l'interdiction totale de l'utilisation de ces molécules en Europe (2007).

Des alternatives à ces traitements chimiques sont en cours de développement ou en cours d'évaluation. Récemment, des itinéraires techniques utilisant une sélection de plantes antagonistes vis-

à-vis de *X. index* ont été expérimentés dans une parcelle naturellement infectée. La culture de tagètes (*Tagetes minuta*) et de vesce velue (*Vicia villosa*) a permis de réduire la population naturelle de *X. index* au vignoble comparé à du sol nu (Villate, 2008). Ces résultats préliminaires intéressants doivent être validés au niveau d'autres sites d'expérimentation et accompagnés par une replantation de pieds de vigne après la culture des plantes antagonistes.

D'autres moyens de lutte consistent à perturber les interactions nématode/plante ou virus/plante. En effet, une résistance totale à *X. index* a été identifiée chez la variété sauvage de vigne *Muscadinia rotundifolia*. L'introgession de cette résistance dans des variétés de porte-greffes cultivés a été réalisée. Ce nouveau porte-greffe disponible sous le nom de Nemadex présente une tolérance au GFLV très intéressante mais pas totale et dont les qualités culturales doivent encore être améliorées (Bouquet *et al.*, 2003; Esmenjaud and Bouquet, 2009). Enfin, des travaux concernant le développement de vignes transgéniques pour les rendre résistantes au virus sont entrepris au laboratoire. Cette stratégie de lutte dérivée du pathogène a permis de produire des plantes résistantes à des virus (pour revue : Fuchs and Gonsalves, 2007). Dans le cas de la vigne, plusieurs lignées transgéniques ont été étudiées. Dans certains cas elles retardent l'apparition de la maladie, mais aucune résistance forte au GFLV n'a été constatée (Gambino *et al.*, 2010; Hemmer *et al.*, 2009).

À l'heure actuelle, aucune stratégie de lutte efficace, ni économiquement viable n'est disponible pour éradiquer la maladie du court noué de la vigne. Cette situation renforce l'importance de l'acquisition de connaissances des mécanismes moléculaires de la transmission du virus par son vecteur, afin de pouvoir en dégager des stratégies de lutttes innovantes.

2. Cycle viral du *Grapevine fanleaf virus*

Le GFLV est l'un des népovirus pour lequel les connaissances sur la structure et l'expression du génome sont les plus avancées. La majorité des données ont été acquises avec la souche F13 du GFLV, isolée à partir d'une vigne cultivée dans un vignoble de Frontignan en France (Vuittenez *et al.*, 1964). La séquence du génome a été déterminée (Ritzenthaler *et al.*, 1991; Serghini *et al.*, 1990) ainsi que celle de l'ARN satellite (Fuchs *et al.*, 1989). Des ADNc complets correspondant aux ARN génomiques, ainsi qu'à l'ARN satellite, ont été produits. A partir de ces clones complets, des ARN ayant les mêmes propriétés biologiques que les ARN viraux peuvent être synthétisés *in vitro* (Viry *et al.*, 1993). Le génome du GFLV est constitué de deux molécules d'ARN simple brin de polarité positive, dont l'extrémité 5' est liée de manière covalente à une protéine virale (1C^{VPg}) (Pinck *et al.*, 1991) et l'extrémité 3' est polyadénylée (Pinck *et al.*, 1988). Les deux ARN sont nécessaires pour déclencher une infection systémique *in planta* (Viry *et al.*, 1993). Certains isolats, comme la souche F13 (Fuchs *et al.*, 1989), présentent un ARN3 satellite. Chaque ARN est traduit en une polyprotéine maturée en plusieurs protéines fonctionnelles (fig. 2.4). La polyprotéine P1 (253 kDa), codée par

l'ARN1 est clivée en 5 protéines impliquées dans la maturation et dans la réplication. La polyprotéine P2 (122 kDa) codée par l'ARN2, est clivée en 3 protéines : la protéine 2A^{HP} (28 kDa) ayant pour fonction présumée de guider l'ARN-2 vers les complexes de réplication cellulaire, la 2B^{MP} (38 kDa) qui correspond à la protéine de mouvement, et la 2C^{CP} (56 kDa) ou protéine de capsid. Dans le manuscrit, la protéine de capsid sera également référencée comme la CP et celle de mouvement comme MP. Le satellite code pour une protéine de 37 kDa essentielle à la réplication de l'ARN3 (Hans *et al.*, 1992).

Le GFLV se réplique au sein de corps d'inclusions périnucléaires nommés compartiment viral ou viroplaste, constitués de prolifération membranaire dérivant du réticulum endoplasmique (Ritzenthaler *et al.*, 2002). La morphologie du compartiment viral dit « en rosette » est très similaire de celui retrouvé dans les cellules animales infectées par le *Poliiovirus*. Des études en protoplastes ont montré que l'ARN1 est capable de se répliquer de manière indépendante de l'ARN2 (Viry *et al.*, 1993).

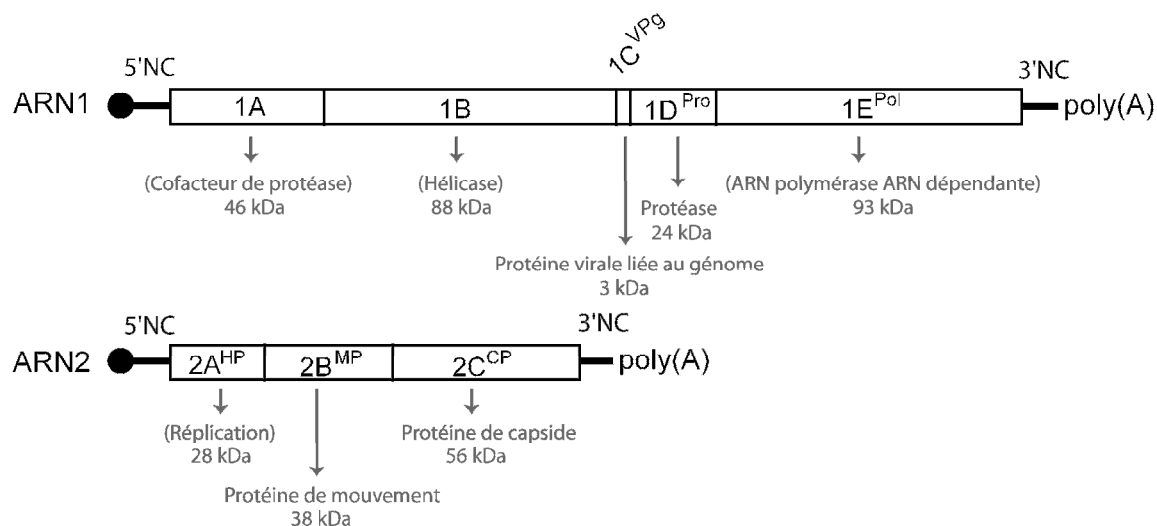


Fig. 2.4. Structure et expression des ARN 1 et 2 du *Grapevine fanleaf virus*. Les deux ARN comportent une longue phase de lecture ouverte représentée par des rectangles. Les séquences non codantes sont représentées par des lignes pleines et la VPg par un rond noir. Les parenthèses sont associées aux protéines pour lesquelles la fonction suspectée n'a pas été démontrée expérimentalement.

L'ARN1 contient les trois gènes - Hélicase - Protéase - et ARN Polymérase - dont l'organisation en triade est commune chez les virus appartenant à l'ordre des *Picornavirales* (Le Gall *et al.*, 2008). Le rôle de la protéine 1D^{Pro} dans la maturation de la polyprotéine a été mis en évidence par des études de génétique inverse associées à des expérimentations de traduction *in vitro* (Margis and Pinck, 1992; Margis *et al.*, 1994; Margis *et al.*, 1991). Quant aux protéines 1A, 1B^{hel} et 1E^{Pol}, leur fonction a été proposée suite à des études de comparaison de séquences (Ritzenthaler *et al.*, 1991).

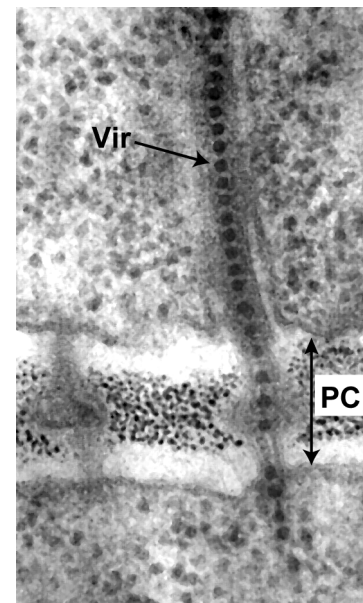
L'ARN2 est requis pour le mouvement du virus de cellule à cellule. La 2B^{MP} forme des tubules par oligomérisation de la protéine à l'intérieur des plasmodesmes, par lesquels les virions

traversent la paroi pectocellulosique des cellules végétales (fig. 2.5). L'adressage de la MP aux plasmodesmes est dépendant du système de sécrétion (Laporte *et al.*, 2003) et son encrage se produit via la reconnaissance spécifique d'une famille de protéines de plasmodesmes nommées PDLP (Plasmodesmata Located Protein) (Thomas *et al.*, 2008 ; Amari *et al.*, 2010 article soumis).

La spécificité d'interaction entre la 2C^{CP} et la 2B^{MP} est nécessaire au mouvement du virus. Des expériences de génétique inverse échangeant les séquences codantes pour ces protéines entre GFLV et ArMV produisent des chimères capables de se répliquer, d'être traduites et clivées par la protéase virale du GFLV, mais incapable de déclencher une infection systémique (Belin *et al.*, 2001; Belin *et al.*, 1999). Les 9 derniers résidus de la 2B^{MP} semblent jouer un rôle déterminant dans l'interaction CP-MP. En effet, lorsque ces résidus sont remplacés par leur équivalent ArMV, l'infektivité du virus chimère est perdue *in planta* (Belin *et al.*, 1999). La protéine 2A^{HP} quant à elle, participe à la réplication du génome viral, puisque elle est associée au complexe réplicatif et qu'elle est nécessaire à la réplication de l'ARN2 (Gaire *et al.*, 1999).

La CP du GFLV est une protéine de 504 acides aminés. C'est la seule protéine structurale du génome. Elle est capable de s'auto-assembler pour former des virions sans la présence d'ARN (Quacquarelli *et al.*, 1976). Comme je l'ai déjà mentionné précédemment (chapitre 1, partie 3.4.2), la CP est nécessaire et suffisante à la transmission spécifique du GFLV par *X. index* (Andret-Link *et al.*, 2004b; Belin *et al.*, 2001).

Fig. 2.5. Transport du GFLV à travers la paroi pectocellulosique d'une cellule végétale infectée. Coupe longitudinale d'un tubule qui connecte deux cellules adjacentes au travers de la paroi pecto-cellulosique (PC). La présence de virions (Vir) à l'intérieur du tubule est indiquée par une flèche.



Objectifs du travail de thèse

L'objectif de ma thèse est de mieux comprendre le mécanisme moléculaire de la transmission du GFLV par *X. index*, en caractérisant les déterminants viraux de la transmission présents au niveau de la capsid du virus.

Afin de répondre à la problématique posée nous avons postulé que les acides aminés de la protéine de capsid (CP) impliqués dans la transmission doivent être en surface du virion, afin de pouvoir interagir avec des récepteurs potentiels de l'appareil alimentaire du nématode. Nous avons également postulé qu'ils devaient être différents de ceux de l'ArMV, virus qui partage des caractéristiques communes importantes avec le GFLV mais qui est transmis spécifiquement par le nématode *X. diversicaudatum*. La nécessité de l'ArMV, comme modèle de comparaison, provient principalement du fait qu'aucun isolat du GFLV affecté dans sa transmission par *X. index* n'avait jamais été identifié. Au début de la thèse, 3 régions potentielles avaient été identifiées (Andret-Link, 2005) sur la base d'un modèle tridimensionnel du GFLV reconstruit par homologie avec la structure cristallographique du *Tobacco ringspot virus* (TRSV), permettant de localiser les résidus présents à la surface des virions. Pour tester l'implication des différentes régions dans la transmission, ces dernières avaient été échangées entre GFLV et ArMV, et l'infectivité des chimères analysées en plantes herbacées. La majorité des mutations de la CP ont donné lieu à des mutants non infectieux sur plantes et par conséquent l'impossibilité de les tester en transmission. Ainsi, la fonction dans la transmission d'une seule région de la capsid, comprenant 11 résidus, pouvait être analysée. J'ai donc cherché à obtenir de nouveaux mutants infectieux, en ciblant deux nouvelles régions toujours sur la base du modèle disponible. Le premier axe de ma thèse a ainsi consisté à compléter la collection de virus chimères GFLV/ArMV ciblant ces 5 régions d'intérêt de la CP, à étudier l'impact des mutations sur les propriétés biologiques des virus et à évaluer la transmissibilité des mutants systémiques en plante entière par *X. index* et *X. diversicaudatum*. Pour cela, nous avons mis au point un test de transmission par le vecteur *X. diversicaudatum*, outil qui n'était pas disponible au début de ma thèse. Ce test a permis de montrer que la CP de l'ArMV détermine aussi sa spécificité de transmission par son nématode vecteur.

D'autre part, les résultats d'infectivité des constructions chimères indiquaient que l'introduction de mutations dans la protéine de capsid (CP) du GFLV était délicate, et nécessitait probablement une mutagenèse encore plus ciblée. Résoudre la structure cristallographique du GFLV a ainsi représenté un enjeu majeur du travail de thèse, car il pouvait permettre à la fois de localiser les résidus déjà mutés dans les différentes régions de la capsid et de poursuivre l'identification des

résidus de surface de la capsid. La suite du travail a donc consisté à déterminer la structure tridimensionnelle de la particule virale du GFLV.

Le deuxième axe du travail de thèse est basé sur la découverte d'un isolat naturel du GFLV faiblement transmis par *X. index*. Comme précisé précédemment, beaucoup d'avancées sur la transmission des phytovirus ont été possibles grâce à la caractérisation d'isolats variant dans leur efficacité de transmission par leurs espèces vectrices. Au cours de mes essais de transmission, j'ai pu identifier un variant naturel du GFLV faiblement transmis par *X. index*. Ce variant possède une mutation ponctuelle (substitution d'un acide aminé) dans la CP. J'ai caractérisé ce variant naturel au niveau fonctionnel et structural.

Enfin, dans le troisième axe, j'ai abordé la recherche de déterminants viraux impliqués dans la spécificité de transmission des couples GFLV-*X. index* et ArMV-*X. diversicaudatum*. Pour cela, j'ai comparé les structures résolues du GFLV avec celle de l'ArMV obtenues par reconstruction 3D à partir d'images de cryomicroscopie électronique à transmission (cryo-MET) sur particules virales purifiées.

Au final, ce travail a combiné des approches expérimentales complémentaires de structure (cristallographie et cryo-MET), et des tests de fonctions biologiques (encapsidation et transmission), pour déboucher d'une part sur la résolution du GFLV sauvage à 3 Å, du variant faiblement transmissible à 2,7 Å, et de l'ArMV à 13,5 Å et d'autre part sur l'identification de deux régions de la CP impliqués dans la transmission.

Les différents chapitres de la thèse rapportent les 3 axes de recherche qui ont été utilisés pour réaliser mon objectif. Le chapitre 3 expliquera la mise au point du test de transmission par le vecteur *X. diversicaudatum*, et son utilisation dans la démonstration que l'ArMV utilise la même stratégie de vection que le GFLV. Le chapitre 4 présentera le premier axe de recherche, avec l'évaluation du rôle des 5 régions de la capsid sélectionnées d'après leur prédiction à la surface de la capsid et différentes entre GFLV et ArMV. Les données relatives au travail basé sur le deuxième axe de recherche, c'est-à-dire la caractérisation fonctionnelle et structurale du variant GFLV-TD et sa comparaison avec la structure du GFLV sauvage, figureront dans le chapitre 5. Enfin, le chapitre 6 détaillera l'obtention de la structure résolue de l'ArMV et sa comparaison avec le GFLV.

Chapitre 3

*La CP détermine également la spécificité de transmission du virus par *Xiphinema diversicaudatum**

Une des approches de l'identification des déterminants viraux impliqués dans la transmission du GFLV repose sur la comparaison de la vection du GFLV, de l'ArMV et de chimères construites avec ces 2 virus. Ces chimères sont obtenues en effectuant des échanges de séquences GFLV par leur équivalent ArMV au niveau d'un ADNc codant pour l'ARN2 du GFLV. Cependant, jusqu'à présent, seule la perte ou le gain de transmissibilité de ces chimères par *X. index* pouvait être évaluée (Andret-Link *et al.*, 2004b; Belin *et al.*, 2001) car le laboratoire ne disposait pas d'élevage de *X. diversicaudatum*. Pour pouvoir conclure avec rigueur sur l'implication des déterminants viraux dans la spécificité de la transmission, il est essentiel de pouvoir évaluer la transmissibilité des virus sauvages et des virus chimériques à la fois par *X. index* et par *X. diversicaudatum*.

Après avoir obtenu une multiplication en masse du nématode *X. diversicaudatum*, j'ai mis au point une procédure de transmission de l'ArMV par *X. diversicaudatum* en conditions contrôlées. Cette procédure, associée au test de transmission par *X. index*, a été utilisée pour évaluer la transmissibilité de virus chimères possédant une capsid de nature ArMV. Ce travail est présenté dans l'article 1 intégré dans ce chapitre. Il a permis de montrer que les déterminants de la spécificité de transmission de l'ArMV sont localisés, tout comme pour le GFLV, sur le gène codant pour la CP.

1. Caractéristiques de la population de *Xiphinema diversicaudatum*

1. 1 Un peu d'écologie concernant *X. diversicaudatum*

X. diversicaudatum, le vecteur naturel de l'ArMV, est un grand nématode ectoparasite filiforme. Au stade adulte, il mesure entre 4 à 6,2 mm de long. Contrairement à toutes les autres espèces de *Xiphinema* vecteurs de virus qui ont un mode de reproduction majoritairement parthénogénétique, *X. diversicaudatum* possède un mode de reproduction sexuée. Ce mode de

reproduction est potentiellement un obstacle à l'établissement rapide d'élevages de masse. Il possède les mêmes capacités de rétention et de transmission de virus que *X. index* (Mc Namara, 1980). Ce nématode est très polyphage et présente une gamme d'hôtes naturels beaucoup plus étendue que celle de *X. index*. Alors que *X. index* est inféodée naturellement à la vigne, les hôtes naturels de *X. diversicaudatum* sont pour une grande partie des arbres fruitiers et des plantes à petits fruits (Taylor and Brown, 1997).

1. 2 Multiplication de la population de *X. diversicaudatum*

Le travail visant à obtenir une population de *X. diversicaudatum* suffisamment nombreuse pour effectuer des tests de transmission a été initiée par le laboratoire avant mon arrivée dans le projet, et a été poursuivie pendant ma thèse. Je vais brièvement décrire la caractérisation et l'obtention de cette population.

Une population de *X. diversicaudatum* établie depuis 1994 sur prunier (*Prunus domestica*) a été offerte par l'INRA de Sophia-Antipolis (Daniel Esmenjaud). À partir d'un échantillon de sol de cet élevage initial, des nématodes ont été isolés puis (i) caractérisés sous une loupe binoculaire en utilisant les clés de détermination de l'espèce *Xiphinema* (Luc M. and Dalmasso A., 1975) (ii) identifiés par PCR en utilisant des amorces spécifiques de l'espèce *X. diversicaudatum* (Wang *et al.*, 2003) (iii) analysés par RT-PCR pour rechercher la présence d'ArMV en utilisant une méthode développée dans notre laboratoire (Demangeat *et al.*, 2004). L'étude des caractères biométriques d'individus pris au hasard confirme que ce sont des nématodes appartenant à l'espèce *X. diversicaudatum*. L'analyse moléculaire par PCR réalisée à partir de l'ADN extrait de nématodes avec des amorces spécifiques de l'espèce *X. diversicaudatum* permet d'amplifier un seul fragment d'ADN de la taille attendue indiquant que tous les individus analysés sont bien des *X. diversicaudatum*. Une RT-PCR effectuées sur trois lots de 30 nématodes à l'aide d'amorces spécifiques de la partie 3' de l'ARN2-U de l'ArMV-S n'a pas révélé la présence d'ArMV dans les nématodes.

Cette population de *X. diversicaudatum* active n'hébergeant pas d'ArMV, a donc été utilisée pour une multiplication à grande échelle. Des lots d'une centaine d'individus ont été déposés au niveau de racines de pétunia (*Petunia hybrida*), vigne (*Vitis vinifera*), fraisiers (*Fragaria ananassa*) et framboisiers (*Rubus idaeus*). Dans nos conditions de travail et après deux années d'élevage en serre, le framboisier s'est révélé être le meilleur hôte permettant d'avoir une multiplication importante de *X. diversicaudatum* (100 à 360 fois la population initiale).

Les *X. diversicaudatum* utilisés au cours de mon travail sont donc issus d'élevages sur framboisiers pour lesquels je vérifie l'absence d'ArMV avant de procéder au test de transmission.

2. Choix des conditions expérimentales de transmission de virus par *X. diversicaudatum*

L'ensemble des paramètres utilisés lors de nos tests de transmission avec *X. diversicaudatum* a été fixé par rapport aux données bibliographiques et à l'expérience du laboratoire acquise avec les tests de transmission par *X. index*. Ce dernier se déroule en 2 étapes : (i) une période d'acquisition (AAP pour « Acquisition Access Period ») de 6 semaines pendant laquelle environ 200 à 300 nématodes avirulifères sont mis en contact avec des racines d'une plante source infectée par un virus de notre choix, (ii) une étape de transmission (IAP pour « Inoculation Access Period ») au cours de laquelle les nématodes préalablement au contact d'une plante source sont déposés au niveau des racines d'une plante piège. Après 6 semaines de contact, la présence du virus est recherchée par ELISA au niveau des racines de la plante piège. La descendance des virus transmis est caractérisée par IC RT-PCR et séquençage. Sur la base de ce test de transmission, plusieurs choix ont été effectués pour mettre en place un test de transmission avec *X. diversicaudatum*

Dans la littérature, la transmission de l'ArMV par *X. diversicaudatum* a déjà été réalisée en utilisant entre 50 à 100 nématodes lors de l'AAP puis entre 10, 5, 2 voire 1 seul nématode par plante piège au cours de IAP (Brown, 1986a; Valdez, 1972; Valdez *et al.*, 1974). Avant mon arrivée au laboratoire, des procédures similaires ont été évaluées mais sans succès de transmission. Cela provient probablement du fait que ces procédures impliquent la manipulation de nématodes isolés du sol, technique que le laboratoire ne maîtrise pas assez efficacement. Nous avons donc choisi de procéder de la même manière qu'avec *X. index*, c'est à dire d'utiliser des échantillons de sol contenant environ 200 nématodes qui sont déposés au niveau des racines de la plante source infectée puis transférés au niveau des racines de la plante piège saine.

Les tests de transmission par *X. diversicaudatum* décrits dans la littérature sont effectués en utilisant *Chenopodium quinoa* et *Petunia hybrida* (Brown, 1986a; Valdez, 1972), qui sont à la fois des plantes hôtes pour le GFLV, l'ArMV et pour *X. diversicaudatum*. *C. quinoa* est utilisé comme plante source et piège, alors que *P. hybrida* est majoritairement utilisée comme plante piège. Par ailleurs, *X. diversicaudatum* est capable d'induire la formation de galles au niveau des racines d'*Arabidopsis thaliana* (Barthels *et al.*, 1997) suggérant son utilisation potentielle pour des tests de transmission. L'utilisation de cette espèce végétale pourrait considérablement élargir les perspectives d'études menées dans l'équipe, grâce à la disponibilité de lignées de plantes mutantes et de protocoles efficaces de transformation génétique.

Au laboratoire, les conditions expérimentales d'inoculation de *C. quinoa* avec le GFLV et l'ArMV sont bien maîtrisées. En revanche, le laboratoire a peu d'expérience avec *A. thaliana* et *P. hybrida*. Nous avons donc utilisé des jus bruts de *C. quinoa* infectés par le GFLV-F13 (Vuittenez *et al.*, 1964) ou l'ArMV-S (Huss *et al.*, 1989), pour inoculer mécaniquement des plantules de *P. hybrida*

et *A. thaliana* (écotype Col0) (Tableau 3.1). Dans nos conditions expérimentales, environ 50% des plantes de *P. hybrida* sont infectées par l'ArMV alors qu'aucune plante n'a été infectée par le GFLV. Le même résultat a été obtenu en changeant certains paramètres d'inoculation, comme le stade de développement de la plante ou le tampon de broyage. Moins de 50% des plantes d'*A. thaliana* ont été infectées par le GFLV ou l'ArMV. (Tableau 3.1). Ce faible taux de plantes infectées pour *P. hybrida* et *A. thaliana*, ne nous permettent pas d'utiliser ces deux espèces pour des tests de transmission. J'ai donc utilisé uniquement *C. quinoa* comme plante source et plante piège dans mes tests de transmission.

Table 3.1. Test d'infectivité du GFLV et de l'ArMV sur trois espèces végétales pouvant être utilisées comme plante source dans les tests de transmission par *X. diversicaudatum*.

	Plantes hôtes ^a		
	<i>Petunia hybrida</i>	<i>Arabidopsis thaliana</i>	<i>Chenopodium quinoa</i>
Nombre de plantes infectées par le ArMV-S / nbre de plantes inoculées ^a	8/17 (47%)	21/49 (43%)	21/30 (70%)
Nombre de plantes infectées par le GFLV-F13 / nbre de plantes inoculées ^a	0/17 (0%)	19/49 (39%)	27/30 (90%)

^a Plantules inoculées mécaniquement au stade 4 feuilles d'après Viry *et al.*, 1993, dans deux tests indépendants. La présence de virus est détectée dans les feuilles non inoculées 15 jours après inoculation (dpi) par DAS-ELISA à l'aide d'anticorps polyclonaux dirigés contre le GFLV et l'ArMV. La plante est considérée comme infectée si l'absorbance à 405 nm est 3 fois supérieure à celle de feuilles saines après 120 minutes d'hydrolyse du substrat.

Les deux phases du test de transmission sont de 6 semaines chacune (Andret-Link *et al.*, 2004b; Belin *et al.*, 2001). Ces 2 périodes relativement longues peuvent potentiellement être réduites. Dans la littérature, plusieurs tests de transmission de nepovirus par nématodes ont été effectués avec succès en utilisant une période d'AAP et d'IAP inférieure à celle utilisée au laboratoire pour *X. index*. L'acquisition (AAP) du TRSV par *X. americanum* peut se faire en 10 jours (McGuire, 1973), et celle du RpSRV par *L. elongatus* en 4 semaines (Harrison *et al.*, 1974a). Brown (1986) utilise une procédure de 4 semaines d'AAP et 4 semaines pour l'IAP pour la transmission de l'ArMV par *X. diversicaudatum*. J'ai donc choisi de tester une période de 4 semaines d'AAP et d'IAP.

Finalement, l'aptitude de la population de *X. diversicaudatum* à transmettre l'ArMV-S a été évaluée en conditions contrôlées d'après le protocole suivant : environ 200 nématodes par plante, 4 semaines d'AAP, 4 semaines d'IAP, *C. quinoa* pour la plante source (infectée par le virus de notre choix) et *C. quinoa* sain pour la plante piège. La transmissibilité de l'ArMV-S et du GFLV-F13 est présentée dans le tableau 3.2. L'ArMV-S est transmis par *X. diversicaudatum*, avec une efficacité de 100 %. Cette transmission est spécifique puisque le GFLV-F13 n'est pas transmis par *X. diversicaudatum* dans les mêmes conditions.

Table 3.2. Transmissibilité du GFLV et de l'ArMV par *X. diversidaucatum*

	Transmission ^a	
GFLV-F13	0/11	0 %
ArMV-S	11/11	100 %
Témoin sain	0/11	0 %

^a Le taux de transmission est calculé d'après le nombre de plantes pièges ayant réagit positivement en DAS-ELISA en utilisant des anticorps polyclonaux dirigés contre le GFLV et l'ArMV, sur le nombre de plantes pièges testées.

3. Adaptation du test de transmission par *X. index*

La transmission du GFLV par *X. index* se déroule en deux périodes de 6 semaines. Cette procédure de transmission du GFLV utilise *Nicotiana benthamiana* comme source de virus et une plantule de vigne est généralement utilisée comme plante piège (Andret-Link *et al.*, 2004b; Belin *et al.*, 2001). Sur la base des résultats obtenus avec *X. diversicaudatum*, j'ai adapté la procédure de transmission du GFLV par *X. index* en diminuant les phases AAP et IAP à 4 semaines chacune. J'ai également remplacé la vigne, généralement utilisée au stade de jeunes boutures comme plante piège, par *N. benthamiana*. Dans ces conditions, la transmission du GFLV est réalisée par *X. index* avec une efficacité équivalente à celle obtenue avec la procédure précédemment utilisée au laboratoire (Tableau 3.3).

Table 3.3. Transmissibilité du GFLV et de l'ArMV par *X. index*

	Transmission ^a	
GFLV-F13	10/10	100 %
ArMV-S	0/13	0 %
Témoin sain	0/5	0 %

^a Le taux de transmission est calculé d'après le nombre de plantes pièges ayant réagit positivement en DAS-ELISA en utilisant des anticorps polyclonaux dirigés contre le GFLV et l'ArMV, sur le nombre de plantes pièges testées.

4. Article 1 : The coat protein determines the specificity of virus transmission by *Xiphinema diversicaudatum*.

Journal of Plant Pathology 2010, 92 (1) : 275-279.

SHORT COMMUNICATION

**The coat protein determines the specificity of virus transmission by
*Xiphinema diversicaudatum***

A. Marmonier^{1†}, P. Schellenberger^{1,2†}, D. Esmenjaud³, C. Schmitt-Keichinger², C. Ritzenthaler², P. Andret-Link¹, O. Lemaire¹, M. Fuchs⁴ and G. Demangeat¹

[†]These authors have contributed equally to this work.

¹UMR 'Santé de la Vigne et Qualité du Vin', INRA/UDS, 28 rue de Herrlisheim, 68021 Colmar, France.

²Institut de Biologie Moléculaire des Plantes, CNRS/UDS, 12 rue du Général Zimmer, 67084 Strasbourg, France.

³UMR 'Interactions Biotiques et Santé Végétale', INRA 1064-UNSA-CNRS 6192, 400 route des Chappes, 06903 Sophia Antipolis, France.

⁴Department of Plant Pathology, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456, USA.

Corresponding author: G. Demangeat . E-mail: gerard.demangeat@colmar.inra.fr

SUMMARY

Arabis mosaic virus (ArMV) and *Grapevine fanleaf virus* (GFLV), two closely related members of the genus *Nepovirus*, family *Secoviridae*, are responsible for fanleaf degeneration disease of grapevines. ArMV and GFLV are specifically transmitted by the ectoparasitic soil-borne nematodes *Xiphinema diversicaudatum* and *X. index*, respectively. Previous work has shown that the coat protein determines the transmission specificity of GFLV. No information is available on the determinants of the specific transmission of ArMV by *X. diversicaudatum*. We addressed this issue by testing the transmissibility of three chimeric viruses comprising GFLV RNA-1 and GFLV/ArMV chimeric RNA-2 molecules coding a coat protein from ArMV. The three recombinant viruses and wild type ArMV, unlike GFLV, were transmitted by *X. diversicaudatum* but not by *X. index*. In addition, RT-PCR assays indicated that ArMV and the three chimeric viruses seemed to be retained by *X. diversicaudatum* but not by *X. index* following nematode feeding on infected plants. Conversely, GFLV was retained by *X. index* but not by *X. diversicaudatum*. These results are consistent with the notion that the coat protein carries the determinants for the specific transmission of ArMV by *X. diversicaudatum* and that transmission specificity of both ArMV and GFLV is associated with specific virus retention by nematode vectors.

Key words: Nepovirus, ELISA, GFLV, RT-PCR, nematodes, *Xiphinema index*, recombinant viruses.

ARTICLE

Arabis mosaic virus (ArMV) belongs to subgroup A of the plant virus genus *Nepovirus* in the family *Secoviridae* (Le Gall *et al.*, 2008; Sanfaçon *et al.*, 2009). ArMV has polyhedral particles and is serologically distantly related to *Grapevine fanleaf virus* (GFLV). Both viruses are responsible for fanleaf degeneration, one of the most severe viral diseases of grapevines, which causes serious economic losses by substantially reducing yield and affecting fruit quality (Andret-Link *et al.*, 2004a). ArMV and GFLV are specifically transmitted from plant to plant by two different ectoparasitic root-feeding *Xiphinema* nematode species. ArMV is specifically transmitted by *Xiphinema diversicaudatum* and GFLV is exclusively vectored by *X. index* (Andret-Link *et al.*, 2005).

The genomes of ArMV and GFLV are similarly composed of two single-stranded positive-sense RNAs, called RNA-1 and RNA-2. Each genomic RNA codes for a polyprotein from which functional proteins are released by proteolytic processing. RNA-1 codes for the proteins implicated in RNA replication and for the viral proteinase (Andret-Link *et al.*, 2004b, Wetzel *et al.*, 2008). RNA-2 encodes three final cleavage products: the N-proximal 2A^{HP} homing protein, the central 2B^{MP} movement protein, and the C-proximal 2C^{CP} capsid protein (Andret-Link *et al.*, 2004b, Wetzel *et al.*, 2002) (Fig. 1). The RNA-2-encoded proteins of GFLV and ArMV share 68-78% similarity at the

amino acid level (Vigne *et al.*, 2008). Full-length cDNA clones of RNA-1 and RNA-2 of GFLV strain F13 are available for *in vitro* synthesis of infectious transcripts (Viry *et al.*, 1993).

Early studies on nepovirus transmission with pseudorecombinants of *Tomato black ring virus* (TBRV) and *Raspberry ringspot virus* (RpRSV) isolates showed that viral determinants of transmission map to RNA-2 and suggested a critical role for protein 2C^{CP} in nematode transmission (Harrison *et al.*, 1974; Harrison and Murrant, 1977). More recently, recombinant RNA-2 molecules of GFLV strain F13 (Vuittenez *et al.*, 1964) and ArMV strain S (Huss *et al.*, 1989) were engineered and their transmissibility by *X. index* was tested in the presence of GFLV RNA-1 (Belin *et al.*, 2001; Andret-Link *et al.*, 2004b). Transmission was achieved for recombinant viruses carrying a protein 2C^{CP} of GFLV origin, regardless of the origin of proteins 2A^{HP} and 2B^{MP}. It was therefore concluded that protein 2C^{CP} is the sole viral determinant for the specific transmission of GFLV by *X. index* (Andret-Link *et al.*, 2004b).

No information is available on the ArMV determinants of transmission specificity by *X. diversicaudatum*. We addressed this issue by investigating the transmissibility of chimeric GFLV/ArMV viruses encoding a protein 2C^{CP} of ArMV origin by *X. diversicaudatum* and *X. index*. As no infectious transcripts of ArMV are available, we carried out inverse genetics experiments with full-length cDNA clones of GFLV in which ArMV gene segments were introduced.

Based on our previous findings on GFLV (Andret-Link *et al.*, 2004b), we hypothesized that protein 2C^{CP} carries the determinants for the specific transmission of ArMV by *X. diversicaudatum*. To determine the role of ArMV protein 2C^{CP} in transmission, we used three recombinant viruses, plus synthetic wild type GFLV strain F13 (Vuittenez *et al.*, 1964) and wild type ArMV strain S (Huss *et al.*, 1989). The three recombinant viruses contained GFLV RNA-1 and one of the following three recombinant RNA-2 molecules: 2AB_(HMST)C_A, 2AB_{A(HLST)}C_A or 2AB_{A(EMSV)}C_A (Fig 1). The number 2 refers to RNA-2; capital letters refer to the GFLV RNA-2-encoded genes, e.g. A for gene 2A^{HP}, B for gene 2B^{MP} and C for gene 2C^{CP}; A as subscript indicates the ArMV origin of the corresponding gene, B_A and C_A refer to the ArMV origin of genes 2B^{MP} and 2C^{CP}, respectively; and letters as subscripts in parenthesis indicate amino acids of protein 2B^{MP} at position -9, -6, -4 and -2, respectively, upstream of the R/G cleavage site between protein 2B^{MP} and 2C^{CP}. The three chimeric RNA-2 molecules had protein 2A^{HP} of GFLV origin, protein 2C^{CP} of ArMV origin and protein 2B^{MP} of either ArMV or GFLV origin. The engineering of constructs 2AB_{A(HLST)}C_A and 2AB_{A(EMSV)}C_A was described previously (Andret-Link *et al.*, 2004b). Construct 2AB_(HMST)C_A was developed by PCR, using construct 2ABC_A (Belin *et al.*, 1999), forward primer 5'-¹⁰⁰⁸ATGGGAAGACTACCGGTGG¹⁰²⁶-3' and reverse primer 5'-²⁰⁰⁹TCACTTGAGGCTCATCCCAGAAATGAGCTCAACCACGAGAGGGCTAGCCGGCAGAGGGA²⁰⁶⁶-3' carrying appropriate mutations. Nucleotides highlighted in grey in the latter primer correspond to modifications relative to GFLV-F13 RNA2 (GenBank accession No. NC_003623) in order to mutate amino acids in position -2, -6, and -9 at the C-terminus of protein 2B^{MP} upstream of

the R/G cleavage site (Fig. 1). These residues were shown to be critical to maintain a functional interaction between proteins 2B^{MP} and 2C^{CP} for cell-to-cell and long distance movement (Belin *et al.*, 1999; Andret-Link *et al.*, 2004b). Amplified mutagenic PCR fragment were digested with *Age*I and *Ngo*MIV and cloned into the corresponding sites in 2ABC_A to yield 2AB_(HMST)C_A.

Transcripts of GFLV RNA-1 and chimeric or wild type GFLV RNA-2 were obtained by *in vitro* transcription with the T7 RNA polymerase, as described by Andret-Link *et al.* (2004). Transcript infectivity was evaluated by mechanical inoculation of *Chenopodium quinoa*, a systemic herbaceous host for GFLV-F13 and ArMV-S (Andret-link *et al.* 2004). Plant infection by synthetic GFLV, wild type ArMV and chimeric viruses was determined by monitoring symptom development and testing non-inoculated apical leaves by DAS-ELISA using γ -globulins specific to GFLV or ArMV produced at INRA-Colmar (Walter *et al.*, 1987). Viruses were propagated on *C. quinoa* and transferred from *C. quinoa* to *Nicotiana benthamiana* by mechanically inoculation.

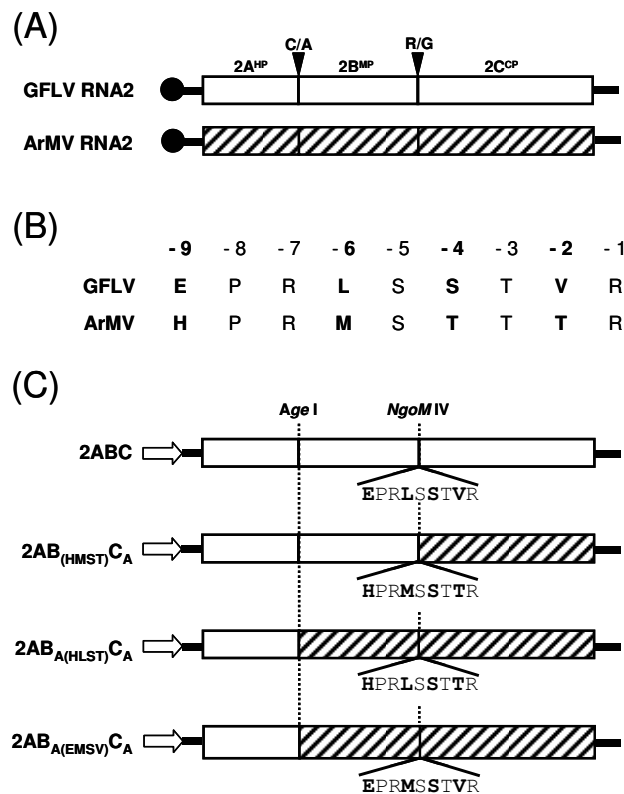


Fig. 1. (A) Schematic representation of RNA-2 of GFLV strain F13 and ArMV strain S. Boxes denote coding sequences and black lines symbolize non-coding GFLV sequences. Open boxes and hatched boxes represent GFLV and ArMV sequences, respectively. The viral genome linked protein (VPg) is shown by a black circle and processed proteins (2A^{HP}, 2B^{MP} and 2C^{CP}) are indicated on top of the boxes. Black arrow heads indicate the proteolytic cleavage sites C/A and R/G. (B) Alignment of the nine C-terminal amino acids of GFLV and ArMV protein 2B^{MP}. Divergent amino acids are in bold. (C) Schematic representation of the three chimeric GFLV/ArMV RNA-2 molecules derived from 2ABC that were used in this study. The ArMV origin of a gene is indicated by an A as subscript. Letters in parenthesis correspond to amino acids in position -2,- 4, -6, and -9 at the C-terminus of protein 2B^{MP}. The T7 promoter is represented by an arrow.

For each combination of transcripts, 2-3 plants were inoculated in two independent experiments. Similar to transcripts of wild type GFLV RNA-2, transcripts of the three chimeric RNA-2 induced systemic plant infection 10 days post-inoculation when associated with GFLV-F13 RNA-1 transcripts, as shown by mosaic symptoms in apical leaves. Infection was confirmed by DAS-ELISA with significantly higher absorbance readings in ArMV- (OD_{405nm} values of 0.7 ± 0.09) and GFLV- (OD_{405nm} values of 1.5 ± 0.5) infected plants compared to healthy controls (OD_{405nm} values of 0.07 ± 0.005) after 90 min of substrate hydrolysis.

These results confirmed the infectivity of constructs $2AB_{A(HLST)}C_A$ and $2AB_{A(EMSV)}C_A$ reported previously (Andret-Link *et al.* 2004) and showed infectivity of the new construct $2AB_{A(HMST)}C_A$. By replacing specific GFLV residues by their ArMV amino acids counterparts in the C-terminus of GFLV- $2B^{MP}$, systemic infection was achieved with construct $2ABC_A$, most probably because a functional compatibility between proteins $2B^{MP}$ and $2C^{CP}$ was restored. A specific interaction between the protein $2C^{CP}$ and C-terminal part of protein $2B^{MP}$ was observed for other GFLV/ArMV chimeric viruses and also for other viruses employing tubule-guided movement (van Lent and Schmitt-Keichinger, 2006).

The transmissibility of the three chimeric viruses was tested with *X. index* and *X. diversicaudatum*. A population of aviruliferous *X. index* from Fréjus (France) was reared on fig (*Ficus carica* L.) plants (Demangeat *et al.*, 2004) and an aviruliferous population of *X. diversicaudatum* from Lavalette (France) was reared on raspberry (*Rubus idaeus* L.). Rearings of both nematodes species were kept in 10-litre containers in a greenhouse under controlled temperature (20-25°C) and light (16 h photoperiod) conditions. A two-step transmission procedure was used essentially as described earlier (Andret-Link *et al.*, 2004b). Briefly, 200 aviruliferous nematodes were allowed to feed on the roots of virus source plants for a four-week acquisition access period (AAP). Virus source plants were two-week-old infected *C. quinoa* and *Nicotiana benthamiana* for transmission assays with *X. diversicaudatum* and *X. index*, respectively. Prior to the transmission assays, the infectious status of all *C. quinoa* and *N. benthamiana* source plants was verified by DAS-ELISA using specific GFLV and ArMV antibodies (data not shown). Then, nematodes were exposed to roots of two week-old healthy *C. quinoa* and *N. benthamiana* bait plants for a four-week inoculation access period (IAP). Transmission assays were performed in a growth chamber in 0.4-litre pots under controlled temperature (22°C) and light conditions (16 h photoperiod and $60 \mu\text{Einstein/m}^2/\text{s}$). The successful transmission of viruses by nematodes was verified in the roots of bait plants by DAS-ELISA using specific GFLV and ArMV antibodies (Walter *et al.*, 1987).

The three chimeric viruses were transmitted by *X. diversicaudatum* but not by *X. index* (Table 1). As expected, synthetic GFLV and wild type ArMV were transmitted by *X. index* and *X. diversicaudatum*, respectively. GFLV was not transmitted by *X. diversicaudatum*, neither was ArMV transmitted by *X. index*, confirming the transmission specificity of the two nepoviruses by their respective nematode vector (Table 1). Also, no control bait plant reacted positively to GFLV or ArMV

in DAS-ELISA when nematodes had been in contact with mock-inoculated plants, confirming that the two nematode populations were aviruliferous at the beginning of the experiment (Table 1).

These results showed that protein 2C^{CP} carries the ArMV determinants of transmission specificity. The fact that chimeric virus 2AB_(HMST)C_A with protein 2B^{MP} of GFLV origin was transmitted by *X. diversicaudatum* as well as constructs 2AB_{A(HLST)}C_A and 2AB_{A(EMSV)}C_A with protein 2B^{MP} of ArMV origin indicated also that protein 2B^{MP} is not involved in transmission specificity. This study did not address the involvement of protein 2A^{HP} in transmission specificity. However, by analogy with our previous findings with GFLV (Belin *et al.*, 1999; Andret-Link *et al.*, 2004b), protein 2A^{HP} can reasonably be excluded from transmission. This hypothesis is validated by the *X. diversicaudatum*-mediated transmission of construct 2AB_(HMST)C_A which carries a 2A^{HP} of GFLV origin.

Table 1. Transmission of wild type ArMV, synthetic GFLV and recombinant viruses by *X. diversicaudatum* and *X. index*.

Virus	<i>X. diversicaudatum</i>		<i>X. index</i>	
	No. of infected bait plants /no. of source plants ^a	Transmission rate (%)	No. of infected bait plants /no. of source plants ^b	Transmission rate (%)
ArMV-S	4/4	100	0/9	0
2ABC (GFLV)	0/4	0	11/15	73
2AB _(HMST) C _A	4/5	80	0/9	0
2AB _{A(HLST)} C _A	5/5	100	0/9	0
2AB _{A(EMSV)} C _A	4/5	80	0/9	0
Mock	0/5	0	0/5	0

^aData represent the number of recipient plants that reacted positively to ArMV in DAS-ELISA over the total number of bait plants tested. Positive bait plants had OD_{405nm} values of 0.8±0.5 compared to 0.045±0.015 for healthy controls with γ -globulins specific to ArMV after 90 min of substrate hydrolysis.

^bData represent the number of recipient plants that reacted positively to GFLV in DAS-ELISA over the total number of bait plants tested. Positive bait plants had OD_{405nm} values of 1.3±1.0 compared to 0.06±0.015 for healthy controls with γ -globulins specific to GFLV after 90 min of substrate hydrolysis.

The specificity of virus transmission is characterized by the nematode's ability to retain virus particles at specific sites within its feeding apparatus and subsequently release them into a healthy plant (Brown and Weischer, 1998). To better characterize the specificity of ArMV transmission by *X. diversicaudatum*, virus retention was investigated in nematodes collected at the end of the AAP and IAP. Nematodes were collected and tested by RT-PCR with total RNA and appropriate primer pairs specific to GFLV or ArMV (Demangeat *et al.* 2004). The RT reaction was performed with 50 pmoles of reverse primer CNR (5'-GGCAAGTGTGTCCAAAGGAC-3') hybridizing to ArMV and GFLV in a final volume of 20 μ l. The PCR reaction was performed in a final volume of 25 μ l containing 5 μ l of the RT reaction, 25 pmoles of reverse primer CNR and 25 pmoles of the ArMV-specific forward primer AF (5'-GTTACATCGAGGAGGATG-3') or the GFLV-specific forward primer GF (5'-ATGTGGAAGAGGACGGAAGT-3'). The PCR reactions were cycled in a Master thermocycler

(Eppendorf, Germany) to obtain a 860-bp GFLV-specific fragment or a 849-bp ArMV-specific fragment.

At the end of the AAP, viral specific amplicons were obtained from all *X. diversicaudatum* and *X. index* that were in contact with virus source plants (Fig. 2, panels A and C, lanes 3-6 and 9), indicating that the two nematode species ingested virus from infected plants. At the end of the IAP, ArMV-specific amplicons were obtained only from *X. diversicaudatum* that had been in contact with source plants infected either by ArMV or chimeric viruses harboring protein 2C^{CP} of ArMV origin (Fig. 2, panel B, lanes 3-6). On the contrary, no GFLV-specific amplicon was obtained from *X. diversicaudatum* exposed to GFLV-infected plants at the end of the IAP (Fig. 2, panel B, lane 9). Similarly, no ArMV-specific RT-PCR product was amplified from *X. index* exposed to ArMV-infected plants (Fig. 2, panel D, lanes 3-6). As expected, a GFLV-specific amplicon was efficiently detected from *X. index* that fed on GFLV-infected plants (Fig. 2, panels C and D, lane 9).

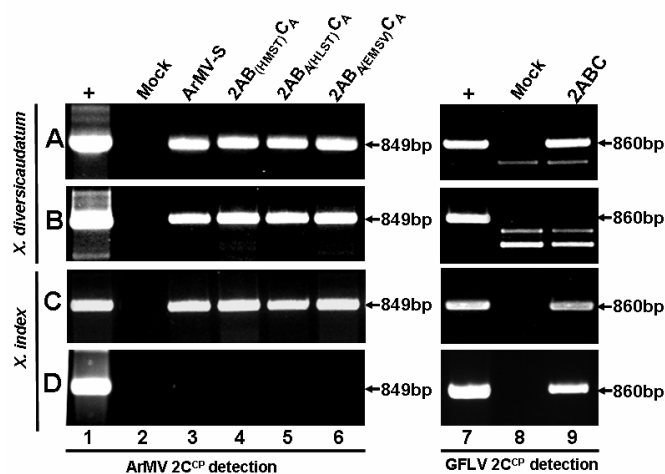


Fig. 2. Analysis by electrophoresis on 1.5% agarose gels and ethidium bromide staining of RT-PCR products obtained for GFLV, ArMV and recombinant viruses in nematodes after the acquisition access period (A and C) and inoculation access period (B and D). Batches of 30 *X. diversicaudatum* (A and B) or *X. index* (C and D) exposed to source plants infected with wild type ArMV-S (lane 3), recombinant viruses 2AB_(HMST)C_A (lane 4), 2AB_(HLST)C_A (lane 5), 2AB_(EMSV)C_A (lane 6), or synthetic GFLV (lane 9) were randomly selected and tested. Lanes 2 and 8 correspond to DNA products obtained from *X. diversicaudatum* or *X. index* that were in contact with mock inoculated source plants, respectively. Lanes 1 and 7 are positive controls (+) where RT-PCR products were obtained from purified ArMV and GFLV preparations, respectively. Total RNA was extracted from nematodes as described previously (Demangeat *et al.*, 2004) and PCR reactions were performed with ArMV (lanes 1-6) or GFLV (lanes 7-9) specific primers. Positive controls (+) were obtained with 2 µg of purified GFLV-F13 or ArMV-S.

These results are consistent with the notion that ArMV and GFLV seem to be retained by their respective nematode vector after feeding on infected plants. A specific retention of ArMV and GFLV is probably associated with receptors present in the cuticular lining of the alimentary tract of *X. diversicaudatum* and *X. index*, respectively. This situation is consistent with electron microscopy observations for which a strong correlation was seen between the absence of virus particles adsorbed on the cuticular lining of nematodes and a lack of virus transmission (Brown and Trudgill, 1983;

Brown *et al.*, 1998). The fact that wild-type and chimeric viruses were detected in both nematode species after the AAP indicated a non specific acquisition and ingestion.

Our study showed that the coat protein acts as the determinant of ArMV transmission specificity and this specificity seems to be linked to the retention capacity of the virus within nematodes. This situation may be valid for subgroup A nepoviruses as identical observations were reported for GFLV previously (Andret-Link *et al.*, 2004b). The availability of infectious RNA-2 transcripts encoding a protein 2C^{CP} of GFLV or ArMV origin and the availability of aviruliferous specimens of two distinct *Xiphinema* species represent major technical breakthroughs that are putting us in a unique position to decipher the nepovirus capsid domain(s) and/or residue(s) that interact with the feeding apparatus of the nematode vectors. Transmission assays of recombinant GFLV/ArMV viruses by both *Xiphinema* species provide also opportunities to study determinants involved in vector transmission *per se*, and not only in vector specificity, by comparing the gain or loss of transmissibility by both nematode vectors.

ACKNOWLEDGEMENTS

This work was partially supported by a competitive grant from the European Commission (QLK3-CT-2002-02140) and the French Ministère de l'Alimentation, de l'Agriculture et de la Pêche. (CASDAR project 440). P. Schellenberger and P. Andret-Link were supported by a fellowship from Département Santé des Plantes et Environnement (SPE-INRA) and the Conseil Interprofessionnel des Vins d'Alsace (CIVA).

REFERENCES

- Andret-Link P., Laporte C., Valat L., Ritzenthaler C., Demangeat G., Vigne E., Laval V., Pfeiffer P., Stussi-Garaud C., Fuchs M., 2004a. *Grapevine fanleaf virus*: Still a major threat to the grapevine industry. *Journal of Plant Pathology* **86**: 183-195.
- Andret-Link P., Schmitt-Keichinger C., Demangeat G., Komar V., Fuchs M., 2004b. The specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index* is solely determined by the viral coat protein. *Virology* **320**: 12-22.
- Andret-Link P., Fuchs M., 2005. Transmission specificity of plant viruses by vectors. *Journal of Plant Pathology* **87**: 153-165.
- Belin C., Schmitt C., Gaire F., Walter B., Demangeat G., Pinck L., 1999. The nine C-terminal residues of the grapevine fanleaf nepovirus movement protein are critical for systemic virus spread. *Journal of General Virology* **80**: 1347-1356.
- Belin C., Schmitt C., Demangeat G., Komar V., Pinck L., Fuchs M., 2001. Involvement of RNA2-encoded proteins in the specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index*. *Virology* **291**: 161-171.

- Brown D.J.F., Trugdill D.L., 1983. Differential transmissibility of arabis mosaic virus and strawberry latent ringspot virus by three populations of *Xiphinema diversicaudatum* (Nematoda : Dorylamoidea). *Revue de Nematologie* **6**: 229-238.
- Brown D.J.F., Weischer B., 1998. Specificity, exclusivity and complementarity in the transmission of plant viruses by plant parasitic nematodes: an annotated terminology. *Fundamental and Applied Nematology* **21**: 1-11.
- Brown D.J.F., Kunz P., Grunder J., Robertson W.M., 1998. Differential transmission of cherry rosette nepovirus by populations of *Longidorus arthensis* (Nematoda: Longidoridae) with a description of the association of the virus with the odontostyle of its vector. *Fundamental and Applied Nematology* **21**: 673-677.
- Demangeat G., Komar V., Cornuet P., Esmenjaud D., Fuchs M., 2004. Sensitive and reliable detection of *Grapevine fanleaf virus* in a single *Xiphinema index* nematode vector. *Journal of Virological Methods* **122**: 79-86.
- Harrison B.D., Murrant A.F., 1977. Nematode transmissibility of pseudorecombinant isolates of tomato black ring virus. *Annals Applied of Biology* **86**: 209-212.
- Harrison B.D., Murrant A.F., Mayo M.A., Roberts I.M., 1974. Distribution of determinants for symptom production, host range and nematode transmissibility between the two RNA components of raspberry ringspot virus. *Journal of General Virology* **22**: 233-247.
- Huss B., Walter B., Fuchs M., 1989. Cross-protection between arabis mosaic virus and grapevine fanleaf virus isolates in *Chenopodium quinoa*. *Annals Applied of Biology* **114**: 45-60.
- Le Gall O., Christian P., Fauquet C.M., King A.M.Q., Knowles N.J., Nakashima N., Stanway G., Gorbalenya A.E., 2008. *Picornavirales*, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Archives of Virology* **153**: 715-727.
- Sanfaçon H., Wellink J., Le Gall O., Karasev A., Van der Vlugt R., Wetzel T., 2009. Secoviridae: a proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus *Torradorvirus*. *Archives of Virology* **154**: 899-907.
- Vigne E., Marmonier A., Fuchs M., 2008. Multiple interspecies recombination events within RNA2 of *Grapevine fanleaf virus* and *Arabis mosaic virus*. *Archives of Virology* **153**: 1771-1776.
- Viry M., Serghini M. A., Hans F., Ritzenthaler C., Pinck M. and Pinck L., 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. *Journal of General Virology* **74**: 169-174.
- Vuittenez A., Munck M.C., Kuszala J., 1964. Souches de virus à haute agressivité isolées de vignes atteintes de dégénérescence infectieuse. *Etudes de Virologie Appliquées* **5**: 69-78.
- Van Lent J.M.W., Schmitt-Keichinger C., 2006. Viral movement proteins induce tubule formation in plant and insect cells. In: Volkmann D., Baluska F., Barlow P.W. (eds). *Cell-cell Channels*, pp. 160-175. Springer-Landes Biosciences, Georgetown, WDC, USA.
- Walter B., Etienne L., 1987. Detection of the *Grapevine fanleaf viruses* away from the period of vegetation. *Journal of Phytopathology* **120**: 355-364.
- Wetzel T., Fuchs M., Bobko M., Krczal G., 2002. Size and sequence variability of the *Arabis mosaic virus* protein 2A. *Archives of Virology* **147**: 1643-1653.
- Wetzel T., Chisholm J., Bassler A., Sanfaçon H., 2008. Characterization of proteinase cleavage sites in the N-terminal region of the RNA1-encoded polyprotein from *Arabis mosaic virus* (subgroup A nepovirus). *Virology* **375**: 159-169.



Chapitre 4

*Régions de la protéine de capsid du GFLV impliquées dans la transmission par son nématode vecteur *X. index**

1. Présentation

L'article ci-après (article 2) présente le travail effectué dans le premier axe de recherche et fait directement suite au travail de thèse de P. Andret-Link (Andret-Link, 2005). La stratégie d'identification des régions de la CP du GFLV déterminant la transmission est basée sur l'hypothèse que les résidus nécessaires à l'interaction entre le nématode et les particules virales sont présents à la surface du virion et différents entre le GFLV et l'ArMV. Sur la base de cette hypothèse, trois régions putatives ont été identifiées en utilisant un modèle 3D de la particule virale du GFLV élaboré à partir de la structure résolue du *Tobacco ringspot virus* (TRSV). Cette stratégie avait été initiée par P. Andret-Link (2005) lors de son travail de thèse et avait abouti à la construction de capsides chimères GFLV/ArMV ciblant les trois régions de la capsid identifiées. Ces trois régions sont nommées dans l'Article 2 R2 (11 résidus), R4 (7 résidus), et R5 (9 résidus). Seule la substitution de la région R2 avait permis d'obtenir une chimère infectieuse *in planta*, condition obligatoire pour faire des tests de transmission par nématodes.

Deux autres régions R1 (7 résidus) et R3 (4 résidus) (voir article 2) sont également exposées à la surface de la capsid d'après le modèle 3D, très divergentes entre le GFLV et l'ArMV et n'avait pas pu être prise en compte au cours du travail précédent. Ces deux régions remplissant les critères de sélection, il était donc intéressant de caractériser leur implication dans la transmission.

Dans ce travail, j'ai réalisé toutes les chimères ciblant les régions R1 et R3 et complété les constructions ciblant les régions R2, R4, R5 produites au cours du travail de thèse de P. Andret-Link. Des analyses préliminaires des propriétés biologiques des chimères et un test de transmission par *X. index* avaient déjà été conduits. J'ai confirmé les résultats, complété l'étude en ajoutant à l'analyse les recombinants R1 et R3, et évalué la transmissibilité des chimères par *X. diversicaudatum* (voir chapitre 3). L'échange de la région R1 a abouti à un virus chimérique infectieux *in planta*, contrairement à l'échange de R3 qui s'est avéré létal pour le virus. Il a ainsi été possible d'évaluer l'implication de deux régions (R1 et R2) de la CP du GFLV dans le processus de la transmission.

2. Article 2 : A stretch of 11 amino acids in the β B - β C loop of the coat protein of *Grapevine fanleaf virus* is essential for transmission by the nematode *Xiphinema index*

Article soumis à Journal of Virology

**A stretch of 11 amino acids in the β B - β C loop of the coat protein of
Grapevine fanleaf virus is essential for transmission by the nematode
*Xiphinema index***

Pascale Schellenberger,^{1,2} Peggy Andret-Link,¹ Corinne Schmitt-Keichinger,² Marc Bergdoll,²
Aurélie Marmonier,¹ Emmanuelle Vigne,¹ Olivier Lemaire,¹ Marc Fuchs,³ Gérard Demangeat¹
and Christophe Ritzenthaler^{2*}

¹Institut National de la Recherche Agronomique, UMR 1131 INRA/Université de Strasbourg, 28 rue de Herrlisheim, 68021 Colmar, France. ²Institut de Biologie Moléculaire des Plantes du CNRS, Université de Strasbourg, 12 rue du Général Zimmer, 67084 Strasbourg, France. ³Department of Plant Pathology, Cornell University, New York State Agricultural Experiment Station, Geneva, NY 14456, USA.

*Corresponding author: Mailing address: Institut de Biologie Moléculaire des Plantes du CNRS, Université de Strasbourg, 12 rue du Général Zimmer, 67084 Strasbourg, France, Phone +33 3 88 41 72 57, Fax + 33 3 88 61 44 42, Email: ritzenth@unistra.fr

Running title: Coat protein determinants of GFLV transmission

Keywords: *Grapevine fanleaf virus*, *Arabidopsis mosaic virus*, *Comovirinae*, coat protein determinants, nematode transmission, 3D structure modeling, chimeric coat protein, RNA protection assay.

ABSTRACT

Grapevine fanleaf virus (GFLV) and *Arabidopsis mosaic virus* (ArMV) from the genus *Nepovirus*, family *Secoviridae* cause a severe degeneration of grapevines. GFLV and ArMV have a bipartite RNA genome and are transmitted specifically by the ectoparasitic nematodes *Xiphinema index* and *Xiphinema diversicaudatum*, respectively. The transmission specificity of both viruses maps to their respective RNA2-encoded coat protein (CP). To further delineate the GFLV CP determinants of transmission specificity, three dimensional (3D) homology structure models of virions and CP subunits were constructed based on the crystal structure of *Tobacco ringspot virus*, the type member of the genus *Nepovirus*. The 3D models were examined to predict amino acids that are exposed at the external virion surface, highly conserved among GFLV isolates but divergent between GFLV and ArMV. Five short amino acid stretches that matched these topographical and sequence conservation criteria were selected and substituted in single and multiple combinations by their ArMV counterparts in a GFLV RNA2 cDNA clone. Among the 21 chimeric RNA2 molecules engineered, transcripts of only three of them induced systemic plant infection in the presence of GFLV RNA1. Nematode transmission assays of the three viable recombinant viruses showed that swapping a stretch of i) 11 residues in the β B- β C loop near the icosahedral 3-fold axis abolished transmission by *X. index* but was insufficient to restore transmission by *X. diversicaudatum*, ii) seven residues in the β E- α B loop had a neutral effect on transmission by the two *Xiphinema* species. This study provided new insights into GFLV CP determinants of nematode transmission.

INTRODUCTION

The transmission of a virus from one plant to another is a key step in the virus life cycle. For a majority of plant viruses, transmission is dependant upon vectors, mostly arthropods and soil-inhabiting nematodes and plasmodiophorids (15, 17, 28). A successful transmission depends on the competency of the vector to acquire virions, bind them to specific retention sites and release them in a recipient plant. Virus transmission by a vector is often characterized by some degree of specificity; it can be broad or narrow but it is a prominent feature for numerous plant viruses and their vectors (2). The viral determinants engaged in vector transmission involve the coat protein (CP) that either binds directly to vector ligands or associates with non-structural viral proteins, like a helper component (HC), to create a molecular bridge between virions and the vector (28, 49). For both strategies, only a very limited number of amino acids on either the CP or the HC contribute to the transmission process and to vector specificity.

For the aphid-transmitted *Cucumber mosaic virus* (CMV, genus *Cucumovirus*, family *Cucumoviridae*), the CP is the sole viral determinant of transmission (12, 14) and a negatively charged

loop of eight residues located on the surface of virions is directly involved in vector interactions (21, 29). For potyviruses such as *Tobacco vein mottling virus* (TVMV), the CP and HC are both involved in transmission by aphids. The amino acid motifs DAG located near the N-terminus of the CP, KITC located near the N-terminus of the HC-Pro and PTK located into the central part of the HC-Pro, are essential for transmission (40). For *Cauliflower mosaic virus* (CaMV, genus *Caulimovirus*, family *Caulimoviridae*), a single amino acid change in protein P2, which acts as HC, is sufficient to abolish aphid transmission (27). In the case of nematode-transmitted viruses, a single mutation in the HC protein 2b of *Pea early-browning virus* (PEBV, genus *Tobravirus*, unassigned family) is sufficient to prevent transmission (43). Similarly, a 15 amino acid deletion at the C-terminus of the CP abolishes transmission of PEBV (25). For members of the genus *Nepovirus* in the family *Secoviridae*, little information is available on viral determinants that govern nematode transmission (23). Transmission assays of pseudorecombinants between *Raspberry ringspot virus* (RpRSV) and *Tobacco black ring virus* (TBRV) showed the involvement of RNA2-encoded proteins and suggested a role of the CP in transmission specificity (reviewed in 3). Based on sequence alignments, the residue in position 219 of the RpRSV CP was proposed to be exposed at the surface of particles and involved in nematode transmission (38). However, no experimental evidence is available to confirm these hypotheses. Thus, our knowledge of the viral determinants of nepovirus transmission is rather scarce.

Grapevine fanleaf virus (GFLV) and *Arabidopsis mosaic virus* (ArMV), two closely related species of the genus *Nepovirus*, family *Secoviridae* (36, 37), are responsible for a severe degeneration of grapevines that occurs in most vineyards worldwide (3). GFLV and ArMV have a bipartite RNA genome and share a similar genome organization, as well as a relatively high level of sequence similarity (44). The RNA1 of both viruses encodes replication and protein maturation functions, and the RNA2 encodes a protein involved in RNA2 replication, which is referred to as a homing protein ($2A^{HP}$), the movement protein ($2B^{MP}$) and the coat protein ($2C^{CP}$) (3, Fig. 3). The nematode transmission of GFLV and ArMV illustrates the type of extreme specificity that can exist between vectors and viruses (10). GFLV is transmitted exclusively by *Xiphinema index* and ArMV is transmitted specifically by *X. diversicaudatum*. The transmission specificity of GFLV (4) and ArMV (26) maps to their respective $2C^{CP}$ but no information is available on $2C^{CP}$ residues involved in vector interactions.

The main objective of our study was to delineate $2C^{CP}$ amino acids involved in GFLV transmission by *X. index* and advance our knowledge of molecular features of nepovirus transmission. Three dimensional (3D) homology-based models of GFLV virions and CP subunits were constructed using the 3.5 Å crystal structure of *Tobacco ringspot virus* (TRSV) as a template (11). TRSV is the type member of the genus *Nepovirus* to which GFLV and ArMV belong (37) and is transmitted by the nematode *X. americanum*. From our structural models, residues exposed at the external capsid surface, highly conserved among GFLV isolates but divergent between GFLV and ArMV isolates were identified and selected for a reverse genetics approach to determine their involvement in vector

transmission. Our results indicate that a stretch of 11 residues predicted in the β B- β C loop of the CP B domain is essential for *X. index*-mediated transmission of GFLV. To our knowledge, this represents the first description of structural amino acids essential for nepovirus transmission by nematodes.

MATERIALS AND METHODS

Virus strains. GFLV strain F13 (46) and ArMV strain S (18) were isolated from naturally infected grapevines and propagated in *Chenopodium quinoa*, a systemic host for both viruses. The sequence of the two genomic RNAs of both virus strains is determined (22, 33, 39, Ritzenthaler *et al.*, unpublished). Full-length cDNA clones of GFLV-F13 RNA1 (plasmid pMV13) and RNA2 (plasmid pVec2ABC) are available for *in vitro* synthesis of transcripts (45). ArMV-S was used as control in nematode transmission assays and a cDNA of ArMV-S RNA2-U (22) was used in mutagenesis experiments.

GFLV 3D structure model building. A protein/protein BLAST (1) search for potential structural templates was conducted using the query GFLV-F13 CP sequence (GI25013721) against sequences available in Protein Data Bank. Clustal X (41) was used to generate an alignment of the query sequence (Accession n° NC_003623; 505 amino acids), template sequence (GI3913279; 513 amino acids) and ArMV CP sequence (GI757527; 504 amino acids). The structure described in TRSV 1A6C was used as a template for GFLV model building using MODELLER (35). Hexamers were built rather than isolated monomers because virus capsids are highly constrained by high symmetry, extensive contacts between monomers and the fact that approximately one third of the total accessible surface residues are predictably buried by direct neighbors.

The GFLV CP amino acid stretches presumably involved in transmission by *X. index* were identified by combining information from the GFLV capsid 3D model, including a topographic analysis of surface residues, and from an alignment of the CP sequence of 219 GFLV and 20 ArMV isolates (Table S1). Sequences were aligned with Align X (Vector NTI).

Development of chimeric GFLV RNA2. Plasmid pVec2ABC carrying a full-length cDNA copy of GFLV-F13 RNA2 was used as template to develop chimeric 2C^{CP} genes (8). For cloning purposes, an *Acc65I* restriction site was introduced into pVec2ABC in position 2,678-2,683 (nucleotide positions are given according to the GFLV-F13 RNA2 sequence, GenBank accession n° NC_003623) by site directed PCR-mediated mutagenesis. The plasmid generated was named pVec_{Acc65I}2ABC. Introducing *Acc65I* modified the nucleotide but not the amino acid composition of the wild type GFLV-F13 2C^{CP}.

Previous work showed that functional GFLV/ArMV RNA2 recombinants can be produced, providing a functional interaction between proteins 2B^{MP} and 2C^{CP} is maintained for cell-to-cell movement (7, 8). Mutations affecting regions R1 (nts 2,284-2,302), R2 (nts 2,609-2,640), R4 (nts 2,819-2,842) and R5 (nts 2,937-2,965) of the GFLV 2C^{CP} gene were produced by primer overlap extension mutagenesis (16), using primers described in Table S2. The mutation affecting region R3 (2,666-2,677) only required one PCR step because of the presence of the *Acc65I* restriction site into the reverse primer. Each PCR reaction was carried out as described previously (4), except that the annealing time was reduced to 45 sec at 60°C and the elongation time was adapted to the size of the PCR products. DNA amplicons obtained by PCR were digested with appropriate restriction enzymes (Table S2), and cloned into the corresponding sites in pVec_{Acc65I}2ABC to yield pVec_{Acc65I}2ABC_{G1}, pVec_{Acc65I}2ABC_{G2}, pVec_{Acc65I}2ABC_{G3}, pVec_{Acc65I}2ABC_{G4} and pVec_{Acc65I}2ABC_{G5} (G as subscript indicates the GFLV origin of the

2C^{CP} and 1 to 5 the mutated regions R1 to R5, respectively). For simplicity, transcripts and recombinant viruses derived from this series of constructs were referred to as G1, G2, G3, G4 and G5.

To produce mutants carrying combinations of regions R1, R2 and R3, chimeric plasmids pVec_{Acc65I}2ABC_{G1} and pVec_{Acc65I}2ABC_{G2} were used as matrix for a second primer overlap - or single - PCR mutagenesis and the resulting PCR fragments were cloned into pVec_{Acc65I}2ABC, using *Ngo*MIV and *Acc65I* to produce pVec_{Acc65I}2ABC_{G12}, pVec_{Acc65I}2ABC_{G13} and pVec_{Acc65I}2ABC_{G23} (12, 13 and 23 as subscript indicate the combination of the mutated regions 1 and 2, 1 and 3, and 2 and 3, respectively). For simplicity, transcripts and recombinant viruses derived from this second series of constructs were referred to as G12, G13 and G23.

To produce mutants carrying further combinations of swapped regions, mutant plasmids pVec_{Acc65I}2ABC_{G1}, pVec_{Acc65I}2ABC_{G2}, pVec_{Acc65I}2ABC_{G3}, pVec_{Acc65I}2ABC_{G4} and pVec_{Acc65I}2ABC_{G5} were digested with appropriate restriction enzymes (Table S3) and restriction fragments containing the mutated region(s) was cloned sequentially into plasmids encoding a GFLV 2C^{CP} gene already carrying a mutation (Table S3). For simplicity, transcripts and recombinant viruses derived from this third series of constructs were referred to as G24, G25, G45 and G245.

Plasmid pVec2AB_{U-4-9}C was engineered earlier (4) from pVec2AB_{U9}C (8) to produce chimeric RNA2 constructs for which gene 2C^{CP} is of GFLV origin and gene 2B^{MP} is of ArMV origin, except amino acids in position -2 and -6 downstream of the R/G cleavage site that remain of GFLV origin (4) in order to maintain a functional interaction between proteins 2B^{MP} and 2C^{CP} for systemic plant infection (8). Plasmid pVec2AB_{U-4-9}C was renamed in this study pVec2AB_AC_G (A as subscript indicates the chimeric nature of protein 2B^{MP}). The complete CP coding sequence of G1, G2, G4, G12, G24, G25, G45 and G245 was cloned into plasmid pVec2AB_AC_G using appropriate restriction enzymes (Table S3) to produce pVec2AB_AC_{G2}, pVec2AB_AC_{G4}, pVec2AB_AC_{G12}, pVec2AB_AC_{G24}, pVec2AB_AC_{G25}, pVec2AB_AC_{G45} and pVec2AB_AC_{G245}, respectively. For simplicity, transcripts and recombinant viruses derived from this third series of constructs were referred to as AG1, AG2, AG4, AG12, AG24, AG25, AG45 and AG245.

Plasmid pVec2AB_{U-6}C_U (4) was used in this study to introduce CP gene region R2 of GFLV origin in an ArMV CP gene and renamed pVec2AB_AC_A (A as subscript indicates the chimeric nature of protein 2B^{MP} and the ArMV nature of protein 2C^{CP}). Mutagenesis of CP region R2 and cloning into pVec2AB_AC_A to yield pVec2AB_AC_{A2} was performed as described in Table S2. For simplicity, transcripts derived from this construct are referred to as AA2.

The integrity of the 21 GFLV chimeric RNA2 clones produced for this study was verified by DNA sequencing.

***In vitro* transcription of GFLV cDNA clones and mechanical inoculation of plants.** GFLV RNA1 and RNA2, and GFLV chimeric RNA2 were obtained by *in vitro* transcription (8) and were mechanically inoculated to *C. quinoa*. Recombinant viruses were propagated in *Nicotiana benthamiana* by using crude sap of infected *C. quinoa* leaves as inoculum. Virus infection was assessed in uninoculated apical leaves of test plants two to three weeks post-inoculation by double-antibody sandwich (DAS)-enzyme-linked immunosorbent assay (ELISA) with γ -globulins specific to GFLV and ArMV (47). Samples were considered positive if their optical density (OD_{405nm}) readings were at least three times those of healthy controls after 120 min of substrate hydrolysis.

Functionality of GFLV chimeric RNA2 in *Chenopodium quinoa* protoplasts. The functionality of chimeric RNA2 transcripts was evaluated by transfecting *C. quinoa* protoplasts in the presence of GFLV RNA1 transcripts. Viral RNA replication and encapsidation were assessed by RNase protection experiments, as

described previously (8), except that riboprobes used for northern hybridization corresponded to nts 4,623-5,810 of RNA1 (5' half of gene 1E^{POL}) and nts 233-1,006 of RNA2 (gene 2A^{HP}).

Nematode transmission tests. Nematode transmission assays were performed with *N. benthamiana* for *X. index* (4) and *C. quinoa* for *X. diversicaudatum* (26). Transmission assays used a minimum of 5-6 plants and were repeated at least once. Synthetic GFLV, *i.e.* RNA derived from pMV13 and pVec_{Acc651}2ABC, referred to as 2ABC and wild type ArMV-S were used as controls in transmission assays with *X. index* and *X. diversicaudatum*. The presence of GFLV and ArMV was assessed in leaves of source plants by DAS-ELISA (4, 26).

Detection of GFLV and ArMV in nematodes. The presence of GFLV and ArMV was verified in nematodes by reverse-transcription (RT)-polymerase-chain reaction (PCR) with a GFLV/ArMV consensus reverse primer EVAGR (5'-GGCAAGTGTGTCCAAAGGAC-3') and a GFLV-specific forward primer GF (5'-ATGTGGAAGAGGACGGAAGT-3') or a ArMV-specific forward primer AF (5'-GTTACATCGAGGAGGATG-3') to amplify 860-bp and 849-bp specific fragments, respectively (13).

Characterization of GFLV chimeric RNA2 progeny. The progeny of chimeric GFLV RNA2 was characterized in infected plants by immuno-capture (IC)-RT-PCR (4, 7, 8) with primers 208 (Table S2) and 398 (5'-TGGGARARYRTNGAGGAAC-3') followed by direct sequencing (4). Sequences were analyzed with ContigExpress (Vector NTI).

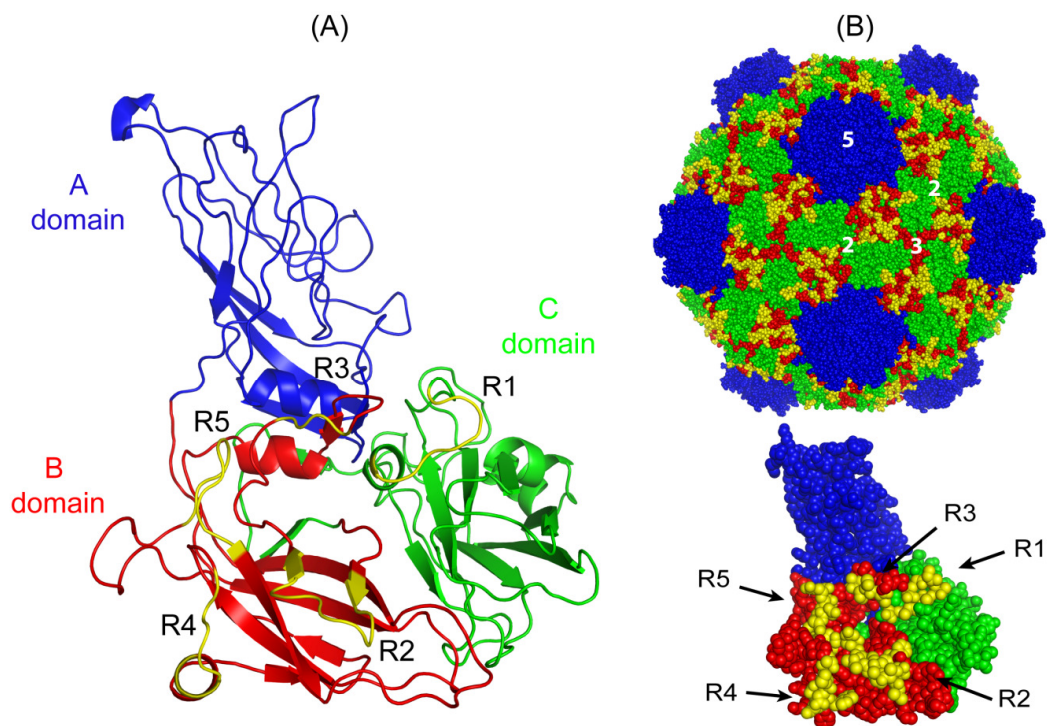


FIG. 1. Homology-based model of GFLV and 2C^{CP} regions targeted for mutagenesis. (A) Ribbon representation of the GFLV CP deduced from the TRSV crystal structure (PDB ID: 1A6C). The predicted A, B and C structural domains are indicated in blue, red and green, respectively. The five regions selected for mutational analysis due to their predicted surface location and divergence between GFLV and ArMV appear in yellow. Regions R2, R3, R4 and R5 are located in the B domain whereas region R1 is located in the C domain. (B) A space-filling model of GFLV particle quaternary structure viewed down the icosahedral 2-fold axis (top) and GFLV 2C^{CP} subunit tertiary structure (bottom) showing the close proximity of regions R1 to R5. The icosahedral 5-fold (5) 3-fold (3) and 2-fold (2) axes are labeled for one icosahedral asymmetric unit (top). Figures were generated using PYMOL (www.pymol.org).

RESULTS

Construction of a GFLV homology-based 3D model and search for surface residues potentially involved in *X. index* transmission. In the absence of a resolved structure for GFLV and ArMV, a 3D structural model of GFLV virions was constructed by examining Protein Data Bank for potential structural templates using the query GFLV-F13 CP sequence. The closest structural template to the GFLV CP was the CP of TRSV (PDB-ID: 1A6C) with 25% sequence identity and 43% sequence similarity. While the crystal structure of TRSV had a score of 219 positive residues out of 513, the second hit, i.e. the crystal structure of Tha0252 from *Thermus thermophilus* Hb8 (2DKF), only had a score of 29 for 62 amino acids, making the latter unsuitable for homology modeling. To identify GFLV 2C^{CP} amino acids involved in the transmission by *X. index*, we hypothesized that candidate residues are likely exposed at the external surface of virions, different between GFLV and ArMV isolates but highly conserved among GFLV isolates. The latter criterion was not considered absolute because no information is available on nematode transmissibility of most of the 219 GFLV and 20 ArMV variants for which CP sequences are available in GenBank (Table S1).

The crystal structure of TRSV (PDB ID: 1A6C) (11) was used to map putative surface residues and divergence in deduced protein sequences among the CP of TRSV, GFLV and ArMV (Fig. S1). Also, 3D homology models of GFLV virion and CP subunit were constructed from the crystal structure of TRSV (Fig. 1). The virion icosahedral lattice is of a pseudo T = 3 symmetry and one CP consists of 3 trapezoid shaped β barrel domains referred to as the C, B and A domains from the N- to the C-terminus, as shown previously for TRSV (11). Out of the 156 non-conserved residues between GFLV-F13 and ArMV-S, 59 amino acids were predicted to be exposed at the external surface of the CP (highlighted in grey in Fig. S1) and nearly half of those (27 out of the 59) clustered within the central B domain. Considering this high number of divergent surface residues, we restricted our reverse genetics approach to areas with the highest concentration of divergent residues rather than to single isolated residues. Five regions termed R1 to R5 ranging from 4 to 11 residues and displaying between 0 (region R3) to 57% (region R4) amino acid sequence identity between GFLV and ArMV were identified and selected for mutagenesis experiments (Fig. 2). The five candidate regions all located to predicted loops connecting β sheets within the B domain, except region R1 that mapped to the C domain (Fig. 1). The five CP regions showed a high degree of amino acid conservation (Fig. 2). When displayed in GFLV CP tertiary and quaternary structure models, the five regions mapped close together, with R1 closest to R3 and R2 closest to R5 (Fig. 1). The five regions were located also alongside a canyon-like structure between the 5-fold and 3-fold axes of virions (Fig. 1). Importantly, no direct contact of the selected regions between the different subunits could be deduced from our 3D models, suggesting that the selected regions are unlikely to be involved in CP-CP interactions necessary for virion assembly or stability.

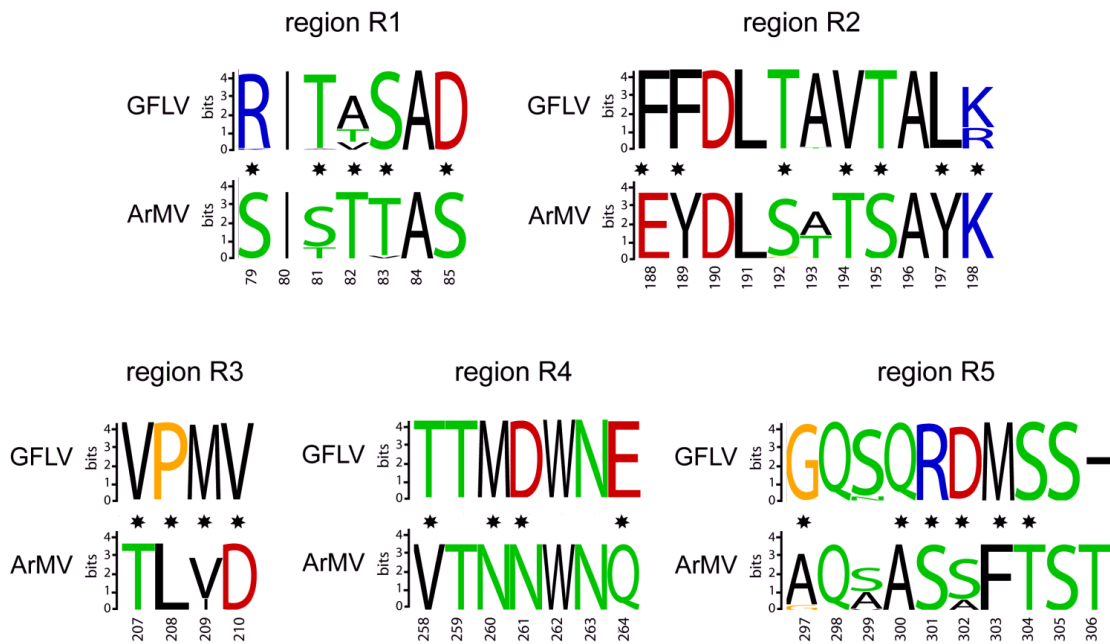


FIG. 2. Logo graphical representations of the amino acid sequence conservation within CP regions R1 to R5 of GFLV and ArMV. The sequences of 219 GFLV and 20 ArMV variants (Table S1) were used to create a logo graphical representation. The different residues at each position in GFLV CP amino acid sequences (top of each panel) and ArMV CP amino acid sequences (bottom of each panel) are scaled according to their frequency. The CP amino acid positions are given on the X-axis of each panel. Black stars indicate residues that differ between GFLV and ArMV. The black line at the end of GFLV region R5 (bottom right panel) represents a gap in the GFLV and ArMV 2C^{CP} sequence alignment. Residues are colored according to their biochemical properties: positively charged in blue; negatively charged in red; hydrophobic in black; polar uncharged in green and the proline, glycine and cysteine in orange. CP sequence alignments were created using AlignX (VectorNTI) and logo representations were obtained using Weblogo (<http://weblogo.berkeley.edu/logo.cgi>).

Engineering of recombinant viruses. A reverse genetics approach was used to investigate the involvement of GFLV candidate CP regions R1 to R5 in nematode transmission. Chimeric CP genes were obtained by substituting GFLV residues within regions R1 to R5 by their ArMV counterparts using site-directed mutagenesis of full-length infectious cDNA clones of GFLV RNA2 (45). Regions were substituted individually, in pairs or in triple combination (Fig. 3B). In addition, to circumvent possible defects in systemic movement of recombinant viruses *in planta* (7, 8), recombinants carrying single or combinations of ArMV CP regions R1 to R5 were also generated in an infectious cDNA clone of GFLV RNA2 for which the GFLV 2B^{MP} gene was replaced by its ArMV counterpart (Fig. 3C), except for Leu³⁴³ (in position -6 downstream of the R/G cleavage site between proteins 2B^{MP} and 2C^{CP}) and Val³⁴⁷ (position -2 downstream of the R/G cleavage site) of protein 2B^{MP} that remained of GFLV origin to maintain infectivity (4, 8). Altogether, 21 CP recombinants were generated, 12 in a full GFLV RNA2 genetic background (Fig. 3B), eight in a partial GFLV RNA2 genetic background that included an ArMV protein 2B^{MP} (Fig. 3C), and one in a partial GFLV RNA2 genetic background that included ArMV proteins 2B^{MP} and 2C^{CP} (Fig. 3D).

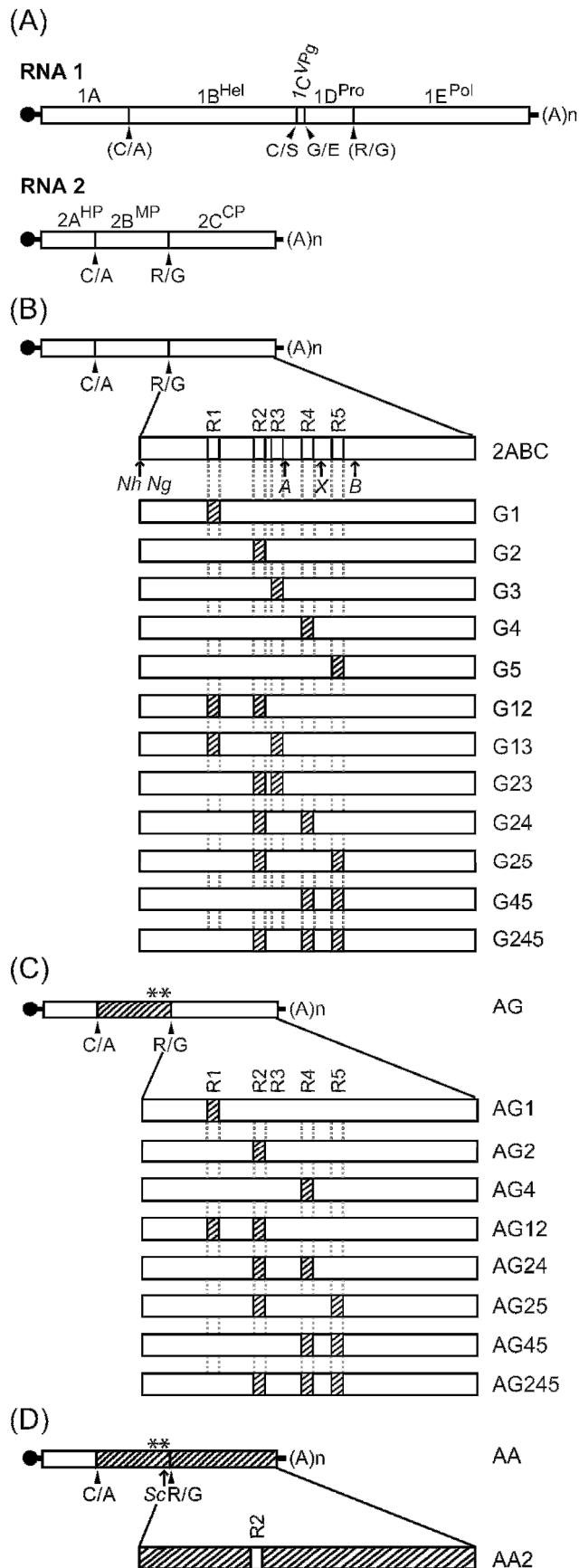


FIG. 3. Schematic representation of GFLV chimeric RNA2 used in this study. (A) Genetic organization of GFLV-F13 RNA1 and RNA2. Open boxes indicate open reading frames. The 5' and 3' untranslated regions are denoted by a single line and the VPg is represented by a black circle. Arrowheads point out the polyprotein cleavage sites. Dipeptides shown in parenthesis were suggested based on sequence analysis but have not been demonstrated experimentally. The name of the processed proteins is given on top of the boxes: Hel = putative helicase, VPg = viral protein genome linked, Pro = protease, Pol = polymerase, HP = putative homing protein, MP = movement protein and CP = coat protein. (B-D) Chimeric RNA2 molecules generated by swapping regions R1 to R5 between GFLV and ArMV in a (B) full GFLV RNA2 genetic background, (C) partial GFLV RNA2 genetic background for which protein 2B^{MP} is of ArMV origin, except amino acids 343 and 347 (-6 and -2 upstream of the R/G cleavage site between proteins 2B^{MP} and 2C^{CP}) which are of GFLV origin (asterisks above protein 2B^{MP}), or (D) partial GFLV RNA2 genetic background where proteins 2C^{CP} and 2B^{MP} are of ArMV origin, except aminoacids 343 and 347. Hatched boxes designate ArMV sequences. Restriction sites used for cloning are positioned with arrows: *Nh* = *NheI* (2,050-2,055), *Ng* = *NgomI*V (2,054-2,059), *A* = *Acc65I* (2,678-2,683), *X* = *XmaI* (2,852-2,857), *B* = *BglIII* (3,055-3,060), *Sc* = *SacI* (2,023-2,028).

Infectivity assays of recombinant viruses *in planta*. The infectivity of *in vitro* synthesized GFLV RNA1 and chimeric RNA2 transcripts, and the ability of corresponding recombinant viruses to cause systemic plant infection is essential to obtain infected plant roots for nematode transmission assays. Out of the 21 GFLV recombinants tested, only recombinants G1, AG1, G2, and AG2 caused a systemic infection *in planta* with typical GFLV symptoms and DAS-ELISA positive uninoculated apical leaves (Table 1). Recombinants exhibited a 2-5 day delay in symptom development compared to synthetic wild type GFLV, *i.e.* 2ABC. However, this difference nearly vanished upon mechanical re-inoculation of *C. quinoa* with crude sap from initially infected plant tissue, except for recombinant G2 which, for unknown reasons, remained recalcitrant to re-inoculation in *C. quinoa* and *N. benthamiana* and was therefore abandoned from further studies. Recombinant viruses G1, AG1 and AG2 were transferred successfully to *N. benthamiana* by mechanical inoculation and the genetic stability and integrity of their RNA2 progeny was confirmed by IC-RT-PCR and sequencing. Importantly, neither reversion to wild type GFLV sequences nor compensatory mutations outside gene 2C^{CP} were ever observed.

Effect of CP mutations on encapsidation. The defect in *in planta* infectivity observed for most recombinant viruses engineered in this study could result from deficiencies in replication, encapsidation and/or cell-to-cell movement. To discriminate between these hypotheses, *C. quinoa* protoplasts were transfected with viral transcripts and total RNA extracted 72 hours post-transfection were analyzed by northern hybridization using GFLV RNA1 and RNA2 riboprobes. Results showed that all chimeric RNA were competent for replication (Table 1). Following replication in protoplasts, the ability of chimeric protein 2C^{CP} to encapsidate viral RNA was assessed under conditions that allows to distinguish RNase-protected, encapsidated viral RNA (P condition) from total, encapsidated and nonencapsidated viral RNA (T condition) (8). As expected, under T condition, progeny viral RNA was detected for all GFLV RNA1 and RNA2 combinations (Fig. 4 and Table 1), indicating that loss of infectivity was not due to replication. Under P condition, the only single recombinants in which progeny viral RNA could be detected were those affecting regions R1, R2 and R4, independently of the origin of the 2B^{MP} gene (G1, G2, G4, AG1, AG2 and AG4, Fig. 4 and Table 1), indicating encapsidation of the corresponding viral genomic RNA. In contrast, no viral RNA protection was observed for recombinants G3 and G5 as for mutant 2AB for which the CP gene is deleted (Fig. 4). RNase protection assays further indicated a deficiency in viral RNA encapsidation for all double or triple recombinants, including the ones combining regions that individually could be exchanged with no effect on encapsidation, *i.e.* G12, G24 and AG24 (Fig. 4 and Table 1). Finally, no protected RNA could be detected for recombinant AA2 (Fig. 4), showing that even if the ArMV region R2 could be placed in a GFLV 2C^{CP} protein without effect on encapsidation (recombinant G2 and AG2), the reverse swapping was deleterious.

TABLE 1. Infectivity of chimeric RNA2 transcripts in the presence of GFLV RNA1

Transcripts	Plant infection		Protoplast infection	
	Symptoms ^a	DAS-ELISA ^b	Replication ^c	Encapsidation ^d
2ABC	+	+	+	+
G1	+	+	+	+
G2	+ ^e	+ ^e	+	+
G3	-	-	+	-
G4	-	-	+	+
G5	-	-	+	-
G12	-	-	+	-
G13	-	-	+	nd
G23	-	-	+	-
G24	-	-	+	-
G25	-	-	+	-
G45	-	-	+	-
G245	-	-	+	-
AG	+	+	+	nd
AG1	+	+	+	nd
AG2	+	+	+	+
AG4	-	-	+	+
AG12	-	-	+	nd
AG24	-	-	+	-
AG25	-	-	+	-
AG45	-	-	+	-
AG245	-	-	+	-
AA	+	+	+	nd
AA2	-	-	+	-

^aSymptoms of mosaic, vein yellowing and deformation were observed on apical leaves of *C. quinoa* after mechanical inoculation with transcripts of GFLV RNA1 and chimeric RNA2. (+) = symptoms; (-) = no symptoms.

^bPlants tested positive (+) or negative (-) for viruses by DAS-ELISA in three independent experiments 15 days after mechanical inoculation of GFLV RNA1 and chimeric RNA2 transcripts. OD_{405nm} values of infected plants were 0.8 ± 0.2 compared to 0.07 ± 0.005 for healthy controls after 120 min of substrate hydrolysis. Positive plants for 2ABC had OD_{405nm} values of 1.5 ± 0.2 compared to 0.07 ± 0.005 for healthy controls after 120 min of substrate hydrolysis.

^cReplication was determined in *C. quinoa* protoplasts by northern hybridization analysis after transfection with GFLV RNA1 and chimeric RNA2 transcripts.

^dEncapsidation was determined by northern hybridization analysis following RNase protection assays on *C. quinoa* protoplasts transfected with GFLV RNA1 and chimeric RNA2. nd = not determined.

^eSystemic infection was only observed in a few plants and recombinant G2 could not be transferred from the initially infected *C. quinoa* to healthy *C. quinoa* or *N. benthamiana* by mechanical inoculation.

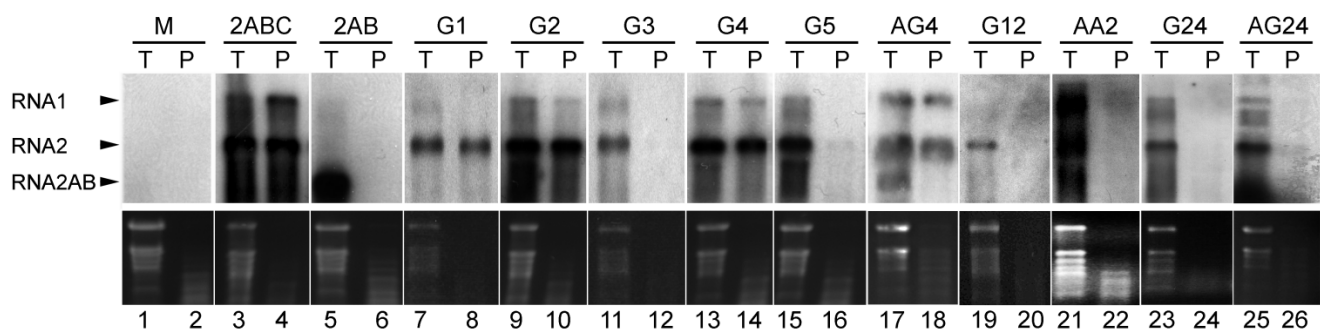


FIG. 4. Northern blot analysis of viral RNA from transfected *C. quinoa* protoplasts. Total RNA was extracted to yield both encapsidated and nonencapsidated RNAs (lanes T with odd numbers) or RNase-protected RNAs (lanes P with even numbers). Protoplasts were transfected with transcripts of GFLV RNA1 and transcripts of RNA2: 2ABC (3 and 4); RNA2 deleted of the entire CP gene, 2AB (5 and 6); chimeric RNA2 with single region R1 to R5 of ArMV origin: G1 (7 and 8), G2 (9 and 10), G3 (11 and 12) G4 (13 and 14), G5 (15 and 16) and AG4 (17 and 18); chimeric RNA2 with combined substitutions of ArMV origin: G12 (19 and 20), G24 (23 and 24) and AG24 (25 and 26); or chimeric RNA2 with region R2 of GFLV origin in a ArMV 2C^{CP}: AA2 (21 and 22). M: RNA extracted from mock inoculated protoplasts (1 and 2). Viral RNAs were detected using digoxigenin-labelled riboprobes complementary to GFLV genes 1E^{Pol} and 2A^{HP} with the positions of GFLV RNAs indicated with arrowheads on the left (upper panel) and ribosomal RNA stained with ethidium bromide (lower panel).

Effect of CP mutations on transmission by nematodes. The transmissibility of recombinant viruses that induced systemic plant infection, *i.e.* G1, AG1 and AG2, by *X. index* and *X. diversicaudatum* was evaluated (4, 7, 26). Recombinant AG2 was not transmitted by any of the two *Xiphinema* species whereas recombinants G1 and AG1, like synthetic GFLV, were transmitted by *X. index* but not by *X. diversicaudatum* (Table 2). ArMV-S was the only virus transmitted by *X. diversicaudatum* (Table 2). Transmission rates varied between 67% and 100%, well in agreement with previous reports on GFLV and ArMV transmission rates (4, 7, 26). Analysis of the RNA2 progeny of recombinants G1 and AG1 in infected bait plants by IC-RT-PCR and sequencing indicated no change in nucleotide sequence relative to the corresponding cDNA clones, confirming the genetic stability of the recombinants. These results are consistent with the notion that CP region R2 is involved in GFLV transmission by *X. index* and CP region R1 is not essential for transmission specificity.

Effect of CP mutations on virus acquisition and retention by nematodes. The deficiency in nematode-mediated virus transmission could result from a lack of virus acquisition during the AAP or a failure of vectors to bind and release virus particles during the IAP. To verify that nematodes fed on infected roots and ingested viruses, nematode specimens were randomly collected after the AAP and tested for virus presence by RT-PCR. Both vectors, *X. index* and *X. diversicaudatum*, ingested recombinants G1, AG1, AG2, synthetic GFLV and wild type ArMV independently of their transmission competency (Fig. 5). These results rule out the possibility that transmission failed due to a lack of virus acquisition by nematodes. Similarly, nematodes were tested for virus presence after the IAP to examine retention competency. Recombinants G1 and AG1 were detected in *X. index* while

recombinant AG2 was undetectable in the two nematode species (Fig. 5). In agreement with the specific transmission of GFLV and ArMV (7), synthetic GFLV was only detected in *X. index* and wild type ArMV in *X. diversicaudatum* (Fig. 5). These RT-PCR results were consistent with our transmission data (Table 2) and confirmed that CP region R2 is required for GFLV retention in *X. index* whereas CP region R1 does not seem essential for the specific virus retention by nematodes.

In summary, CP region R2 is likely exposed at the external surface of virions where it functions as a determinant for GFLV transmission by *X. index*. Unfortunately, we were not able to test CP region R2 in the transmission of ArMV by *X. diversicaudatum*, due to defect in encapsidation of the reciprocal recombinant AA2 (Table 1, Fig. 4). The smaller CP gene region R1, which is probably also located at the external surface of particles, has a neutral effect on GFLV transmission.

TABLE 2. Transmissibility of GFLV recombinants by *Xiphinema* spp.

Virus	Transmission by <i>X. index</i> ^a		Transmission by <i>X. diversicaudatum</i> ^a	
2ABC	11/14	(79%)	0/23	(0%)
ArMV-S	0/30	(0%)	20/26	(77%)
G1	6/9	(67%)	0/15	(0%)
AG1	7/7	(100%)	0/15	(0%)
AG2	0/20	(0%)	0/26	(0%)
Healthy	0/19	(0%)	0/14	(0%)

^aTransmission rates are given as the number of bait plants that reacted positively to GFLV or ArMV antibodies in DAS-ELISA over the total number of plants tested. 200 nematodes were used per virus source plant.

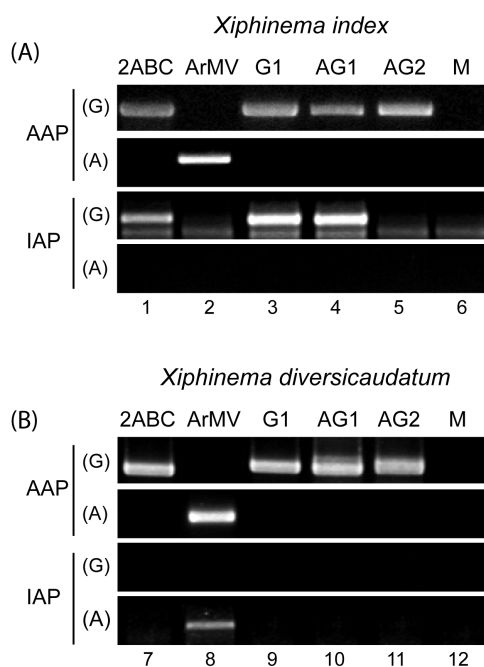


Fig. 5. Virus detection in *Xiphinema index* and *Xiphinema diversicaudatum* at the end of the AAP and IAP. Thirty (A) *X. index* and (B) *X. diversicaudatum* specimens exposed to source plants infected with synthetic GFLV 2ABC (1 and 7), wild type ArMV-S (2 and 8), recombinant G1 (3 and 9), recombinant AG1 (4 and 10), recombinant AG2 (5 and 11) or mock inoculated plants (M, 6 and 12) were randomly collected and tested by RT-PCR with (G) GFLV or (A) ArMV specific primers. DNA products were analyzed by electrophoresis on 1.5% agarose gels.

DISCUSSION

Limited information is available on viral determinants responsible for nematode-mediated transmission of nepoviruses. Previous studies showed that the specific transmission of GFLV and ArMV by *X. index* and *X. diversicaudatum*, respectively, is determined by their respective CP (4, 26). The identification of CP residues involved in transmission has been hampered by unsuccessful attempts at isolating natural GFLV or ArMV isolates that are deficient in nematode transmission and by a lack of resolved virion structure. In this study, we investigated the contribution of five GFLV CP regions predicted at the external surface of virions in transmission by *X. index* following a reverse genetics approach based on the use of GFLV/ArMV chimeric RNA2 molecules. Our results were consistent with the 11 residues of region 2 (2C^{CP} amino acids 188-198) being essential for GFLV transmission by *X. index*. To the best of our knowledge, this is the first report on a CP domain involved in nepovirus transmission by nematodes.

The prediction of the five regions of interest at the surface of the virion was assisted by the construction of a homology based 3D model to minimize the loss of encapsidation that was previously limiting our mutational approach (8). This appeared to be partly successful as three regions (R1, R2 and R4) of the ArMV 2C^{CP} could be inserted in the GFLV 2C^{CP} without affecting encapsidation, showing that the model indeed was decisive to identify residues exposed at the surface of the virus. The usefulness of the model is further shown by the identification of two regions with a contact function; R2 involved in vector transmission and R4 probably involved in cell to cell movement *via* 2C^{CP}/2B^{MP} interactions, as deduced from the absence of plant infection of the encapsidated mutants G4 and AG4. If confirmed, this would constitute the first hint in the identification of 2C^{CP} residues involved in 2B^{MP}/2C^{CP} interactions for cell-to-cell movement of GFLV. However, the biological validation of our model is only partial as two regions (R3 and R5) could not be exchanged without affecting encapsidation, suggesting a possible contribution in interactions with viral RNAs or subunit interactions required for virion assembly and/or stability. This also suggests that our predicted structure of GFLV likely not completely matches the real structure. To overcome these uncertainties on the surface residues, the crystallographic structure of GFLV would be needed, although it might not be sufficient to unravel transmission determinants, because, by analogy with other icosahedral plant viruses, the dynamics of virus-vector interactions could depend on conformational changes (19, 34).

The defect in encapsidation of some recombinant GFLV variants was not associated with the size of the substituted 2C^{CP} region or the degree of amino acid divergence. Indeed, recombinants carrying a mutated CP region R3 with four residues were deficient in encapsidation whereas recombinants affecting CP region R2 spanning over 11 residues encapsidated the corresponding viral RNAs. Also, all CP regions contained conservative and non-conservative amino acid changes. Thus, defects in encapsidation are more likely due to subtle structural changes and variations in

thermodynamic properties of CP subunits that affect/or not the assembly of virions. These subtle changes also likely have a cumulative effect as mutations that are individually not critical for encapsidation (G1, G2 or G4) become deleterious when combined (G12 and G24). Whether these explanations also account for the very poor infectivity of recombinant G2 is unknown.

Assuming that conformational changes occur during transmission, as demonstrated for *Cucumber necrosis virus* (CNV, genus *Tombusvirus*, family *Tombusviridae*, 34), it is not known if the 11 residues in the β B- β C loop ensure a direct contact between GFLV virions and vector ligands or whether they are involved in a critical conformational structure. Regardless, GFLV CP region R2 is essential but may not be sufficient for transmission. This hypothesis is supported by the fact that we could not rule out the implication of CP regions R3, R4 and R5 in transmission due to loss of infectivity of the corresponding recombinants. It is possible that additional, yet unidentified CP residues outside region R2, contribute to transmission. Also, GFLV CP region R2 *per se* conferred a loss of transmission by *X. index* but no gain of transmission by *X. diversicaudatum*, suggesting a lack of reciprocity in the nature and/or location of CP transmission determinants between GFLV and ArMV. The failure of recombinant AA2 to encapsidate viral RNA - while recombinant AG2 was competent for encapsidation - further supports this view since substitutions affecting *a priori* equivalent CP sub-domains did not have the same effect on virus encapsidation. However, we have to keep in mind that CP region R2 of ArMV may not adopt the same conformation within the GFLV CP relative to the ArMV CP, and thus may not be recognized by nematode factors that confer transmission specificity.

Out of the 11 amino acids that form CP region R2, Asp¹⁹⁰, Leu¹⁹¹, Ala¹⁹³ and Ala¹⁹⁶ (Fig. 2) are identical in nearly all GFLV and ArMV CP sequences available in GenBank. Therefore, these four amino acids are unlikely transmission determinants. Positions 189 (Phe/Tyr), 192 (Thr/Ser), 195 (Thr/Ser) and 198 (Arg/Lys) are occupied by different but biochemically related amino acids in GFLV and ArMV (Fig. 2). Whether those subtle differences are sufficient to contribute to transmission and/or to differential vector specificity needs to be assessed. Effects on virus transmission upon substitution by biochemically related amino acids have been reported in a few studies. For example, in the rod shaped TVMV, the replacement of a single polar CP residue (Asp for Glu) induces a loss of aphid transmission (6). Similarly, substitutions of Gln⁶ in CaMV P2 affect considerably the transmission efficiency and vector species range (27). Although the impact of single mutations could not be predicted based on the biophysical properties of the CaMV amino acid exchanged, the least impact on transmission was observed for the most conservative Gln to Asn (Q6N) and Gln to Thr (Q6T) substitutions (27). The most highly divergent amino acids of CP region R2 between GFLV and ArMV are residues 188 (Phe, aromatic in GFLV and Glu, acidic in ArMV), 194 (Val, hydrophobic in GFLV and Thr, polar in ArMV) and 197 (Leu, hydrophobic in GFLV and Tyr, aromatic in ArMV). In our 3D GFLV models, these residues appear well exposed at the surface of virions (Fig. S2) and represent prime candidates to act as transmission determinants. The effect of single, non-conservative, amino

acid changes on transmission have been reported for many viruses vectored in a non circulative manner by aphids (5, 6, 27, 30, 31), oipidium (20) and nematodes (24). Therefore, it is tempting to predict that GFLV transmission determinants may be restricted to only a few, if not a single residue within CP region R2. However, it should be noted that a single amino acid change is represented 180 times in a virus with T = 3 symmetry such as CMV and only 60 times in a virus with pseudo T = 3 symmetry such as GFLV, reducing at least theoretically, the functional impact of single residue changes in the latter virus. Future work will aim at fine mapping of GFLV CP region R2 with priority given to residues Phe¹⁸⁸, Val¹⁹⁴ and Leu¹⁹⁷ (Fig. S2).

According to our 3D homology-based GFLV models, CP region R2 overlaps the β B- β C loop within the B domain and sits closest to the β H- β I loop situated near the icosahedral 3-fold axis (Fig. S2). Remarkably, negatively charged residues of the CMV CP β H- β I loop play an essential role in transmission by aphids (21). Similarly, most residues critical for efficient CNV transmission by *Oipidium* spp. are located near a cavity formed by the particle 3-fold axis (20). It is therefore tempting to speculate that viral determinants necessary for nematode transmission of nepoviruses, aphid transmission of cucumoviruses and *Oipidium* transmission of tombusviruses share common features and may be structurally related. Also, the BC and HI loops close to the 3-fold axis have been described to form a ring and suggested to be of importance for TRSV transmission (11). For CPMV, this same region constitutes a major antigenic site, demonstrating its accessibility (11).

GFLV recombinants carrying the CP region R1 predicted in the β E- α B external loop did not affect the transmissibility of recombinant viruses AG1 and G1 by *X. index* and *X. diversicaudatum*. Among the five CP regions identified from our GFLV 3D models, CP region R1 is the only one that is located within domain C. It is also the most distant from CP region R2. Since recombinants AG1 and G1 are infectious *in planta*, CP region R1 is likely not involved in interactions between proteins 2C^{CP} and 2B^{MP} or between virions and protein 2B^{MP} that are required for tubule-guided cell-to-cell GFLV movement through plasmodesmata (32).

Recombinant AG2 failed to be transmitted by *X. index* and was not detected in *X. index* or *X. diversicaudatum* at the end of the IAP, despite its presence in both nematode species during the AAP. Therefore, as hypothesized for GFLV (7) and ArMV (7, 26) and demonstrated by *in vitro* binding assays for CaMV (9), TVMV (42) and CNV (9, 20, 42), a deficiency in transmission likely results from the failure of the virus to bind specific nematode factors. Similarly to the work performed in *X. americanum* with transmissible and non-transmissible nepoviruses (48), it would be interesting to determine whether recombinant AG2 is absent from the lining of the lumen of the stylet extension and the anterior oesophagus of *Xiphinema* specimens where virus retention occurs (10). If so, recombinant AG2 would be invaluable to develop an *in vitro* or *ex vivo* virus-vector binding test and unravel factor(s) involved in GFLV and ArMV specific recognition by *X. index* and *X. diversicaudatum*, respectively.

ACKNOWLEDGEMENTS

This work was partially supported by a competitive grant from the European Commission (QLK3-CT-2002-02140) and the Ministère de l'Alimentation, de l'Agriculture et de la Pêche (CASDAR project 440). P. Schellenberger and P. Andret-link were supported by a fellowship from INRA, Département Santé des Plantes et Environnement and the Conseil Interprofessionnel des Vins d'Alsace.

REFERENCES

1. **Altshul, S. F., W. Gish, W. Miller, E. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
2. **Andret-Link, P., and M. Fuchs.** 2005. Transmission specificity of plant viruses by vectors. *J. Plant Pathol.* **87**:153-165.
3. **Andret-Link, P., C. Laporte, L. Valat, C. Ritzenthaler, G. Demangeat, E. Vigne, V. Laval, P. Pfeiffer, C. Stussi-Garaud, and M. Fuchs.** 2004. *Grapevine fanleaf virus*: Still a major threat to the grapevine industry. *J. of Plant Pathol.* **86**:183-195.
4. **Andret-Link, P., C. Schmitt-Keichinger, G. Demangeat, V. Komar, and M. Fuchs.** 2004. The specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index* is solely determined by the viral coat protein. *Virology* **320**:12-22.
5. **Atreya, P. L., C. D. Atreya, and T. P. Pirone.** 1991. Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. *Proc. Natl. Acad. Sci. U S A* **88**:7887-7891.
6. **Atreya, P. L., J. J. Lopez-Moya, M. Chu, C. D. Atreya, and T. P. Pirone.** 1995. Mutational analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids. *J. Gen. Virol.* **76**:265-270.
7. **Belin, C., C. Schmitt, G. Demangeat, V. Komar, L. Pinck, and M. Fuchs.** 2001. Involvement of RNA2-encoded proteins in the specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index*. *Virology* **291**:161-171.
8. **Belin, C., C. Schmitt, F. Gaire, B. Walter, G. Demangeat, and L. Pinck.** 1999. The nine C-terminal residues of the grapevine fanleaf nepovirus movement protein are critical for systemic virus spread. *J. Gen. Virol.* **80** 1347-1356.
9. **Blanc, S., J. J. Lopez-Moya, R. Wang, S. Garcia-Lampasona, D. W. Thornbury, and T. P. Pirone.** 1997. A specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus. *Virology* **231**:141-147.
10. **Brown, D. J. F., W. M. Robertson, and D. L. Trudgill.** 1995. Transmission of viruses by plant nematodes. *Annu. Rev. Phytopathol.* **33**:223-249.
11. **Chandrasekar, V., and J. E. Johnson.** 1998. The structure of tobacco ringspot virus: a link in the evolution of icosahedral capsids in the picornavirus superfamily. *Structure* **6**:157-171.
12. **Chen, B., and R. I. Francki.** 1990. Cucumovirus transmission by the aphid *Myzus persicae* is determined solely by the viral coat protein. *J. Gen. Virol.* **71**:939-944.

-
13. **Demangeat, G., V. Komar, P. Cornuet, D. Esmenjaud, and M. Fuchs.** 2004. Sensitive and reliable detection of grapevine fanleaf virus in a single *Xiphinema index* nematode vector. *J. Virol. Methods* **122**:79-86.
 14. **Gera, A., G. Loebenstein, and B. Racciah.** 1979. Protein coats of two strains of Cucumber mosaic virus affect transmission by *Aphis gossypii*. *Phytopathology* **69**:396-399.
 15. **Gray, S. M., and N. Banerjee.** 1999. Mechanisms of arthropod transmission of plant and animal viruses. *Microbiol. Mol. Biol. Rev.* **63**:128-148.
 16. **Ho, S. N., H. D. Hunt, R. M. Horton, J. K. Pullen, and L. R. Pease.** 1989. Site-directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene* **77**:51-59.
 17. **Hogenhout, S. A., D. Ammar el, A. E. Whitfield, and M. G. Redinbaugh.** 2008. Insect vector interactions with persistently transmitted viruses. *Annu. Rev. Phytopathol.* **46**:327-359.
 18. **Huss, B., B. Walter, and M. Fuchs.** 1989. Cross-protection between *arabis mosaic virus* and *grapevine fanleaf virus* isolates on *Chenopodium quinoa*. *Ann. appl. Biol.* **114**:45-60.
 19. **Kakani, K., R. Reade, and D. Rochon.** 2004. Evidence that vector transmission of a plant virus requires conformational change in virus particles. *J. Mol. Biol.* **338**:507-517.
 20. **Kakani, K., J. Y. Sgro, and D. Rochon.** 2001. Identification of specific cucumber necrosis virus coat protein amino acids affecting fungus transmission and zoospore attachment. *J. Virol.* **75**:5576-5583.
 21. **Liu, S., X. He, G. Park, C. Josefsson, and K. L. Perry.** 2002. A conserved capsid protein surface domain of *Cucumber mosaic virus* is essential for efficient aphid vector transmission. *J. Virol.* **76**:9756-9762.
 22. **Loudes, A. M., C. Ritzenthaler, M. Pinck, M. A. Serghini, and L. Pinck.** 1995. The 119 kDa and 124 kDa polyproteins of arabis mosaic nepovirus (isolate S) are encoded by two distinct RNA2 species. *J. Gen. Virol.* **76** 899-906.
 23. **MacFarlane, S.** 2003. Molecular determinants of the transmission of plant viruses by nematodes. *Mol. Plant Pathol.* **4**:211-215.
 24. **MacFarlane, S. A., and D. J. Brown.** 1995. Sequence comparison of RNA2 of nematode-transmissible and nematode-non-transmissible isolates of pea early-browning virus suggests that the gene encoding the 29 kDa protein may be involved in nematode transmission. *J. Gen. Virol.* **76**:1299-304.
 25. **MacFarlane, S. A., C. V. Wallis, and D. J. Brown.** 1996. Multiple virus genes involved in the nematode transmission of pea early browning virus. *Virology* **219**:417-422.
 26. **Marmonier, A., P. Schellenberger, D. Esmenjaud, C. Schmitt- Keichinger, C. Ritzenthaler, P. Andret-Link, O. Lemaire, M. Fuchs, and G. Demangeat.** 2010. The coat protein determines the specificity of virus transmission by *Xiphinema diversicaudatum*. *J. Plant Pathol.* **92**: 275-279
 27. **Moreno, A., E. Hebrard, M. Uzest, S. Blanc, and A. Fereres.** 2005. A single amino acid position in the helper component of cauliflower mosaic virus can change the spectrum of transmitting vector species. *J. Virol.* **79**:13587-13593.
 28. **Ng, J. C., and B. W. Falk.** 2006. Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annu. Rev. Phytopathol.* **44**:183-212.
 29. **Ng, J. C., C. Josefsson, A. J. Clark, A. W. Franz, and K. L. Perry.** 2005. Virion stability and aphid vector transmissibility of Cucumber mosaic virus mutants. *Virology* **332**:397-405.

30. **Perry, K. L., L. Zhang, and P. Palukaitis.** 1998. Amino acid changes in the coat protein of cucumber mosaic virus differentially affect transmission by the aphids *Myzus persicae* and *Aphis gossypii*. *Virology* **242**:204-210.
31. **Perry, K. L., L. Zhang, M. H. Shintaku, and P. Palukaitis.** 1994. Mapping determinants in *cucumber mosaic virus* for transmission by *Aphis gossypii*. *Virology* **205**:591-595.
32. **Ritzenthaler, C., M. Pinck, and L. Pinck.** 1995. Grapevine fanleaf nepovirus P38 putative movement protein is not transiently expressed and is a stable final maturation product *in vivo*. *J. Gen. Virol.* **76**:907-915.
33. **Ritzenthaler, C., M. Viry, M. Pinck, R. Margis, M. Fuchs, and L. Pinck.** 1991. Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. *J. Gen. Virol.* **72**:2357-2365.
34. **Rochon, D., K. Kakani, M. Robbins, and R. Reade.** 2004. Molecular aspects of plant virus transmission by olpidium and plasmodiophorid vectors. *Annu. Rev. Phytopathol.* **42**:211-241.
35. **Sali, A., and T. L. Blundell.** 1993. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **234**:779-815.
36. **Sanfaçon, H.** 2008. Nepoviruses. *Encyclopedia of Virology, Third Edition*, B.W.J. Mahy and M.H.V. Van Regenmortel (eds.) **2**:405-413.
37. **Sanfaçon, H., J. Wellink, O. Le Gall, A. Karasev, R. van der Vlugt, and T. Wetzel.** 2009. Secoviridae: a proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus *Torradovirus*. *Arch. Virol.* **154**:899-907.
38. **Scott, S. W., M. T. Zimmerman, A. T. Jones, and O. Le Gall.** 2000. Differences between the coat protein amino acid sequences of English and Scottish serotypes of *Raspberry ringspot virus* exposed on the surface of virus particles. *Virus Res.* **68**:119-126.
39. **Serghini, M. A., M. Fuchs, M. Pinck, J. Reinbolt, B. Walter, and L. Pinck.** 1990. RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location. *J. Gen. Virol.* **71**:1433-1441.
40. **Syller, J.** 2006. The roles and mechanisms of helper component proteins encoded by potyviruses and caulimoviruses. *Physiol. Mol. Plant Pathol.* **67**:119-130.
41. **Thompson, J. D., D. G. Higgins, and T. J. Gibson.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673-4680.
42. **Uzest, M., D. Gargani, M. Drucker, E. Hebrard, E. Garzo, T. Candresse, A. Fereres, and S. Blanc.** 2007. A protein key to plant virus transmission at the tip of the insect vector stylet. *Proc. Natl. Acad. Sci. U S A* **104**:17959-17964.
43. **Vellios, E., D. J. Brown, and S. A. MacFarlane.** 2002. Substitution of a single amino acid in the 2b protein of *Pea early-browning virus* affects nematode transmission. *J. Gen. Virol.* **83**:1771-1775.
44. **Vigne, E., A. Marmonier, and M. Fuchs.** 2008. Multiple interspecies recombination events within RNA2 of Grapevine fanleaf virus and Arabis mosaic virus. *Arch. Virol.* **153**:1771-1776.

-
45. **Viry, M., M. A. Serghini, F. Hans, C. Ritzenthaler, M. Pinck, and L. Pinck.** 1993. Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. *J. Gen. Virol.* **74**:169-174.
 46. **Vuittenez, M., M. C. Munck, and J. Kuszala.** 1964. Souches de virus à haute agressivité isolées de vignes atteintes de dégénérescence infectieuse. *Etudes Virol. Appliquée* **5**:69-78.
 47. **Walter, B., and L. Etienne.** 1987. Detection of the Grapevine Fanleaf Viruses Away from the Period of Vegetation. *J. Phytopathol.* **120**:355-364.
 48. **Wang, S., R. C. Gergerich, S. L. Wickizer, and K. S. Kim.** 2002. Localization of transmissible and nontransmissible viruses in the vector nematode *Xiphinema americanum*. *Phytopathol.* **92**:646-653.
 49. **Ziegler-Graff, V., and V. Brault.** 2008. Role of vector-transmission proteins. *Methods Mol. Biol.* **451**:81-96.

SUPPLEMENTARY TABLES AND FIGURES

TABLE S1. GenBank accession numbers of the 219 GFLV and 20 ArMV variants used in this study for homology-based identification of GFLV 2C^{CP} amino acids potentially involved in transmission specificity by *X. index*.

Accession numbers in March 2010					
	GFLV			ArMV	
253993132	38604271	38604182	148750534	114325246	194021637
253993129	38604269	38604180	148750532	114325244	167861136
253993127	38604267	38604177	148750530	114325242	148750576
253993125	38604265	38604175	148750528	114325240	157679076
253993123	38604263	38604173	148750526	114325238	157679074
253993121	38604261	161015524	148750524	114325236	116744378
253993119	38604259	161015522	148750522	114325234	221018
253993117	38604257	557017	148750520	114325232	249578
39546456	38604255	38604339	148750518	114325230	14280097
39546454	38604253	38604337	148750516	114325228	50511901
39546452	38604250	154354404	148750574	114325226	50399853
39546450	38604248	253993115	60317523	114325224	60663
38604341	38604246	98991139	60317511	114325222	757527
38604335	38604244	249579	60317497	114325220	757525
38604332	38604242	197310914	60317487	114325218	130397
38604329	38604240	197310906	60317475	25013721	1095725
38604327	38604238	197310900	86450424	20153344	1095724
38604325	38604236	197310898	86450421	1335392	50401101
38604323	38604234	197310896	86450418	60665	50400864
38604321	38604232	197310894	86450415	14575516	148750574
38604319	38604230	197310892	86450412	238560636	
38604317	38604228	207339142	86450409	238560634	
38604315	38604226	207339140	86450406	238560632	
38604313	38604224	207339138	86450403	238560630	
38604311	38604222	207339136	86450400	238560628	
38604309	38604220	207339134	86450397	238560626	
38604307	38604218	148750568	86450394	238560624	
38604305	38604216	148750570	86450391	238560622	
38604303	38604214	148750572	86450388	238560620	
38604301	38604212	148750566	86450385	270052227	
38604299	38604210	148750564	86450382	270052225	
38604297	38604208	148750562	86450379	270052223	
38604295	38604206	148750560	114325272	270052221	
38604293	38604204	148750558	114325268	270052219	
38604291	38604202	148750556	114325266	270052217	
38604289	38604200	148750554	114325264	270052215	
38604287	38604198	148750552	114325262	270052213	
38604285	38604196	148750550	114325260	4377575	
38604283	38604194	148750548	114325258	14280095	
38604281	38604192	148750546	114325256	50403683	
38604279	38604190	148750542	114325254	50401102	
38604277	38604188	148750540	114325252	148750576	
38604275	38604186	148750538	114325250	42669117	
38604273	38604184	148750536	114325248		

GFLV-F13 and ArMV-S accession numbers are in bold.

TABLE S2. Mutagenesis strategies of GFLV and ArMV fragments to produce genes 2C^{CP} with mutated regions.

Mutant ^a	Matrix	Mutagenic PCR primers			Flanking PCR primers			Restriction sites ^c	Plasmid ^d
		Name	Direction-position ^b	Sequence	Name	Direction-position ^b	Sequence		
G1	pVec _{Acc651} 2ABC	mutR1F	F - 2268	5' <u>ACCGGATAACTAGTAGTATTCCACTACTGCTA</u> <u>GCCCTGTATACACCTTG</u> 3'	115	F - 1977	5' <u>CTGTGAGGATT</u> <u>GATAGAAACG</u> 3'	<i>Ngo</i> MIV	pVec _{Acc651} 2ABC
		mutR1R	R - 2317	5' <u>CAAGGTGTATACAGGGCTAGCA</u> <u>GTAGTGGAAA</u> <u>TACTACTAGTTATCCGGT</u> 3'	18	R - 3083	5' <u>ATCCACCCATAC</u> <u>GAAATAGTC</u> 3'	<i>Acc</i> 65I	
G2	pVec _{Acc651} 2ABC	mutR2F	F - 2591	5' <u>CTTACTTGCCCTCCCATAGAATATGACTTAAGT</u> <u>GCAACCAGTGCCTACA</u> <u>AGAGTGTGGGCTGACA</u> <u>TTG</u> 3'	61	F - 1529	5' <u>ATGACGCGCGC</u> <u>CTTTATTGG</u> 3'	<i>Nhe</i> I	pVec _{Acc651} 2ABC
		mutR2R	R - 2659	5' <u>CAATGTCAGCCCAGCACTTGTAG</u> <u>GCACTGG</u> <u>TTGCACTTAA</u> <u>GTCATATTCTATGGAGGGCAAGT</u> <u>AAG</u> 3'					
G4	pVec _{Acc651} 2ABC	mutR4F	F - 2801	5' <u>GTTAGTGAGTGGAACGGGTC</u> <u>ACTAACA</u> <u>ACTG</u> <u>GAATCAGCTGTTTAAGTATCCCGGGG</u> 3'	61	F - 1529	5' <u>ATGACGCGCGC</u> <u>CTTTATTGG</u> 3'	<i>Nhe</i> I	pVec _{Acc651} 2ABC
		mutR4R	R - 2858	5' <u>CCCGGGATACTTAAACAGCTG</u> <u>ATTCCAGTTGT</u> <u>TAGTGACCCCGTTCCTCACTA</u> <u>CAAC</u> 3'					
G5	pVec _{Acc651} 2ABC	mutR5F	F - 2921	5' <u>GCCAGATTGCTTGCTG</u> <u>CGCAATCAGCAAGCTC</u> <u>TTTTACGAGT</u> ACGT <u>TGAATTTTATGCAATAGCAG</u> <u>GACC</u> 3'	61	F - 1529	5' <u>ATGACGCGCGC</u> <u>CTTTATTGG</u> 3'	<i>Nhe</i> I	pVec _{Acc651} 2ABC
		mutR5R	R - 2991	5' <u>GGTCCTGCTATTGCATAAAAATTCA</u> ACGT <u>ACTC</u> <u>GTAAAAGAGCT</u> <u>TGCTGATTGCGCAGCAAGCAAT</u> <u>CTGGC</u> 3'					
G3	pVec _{Acc651} 2ABC	mutR3R	R - 2649	5' <u>TTATAAACCTTAGTGGTACC</u> <u>GTCAACAAGAGTT</u> <u>TGCCCAATGTACGCC</u> 3'			NA		
G12	pVec _{Acc651} 2ABC _{G2}	mutR1F			115			<i>Ngo</i> MIV <i>Acc</i> 65I	pVec _{Acc651} 2ABC
		mutR1R			18				
G13	pVec _{Acc651} 2ABC _{G1}	mutR3R			115		NA		
G23	pVec _{Acc651} 2ABC _{G2}	mutR3R			18		NA		
AA2	pVec2AB _A C _A	mutR2GF*	F - 2684	5' <u>GACTATCCCACC</u> <u>ACTGTTC</u> <u>TTGATTTAACGGC</u> <u>CGTCACGGCC</u> <u>CTTAGGAGTGTGTCCTGTATTG</u> <u>GGG</u> 3'	20*	F - 1095	5' <u>TGCTGTGCTGAT</u> <u>GGTAGGACTACCG</u> <u>GTGGGC</u> 3'	<i>Sac</i> I	pVec _{Acc651} 2AB _A C _A
		mutR2GR*	R - 2753	5' <u>CCCCAATAACAGGGACACTCCTA</u> <u>AGGGCCG</u> <u>TGACGGCCGTTAA</u> <u>ATCAAGAACAGTGGTGGGA</u> <u>TAGTC</u> 3'	263	R - 6400**	5' <u>GGGCGCGTCAG</u> <u>CGGTGTTGGCG</u> <u>GG</u> 3'	<i>Sal</i> I	

^aMutants were produced using overlapping PCRs, except for mutant G3, G13 and G23 that were generated by a single PCR, and subsequently cloned into one of two plasmids using restriction enzymes. Mutagenic nucleotides are underlined. The restriction site *Acc*65I in mutR3R (italics) and the nucleotide insertion in mutR5F and mutR5R (bold) are indicated.

^bPrimer positions are given as the position of the 5' first nucleotide according to the GFLV-F13 RNA2 sequence (NC_003623), except for *positions given according to the ArMV-S RNA2-U sequence (X81814) and **position given according to the vector sequence. F = forward and R = reverse orientation; NA: not applicable.

^cRestriction site used for cloning. The *Sal*I site sits in the vector downstream of the viral sequence.

^dPlasmids used for cloning.

TABLE S3. Subcloning strategies of GFLV and ArMV fragments to produce genes 2C^{CP} with single or multiple mutated regions.

Mutant ^a	Restriction fragment ^b	Plasmid ^c	Restriction sites ^d	
G45	G4	G5	<i>NheI</i>	<i>XmaI</i>
G24	G4	G2] <i>Acc65I</i>	<i>BglII</i>
G25	G5			
G245	G45			
AG1	G1] AG] <i>NgoMIV</i>	<i>SalI</i>
AG2	G2			
AG4	G4			
AG12	G12			
AG24	G24			
AG25	G25			
AG45	G45			
AG245	G245			

^aMutants generated with single, dual or triple mutated regions

^bRestriction fragments carrying mutations

^cPlasmid used for cloning

^dRestriction enzymes used for cloning

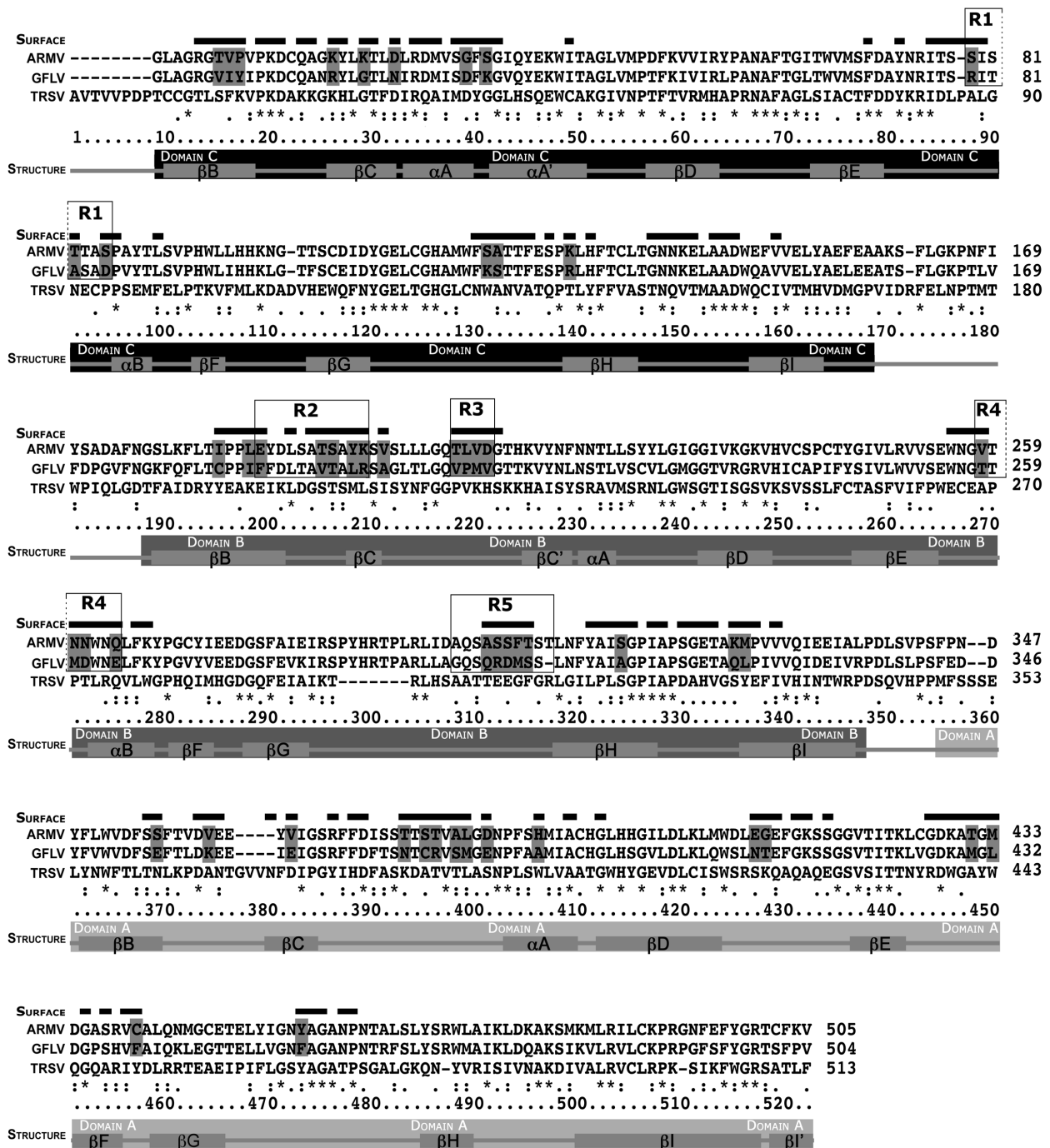


FIG. S1. Map of GFLV surface residues potentially involved in *X. index* transmission. Clustal X alignment of the CP residues of TRSV (GI 3913279), GFLV-F13 (NC_003623) and ArMV-S (X81814). Secondary structure of the TRSV atomic model is shown below the alignment. Surface residues of TRSV denoted previously (11) are indicated by horizontal bars above the sequences. Predicted surface residues that are different between GFLV and ArMV are highlighted in grey. Boxes delineate regions 1 (R1), 2 (R2), 3 (R3), 4 (R4), and 5 (R5) used for inverse genetics studies because corresponding predicted surface residues are poorly conserved between GFLV and ArMV.

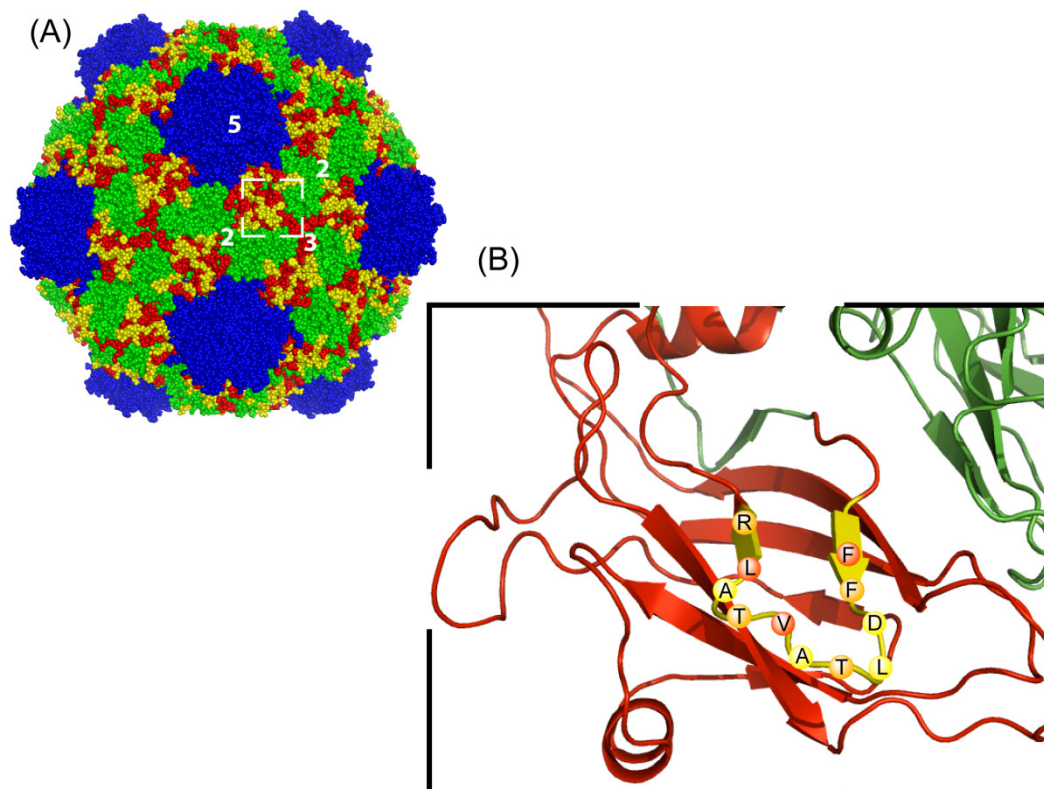


FIG. S2. Predicted 3D-model of GFLV illustrating the position of CP region R2. (A) A space-filling model of the particle quaternary structure viewed down the icosahedral 2-fold axis with R2 appearing in yellow, near the 3-fold axis. The A, B and C structural domains are indicated in blue, red and green, respectively. The icosahedral 5-fold (5) 3-fold (3) and 2-fold (2) axis are labeled for one icosahedral asymmetric unit. (B) Enlargement of the B domain, represented as a ribbon model with the amino acid sequence from Phe¹⁸⁸ to Arg¹⁹⁸ of GFLV isolate F13. The position of amino acids in region R2 is depicted as spheres. Yellow spheres indicate residues that are conserved between GFLV and ArMV. Orange and red spheres indicate residues with similar or different biochemical properties between GFLV and ArMV, respectively. Figures were generated using PYMOL (www.pymol.org).



Chapitre 5

Étude structurale de la capsidie du GFLV et d'un variant faiblement transmis par X. index

1. Introduction - Stratégies de détermination des structures de capsidie de virus à haute résolution

1. 1 Les stratégies

Une capsidie virale est un grand complexe macromoléculaire. C'est pourquoi, en plus de la technique la plus répandue pour déterminer une structure qui est la cristallographie aux rayons X, les virus se prêtent très bien à la cryomicroscopie électronique à transmission (cryo-MET). En effet, à l'heure actuelle, les traitements d'images de MET ne permettent pas de résoudre une structure de protéines inférieure à 100 kDa (Cheng and Walze, 2009; Saibil, 2000). L'étude de structure de virus icosaédriques possède un autre avantage : la haute symétrie des particules virales. Cette propriété est largement utilisée dans l'analyse structurale des capsidies icosaédriques par cryo-MET et en cristallographie. Elle permet, entre autres, de réduire le rapport signal sur bruit. Les deux procédés sont basés sur des principes différents, mais ils peuvent, comme nous allons le voir, être utilisés de manière complémentaire en biologie structurale.

Toutes les techniques d'obtention de structure 3D à partir d'images de ME reposent sur le même principe. Le microscope électronique produit une projection 2D de l'objet observé et en combinant les projections 2D de l'ensemble des différentes orientations de l'objet, le volume 3D originel est reconstitué. De ce principe découle le terme majoritairement utilisé pour qualifier la technique : la reconstruction 3D. Il faut également souligner que la technique de cryo-MET permet d'observer les protéines dans un état hydraté-congelé (fixées dans de la glace amorphe) ce qui a pour conséquence de préserver leur structure dans un état quasi-natif.

En cristallographie, c'est la diffraction des rayons X par un cristal de protéine qui permet d'obtenir la structure 3D. Un cristal est un état solide ordonné et la molécule qui le compose y est empilée de façon régulière, périodique dans les trois dimensions de l'espace. La maille cristalline élémentaire constitue le motif de base répété par translation à travers tout le cristal. Cet empilement

répété va créer une diffraction intense et mesurable des rayons X, selon le plan d'après lequel le cristal est positionné dans le faisceau de lumière. Seuls les électrons des atomes diffractent les rayons X. C'est pourquoi, à partir de l'intensité de diffraction et après reconstitution de leur phase, le contenu de la maille va pouvoir être calculé sous forme de carte de densité électronique. Par la suite, la position de chaque atome est modélisée à l'intérieur de cette carte, ce qui permet de reconstituer un modèle atomique de la protéine. Contrairement à la microscopie électronique, en cristallographie, il est nécessaire de reconstituer les phases du signal enregistré pour obtenir la carte de densité électronique. La détermination de la phase peut se faire soit de manière expérimentale, par exemple par la mesure de la diffraction de dérivé d'atome lourd présent dans le cristal, soit par simulation numérique à partir d'une structure cristallographique connue. On parle alors de remplacement moléculaire. Il est également possible d'utiliser les données de microscopie électronique afin de phaser les données obtenues aux rayons X (Baker *et al.*, 1999; Navaza, 2008). Ce procédé a été utilisé pour la première fois avec le *Cowpea chlorotic mosaic virus* (CCMV) (Speir *et al.*, 1995) et d'autres exemples ont suivi comme celui du *Bluetongue virus* (Grimes *et al.*, 1997).

La détermination structurale de complexes macromoléculaires peut s'effectuer en plaçant dans une reconstruction 3D faite sur un complexe de protéines, les modèles atomiques calculés par cristallographie aux rayons X de chaque composant du complexe (Rossmann *et al.*, 2005). On obtient ainsi une structure dite « pseudo atomique » qui combine l'information à haute résolution de cristallographie aux rayons X et la plus faible résolution généralement obtenue par des données de cryo-ME. Dans le cas des capsides icosaédriques cette technique est beaucoup utilisée pour étudier les interactions capsidique / récepteur ou capsidique / anticorps. De nombreux exemples sont retrouvés parmi les *Picornaviridae* (Belnap *et al.*, 2000; He *et al.*, 2000; Kolatkar *et al.*, 1999; Olson *et al.*, 1993; Xiao *et al.*, 2001; Zhang *et al.*, 2008). Elle est également utile pour positionner les protéines de capsidique dans l'unité asymétrique du virus (Fabry *et al.*, 2005; Olland *et al.*, 2001) et les protéines d'enveloppes des virions (Ferlenghi *et al.*, 2001). Elle permet aussi d'obtenir les structures complètes de phage dont les structures sont souvent composées d'éléments supplémentaires non icosaédriques.

1. 2 Les résolutions obtenues

La limite de résolution dépend directement de la longueur d'onde utilisée pour observer un objet. En théorie, la distance entre deux points observés ne peut être inférieure à $\lambda/2$.

En microscopie électronique, la longueur d'onde de l'électron varie en fonction de la tension accélératrice du microscope permettant de propulser les électrons. Les tensions d'accélération des MET varient entre 100 et 400 kV ce qui produit des électrons possédant une longueur d'onde comprise entre $3,7 \times 10^{-3}$ nm et $1,64 \times 10^{-3}$ nm. En théorie, le plus petit détail observable est donc largement inférieur à l'Angström. En pratique, la résolution est limitée principalement par le fait que l'image

obtenue contient des défauts provoqués par le microscope. L'homogénéité des protéines et la qualité des purifications sont également des points essentiels pour obtenir une reconstruction 3D à haute résolution. Actuellement, les meilleures reconstructions 3D de virus descendent sous les 5 Å de résolution : 3,80 Å pour le *Rotavirus* (Cheng and Walze, 2009), 3,88 Å pour le *Cytoplasmic polyhedrosis virus* (Yu *et al.*, 2008), 4,50 Å pour la capsid du phage *epsilon15* (Jiang *et al.*, 2008). Ces résolutions donnent accès aux structures secondaires des protéines, qui sont visibles à partir de 7 Å.

Actuellement, la microscopie électronique, la résonance magnétique nucléaire (RMN) et la cristallographie aux rayons X permettent d'atteindre une résolution atomique. Notons que la RMN se limite à des protéines de faibles poids moléculaires, de l'ordre d'une dizaine de kDa (Raman *et al.*, 2010), c'est pourquoi elle n'est pas utilisée dans le cas des virus. Concernant la cristallographie, la limite de résolution ne provient pas de la longueur d'onde utilisée, ni de l'appareillage, surtout depuis le développement des rayonnements synchrotrons et de nouveaux détecteurs de rayons X, mais elle va dépendre de la qualité de l'empilement cristallin. La majorité des structures cristallographiques de capsides a été résolue entre 2 et 4 Å. Seules trois structures ont franchi le seuil des 2 Å, dans le cas de deux virus satellites de phytovirus (1,8 Å et 1,9 Å) (Ban and McPherson, 1995; Larson *et al.*, 1998) et le *Foot-and-mouth disease virus* (FMDV, 1,9 Å) (Fry *et al.*, 1999).

2. Démarche expérimentale

Plusieurs raisons m'ont amenée à travailler pour l'obtention de la structure du GFLV. Elle nous permettrait de localiser les régions ayant été ciblées pour la mutagenèse dirigée et de mieux comprendre les conséquences des mutations effectuées dans la capsid. Plusieurs résidus déterminant la transmission par *X. index* ont été identifiés, mais peuvent-ils interagir directement avec le vecteur ou non ? Par ailleurs, il fallait également vérifier si les régions R3 et R4 étaient correctement délimitées sur la base du modèle 3D du GFLV et savoir si R4 est à la surface des virions pour une interaction potentielle avec la MP. Enfin, dans l'optique d'effectuer de la mutagenèse ponctuelle - entre autres sur la région R2 - et cibler les résidus de surface du virion, la connaissance de la structure cristallographique du virus prend toute son ampleur.

Le travail structural a été mené sur l'isolat F13 du GFLV (Vuittenez *et al.*, 1964), mais également sur un variant du GFLV (GFLV-TD) que j'ai identifié lors des évaluations de transmissibilité de virus chimériques GFLV/ArMV par nématodes. Ce variant est faiblement transmis par *X. index* et non transmis par *X. diversicaudatum*. Étant donné qu'il s'agit de la première découverte d'un variant naturel du GFLV dont la transmission est affectée, son étude est d'un intérêt considérable. Sa caractérisation moléculaire a révélé la présence d'une seule mutation au niveau du gène codant pour la CP. J'ai donc cherché à obtenir sa structure cristallographique et en parallèle

effectuer une étude de mutagenèse inverse, afin de confirmer le rôle du résidu 297 dans la vection du GFLV.

Pour mener à bien cette approche structurale du projet, j'ai suivi deux approches complémentaires, la cristallographie aux rayons X et la reconstruction 3D à partir d'images de cryo-MET. Les études structurales ont été menées en collaboration avec Bernard Lorber, Claude Sauter (IBMC, Strasbourg) et Marc Bergdoll (IBMP, Strasbourg), Stefano Trapani et Patrick Bron (CBS, Montpellier).

L'article 3 présente la stratégie de cristallogenèse que j'ai employée afin d'obtenir des cristaux de qualité suffisante pour l'analyse cristallographique. Le travail a nécessité l'utilisation de plusieurs techniques de cristallogenèse dont la cristallisation en gel. Ce qui m'a donné l'opportunité de participer à la rédaction d'un article de revue sur la cristallisation des protéines et acides nucléiques en présence de gel (Lorber *et al.*, 2009), qui figure en annexe de la thèse. Il décrit avec précision les principes et les avantages de cette technique spécifique de cristallisation en milieu diffusif.

L'ensemble des travaux sur la détermination de la structure cristallographique du GFLV sauvage et du variant GFLV-TD, la caractérisation fonctionnelle du GFLV-TD, ainsi que l'analyse des données, est présenté dans l'article 4 qui suit l'article sur la cristallogenèse des virions.

3. Article 3 : Strategies for the crystallization of Viruses : The case of *Grapevine fanleaf virus*

Article en preparation

**Strategies for the Crystallization of Viruses:
The case of *Grapevine fanleaf virus***

Pascale Schellenberger^{1,2}, Gérard Demangeat¹, Olivier Lemaire¹, Christophe Ritzenthaler², Marc Bergdoll², Vincent Oliéric³, Claude Sauter⁴ and Bernard Lorber^{4*}

1 Institut National de la Recherche Agronomique, UMR 1131 Santé de la Vigne et Qualité du Vin, INRA/UDS, Laboratoire de Virologie et Vecton, 28 rue de Herrlisheim, F-68021 Colmar, France, **2** Institut de Biologie Moléculaire des Plantes (IBMP), CNRS université de Strasbourg, UPR2357, 12 rue du Général Zimmer, F-67084 Strasbourg, France, **3** Swiss Light Source, Paul Scherrer Institute, Villigen, CH-5232, Switzerland, **4** Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, Institut de Biologie Moléculaire et Cellulaire (IBMC), 15 rue René Descartes, F-67084 Strasbourg, France.

* Corresponding author : E-mail: B.Lorber@ibmc-cnrs.unistra.fr

ABSTRACT

Background: *Grapevine fanleaf virus* (GFLV) is a small nematode-transmitted icosahedral plant RNA *Nepovirus* causing major damage to vineyards worldwide. GFLV was crystallized to study the molecular mechanisms underlying the recognition between the surface of the viral capsid and cellular components of its nematode vector or of its host as well as interactions with other viral proteins produced upon infection. Several obstacles had to be overcome to produce crystals having a satisfactory diffraction quality.

Methodology/Principal Findings: Here we report the preparation and characterization of the wild type GFLV strain F13 (GFLV-F13) and a transmission defective variant GFLV-TD in which Gly²⁹⁷ is replaced by Asp. Both variants were crystallized. The geometry and volume of the crystals were optimized by establishing their respective phase diagrams. GFLV-TD is more soluble than the wild type and crystallizes at high virus and PEG-3350 concentrations. Its crystals diffract X-rays at a resolution of 2.7 Å. The less soluble GFLV-F13 crystallizes at virus and PEG concentrations that are approximately half. The diffraction limit of its crystals could be extended to 3 Å by growing them in the presence of a low concentration of agarose gel. Preliminary analyses of the crystallographic data indicate that they are suitable for structure determination by molecular replacement using the related nepovirus *Tobacco ringspot virus* (TRSV) as a model.

Conclusion/Significance: Rigorous control of the quality of virus preparations, together with improvement of crystal quality using phase diagrams and enhancement of crystal stability by growth in agarose gel were the keys for the successful production of crystals of the wild-type virus and a natural mutant. The determination of the structures of the wild-type virus and of its variant is in progress. The reported crystallization methods are also applicable to biological macromolecules including nucleic acids and proteins.

INTRODUCTION

The very first challenge of any structural biology project based on crystallography is the preparation of ordered crystals with minimal defect and excellent diffraction properties. Frequent obstacles to crystallization are the preparation of a sufficient amount of soluble, highly pure and homogenous material and the growth of well ordered crystal. In addition to this, the very large unit cells of crystals made by great macromolecular complexes, such as viruses result in weaker diffraction as compared with small protein crystals. In this case the data collection procedure can be optimized to extend the crystal diffraction limit [1, 2]. On the other hand, the quality of the crystal can be enhanced using various crystallogenes approaches. They include the establishment of a phase diagram [3] and crystallization in diffusive environments like gels [4] or microfluidic channels [5].

We have chosen to determine the structure of *Grapevine fanleaf virus* (GFLV) by crystallography. GFLV is responsible for fanleaf degeneration, the most severe viral disease of grapevines worldwide. This virus reduces plant vigor, crop yield (up to 80%) and fruit quality. It also reduces the lifespan of vineyards and can cause plant mortality [6]. Within the *Secoviridae* family, GFLV belongs to the genus *Nepovirus* [7]. It is specifically transmitted from grapevine to grapevine by the soil-borne nematode *Xiphinema index*. The icosahedral and non-enveloped GFLV virion has a diameter of about 30 nm. Its capsid is composed of 60 identical subunits each encompassing 504 aminoacid residues (M_w 56 kDa, theor. pI 6.2). GFLV has a positive sense bipartite RNA genome called RNA-1 and RNA-2 that are encapsidated separately [8]. Each RNA is covalently attached to the virus-genome-linked protein (VPg) at its 5' terminus [9] and has a 3' polyadenylated terminus. GFLV shares a great similarity with *Tobacco Ringspot Virus* (TRSV) whose structure is known [10]. Obviously, the three-dimensional structure of the GFLV virion would give clues about its architecture and how it distinguishes from other nepoviruses. It may also inform about the molecular mechanisms underlying the recognition of the surface of the viral capsid by cellular components of its vector or of its host as well as by other viral proteins.

Here we report on the strategies and experimental steps that were crucial for the production of well-diffracting GFLV crystals as well as for the analysis of these crystals using synchrotron radiation. Crystallization was performed on the GFLV strain F13 (GFLV-F13) derived from infectious transcripts [11] and on a naturally occurring transmission defective variant which differs by a single amino acid substitution in the capsid protein (GFLV-TD). We show how the rigorous control of the quality of virus preparations, the improvement of crystal quality using phase diagrams, and the enhancement of crystal stability by growth in agarose gel led to the successful production of well-diffracting crystals. These strategies can be adapted for the crystallization and the crystallographic analysis of other viruses as well as of biological macromolecules like nucleic acids and proteins.

RESULTS

Quality control of GFLV preparations. To measure reliably the concentration of the virus samples it was necessary to determine the extinction coefficient of the virus purification. GFLV purification are mixtures of empty capsid shells and of capsids containing nucleic acids of various lengths, leading to 3 populations in density gradients (Top, Middle and Bottom components), that are collected together at the end of the GFLV purification protocol. All virions in these populations have a similar architecture and shape, but differ by their UV absorbance properties [12]. We assumed a theoretical 1/1/1 mixture of the three types of capsids. The calculation based on protein and nucleic acid content gives $E_{260\text{nm}} = 9.0 \text{ mg/ml/cm}$. To confirm this value, we compare the intensity of light scattered by a GFLV solution of known absorbance to that of a *Brome mosaic virus* (BMV) solution of known concentration. We find the same value when we assume that the virions of GFLV have the same diameter as those of BMV for which $E_{260\text{nm}} = 5.0 \text{ mg/ml/cm}$ [13].

In the crystallization process, it is crucial that purified viral particles are free of contaminants and homogeneous in size and shape [14, 15]. For this reason we have systematically performed a full set of analyses on every GFLV preparation. Electrophoresis, light scattering and electron microscopy were routine diagnostic tools to select the best batches for virus crystallization. Our purification protocol yields virions which are pure according to their absorbance properties in UV ($A_{260\text{nm}}/A_{280\text{nm}}$ ratio = 1.80 to 1.97). No co-purified protein is detected in the electrophoretic mobility pattern (**Fig. 1A**). Pure preparations are characterized in denaturant condition by the presence of a single polypeptide chain with molecular weight of ~56 kDa as expected from amino acid sequence [16]. Dynamic light scattering analyses indicate that the hydrated particles have a diameter of $31 \pm 2 \text{ nm}$ and polydispersity $\leq 5\%$ (**Fig. 1B**). Minor batch-to-batch variations in GFLV particle diameter and population polydispersity were detected by dynamic light scattering. Transmission electron microscopy (TEM) analyses reveal that the dehydrated capsids consist of a majority of quasi-spherical and filled shells accompanied by a few empty ones, all with a diameter of 29 to 30 nm (**Fig. 1C**). Careful inspection of micrographs sometimes revealed a number of small-size particles that could be protein aggregates. These less pure samples were not used for crystallization assays. In practice, virus batches which did not fulfill above quality criteria either yielded small ill-shaped crystals or formed only precipitates.

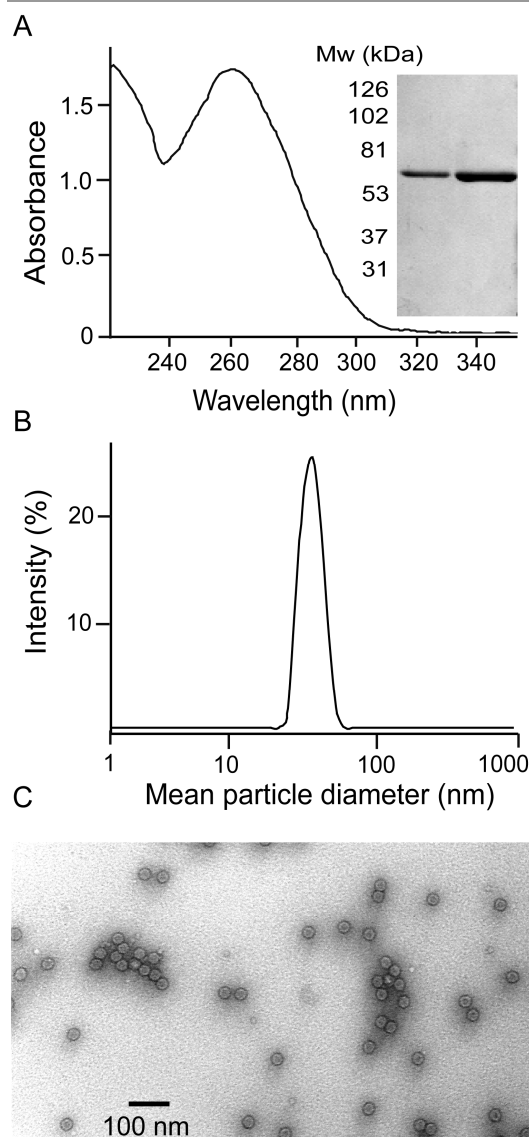


Figure 1. Purity of GFLV preparations. Variants GFLV-F13 and GFLV-TD gave comparable results. (A) UV-absorbance spectrum and (inset) coomassie blue staining after SDS-PAGE separation. Lane 1 and 2 correspond to 5 and 10 µg protein, respectively. (B) Intensity particle size distribution from dynamic light scattering analysis of virus in 0.15 M NaCl. Mean particle diameter is 31.5 nm and polydispersity ~5%. (C) Electron micrograph of virions after negative staining.

Crystal growth improvement using a phase diagram. Crystallization conditions for GFLV-F13 and GFLV-TD were searched using sparse matrices of crystallant solutions. The first crystals of GFLV-TD were found in a drop containing 10% (m/v) PEG-3350, 0.2 M proline and 0.1 M Hepes pH 7.5 equilibrated by vapor diffusion at 20°C. Their dimensions did not exceed 30 µm and their polyhedral habit was ill-defined. They were accompanied by numerous dendrites and spherulites (result not shown) known to grow at very high supersaturations and unsuitable for diffraction analyses [17]. At lower virus and crystallant concentrations the rate of nucleation was decreased, dendrites and precipitate were eliminated, and the growth of large crystals was promoted. This was done by preparing manually sets of 16 drops in squares covering 4 virus and 4 PEG concentrations. Further, drop volume was tripled to make a greater amount of virus available for crystal growth. A typical example of the consequence of the lower virus concentration is displayed [Fig. 2A](#). There are fewer dendrites, no precipitate, and the pyramidal habit of the crystals can be recognized. After a second round of condition refinement at lower PEG concentration, the drops contained few single crystals measuring up to 150 µm across and a few large dendrites ([Fig. 2B](#) central panel). Finally, after minor

adjustments single crystals with a diameter of up to 300 μm were obtained in a solution containing 4 mg/ml GFLV-TD and 7% (m/v) PEG-3350 (**Fig. 2C**).

The best condition found for GFLV-TD gave numerous tiny crystals and dendrites when applied to GFLV-F13 (result not shown). A similar procedure as described above was therefore applied to grow large single crystals with well-developed facets. Best crystal form (truncated monoclinic prisms) of GFLV-F13 reach 200 μm in length and grow at virus and PEG concentrations of 1.8-2.2 mg/ml and 3.0-4.0% m/v, respectively, *i.e.* approximately half of those required to produce best GFLV-TD crystals. Arrows in **Fig. 3** indicate the paths followed during the optimization of the crystal growth of each virus, from initial crystals (I and I') to the best crystals (C and C'). Hatched curves correspond to the hypothetical solubility lines separating the regions above which the viruses do crystallize from those below, were they do not. Experimental solubilities (plotted as open and closed circles) are 1.3 mg/ml and 3.3 mg/ml in the presence of 7% (m/v) PEG-3350 for GFLV-F13 and GFLV-TD, respectively.

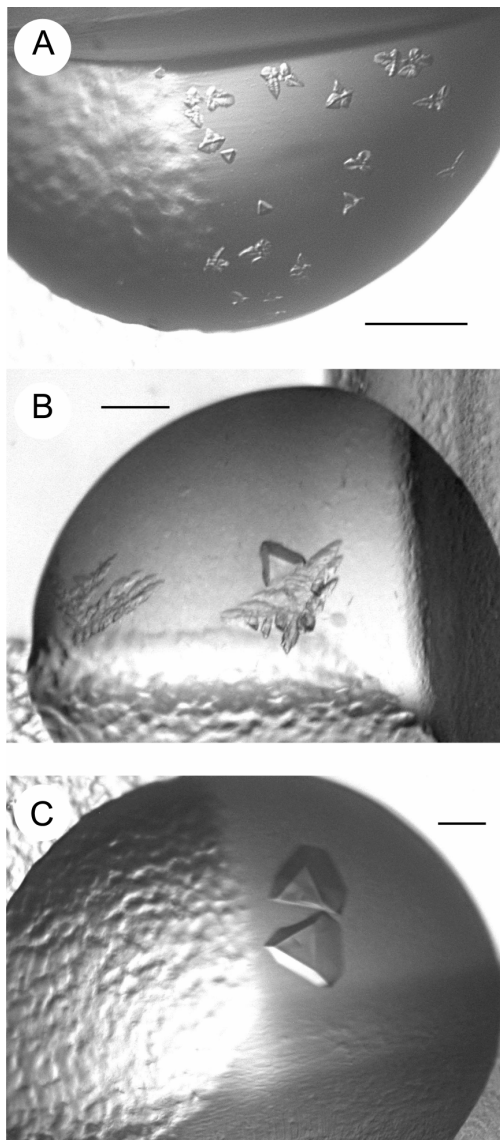


Figure 2. Optimization of GFLV-TD crystallization. (A) Initial assay containing GFLV-TD at 6.5 mg/ml and 10% (m/v) PEG-3350. After 7 days crystallization yields many dendrites and a few truncated pyramidal crystals. (B) Single crystal accompanied by a few dendrites obtained after 7 days at 5 mg/ml virus and 8% (m/v) PEG. (C) Large crystals in clear solution at 4 mg/ml virus and 7% (m/v) on day 10. In all pictures drop volume is 1 μL and scale bar corresponds to 200 μm . A similar optimization protocol was applied to the wild type strain F13.

Crystal quality enhancement by growth in agarose gel. Crystals of GFLV-F13 and GFLV-TD were analyzed on synchrotron X-ray sources. The diffraction limit of those of GFLV-TD grown in solution ranged from 3.5 to 4.5 Å although one sample exhibited well-resolved diffraction patterns with reflections beyond 3 Å (Fig. 4A). In contrast, most GFLV-F13 crystals did not diffract X-rays to a resolution better than 6 Å. Although the absence of ice rings suggested that cryoprotection was sufficient, only one crystal out of two dozen had a diffraction limit of 5.5 Å (Fig. 4B). No enhancement was noticed by varying neither the composition of crystallization and cryoprotectant solutions nor the duration of soaking before crystal flash cooling. The diffraction limit of GFLV-F13 crystals prepared under identical conditions except the addition of 0.2% (m/v) agarose was 3 Å (Fig. 4C). Gel-grown GFLV-F13 crystals are less fragile and display an improvement in net resolution of 2 to 3 Å compared to those produced in the absence of agarose.

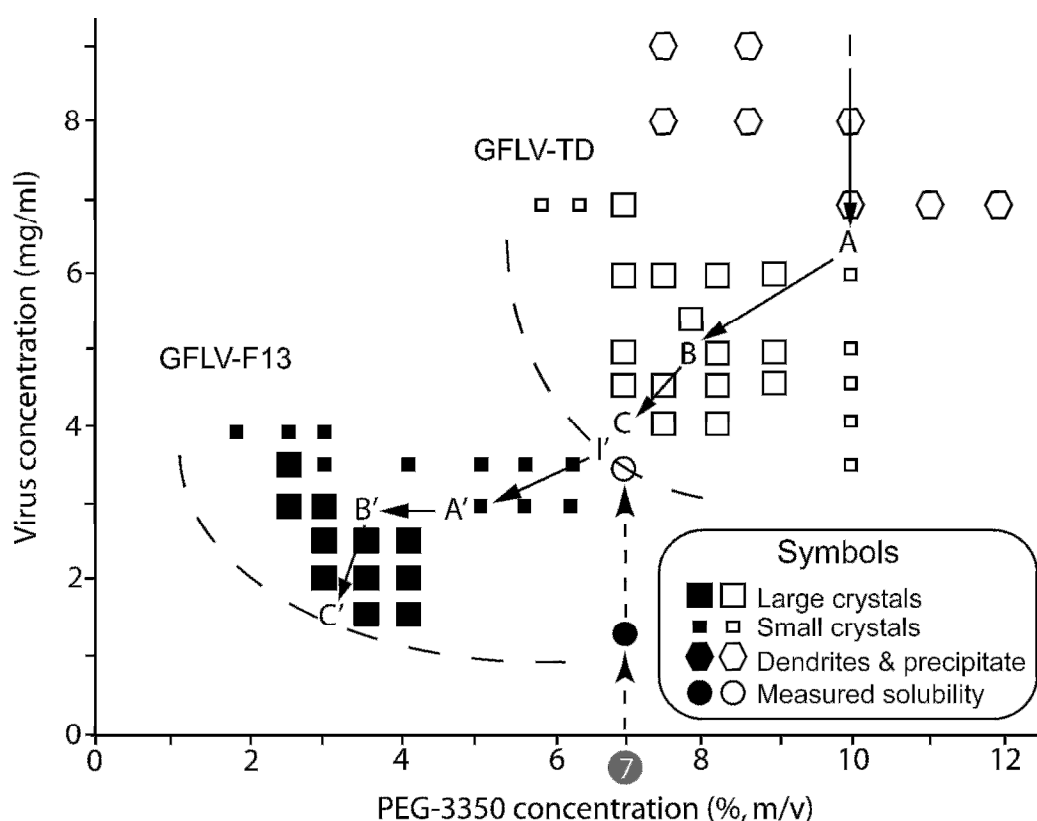


Figure 3. Superimposed phase diagrams of GFLV-F13 and GFLV-TD. Both diagrams were obtained from multiple independent crystallization assays done in 1 μ L sitting drops at 20°C. They are superimposed to facilitate comparison. Symbols stand for the major crystal form observed with a binocular at 50-fold magnification. Solid symbols represent data from GFLV-F13 and open symbols those of GFLV-TD. The compositions of the drop in the initial screen of GFLV-TD (I) and those of drops A, B and C in Fig. 2 are indicated. The path followed during the optimization of GFLV-F13 crystals is indicated by the letters I', A', B', and C'. The values of the experimental solubilities (1.3 and 3.3 mg/ml for GFLV-F13 and GFLV-TD, respectively) are given by round symbols. For each variant, no crystals were observed in samples having compositions located beyond the corresponding dashed solubility curve.

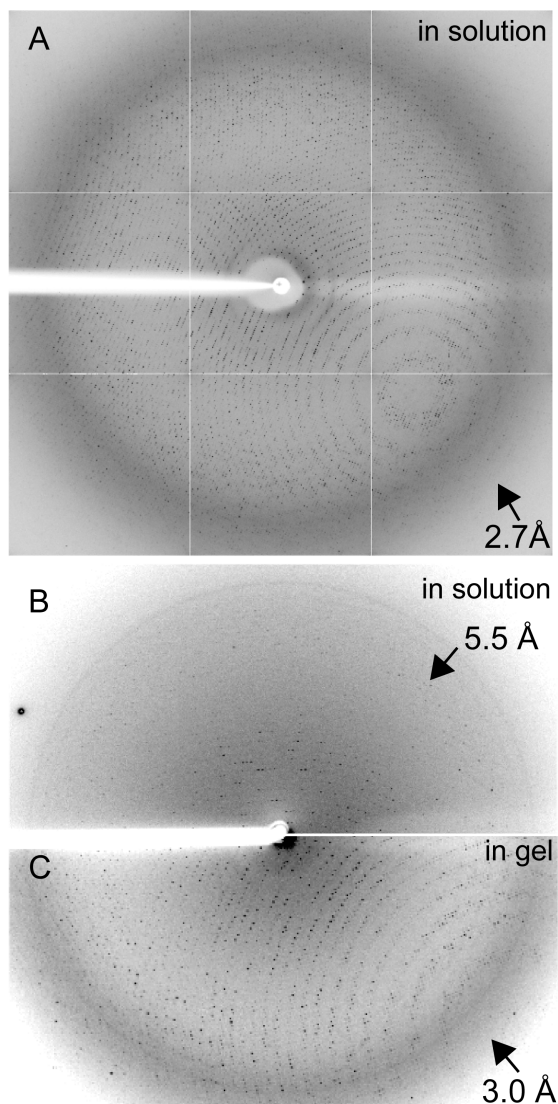


Figure 4. Diffraction limits of GFLV-F13 and GFLV-TD crystals at 100K. From top to bottom: Diffraction patterns of best crystals of GFLV-TD in solution (A) and of GFLV-F13 in solution (B) or in gel (C). Arrows indicate the diffractions limits of 5.5, 3.0, and 2.7 Å, respectively. See Table 1 for details.

Preliminary crystallographic analyses. Table 1 summarizes the results of preliminary crystallographic analyses of solution and gel-grown GFLV crystals. Two crystal forms of GFLV-F13 were identified; a triclinic one (grown and analyzed in solution or in gel) and a monoclinic one (grown and analyzed in gel). The crystal form of GFLV-TD is cubic. Alike many virus crystals those of GFLV exhibit a low mosaicity ($0.07\text{-}0.15^\circ$) even though they were analyzed at cryogenic temperatures. A decrease of the high resolution signal in the $3\text{-}4\text{ \AA}$ resolution shell was seen once 400-500 images were collected. Consequently, data collection strategies were adjusted to limit the effect of radiation damage on datasets. Accordingly, full completeness could not be achieved for low symmetry crystal forms (P1 and P2). However, this should not be detrimental for structure determination owing to the high non-crystallographic redundancy of the capsids made of 60 identical protein subunits. Hence, the icosahedral 532 point group symmetry was clearly detected for each of the three crystal forms. In the cubic ones, the viral 3 fold axis was aligned with the crystallographic 3 fold axis (not shown). The most probable content of the respective asymmetric units is one third of a capsid (or 4

pentamers) in cubic crystals and a complete capsid in both GFLV-F13 crystal forms. This corresponds to a protein content of about 25% in the three cases. For the evolutionary-related TRSV, the solvent surrounding the virions occupies 32% of the unit cell volume, since the genomic RNA fills the interior of the capsid and the mean outer diameter is 140 Å. So, the virions are rather densely packed. Interestingly, cubic GFLV crystals and TRSV crystals [10] have in common two unit cell dimensions and the same crystal density. This supports the idea that both viruses are very similar in shape and that TRSV (PDB accession code 1A6C) [10] is a good starting model for molecular replacement (in progress).

Table 1. X-ray analysis of GFLV crystals

Virus	GFLV-F13	GFLV-F13	GFLV-F13	GFLV-TD
Crystallization medium at 293K	Solution	Gel	Gel	Solution
Data collection at 100K				
Beamline	SLS/X06DA	SLS/X06DA	SLS/X06DA	ESRF/BM30
Detector	X06DA	X06DA	X06DA	FIP-BM30
Wavelength (Å)	1.00	0.979	0.979	0.980
Oscillation (°)	0.3	0.3	0.3	0.25/0.2
Number of frames	600	600	400	400
Statistics				
Space group (number)	P1 (1)	P1 (1)	P2 (3)	P2 ₁ 3 (198)
Unit cell lengths a, b, c (Å)	278.4	279.4	283.8	408.0
	298.3	279.5	295.5	408.0
	291.6	293.3	393.7	408.0
Unit cell angles α, β, γ (°)	102.0	102.4	90.0	90.0
	116.4	116.4	91.5	90.0
	108.2	108.2	90.0	90.0
Crystal mosaicity (°)	0.08	0.10	0.15	0.07
Resolution range (Å)	100 - 5.5	135 - 3.0	100 - 3.7	36 - 2.7
Highest resolution shell (Å)	5.64 - 5.5	3.08 - 3.0	3.8 - 3.7	2.77 - 2.7
No. of observations	394916	2416076	1710375	6205218
No. of unique reflections	198238 (13886)	1214336 (73170)	572178 (19142)	563202 (32448)
Completeness (%)	88.8 (84.1)	88.1 (71.7)	83.1 (76.5)	92.0 (72.0)
Multiplicity	2.0 (2.0)	2.0 (1.9)	3.0 (3.0)	11.0 (3.1)
R _{merge} (%) [†]	13.9 (25.5)	10.0 (35.0)	18.8 (49.5)	12.5 (68.1)
$\langle I/\sigma(I) \rangle$	7.2 (3.4)	9.2 (2.4)	7.3 (2.5)	18 (1.9)
Asymmetric unit content	60-mer	60-mer	60-mer	20-mer
Matthews coefficient (Å ³ Da ⁻¹)	5.2	5.2	4.8	5.0
Protein content (%)	26	26	24	25

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

Values in parentheses are for the highest resolution shell.

DISCUSSION

Many biocrystallography projects fail because of the impossibility to grow suitable crystals. In the case of GFLV-F13 and of its variant GFLV-TD several implementations were necessary to grow and reproduce the crystals and make them diffract at high resolution (3 Å). Once a promising crystallant was identified, two-dimensional phase diagrams covering a range of virus and crystallant concentrations were used to optimize the growth of the crystals for X-ray analyses. The stability of GFLV-F13 crystals during the flash cooling procedure was enhanced by growing them in the presence of a low concentration of agarose gel.

Importance of virus purity and homogeneity. Crystals defects may be generated by contaminants competing with virions during crystal growth. Indeed, virus crystals can incorporate without discrimination a wide panel of impurities from small molecules to larger viruses [18]. Most virus purification protocols cannot separate particles that differs only a little in shape, charge, size or presence of proteins possessing some affinity to the virus. For example, an α -glucosidase from the host plant interacts strongly and co-purifies with the poliovirus *Beet western yellows virus* [19]. More over, seasonal variations in the host plant's physiological state lead to fluctuating levels of contaminants in virus preparations. The fact that we worked exclusively with the purest and most homogeneous virus batches ensured the reproducibility of the crystallization and permitted to establish the phase diagram of each virus variant.

Advantages of phase diagrams and virus solubility. The visual inspection of the result of a small number of individual crystallization assays suffice to get a gross idea of the location of the unsaturated region (where the solution remains clear and no crystal grows), the supersaturated region (where either crystals or precipitate form) and the solubility curve separating them (**Fig. 3**). In the case of the GFLV variants, minor variations in virus and/or crystallant concentration led to significant changes in the crystallization outcome. For this reason, it was necessary to delineate the exact composition of the solutions which gave the best crystallization result.

In our present case, we noticed that both variant do not behave in the same way with regard to their solubility. Indeed, the single mutation from Gly to Asp alters dramatically the crystallization phase diagram of GFLV in a given crystallant. The knowledge of the two phase diagrams and of a few experimental solubility data permits to correlate the addition of a carboxyl group to a 2.5-fold increase in virion solubility. This result suggests that this residue is accessible to the solvent, as previously predicted by the alignment between GFLV and TRSV amino acid sequence and the TRSV crystal structure [10]. The more soluble virions crystallize more easily presumably because supersaturation varies more slowly during drop equilibration when the composition of the drops crosses the nucleation zone [20]. Cubic symmetry is another parameter which may favor packing regularity. In contrast, the

less soluble variant crystallization is more sensitive to supersaturation variations. Slightly too high supersaturations rapidly leads to fast and uncontrolled nucleation, changes in growth mechanism accompanied by the incorporation of lattice defects and ultimately in reduced diffraction properties.

Interestingly, the degree of supersaturation (σ) required to trigger nucleation and sustained growth of GFLV-F13 and GFLV-TD crystals with best geometry and greatest dimensions does not exceed 2. This is slightly less than the values ($2 \leq \sigma \leq 3$) at which *Tomato bushy stunt virus* crystallizes at low PEG-8000 or ammonium sulfate concentration or precipitates at higher ones [3]. Also, phase diagrams indicate that crystals of BMV grow near the solubility curve, *i.e.* at $1.5 \leq \sigma \leq 2$ in the presence of PEG-8000 or PEG-20000 [21]. GFLV is composed of 60 equivalent capsid proteins with a defined position and orientation, following the icosahedral symmetry. This gives the virions many more possibilities - than any asymmetrical macromolecules - to bind readily in the right orientation at the surface of a growing crystal. Further, the greater volume fraction of the virus compared to that of smaller particle makes nucleation possible at much lower concentration [14, 15]. Altogether, this explains why extremely low supersaturations provide the sufficient driving force to nucleate and grow virus crystals. Even if, with regard to very small molecules like salts these supersaturations are still high since the latter can crystallize when the concentration in solution exceeds solubility by just a few percent [22].

Advantage of crystallization in agarose. Crystals of wild-type GFLV-F13 diffracting to high resolution were obtained by adding agarose gel. A similar effect was observed with various proteins [4]. The network of this polysaccharide is entrapped in the channels of the crystals during growth and stabilizes the lattice [23]. As reported previously, the fibers of the gel reinforce the crystals and prevent the decay or loss of diffraction properties upon cryocooling [24, 25]. More over, it provides a convection-free environment suitable for the slow growth of well-ordered crystals.

Perspectives. The strategies implemented for GFLV can be applied to other viruses and biological macromolecules. Phase diagrams and gels can be used with any biological macromolecules (i) to find crystallization conditions producing at low supersaturation single crystals with large volumes and best diffraction properties and (ii) reinforce crystalline lattices before subjected to the thermal shock of cryocooling preceding diffraction analyses.

MATERIALS AND METHODS

Chemicals and biochemicals. Ammonium molybdate was purchased from Merck, molecular weight marker proteins from Sigma (Cat. No. P-1677), bentonit from Alfa Aesar and PEG-20000 from Prolabo. Polyethylene glycol (PEG-3350), INDEX™ sparse matrix and Aqua Sil™ were from Hampton Research (Aliso Viejo, CA).

Low-gelling temperature (T_g 28°C) agarose was a gift from So.bi.gel. (Hendaye, France). All solvents in contact with the virus during crystallization assays were prepared with sterile thrice-distilled water.

Virus strain and propagation. GFLV strain F13 were isolated from naturally infected grapevines and propagated in *Chenopodium quinoa*, a systemic host for the virus [26]. A variant of this strain which had been mechanically inoculated for several years from *C. quinoa* to *C. quinoa*, was tested for transmissibility by *X. index* and a quasi-total loss of transmission was obtained (Schellenberger *et al*, article 4). This naturally occurring variant, identified in the course of this study, was named GFLV-TD. The capsid protein gene was characterized by IC/RT/PCR and sequencing with the method described by Schellenberger *et al*, article 4. The GFLV-F13 strain crystallized originated from *C. quinoa* initially infected with synthetic virus produced from cDNA clone of GFLV strain F13 [11].

Virus purification. Viral particles are purified mainly as described [27] with one additional separation on sucrose gradients. Plants are infected using crude sap from infected *C. quinoa* leaves as inoculum. 14 days post-infection, one kg of fresh infected leaves are harvested and ground at room temperature with 2 volumes of grinding buffer (65 mM sodium phosphate buffer pH 7.2, 10 mM EDTA, 100 mM ascorbic acid). The crude homogenate filtered over cloth is clarified by low speed centrifugation (15 min, 1.900 x g, 15°C) after 15 min contact with 8% (m/v) bentonite. The virus is then precipitated by adding 10% (m/v) PEG-20000 and 1% (m/v) sodium chloride at room temperature. The precipitate centrifuged at low speed centrifugation (10,000 x g, 0.5 h, 5°C) is resuspended in 5 mM sodium phosphate buffer pH 7.0 and 1 mM EDTA. After sedimentation by ultracentrifugation at 140,000 x g for 2 h in 50.2Ti rotor (Beckman) the virus pellets are resuspended in 5 mM sodium phosphate buffer pH 7. Contaminating proteins are removed by centrifugation on a 20 ml sucrose cushion in 5 mM Na phosphate buffer pH 7.0 (140,000 x g, 5 h, 50.2Ti rotor, Beckman). The resulting pellet is suspended in 15 mM sodium phosphate and 5 mM potassium phosphate pH 7.0 and insoluble material removed by low speed centrifugation (4,000 x g, 10 min, 5 °C). The viral particles in the supernatant are then purified by two consecutive sucrose density gradient fractionations. A first gradient is established by thawing a frozen 30% (m/v) sucrose solution followed by 5h centrifugation at 113,000 x g in a SW28 rotor (Beckmann). Its virus-containing fractions are pooled and centrifuged at 290,000 x g for 2 h in 50.2Ti rotor (Beckmann). Pure virions are suspended in 15 mM sodium phosphate and 5 mM potassium phosphate pH 7.0, ultracentrifugated for 2.5 h through a second 60 to 10% (m/v) sucrose gradient at 210,000 x g in SW41 rotor (Beckman), and collected with a peristaltic pump. Finally, they are resuspended in sterile water and filtered through a 0.22 µm pore-size Ultrafree-MC™ membrane (Millex) prior to storage at 4°C.

Characterization of GFLV preparations. The UV-absorbance of virus was measured on 2 µl solution with a Nanodrop® ND-100 spectrophotometer. The protein content of virus preparations was separated by electrophoresis (SDS-PAGE) on 10% (m/v) polyacrylamide gels after denaturation at 95°C in presence of sodium dodecyl sulfate and mercaptoethanol.

Dynamic light scattering analyses were performed using a Zetasizer NanoS instrument from Malvern (Malvern, UK) to estimate mean particle diameter, polydispersity and presence of aggregates. Virus solutions at 0.05 mg/ml in 0.15 M NaCl were transferred in quartz cells. Scattered intensities were recorded at 20°C and data

processed with the DTS software developed by the manufacturer. Experimental data were corrected for solvent viscosity and refractive index. Particle populations are characterized by their mean diameter and their polydispersity (standard deviation on the diameter). Polydispersities inferior to 15% reflect narrow particle size distributions.

Electron microscopy analyses were performed with a Hitachi H-600 electron microscope equipped with a Hamamatsu CCD Advantage HR camera. Three μl virus suspension at 0.1 mg/ml in water was deposited on a nickel grid coated with a polyvinyl formal (Formvar®) film. The excess was blotted after 2 min and 5 μl ammonium molybdate solution at 1% (m/v) added. After 5 s the excess stain was blotted by contact with a piece of filter paper and the grid was air-dried at ambient temperature. Grids were examined at 75 kV.

Crystallization in solution. The initial search of crystallization conditions was done in 96-well Crystalquick® LBR microplates (Greiner BioOne) with 3 round wells per reservoir using the Mosquito™ pipeting station (TTP Labtech, UK). Nanodrops were prepared by mixing 150 nl virus at 8 mg/ml in 50 mM Hepes pH 7.5 with 150 nL reservoir INDEX™ sparse matrix solution and equilibrated by vapor diffusion over 50 μl reservoirs (Ducruix & Giegé, 1999, Sauter et al., 2010). After 7 days, crystals were found in a drop containing 10% (m/v) polyethylene glycol PEG-3350, 0.2 M proline and 0.1 M Hepes pH 7.5. Subsequently, series of 1 μl sitting drops were prepared manually in Crystalquick™ low-profile microplates (Greiner BioOne) and equilibrated against 50 μl reservoirs. A set of 16 drops covered 4 virus concentrations and 4 crystallant concentrations.

Crystallization in gel. Agarose with a gelling temperature $T_g = 28^\circ\text{C}$ was dissolved in distilled water at 2% (m/v) at 100°C . The warm solution was transferred in a syringe, filtered through a 0.45 mm Millex™ membrane, and dispensed in 1 ml vials. At room temperature this stock solution becomes a gel. Before use the gel was liquefied at 90°C in a heating block and kept at 35°C . Appropriate volume of warm agarose solution was added to the mixture of virus and crystallant to reach a final polysaccharide concentration of 0.2% (m/v) (for additional details, see ref. [4]). Crystallization assays containing agarose gel were stored at 20°C

Solubility measurements. Drops composed of 5 μl virus and 5 μl crystallant solution were deposited on 22 mm round glass coverslips treated to be water-repellent and equilibrated over 500 μL reservoirs. Assays were left undisturbed for 3 months at 20°C . After this time, two independent measurements were performed. For each, a 1 μl sample was collected under the binocular with a pipette equipped with a thin tip from a part of the drop that contained neither crystals nor precipitate. After mixing with 5 μl water the absorbance was measured at 260 nm. The derived concentration of the virus in equilibrium with crystals is called the solubility S . Supersaturation (σ) is defined as the quotient of the initial concentration C in the crystallization drop over S , such as $\sigma = C/S$.

X-ray analysis. Prior to flash cooling in a nitrogen stream, crystals were transferred in a cryo-protectant solution. This solution was composed of 100 mM Hepes pH 7.2, 30% (v/v) glycerol, and 1% (m/v) PEG-3350 in excess to the concentration in the crystallization drop. A pyramidal crystal of the GFLV-TD variant measuring 250 μm across was analyzed on the FIP-BM30 beamline at the European Synchrotron Radiation Facility (Grenoble, France) using an ADSC Q315r detector (wavelength 0.98 Å, crystal-to-detector distance 500 mm). A first set of 121 images (oscillation angle 0.25° , exposure 120 s) was merged with a second one (oscillation angle

0.2°, exposure 120 s). The latter set initially encompassed 1100 images, but frames recorded beyond number 400 were discarded because of significant radiation damage (**Table 1**). Crystals of both variants were also irradiated on Beamline X06DA of the Swiss Light Source at the Paul Scherrer Institute (Villigen, Switzerland). Sets of 400 or 600 diffraction images (exposure 2 s, oscillation angle 0.3°) were collected (**Table 1**). Data were processed with *XDS* and *XSCALE* [28]. Indexed intensities were converted to structure factors without any sigma cut-off using *TRUNCATE* and non-crystallographic symmetry elements associated to the viral capsid were searched in the Self-rotation function using *POLARRFN*, two programs from the *CCP4* package [29].

ACKNOWLEDGEMENTS

The authors thank the European Synchrotron Radiation Facility and the Swiss Light Source for the beamtime allocated to this project. In particular, they acknowledge the help of M. Pirrochi and J.L. Ferrer (on beamline FIP-BM30 at ESRF) and of M. Wang (on beamline X06DA at SLS). They are grateful to C. Schulze-Briese and F. Rey for interest in this project.

We thank the staff of the Experimental Unit at INRA Colmar for rearing the plants. P. Schellenberger was supported by a fellowship from Département Santé des Plantes et Environnement (SPE-INRA) and the Conseil Interprofessionnel des Vins d'Alsace (CIVA). This work was supported by grants from the Département Santé des Plantes et Environnement (SPE-INRA), CNRS and the Université de Strasbourg.

Authors contribution

Conceived and designed the experiments: PS and BL. Performed the experiments: PS, GD, MB, BL, VO and CS. Analyzed the data: PS, BL, MB, VO and CS. Wrote the paper: BL, PS, CS, GD, OL, CR.

REFERENCES

1. Gan, L. and J.E. Johnson, *An optimal exposure strategy for cryoprotected virus crystals with lattice constants greater than 1000 Å*. J Synchrotron Radiat, 2008. **15**(3): p. 223-226.
2. Girard, E., et al., *The first crystal structure of a macromolecular assembly under high pressure: CpMV at 330 MPa*. Biophys. J., 2005. **88**(5): p. 3562-3571.
3. Lorber, B., J. Witz, and *An investigation of the crystallogenesi s of an icosahedral RNA plant virus with solubility phase diagrams*. Crystal Growth & Design, 2008. **8**: p. 1522-1529.
4. Lorber, B., et al., *Crystal growth of proteins, nucleic acids, and viruses in gels*. Prog Biophys Mol Biol, 2009. **101**(1-3): p. 13-25.
5. Sauter, C., K. Dhouib, and B. Lorber, *From macrofluidics to microfluidics for the crystallization of biological macromolecules*. Crystal Growth and Design, 2007. **7**(11): p. 2247-2250.
6. Andret-Link, P., et al., *Grapevine fanleaf virus: Still a major threat to the grapevine industry*. J. Plant Pathol., 2004. **86**(3): p. 183-195.
7. Sanfaçon, H., *Nepoviruses*, in *Encyclopedia of Virology* B.W.J. Mahy and M.H.V. Van Regenmortel, Editors. 2008. p. 405-413.
8. Quacquarelli, A., et al., *Properties of Grapevine Fanleaf Virus*. J. Gen. Virol., 1976. **32**: p. 349-360.
9. Pinck, M., et al., *Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus*. FEBS Lett, 1991. **284**(1): p. 117-119.
10. Chandrasekar, V. and J.E. Johnson, *The structure of tobacco ringspot virus: a link in the evolution of icosahedral capsids in the picornavirus superfamily*. Structure, 1998. **6**(2): p. 157-171.
11. Viry, M., et al., *Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs*. J. Gen. Virol., 1993. **74** (2): p. 169-174.
12. Mayo, M.A. and D.G. Robinson, *Nepoviruses: molecular biology and replication*, in *The plant viruses, Polyhedral virions and bipartite RNA genomes*, B.D. Harrison and A.F. Murrant, Editors. 1996, Plenum Press: New York, USA. p. 287-288.
13. Lorber, B., et al., *Formation of two-dimensional crystals of icosahedral RNA viruses*. Micron, 2008. **39**(4): p. 431-446.
14. Ducruix, A. and R. Giegé, *Crystallization of nucleic acids and proteins: A practical approach*, IRL press, Oxford University Press.
15. McPherson, A., *Crystallization of biological macromolecules* 1999, Cold spring Harbor Laboratory Press: Cold Spring Harbor, N.Y.
16. Serghini, M.A., et al., *RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location*. J. Gen. Virol., 1990. **71**: p. 1433-1441.
17. Sunagawa, I., *Crystals, growth, morphology, and perfection*. 2005, Cambridge University Press: Cambridge, U.K.
18. Kuznetsov, Y.G., et al., *The incorporation of large impurities into virus crystals*. Acta Cryst. D, 2005. **61**(3): p. 720-723.
19. Schellenbaum, P., et al., in *Advances in Plant Virology, colloque AAB*. 2003: Montpellier.

-
20. Sauter C, et al., *Crystallization – General methods in International Table for Macromolecular Crystallography*, M. Rossmann and E. Arnold, Editors. 2010, International Union of Crystallography, in press.
 21. Casselyn, M., et al., *Spherical plant viruses: interactions in solution, phase diagrams and crystallization of brome mosaic virus*. Acta Cryst. D, 2001. **57**(12): p. 1799-812.
 22. Boistelle, R. and J.P. Astier, *Crystallization mechanisms in solution*. J. Crystal Growth 1988. **90**: p. 14-30.
 23. Gavira, J.A. and J.M. Garcia-Ruiz, *Agarose as crystallisation media for protein: Trapping of gel fibers into the crystals*. Acta Cryst. D, 2002. **58**: p. 1653-1656.
 24. Biertumpfel, C., et al., *Characterization of crystals of the Hjc resolvase from Archaeoglobus fulgidus grown in gel by counter-diffusion*. Acta Cryst. F, 2005. **61**(7): p. 684-687.
 25. Zhu, D.W., et al., *Growth kinetics, diffraction properties and effect of agarose on the stability of a novel crystal form of Thermus thermophilus aspartyl-tRNA synthetase-1*. Acta Cryst. D, 2001. **57**(4): p. 552-558.
 26. Vuittenez, M., M.C. Munck, and J. Kuszala, *Souches de virus à haute agressivité isolées de vignes atteintes de dégénérescence infectieuse*. Etudes Virol. Appliquée, 1964. **5**: p. 69-78.
 27. Pinck, L., et al., *A satellite RNA in Grapevine Fanleaf Virus Strain F13*. J. Gen. Virol., 1988. **69**: p. 233-239.
 28. Kabsch, W., *XDS in International Tables for Crystallography*, in *Crystallography of Biological Macromolecules*, M.G. Rossmann and E. Arnold, Editors. 2001, Dordrecht: Kluwer Academic Publishers.
 29. *Collaborative Computational Project Number 4* Acta Cryst. D, 1994. **50**: p. 760-763.

4. Article 4 : Structural insights into the molecular mechanisms governing *Grapevine fanleaf nepovirus* transmission by nematodes

Article en preparation



**Structural insights into the molecular mechanisms governing
Grapevine fanleaf nepovirus transmission by nematodes**

Pascale Schellenberger^{1,2}, Claude Sauter³, Bernard Lorber³, Patrick Bron⁴, Stefano Trapani⁴,
Marc Bergdoll², Olivier Lemaire¹, Gérard Demangeat¹ and Christophe Ritzenthaler²

¹Institut National de la Recherche Agronomique, INRA/UDS UMR 1131, 28 rue de Herrlisheim, 68021 Colmar cedex, France. ²Institut de Biologie Moléculaire des Plantes, CNRS/UDS UPR 2357, 12 rue du Général Zimmer, 67084 Strasbourg Cedex, France. ³Institut de Biologie Moléculaire et Cellulaire, CNRS/UDS UPR 9002, 15 rue R. Descartes 67084 Strasbourg Cedex, France. ⁴Centre de Biochimie Structurale CNRS/Université de Montpellier II UMR 5048 - INSERM UMR 554, 29 rue de Navacelles, 34090 Montpellier Cedex, France.

ABSTRACT

Many animal viruses and the vast majority of plant viruses depend on vectors for transmission to their natural hosts. The major virus of grapes worldwide, *Grapevine fanleaf virus* (GFLV) is transmitted in a highly specific manner by the nematode *Xiphinema index*. However, the molecular rules governing this specificity remain unknown. In order to elucidate these rules, we determined the crystal structures of a wild-type virus and of a poorly transmitted variant at 3.0 Å and 2.7 Å resolutions, respectively. The variant differed from its parental strain by a single Gly₂₉₇Asp mutation in a loop exposed at the outer surface of the virus. Structural comparisons did not reveal any significant conformational changes. Therefore, the loss of transmission resulted from the addition of a negatively charged side chain at the surface of each subunit in the viral particle. These data provide the first structural insights into how a plant virus interacts with nematode vectors and reveals a putative nematode binding pocket within the CP B domain.

INTRODUCTION

It has long been recognized that transmission is a specific process allowing defined vectors to transmit certain viruses and comparative studies of transmissible and non-transmissible virus isolates initially allowed the identification of determinants governing this process [1-7]. Numerous virus infecting mammals and nearly all plant viruses use vectors, mainly arthropods, to propagate. Structural capsid proteins (CP) that make up the viral capsid are important for the interaction with the vector and contain such determinants. It has been shown for non-circulative viruses, *i.e.* viruses that do not replicate in their vectors, that CPs may require other viral proteins, referred to as helper component (HC), to bind to the vector (for reviews, [8-11]). Protein motifs in CPs or HCs have been identified for a broad range of non-circulative viruses, in particular those vectored by insects like aphids. For instance, rod shaped potyviruses possess conserved stretches of amino acids necessary for aphid transmission: a single DAG motif near the N-terminus of the CP and two motifs in the multifunctional protein called HC-Pro with a KITC sequence for vector binding and a PTK sequence for capsid binding [12-15]. Recently, two additional single residue sites have been characterized on CP (H²⁵⁶) and on HC-Pro (R⁴⁵⁵) that are important for direct CP-HC-Pro interaction [16]. Point mutations in all these motifs have a negative effect on transmission rates [13, 16-18]. However, the resolution of recent Cryo-TEM 3D reconstructions of potyviruses [19] and their HC-Pro partners [20, 21] is not yet sufficient to precisely pinpoint these motifs on the CPs and HC-Pro. In the case of the icosahedral *Cauliflower mosaic virus* (CaMV, family *Caulimoviridae*), transmission is even more complex since it relies on three viral proteins, the CP named P4 and two HCs (P2 and P3). P2 binds simultaneously to the aphid vector [22] and to P3 [23] via its N- and C-termini, respectively. The HC interaction linking

P2 and P3 is mediated by coiled-coil domains involving the N-terminus of P3 which, in turn, binds to the CP through its residues 80-112 [12]. Structural analyses have revealed the existence of different conformations for P3 that may regulate the anchoring on the virions and prevent P2 oligomerization into transmission-defective paracrystals [24, 25]. In contrast to CaMV, the CP is the unique viral element governing aphid transmission of the icosahedral *Cucumber mosaic virus* (CMV, family *Cucumoviridae*) [26]. There the β H- β I loop exposed on the virion surface consisting mainly of negatively charged amino acids and conserved among cucumovirus, is thought to be involved in electrostatic interactions with components in the aphid mouthparts during transmission [27-29].

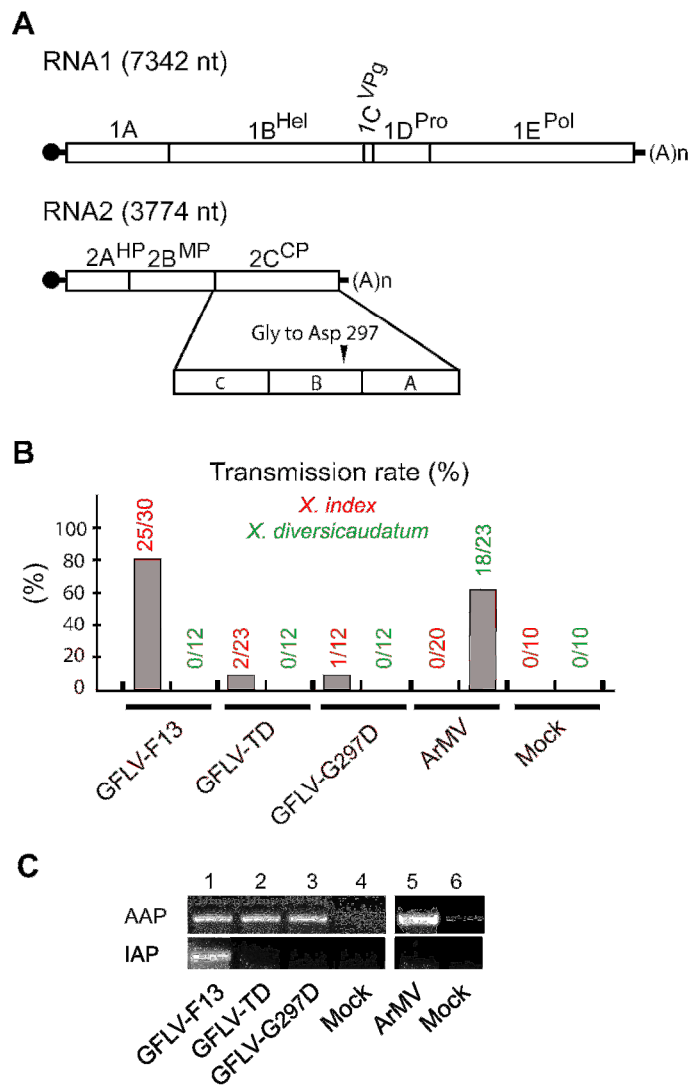
Much less is known at the structure level regarding transmission determinants of soil-borne nepoviruses and tobnaviruses transmitted by nematodes. For rod shaped tobnaviruses, the highly antigenic and partly unstructured C-terminal tail of the CP at the surface of the particles was demonstrated to be necessary, but not sufficient, to promote transmission [30, 31]. In this case, other viral proteins like protein 2b may act as HC [30-32]. The other nematode-transmitted genus, namely *Nepovirus*, belongs to the order *Picornavirales* [33] and consists of icosahedral viruses with pseudo $T = 3$ symmetry made of 60 identical subunits. *Tobacco ringspot virus* (TRSV) is the only *Nepovirus* whose crystal structure is solved [34]. Analysis of its surface residues revealed the presence of two regions conserved within nepovirus genus possibly involved in nematode transmission [34]. However, in the absence of reported strains deficient in nematode transmission the precise determination of transmission determinants has remained a challenge. The two closely related nepoviruses causing the fanleaf degeneration of grapevine provide an appealing system to address this issue. Indeed, *Grapevine fanleaf virus* (GFLV) and *Arabidopsis mosaic virus* (ArMV) which are responsible for one of the most widespread and damaging viral disease of grapevine worldwide [35], are transmitted in an exclusive manner by the nematodes *Xiphinema index* and *X. diversicaudatum*, respectively. The production of infectious cDNA clones of GFLV [36] and GFLV/ArMV recombinant confirmed the CP as the unique determinant for vector transmission and specificity [37, 38]. Thus, a chimeric GFLV encoding a CP of ArMV origin becomes totally recalcitrant to transmission by *X. index* but vectored by *X. diversicaudatum* with the same efficiency than ArMV [39]. Further, the exchange of regions predicted to be at the virion surface and different in GFLV and ArMV allowed to identify a stretch of 11 amino acid in the β B- β C loop of the CP B-domain essential for GFLV transmission by *X. index* (Schellenberger *et al.*, *J. Virol.* submitted).

To better understand the GFLV transmission mechanism and how it specifically interacts with its nematode vector, we determined the crystal structure of wild-type virus (GFLV-F13) and a naturally occurring transmission defective variant (GFLV-TD). The latter differs by a single Gly297Asp mutation from the wild-type GFLV. This study revealed a GFLV architecture close to the TRSV one, consisting of three β -barrel domains linked by extended polypeptides. Remarkably, the GFLV variant retained the same structure as the wild-type, and the residue in position 297 that affects transmission is exposed in the β G- β H loop at the particle surface. The introduction of a negative

charge side chain in every subunit is sufficient to hinder the interaction with the vector. This provides the first structural insights into how a plant virus may interact with nematodes and open new perspectives that will be discussed below.

Figure 1. Involvement of the capsid protein residue 297 in nematode transmission.

(A) Genomic organization of the GFLV. Open boxes indicate open reading frames (ORF). The 5' and 3' untranslated regions are denoted by single lines and the VPg is represented by a black circle. RNA1 and RNA2 code for five (1A-1E) and three (2A-2C) ORFs, respectively. They are translated into one polyprotein which lead to mature viral proteins after processing: helicase (Hel), viral protein genome linked (VPg), protease (Pro), polymerase (Pol), homing protein (HP), movement protein (MP), coat protein (CP). As indicated, the capsid protein is composed of three domains called C, B, A. In the variant GFLV-TD, the Gly residue at position 297 is replaced by Asp. **(B)** Transmission of wild type GFLV (GFLV-F13), viruses with the substitution of the Gly297 by an Asp (GFLV-TD and GFLV-G₂₉₇D) and wild type ArMV by *X. index* and *X. diversicaudatum*. Transmission rate is expressed as the percentage of infected plants with respect to the total tested plant. **(C)** Virus detection in *X. index* at the end of the AAP and the IAP showed that the mutated viruses, as ArMV, were ingested but not retained in nematodes. Thirty nematodes specimens exposed to source plants infected with GFLV-F13 (1), ArMV (5), GFLV-TD (2), GFLV-G₂₉₇D (3), or mock inoculated plants (4, 6) were randomly collected and tested by RT-PCR with GFLV (1-4) or ArMV (5-6) specific primers. DNA products were analyzed by electrophoresis on 1.5% agarose gels.



RESULTS

Identification and characterization of a GFLV variant defective in nematode transmission.

GFLV strain F13 (GFLV-F13) was first isolated from a Southern France vineyard in 1964 [40] and ever since propagated on grapevine either by grafting or nematode transmission, or in the laboratory, by mechanical inoculation on *Chenopodium quinoa*, a systemic herbaceous host for GFLV. Its genome is divided in two plus strand RNA molecules with essential roles in replication for RNA1 and movement and encapsidation functions for RNA2 (**Figure 1A**). The screening for variants impaired in nematode transmission from a mechanically maintained GFLV-F13 collection, led to the identification of a variant named GFLV-TD poorly transmitted by *X. index*, the GFLV vector, and totally recalcitrant to transmission by *X. diversicaudatum*, the ArMV vector (**Figure 1B**). Apart from this difference in transmission, GFLV-TD remained indistinguishable from its parental GFLV-F13 strain in delays of symptoms appearance on *C. quinoa* apical leaves, DAS-ELISA absorbance value and purification yields. RT/PCR analyses performed on nematodes at the end of the monthly Acquisition Access Period (AAP) period, revealed no differences in the ability of GFLV-F13 and GFLV-TD to be ingested by vectors (**Figure 1C, top lane**). However, at the end of the Inoculation Access Period (IAP), only GFLV-F13 remained detectable in *X. index* (**Figure 1C, bottom lane**). It is therefore concluded that the defect in transmission of GFLV-TD is likely due to the incapacity of the virus to be specifically retained on the vector mouthparts.

GFLV CP being the unique determinant required for nematode transmission [37-39], it was therefore analyzed for possible mutation(s). The CP coding sequence of GFLV-TD was sequenced and revealed a single Gly to Asp mutation at position 297 in the CP. To assess whether this mutation was at the origin of the defect in nematode transmission, the same substitution was introduced into the GFLV genome through site-directed mutagenesis of the cDNA infectious clone of GFLV [36]. The *in vitro* generated variant thereof, named GFLV-G₂₉₇D, failed to be efficiently transmitted by *X. index* due to loss of retention of the virus in the vector, therefore mimicking GFLV-TD (**Figure 1B, C**). This confirms the critical function of residue 297 in GFLV transmission by *X. index*. This position is highly conserved within the 219 GFLV CP sequences available from the database (Schellenberger *et al.*, *J. Virol. submitted*), suggesting that the mutation arise as a consequence of the mechanical propagation of the virus under laboratory conditions. In ArMV, position 297 is slightly less conserved and occupied predominantly by an Ala (Schellenberger *et al.*, *J. Virol. submitted*).

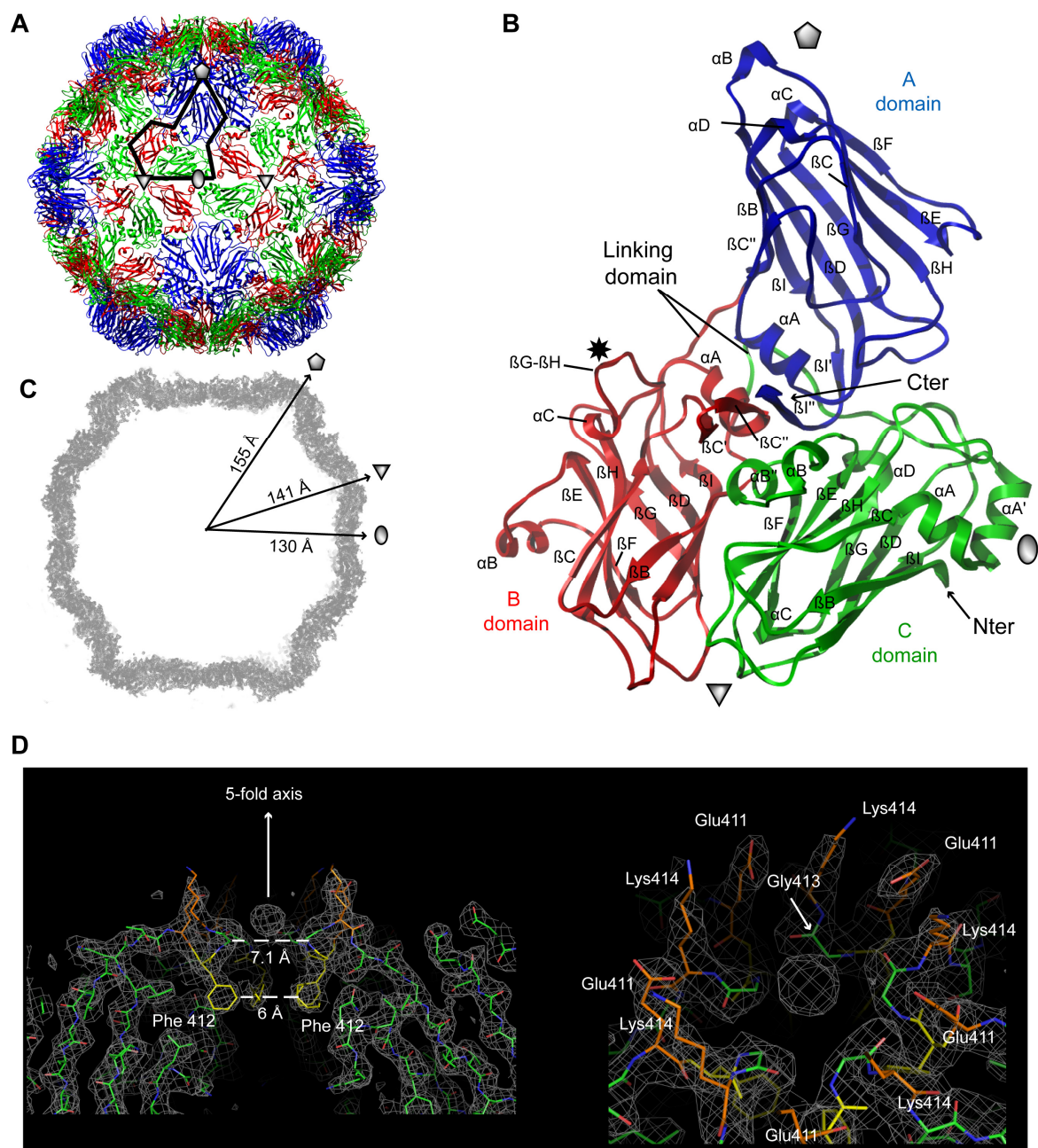


Figure 2. Crystal structure of GFLV-F13 and GFLV-TD. (A) Both structures of GFLV-F13 and -TD are very similar as illustrated by extremely low RMSD values (see Table S1). Thus, only the highest resolution model (GFLV-TD) is represented in this figure. Ribbon diagram of the virus capsid viewed down an icosahedral 2-fold axis normal to the paper plane. Sixties copies of the capsid protein (CP) are arranged in an icosahedral pseudo $T = 3$ symmetry. The black line delineates one CP position. The grey pentagon, triangle and oval symbolize the icosahedral 5-fold, 3-fold and 2-fold icosahedral symmetry axes, respectively. (B) Each CP comprises 3 jelly roll β sandwiches showing similar secondary structure organisation. They are termed C, B and A domains from N- to C-terminus and are depicted in green, red and blue, respectively. A star indicates the position of residue 297. (C) The central section of the $2Fo-Fc$ electron density map (2σ contour level) of GFLV-F13 particle is viewed down a viral twofold axis. The outer radial dimensions along the icosahedral symmetry axes are indicated. (D) Thin slice of $2Fo-Fc$ electron density map (contoured at 1σ) revealing a strong density peak (about 3σ in the $2Fo-Fc$ and 17σ in the $Fo-Fc$ map) on the 5-fold axis of GFLV-TD. The arrow symbolizing the axis points toward at the outersurface of the particle, as well as neighboring charged residues (Lys, Glu), whereas Phe side chains are directed towards the viral cavity. The left panel shows a slightly shifted top view illustrating the organisation of the residues around the fivefold. Five

water molecules bridge the Gly413 carbonyl groups and the Asp411 side chains to the large central ion, possibly a phosphate coming from the crystallization medium.

Structure of GFLV-F13 and GFLV-TD. To gain insight into the structural differences that could explain the defective transmission, we solved the structure of GFLV-F13 and GFLV-TD at 3.0 Å and 2.7 Å, respectively. GFLV-F13 and GFLV-TD complete virions were crystallized as described by Schellenberger *et al.*, *PlosOne submitted*. The asymmetric unit of the GFLV-TD crystal contains 20 subunit of the virus and the one of GFLV-F13 contains the entire virion, *i.e.* 60 subunits. The structure of GFLV-TD was solved by molecular replacement using a cryo-EM model, followed by non crystallographic symmetry (NCS) averaging and refinement at 2.70 Å (**Table 1**). The complete GFLV-TD particle was generated by symmetry operation and used as a model to solve the structure of GFLV-F13 by molecular replacement at 3 Å (**Table 1**).

Irrespective of the virus analyzed, sixty copies of the CP form the capsid pseudo $T = 3$ icosahedral symmetry (**Figure 2A**). The CP folds in 3 jelly-roll β sandwiches termed C, B, and A domains, with two linking peptides between the C-B and the B-A domains (**Figure 2B**). The B and C domains clustered at the 3-fold axis and five A domains clustered at the 5-fold axis forming outersurface protrusions (**Figure 2A**). The particle maximum outer radius, on the 5-fold, 3-fold and two-fold axes, are 155 Å, 141 and 130 Å, respectively (**Figure 2C**). The A domain deviates most from the β sandwich fold of the other domain. For example, the insertion between the β C and the β D strands is extensive and comprises 1 additional strand (**Figure 3**). This is in contrast with closely related comoviruses capsid structures where 2 strands are added at this position [41]. On each 5-fold axis, *i.e.* of the summit of the pentamers, a channel of 7.1 Å in diameter is observed containing an additional electron density suggesting the presence of an ion (**Figure 2D**). However, the distance to the neighboring Lys atoms is incompatible with hydrogen bonding indicating that the ion could be linked by the intermediate of water molecules.

The crystal structures of GFLV-F13 and GFLV-TD were compared. First, the three dimensional structure of GFLV-TD crystal asymmetric unit cell and 20 subunits of GFLV-F13 were superimposed using $C\alpha$ atoms. The superimposition gives an average rmsd of 0.40Å over 10080 $C\alpha$ (all $C\alpha$ matched). In regard to the number of aligned atoms, the rmsd value is exceptionally small. Modest effects of crystal packing interaction were detected on the maximum rmsd values (maximum of 1.9 Å). Those crystal packing effects were the only significant structural differences between the 2 crystal structures. Hence, knowing that each virus icosahedral asymmetric unit (one CP in our case) has equivalent conformation in the capsid, we could compare CPs regardless to their position in the icosahedral lattice, except those involved in crystal packing. Under these conditions, GFLV-F13 and

GFLV-TD CP superimpose almost perfectly (**Figure S1A**) with an average rmsd of 0.12 Å between all C α (**Table S1**).

As expected, GFLV and TRSV have a similar capsid shape. The maximal outer radial distances on the 5, 3 and 2 fold icosahedral symmetry axes are about the same (155, 141 and 130 Å in GFLV versus 155, 137 and 123 Å in TRSV [34]). The overall contact between CP are the same as the one previously described for TRSV [34], with the β H- β I and β B- β C loops of the B and C domain involved on the 3-fold axis subunit interface and the α A' helice of the C domain on the 2-fold axis subunit interface (**Figure S2**). The 3 jelly-roll folding of GFLV and TRSV CP are nearly identical to each other, except the presence of more α helices and 2 supplementary β sheets for the GFLV structure (**Figure 3**). Although GFLV and TRSV CP exhibit quasi-similar folding, the superimposition of both CP is slightly shifted in several areas (**Figure S1B**). The superimposition of the C, B and A domain independently, demonstrates the A domain to be most divergent and the C domain the most conserved (**Table S1**). Despite these many commonalities, several differences between TRSV and GFLV exist. One of the most striking differences affects the β G- β H loop that exhibits strong folding differences between the two crystal structures (**Figure S2B**). In GFLV, the loop is much less protruding, thereby changing the CP topology. Also, the N-terminus tail facing the interior of the capsid in TRSV is absent from the GFLV structure well in agreement with the size difference of the two CPs (504 in GFLV vs 513 in TRSV). In a previous study we produced a 3D model of GFLV with the aim to identify surface residues important for GFLV transmission (Schellenberger *et al. J. Virol. submitted*). Considering this model was based on the resolved structure of TRSV, it displayed approximately the same differences with GFLV (**Figure S2C**) than those observed between TRSV and GFLV (**Figure S2C**).

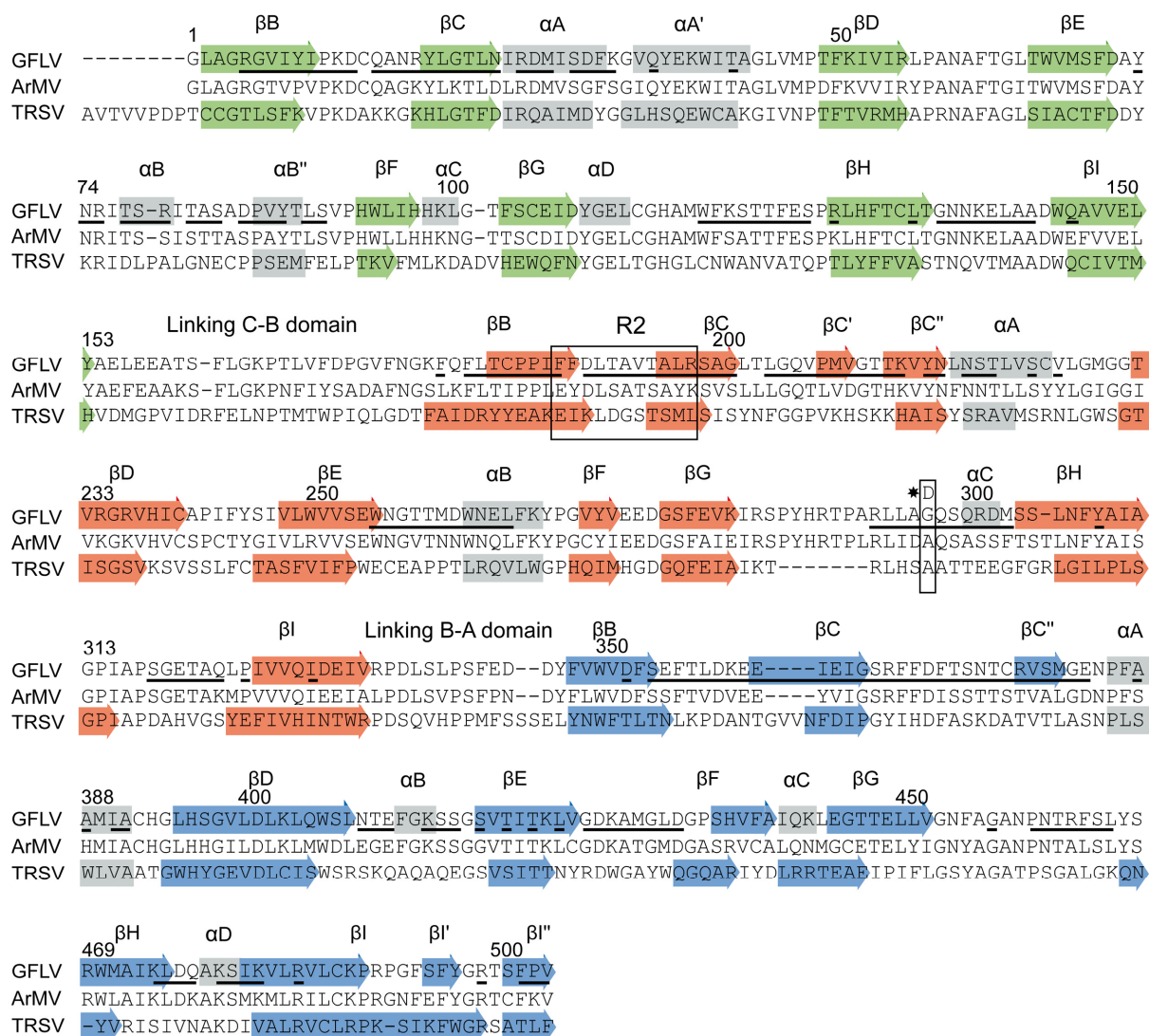


Figure 3. Sequence alignment of GFLV, ArMV and TRSV. Secondary structures observed in GFLV and TRSV crystal structures are indicated on the sequences by arrows (β strands) and grey blocks (α helices). The same color code as in Figure 2 is used for strands to indicate the CP domain they belong to: green, red and blue for the C, B, and A domain respectively. Underlined residues occur on the outer surface of the GFLV capsid. The sequence alignment was created with Clustal X. The Asp at position 297 in the GFLV-TD CP is indicated by a star.

Transmission determinants. Residues important for transmission are likely exposed at the outer surface of the particles. Therefore, we mapped those of GFLV-F13 and GFLV-TD highlighting 208 critical amino-acids (underlined in Figure 3). Remarkably, among those, residue 297 that lies at the outermost portion of the β G- β H loop of the B domain (Figure 2B, Figure S3A) and differs between both strains, appears as one of the most accessible residue to the solvent. The conformation of the β G- β H loop from GFLV-F13 and GFLV-TD is nearly identical in both CP structures with an average distance of only 0.18 Å (Figure 4). Folding modification of the loop can therefore not account for the loss of transmissibility of GFLV-TD. Sequence information together with the electron density map at 1.2 σ unambiguously assess the presence of the Asp side chain in GFLV-TD otherwise absent in

GFLV-F13 (**Figure 4**). Considering the overall lack of structural difference between both viruses, including the β G- β H loop (**Figure S1A** and **Figure 4**), it is concluded that the addition of a carboxylic negatively charged side chain is *per se* sufficient to trigger the loss of transmission of GFLV-TD.

Interestingly, within the CP crystal structure, residue 297 is situated nearby the β B- β C loop of the CP B domain (22 Å and 31 Å, min and max distances, **Figure S3A**) previously defined as a transmission determinant (region R2, Schellenberger *et al. J. Virol. submitted*). On the quaternary structure, R2 is close to the 3-fold axis while Gly²⁹⁷ stands in the middle of a canyon between the 3-fold and the 5-fold protuberances (**Figure S3B**). In agreement with our predictions (Schellenberger *et al. J. Virol. Submitted*), the crystal structure of GFLV confirmed R2 to be well exposed at the capsid outer surface with 10 residues out of 11 accessible to the solvent including all 6 amino acids of the β B- β C loop (Figure 3). The remaining 5 locate to secondary structures, either to the β B strand, (Phe¹⁸⁸ and Phe¹⁸⁹) or to the neighboring β C strand (Ala¹⁹⁶ Leu¹⁹⁷ and Arg¹⁹⁸) (**Figure 3**). When compared to TRSV, both β B- β C loops adopt the same conformation (**Figure S1B**), with an rmsd of 1.45 Å between all C α .

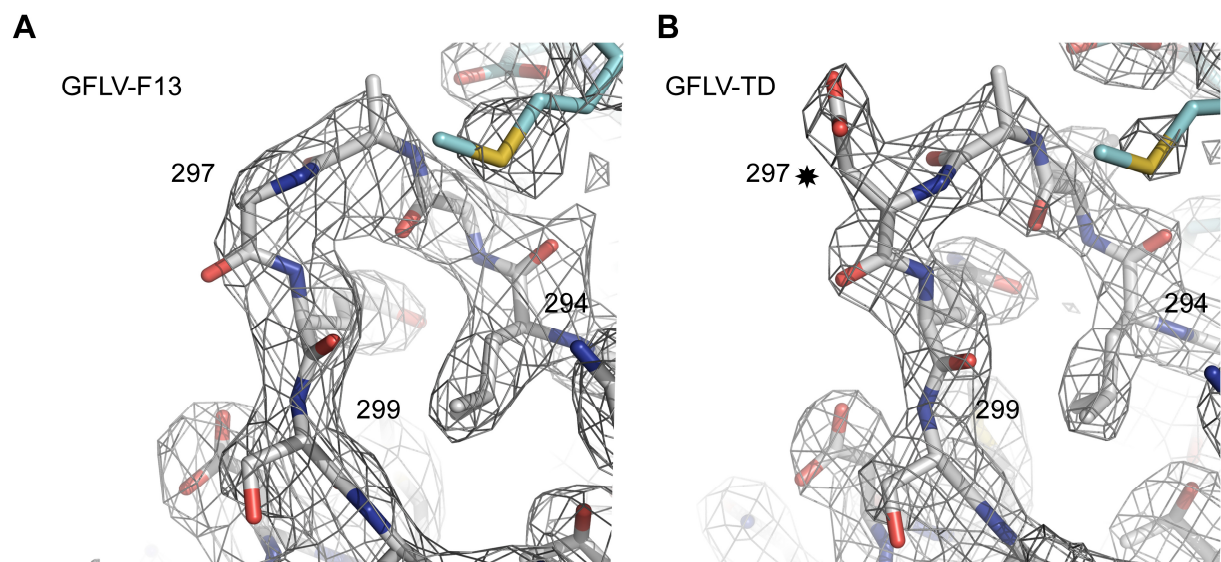


Figure 4. Close-up view of the protruding β G- β H loop in GFLV-F13 (A) and GFLV-TD (B). The conformation of the loop remains the same despite the introduction of an Asp residues at position 297 of the transmission defective variant. *2Fo-Fc* electron density maps are contoured at 1.2 σ .

Role of the viral determinant in transmission specificity. Knowing that position 297 and region R2 (Schellenberger *et al., J. Virol. submitted*) are critical for the transmission of GFLV by *X. index*, we wanted to know if they act together to confer transmission specificity. For this, we exchanged residues in both locations by their ArMV counterparts and searched for an occurrence of transmission by its vector, *X. diversicaudatum*. The substitution Gly₂₉₇Ala, generated a recombinant named GFLV-G₂₉₇A. The substitution of R2 and Gly²⁹⁷ by their ArMV counterpart generated a recombinant named GFLV-R2G₂₉₇A. Both viruses produced infections *in planta* and could therefore be analyzed for transmission.

GFLV-G₂₉₇A, tested positive for transmission by *X. index* but failed to be transmitted by *X. diversicaudatum* (Figure 5). In contrast, GFLV-R2G₂₉₇A in which both regions were replaced by ArMV residues, failed to be vectored by either nematodes species (Figure 5), although the nematodes had ingested virions during AAP (Figure S4). This indicates that the substitution of the CP transmission determinant Gly²⁹⁷ and R2 by ArMV sequences are not sufficient to confer a gain of transmission by *X. diversicaudatum* and suggests the existence of other residues involved in this process. Interestingly, both transmission determinants, Gly²⁹⁷ and R2, are situated at the edge of a pocket formed by the CP B domain, more precisely the βC', βC'', βB, βC strands and several residues within the βG-βH loop (Figure 6).

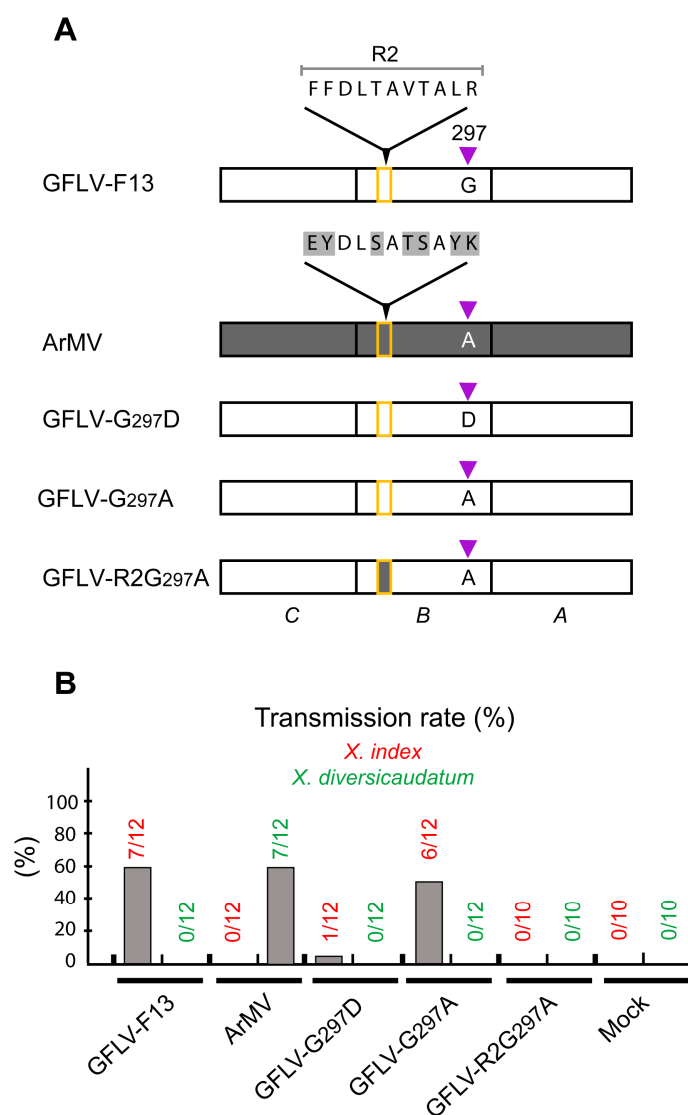


Figure 5. Nematode transmissibility of GFLV CP mutants. (A) The tested constructs were different in their B domain, some of them containing modifications in the R2 region (residues 188 to 198) and/or at position 297. The nature of residue at position 297 is indicated and the R2 region consisted either of the GFLV (open rectangle with orange border) or of ArMV (grey rectangle with orange border) sequences. The sequence differences are highlighted in grey in the enlargement of ArMV R2 region. (B) Transmission rate is expressed as the percentage of infected plants with respect to the total tested plant.

DISCUSSION

The high-resolution structure of the GFLV reveals an overall organization well in agreement with its classification in the order *Picornavirales* [33] with a capsid consisting of 60 capsomers, each containing three distantly related jelly-roll domains that may have arisen by triplication of a single copy present in some ancestor virus and subsequent divergent evolution [34, 42]. GFLV, as a representative of the genus *Nepovirus*, is structurally very similar to TBSV (Figure 3, S1, Table S1). Yet, the superposition of their capsid proteins is not perfect mainly due to differences in the length of surface loops (see loop β G- β H for instance) and the orientation of the subunits within the particles. This is probably the reason why classical molecular replacement using homology models derived from TRSV was not as straightforward as expected, whereas the cryo-EM map of GFLV rapidly led to an unambiguous electron density map. Overall, the resulting structures of the wild-type and a transmission-defective GFLV are almost identical, although they were determined in very different crystalline contexts (Schellenberger *et al.*, article 3). This shows that the particle is quite rigid and phenotypical differences between the two forms not related to conformational changes but must be searched at the level of their physical-chemical properties.

GFLV-TD is a variant of GFLV-F13 that emerged spontaneously in the laboratory after multiple rounds of virus amplification following mechanical inoculation of *C. quinoa* plants. Many examples of loss of virus transmission by vectors have been observed in such experimental conditions [6, 18, 43-46]. However, to our knowledge this is the first isolation and characterization of a spontaneously occurring transmission-deficient strain within the genus *Nepovirus*. Sequencing revealed the occurrence of a single Gly₂₉₇Asp mutation in the CP and reverse genetic confirmed this was sufficient to trigger the severe loss of transmission of GFLV-TD by *X. index*. In addition, we show that the defect in vection is correlated to a lack of retention, whereas the acquisition of the virus by the vector is not affected. Therefore, it can be concluded that Gly²⁹⁷ is a *bona fide* determinant for GFLV transmission and is likely exposed at the virion outer surface.

Single point mutations detrimental to transmission are not rare and often affect highly conserved residues. For instance, single mutations in the β H- β I conserved loop of CMV either reduce or totally abolish the aphid transmission [28]. Also, single mutations in the conserved PTK motif of ZYMV HC-Pro [13], or in the DAG motif of TYMV CP [17, 47] hinder transmission. Within the 219 sequences of GFLV CP available to date in databases isolated from vineyards worldwide (Schellenberger *et al.*, *J. Virol submitted*), Gly²⁹⁷ is almost fully conserved suffering only two exceptions with Asn²⁹⁷ in both cases. Notice that the transmission of these naturally occurring Asn²⁹⁷ variants has not been assessed yet. However, GFLV is rather unique among plant viruses since it is very stable genetically. This is reflected by the high level of sequence conservation within the CP [48]. Therefore, residue conservation may not only be relevant to the function of specific residues but also linked to this limited variability of GFLV whose origin may be related to the fact that grapevine is

essentially propagated by grafting, that grapevine is the unique natural host of GFLV and *X. index* its unique natural vector.

Beyond its high conservation, Gly²⁹⁷ is located within domain B in the β G- β H loop which was already suspected to be involved in transmission for two main reasons: this loop was predicted to be exposed at the surface of the virion and its sequence is conserved from those in other nepoviruses [34]. The present study confirms the surface position of the β G- β H loop and in particular the Gly²⁹⁷. However, we also show that the Gly₂₉₇Ala single mutant (GFLV-G₂₉₇A) – Ala is the most frequent residue at position 297 in the capsid protein of ArMV– does not alter the transmission by *X. index*, contrarily to a Gly₂₉₇Asp mutation.

Several possibilities can be envisaged to explain a lack of transmission with GFLV-G₂₉₇D or GFLV-TD, and not with GFLV-G₂₉₇A or GFLV-F13. A structural effect of Gly²⁹⁷ on the β G- β H loop can be excluded since the same local conformation is observed in GFLV-TD and GFLV-F13. As seen in **Figure 4**, the side chain from Asp²⁹⁷ in GFLV-TD is the unique significant difference between both viruses. Thus, the introduction of this negative charge is much more likely to affect the physico-chemical properties of the particle in solution by modifying the surface electrostatic potential and the overall solvation shell of the virus. This residue is highly exposed at the surface of each subunit, therefore adding 60 negative charges per particle. The striking effect of this mutation on GFLV behavior during crystallization strongly supports this hypothesis. Thus, GFLV-TD particles are 2.5 times more soluble than wild type GFLV-F13 particles in the same crystallization solutions, and they adopt different crystal packing organizations (Schellenberger *et al.*, article 3). It is likely that the alteration of the electrostatic potential and solvation of the virus surface also impacts the binding and the retention of GFLV to the nematode mouthparts, thereby reducing the transmissibility of GFLV-TD.

Another likely scenario would be that residue 297 belongs to an ensemble of outersurface positions that together form the nematode recognition motif responsible for the specific binding of GFLV to *X. index* mouthparts. Of course, this does not prevent the mutational disruption of a single contact involved in virus-vector interaction to be detrimental to transmission and three observations strongly argue in favor of this hypothesis.

First, bioinformatics analyses based on sequence divergence within nepoviruses pointed out the β G- β H loop encompassing residue 297 as well as other surface residues including the β C- β C'' loop in nematode transmission (Schellenberger *et al.*, *J. Virol submitted* [34]). Here we show that β G- β H and β C- β C'' loops are indeed well exposed at the outer surface of the particles, forming protrusions in the CP B domain (purple and cyan in **Figure 6**). Unfortunately, recombinant viruses designed to elucidate the function of the β G- β H and β C- β C'' loops in transmission could not be tested due to defect in encapsidation and thereby loss of infectivity *in planta* (Schellenberger *et al.*, *J. Virol submitted*).

Second, we previously identified 11 residues (region R2) encompassing the β B- β C loop within the B domain as a GFLV transmission determinant (Schellenberger *et al.*, *J. Virol submitted*). The present data show that 10 of them are accessible to the solvent and are positioned at the extremity of a flat area or plateau (orange **Figure 6**) formed by additional, surface exposed residues in the β B and β C strands. Remarkably, the left and right edges of this plateau are occupied by the β G- β H and β C'- β C'' protrusions, respectively (**Figure 6**).

Third, independently of the nature of ligands, specific virus-ligand interactions almost always involve multiple residues. Examples within the order *Picornavirales*, include the CD155 binding pocket of *Poliovirus 1* (PV1) or the heparin binding pockets of *Foot-and-mouth disease virus* (FMDV) [49, 50]. The heparin recognition motif of FMDV is a shallow depression of positive electrostatic charge at the junction of the three major CPs VP1 VP2 and VP3.

These observations together with our structural and functional data support the existence of a nematode recognition motif within the CP B domain located between the 3-fold and the 5-fold axes of GFLV. This motif involves multiple surface residues, in particular those of the β G- β H and β B , β C loop encompassing the Gly²⁹⁷ and residues of region R2, respectively, as well as amino acids of the β C' and β C'' strands (**Figure 6**). Remarkably, this pocket shares several commonalities with the heparin recognition motif of FMDV. Thus, both pockets occupy a very similar position within the icosahedral asymmetric unit (**Figure S5**) and present similar electrostatic features characterized by the presence positively charged residues [49], whereas the surface of GFLV is otherwise essentially negative or neutral (**Figure 7**). Whether the occurrence of negatively charged residues in this pocket is detrimental for GFLV transmission by its vector needs to be determined. In this respect, it should be noted that the two only situations that led to defect in transmission also affected the net electrostatic charge of the putative nematode binding pocket by introducing negative charges: Gly₂₉₇Asp (GFLV-TD, this work, **Figure 7**) and Phe₁₈₈Glu (the first residue in region R2, Schellenberger *et al.*, *J. Virol, submitted*). This provides a new framework for further analyses aimed at precisely defining the function of electrostatic potential and structure of the predicted binding pocket in nematode transmission. Fundamental knowledge of the molecular mechanisms underlying virus transmission is a prerequisite for the creation of novel strategies to control virus transmission and to identify the ligand within the nematode mouthpart.

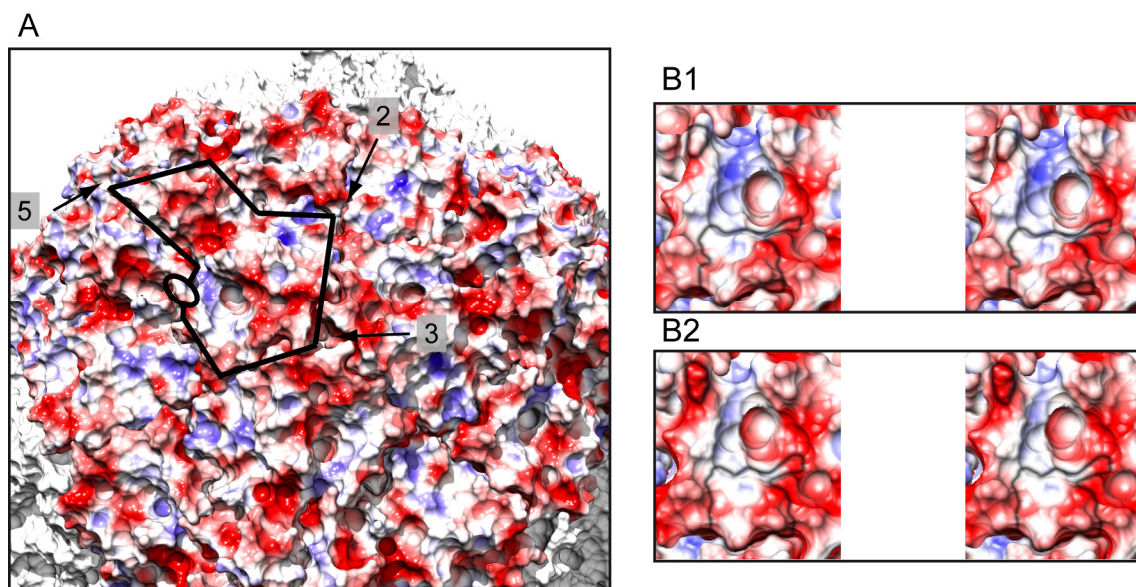


Figure 7. Electrostatic surface potential of GFLV particles. (A) GFLV-F13 capsid with red, blue and white regions representing negative, positive and neutral potentials, respectively. The black line highlights a single CP and arrows denote the 5, 3, and 2-fold axis. Residue 297 is located in the black circle on top of a negatively charged pocket which is enlarged in panels (B1) and (B2). The latter are stereoviews of the pocket in GFLV-F13 (B1) and GFLV-TD (B2), respectively. The transmission determinant region R2 is located at the bottom of the pocket and is denoted by a thin black line in (B1-2). The electrostatic potential was calculated with APBS and visualized from -3 to 3 k/e⁻ with Chimera using a probe of 3 Å to display a smoother surface.

CONCLUSION

Although experimental data are still very sparse, it is widely accepted that effective viral transmission from host to host relies on a strong and specific interaction with its vector. Here, we identified such an interaction area on the surface of a 30 nm icosahedral nepovirus. Furthermore, we show that a single mutation (Gly₂₉₇Asp) in this zone is sufficient to disrupt GFLV propagation from one plant to another. In the absence of any detectable difference in the 3D structures of the wild-type virus and the mutant, it is very likely that the introduction of a negative charge in the 60 subunits forming the mutant capsid is the reason why the nematode *Xiphinema index* is no longer able to carry the virus.

This preliminary mapping of the binding site on the GFLV surface opens new perspectives: first, it provides a starting point for the determination of molecular determinants leading to the GFLV recognition by its vector. It is also a step forward toward the identification of a viral receptor at the nematode level. Getting deeper insight into these viral-vector interaction mechanisms at atomic level will certainly facilitate the development of new strategies to break the viral cycle and to cure infected vineyards in a way that is more compatible with sustainable agriculture.

ACKNOWLEDGMENTS

Acknowledgement to Félix Rey and Gérard Pehau-Arnaudet – Institut Pasteur Paris – for helpful scientific advices and assistance in cryo-MET, respectively. The authors thank the European Synchrotron Radiation Facility and the Swiss Light Source for the beamtime allocated to this project. In particular, they acknowledge the help of M. Pirrochi and J.L Ferrer (on beamline FIP-BM30 at ESRF) and of M. Wang (on beamline X06DA at SLS).

MATERIALS AND METHODS

Virus production. GFLV strain F13 [40] and ArMV strain S [51] were isolated from naturally infected grapevines and propagated in *Chenopodium quinoa*, a systemic host for both viruses. The sequence of the two genomic RNAs of both virus strains is determined [52-54], Ritzenthaler et al., unpublished). Full-length cDNA clones of GFLV-F13 RNA1 and RNA2 are available [36]. They were used in this study to produce RNA molecules by in vitro transcription as previously described by Belin *et al.*, 1999 [55]. Transcripts of either GFLV RNA1 and RNA2 or GFLV RNA1 and mutated RNA2 were mechanically inoculated to *Chenopodium quinoa* [36]. Virus infection was assessed in uninoculated apical leaves of tested plants 2 to 3 weeks post-inoculation by double-antibody sandwich (DAS)-enzyme-linked immunosorbent assay (ELISA) with specific γ -globulins to

GFLV and ArMV [56]. Samples were considered positive if their optical density (OD_{405nm}) readings were at least three times those of healthy controls after 120 min of substrate hydrolysis.

Mutagenesis. Plasmid pVec_{Acc65I}2ABC, carrying the full-length cDNA copy of GFLV RNA2 and the 5' and 3' non-coding sequences, was used as template for the production of CP harboring a single mutation at amino acid 297 of the coat protein (CP) by PCR site directed mutagenesis overlap extension mutagenesis (Schellenberger *et al*; *J. Virol*, submitted). Plasmid pVec_{Acc65I}2ABC_{G2} is a pVec_{Acc65I}2ABC derivative carrying the CP region R2 in position nts 2,609-2,640 (nucleotide positions are given according to the GFLV-F13 RNA2 sequence, GenBank accession no. NC_003623)(Schellenberger *et al*; *J. Virol*, submitted). Residue 297 (corresponding codon nts 2,936-2,938) was mutated into an aspartic acid, using pVec_{Acc65I}2ABC as template, the mutagenic primers mutDF/mutDR and the external primers 18/36 (Table S2). Mutagenic PCR amplified products were digested with both *Acc65I* (nts 2,678-2683) and *BglIII* (nts 3,055-3,060) and cloned into the corresponding sites in pVec_{Acc65I}2ABC to yield pVec_{Acc65I}2ABC_{G297D}. Residue 297 was mutated into Alanine with the mutagenic primers mutAF/mutAR and the external primers 18/36 (Table S2); PCR-amplified products were digested with *Acc65I* and *BglIII*, and cloned into the corresponding sites in pVec_{Acc65I}2ABC and pVec_{Acc65I}2ABC_{G2} to yield pVec_{Acc65I}2ABC_{G297A} and pVec_{Acc65I}2ABC_{G2-G297A}, respectively. Each PCR reaction was carried out as described by Schellenberger *et al*; *J. Virol*, submitted. For simplicity, transcripts and mutant viruses derived from this series of constructs were referred to as GFLV-G_{297D} (plasmid pVec_{Acc65I}2ABC_{G297D}), GFLV-G_{297A} (plasmid pVec_{Acc65I}2ABC_{G297A}), and GFLV-R2G_{297A} (plasmid pVec_{Acc65I}2ABC_{G2-G297A}). The integrity of all GFLV RNA2 clones was verified by DNA sequencing.

Nematode transmission tests. Nematode transmission assays were performed as described by Marmonier *et al.*, 2010, with a 2 step procedure (acquisition access period and inoculation access period) of 4 weeks each, using 200 nematodes per plants (see Figure S6). *Chenopodium quinoa* was used as source plants and bait plants for transmission assays using *Xiphinema diversicaudatum*. *Nicotiana benthamiana* was used as source plants and bait plants for transmission assays using *Xiphinema index*.

Detection of GFLV and ArMV in nematodes. The presence of GFLV and ArMV was verified in total RNA extract from nematodes by reverse-transcription (RT)-polymerase-chain reaction (PCR) as described by Marmonier *et al.*, 2010.

Characterization of GFLV RNA2 progeny. The progeny of GFLV RNA2 CP sequence was characterized in infected plants by immuno-capture (IC)-RT-PCR as described by Andret-Link *et al.*, 2004 [37] except that two cDNA fragments were amplified with primers 397/227 and 115/18 (see Table S2). Sequences were analyzed with ContigExpress (Vector NTI).

Cryo-electron microscopy 3D reconstruction The GFLV viral particles were applied to a quantifoil R 2/2 carbon grid (Quantifoil Micro Tools mbH, Germany), blotted by filter paper, and flash-frozen in liquid ethane to make a vitreous-ice embedded GFLV sample. Electron micrographs were recorded under low-dose conditions at liquid-N₂ temperature with a JEOL 2010 operating at 200 kV microscope. Images were collected at X 50,000 magnification with a defocus range of 1.3 – 2.5 μm. Micrographs were digitized on a Nikon Coolscan 9000 ED

with a step size of 10 μm . The digitized images were coarsened by a factor of 2, resulting in a pixel size corresponding to 4 \AA at the specimen level. The semi-automatic X3D program (J.F. Conway) was used for picking particles. The defocus value was estimated for each micrograph using CTFFIND3 [57], and phases flipped using CTFMIX [58]. Particle origins and orientations were determined and refined using the model-based orientation determination method [59]. The GFLV reconstruction was determined using as starting model the 3D reconstruction of the TRSV filtered at 40 \AA resolution. The density map was calculated by Fourier-Bessel formalism as described [59], and implemented in the EM3DR program. Resolution was estimated using the Fourier shell correlation (FSC) criterion with a cutting level of 0.5 [60]. The final density map computed at 16.5 \AA resolution includes 2424 particles extracted from 8 micrographs.

X-ray structure determination and analysis. Crystals of wild-type GFLV (GFLV-F13) and the transmission defective variant (GFLV-TD) particles were prepared by vapor diffusion at 20°C in sub-microliter sitting drops as described by Schellenberger et al. (2010). Data were collected on crystal-cooled samples (**Table 1**) at FIP-BM30 beamline (ESRF, Grenoble, France) and at X06DA beamline (SLS, Villigen, Switzerland). They were reduced using the XDS package [61].

Diffraction data were phased by molecular replacement using *AMoRe* [62, 63] followed by non-crystallographic symmetry (NCS) averaging and solvent flattening using RAVE [64-67]. Attempts to phase data using TRSV-based homology models generated by Modeller [68] were not successful. In contrast, the 3D EM reconstruction led to a clear molecular replacement solution for the cubic data sets in the 30-15 \AA resolution range. The orientations of the viral particles within the cubic crystal were identified by inspection of the self-rotation function calculated at the highest resolution available (4.5 and 2.85 \AA respectively). Self-rotations corresponding to four differently oriented icosahedral particles were found. Calculation of the translation-function using the correctly oriented 3D EM model showed that four icosahedral particles were present in the unit cell, each one of them sharing one of its 3-fold axis with the crystal. The molecular replacement solutions defined the molecular boundaries (masks) of the particles within the cubic crystals. Based on the icosahedral symmetry of the 3D EM model, the rigid-body operators relating equivalent regions within the molecular boundaries were defined (20 NCS x 3 crystallographic transformations). An iterative procedure of phase extension from 18 \AA to the maximum available resolution was then carried out by using density modification techniques, including NCS map averaging, solvent flattening and intermediate steps where the molecular masks and the NCS operators were refined.

The incorporation of high resolution data finally converged to an experimental map at 2.85 \AA which allowed the rapid rebuilding of GFLV subunit from homology models. The atomic model of GFLV-TD was refined with Phenix [69] with cubic data reprocessed at 2.7 \AA resolution. NCS constraints were applied to the ensemble of monomers in the asymmetric unit except three regions which changed conformation due to packing contacts (Tyr 9, loops 15-19 and 259-265). Water molecules were added after convergence of capsid refinement. Strong peaks in the difference map were examined in Coot [70] to identify solvent molecules around one monomer. They were then transferred by symmetry to other subunits. The structure of the wild-type GFLV particle was solved by MR using the GFLV-TD model and was refined at 3 \AA resolution. No solvent molecule was included, since it was not possible at this resolution to describe a common hydration pattern for the 60 viral subunits in the asymmetric unit. The stereochemical quality (**Table 1**) of final models was assessed with coot and procheck [71]. The totality of the amino acids (504 residues per subunit) was observed in both wild-type and TD GFLV

structures. Atomic coordinates have been deposited in the Protein Databank (accession numbers: xxxx, yyyy). GFLV structures were superimposed with TRSV and derived homology models, and their root mean square deviations (rmsd) were computed with lsqman [72]. Solvent accessible surface was calculated with a probe radius of 1.4 Å, with the program MSMS [73]. The analysis of the surface potential was performed with APBS [74]. Figures were prepared using PyMol (<http://www.pymol.org/>) and chimera [75].

REFERENCES

1. Atreya, C.D., B. Raccach, and T.P. Pirone, *A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus*. *Virology*, 1990. 178(1): p. 161-165.
2. Huet, H., et al., *Mutations in the helper component protease gene of zucchini yellow mosaic virus affect its ability to mediate aphid transmissibility*. *J. Gen. Virol.*, 1994. 75(6): p. 1407-1414.
3. Lopez-Moya, J.J., R.Y. Wang, and T.P. Pirone, *Context of the coat protein DAG motif affects potyvirus transmissibility by aphids*. *J. Gen. Virol.*, 1999. 80 (12): p. 3281-3288.
4. Perry, K.L., L. Zhang, and P. Palukaitis, *Amino acid changes in the coat protein of cucumber mosaic virus differentially affect transmission by the aphids *Myzus persicae* and *Aphis gossypii**. *Virology*, 1998. 242(1): p. 204-210.
5. Perry, K.L., et al., *Mapping determinants in cucumber mosaic virus for transmission by *Aphis gossypii**. *Virology*, 1994. 205(2): p. 591-595.
6. Robbins, M.A., R.D. Reade, and D.M. Rochon, *A cucumber necrosis virus variant deficient in fungal transmissibility contains an altered coat protein shell domain*. *Virology*, 1997. 234(1): p. 138-146.
7. Vellios, E., D.J. Brown, and S.A. MacFarlane, *Substitution of a single amino acid in the 2b protein of *Pea early-browning virus* affects nematode transmission*. *J. Gen. Virol.*, 2002. 83: p. 1771-1775.
8. Ziegler-Graff, V. and V. Brault, *Role of vector-transmission proteins*. *Methods Mol Biol*, 2008. 451: p. 81-96.
9. Hogenhout, S.A., et al., *Insect vector interactions with persistently transmitted viruses*. *Annu. Rev. Phytopathol.*, 2008. 46: p. 327-359.
10. Ng, J.C. and B.W. Falk, *Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses*. *Annu. Rev. Phytopathol.*, 2006. 44: p. 183-212.
11. Andret-Link, P. and M. Fuchs, *Transmission specificity of plant viruses by vectors*. *J. Phytopathol.*, 2005. 87(3): p. 153-165.
12. Syller, J., *The roles and mechanisms of helper component proteins encoded by potyviruses and caulimoviruses*. *Physiol. Mol. Plant. Pathol.*, 2006. 67(3-5): p. 119-130.
13. Peng, Y.H., et al., *Mutations in the HC-Pro gene of zucchini yellow mosaic potyvirus: effects on aphid transmission and binding to purified virions*. *J. Gen. Virol.*, 1998. 79 p. 897-904.
14. Blanc, S., et al., *Mutations in the potyvirus helper component protein: effects on interactions with virions and aphid stylets*. *J. Gen. Virol.*, 1998. 79 p. 3119-3122.
15. Blanc, S., et al., *A specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus*. *Virology*, 1997. 231(1): p. 141-147.

16. Seo, J.K., et al., *Mutational analysis of interaction between coat protein and helper component-proteinase of Soybean mosaic virus involved in aphid transmission*. *Molecular Plant Pathology*, 2010. 11(2): p. 265-276.
17. Atreya, P.L., et al., *Mutational analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids*. *J. Gen. Virol.*, 1995. 76 (Pt 2): p. 265-270.
18. Canto, T., et al., *Different Helper Component Mutations Associated with Lack of Aphid Transmissibility in Two Isolates of Potato Virus Y*. *Phytopathology*, 1995. 85: p. 1519-1524.
19. Kendall, A., et al., *Structure of flexible filamentous plant viruses*. *J Virol*, 2008. 82(19): p. 9546-54.
20. Ruiz-Ferrer, V., et al., *Structural analysis of tobacco etch potyvirus HC-pro oligomers involved in aphid transmission*. *J. Virol.*, 2005. 79(6): p. 3758-3765.
21. Plisson, C., et al., *Structural characterization of HC-Pro, a plant virus multifunctional protein*. *J. Biol. Chem.*, 2003. 278(26): p. 23753-23761.
22. Uzest, M., et al., *A protein key to plant virus transmission at the tip of the insect vector stylet*. *Proc. Natl. Acad. Sci. U S A*, 2007. 104(46): p. 17959-17964.
23. Leh, V., et al., *Interaction between the open reading frame III product and the coat protein is required for transmission of cauliflower mosaic virus by aphids*. *J Virol*, 2001. 75(1): p. 100-106.
24. Hoh, F., et al., *Structural insights into the molecular mechanisms of Cauliflower mosaic virus transmission by its insect vector*. *J Virol*, 2010.
25. Plisson, C., et al., *Structure of the mature P3-virus particle complex of cauliflower mosaic virus revealed by cryo-electron microscopy*. *J. Mol. Biol.*, 2005. 346(1): p. 267-277.
26. Chen, B. and R.I. Francki, *Cucumovirus transmission by the aphid Myzus persicae is determined solely by the viral coat protein*. *J. Gen. Virol.*, 1990. 71: p. 939-944.
27. Smith, T.J., et al., *The structure of cucumber mosaic virus and comparison to cowpea chlorotic mottle virus*. *J. Virol.*, 2000. 74(16): p. 7578-7586.
28. Liu, S., et al., *A conserved capsid protein surface domain of Cucumber mosaic virus is essential for efficient aphid vector transmission*. *J. Virol.*, 2002. 76(19): p. 9756-9762.
29. Ng, J.C., et al., *Virion stability and aphid vector transmissibility of Cucumber mosaic virus mutants*. *Virology*, 2005. 332(1): p. 397-405.
30. Hernandez, C., et al., *Transmission of tobacco rattle virus isolate PpK20 by its nematode vector requires one of the two non-structural genes in the viral RNA 2*. *J. Gen. Virol.*, 1997. 78: p. 465-467.
31. MacFarlane, S.A., C.V. Wallis, and D.J. Brown, *Multiple virus genes involved in the nematode transmission of pea early browning virus*. *Virology*, 1996. 219(2): p. 417-422.
32. Vassilakos, N., et al., *Tobravirus 2b protein acts in trans to facilitate transmission by nematodes*. *Virology*, 2001. 279(2): p. 478-487.
33. Le Gall, O., et al., *Picornavirales, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture*. *Arch. Virol.*, 2008. 153(4): p. 714-727.
34. Chandrasekar, V. and J.E. Johnson, *The structure of tobacco ringspot virus: a link in the evolution of icosahedral capsids in the picornavirus superfamily*. *Structure*, 1998. 6(2): p. 157-171.
35. Andret-Link, P., et al., *Grapevine fanleaf virus: Still a major threat to the grapevine industry*. *J. Plant Pathol.*, 2004. 86(3): p. 183-195.

-
36. Viry, M., et al., *Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs*. J. Gen. Virol., 1993. 74 (2): p. 169-174.
 37. Andret-Link, P., et al., *The specific transmission of Grapevine fanleaf virus by its nematode vector Xiphinema index is solely determined by the viral coat protein*. Virology, 2004. 320(1): p. 12-22.
 38. Belin, C., et al., *Involvement of RNA2-encoded proteins in the specific transmission of Grapevine fanleaf virus by its nematode vector Xiphinema index*. Virology, 2001. 291(1): p. 161-171.
 39. Marmonier, A., et al., *The coat protein determines the specificity of virus transmission by X. diversicaudatum*. J. Plant Pathol., 2010. 92: p. 275-279.
 40. Vuittenez, M., M.C. Munck, and J. Kuszala, *Souches de virus à haute agressivité isolées de vignes atteintes de dégénérescence infectieuse*. Etudes Virol. Appliquée, 1964. 5: p. 69-78.
 41. Lin, T. and J.E. Johnson, *Structures of picorna-like plant viruses: implications and applications*. Adv. Virus. Res., 2003. 62: p. 167-239.
 42. Chen Stauffacher, Z.C., et al., *Protein-RNA interactions in an icosahedral virus at 3.0 Å resolution*. Science, 1989. 245(4914): p. 154-159.
 43. Gera, A., G. Loebenstein, and B. Raccach, *Protein coats of two strains of Cucumber mosaic virus affect transmission by Aphis gossypii*. Phytopathology, 1979. 69: p. 396-399.
 44. Simons, J.N., *Aphid Transmission of a nonaphid-transmissible strain of Tobacco etch Virus*. Phytopathology, 1976. 66: p. 652-654.
 45. Kakani, K., J.Y. Sgro, and D. Rochon, *Identification of specific cucumber necrosis virus coat protein amino acids affecting fungus transmission and zoospore attachment*. J. Virol., 2001. 75(12): p. 5576-5583.
 46. Moreno, A., et al., *A single amino acid position in the helper component of cauliflower mosaic virus can change the spectrum of transmitting vector species*. J. Virol., 2005. 79(21): p. 13587-13593.
 47. Atreya, P.L., C.D. Atreya, and T.P. Pirone, *Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus*. Proc. Natl. Acad. Sci. U S A, 1991. 88(17): p. 7887-7891.
 48. Vigne, E., et al., *Population structure and genetic variability within isolates of Grapevine fanleaf virus from a naturally infected vineyard in France: evidence for mixed infection and recombination*. J. Gen. Virol., 2004. 85(8): p. 2435-2445.
 49. Fry, E.E., et al., *The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex*. EMBO J., 1999. 18(3): p. 543-554.
 50. Zhang, P., et al., *Crystal structure of CD155 and electron microscopic studies of its complexes with polioviruses*. Proc. Natl. Acad. Sci. U S A, 2008. 105(47): p. 18284-18289.
 51. Huss, B., B. Walter, and M. Fuchs, *Cross-protection between arabis mosaic virus and grapevine fanleaf virus isolates on Chenopodium quinoa*. Ann. appl. Virol., 1989. 114: p. 45-60.
 52. Loudes, A.M., et al., *The 119 kDa and 124 kDa polyproteins of arabis mosaic nepovirus (isolate S) are encoded by two distinct RNA2 species*. J. Gen. Virol., 1995. 76 (4): p. 899-906.
 53. Ritzenthaler, C., et al., *Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1*. J. Gen. Virol., 1991. 72 (10): p. 2357-2365.
 54. Serghini, M.A., et al., *RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location*. J. Gen. Virol., 1990. 71: p. 1433-1441.

55. Belin, C., et al., *The nine C-terminal residues of the grapevine fanleaf nepovirus movement protein are critical for systemic virus spread*. J. Gen. Virol., 1999. 80 p. 1347-1356.
56. Walter, B. and L. Etienne, *Detection of the Grapevine Fanleaf Viruses Away from the Period of Vegetation*. J. Phytopathol., 1987. 120(4): p. 355-364.
57. Mindell, J.A. and N. Grigorieff, *Accurate determination of local defocus and specimen tilt in electron microscopy*. J. Struct. Biol., 2003. 142(3): p. 334-347.
58. Conway, J.F. and A.C. Steven, *Methods for reconstructing density maps of 'single' particles from cryoelectron micrographs to subnanometer resolution*. Journal of Structural Biology, 1999. 128(1): p. 106-118.
59. Baker, T.S. and R.H. Cheng, *A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy*. J. Struct. Biol., 1996. 116(1): p. 120-130.
60. van Heel, M. and G. Harauz, *Resolution criteria for three dimensional reconstruction* Optik (Jena), 1986. 73(3): p. 119-122.
61. Kabsch, W., *XDS in International Tables for Crystallography*, in *Crystallography of Biological Macromolecules*, M.G. Rossmann and E. Arnold, Editors. 2001, Dordrecht: Kluwer Academic Publishers.
62. Navaza, J., *AMoRe: an automated package for molecular replacement*. Acta Cryst. A, 1994. 50(2): p. 157-163.
63. Trapani, S. and J. Navaza, *AMoRe: classical and modern*. Acta cryst. D, 2008. 64(1): p. 11-16.
64. Jones, T.A., *Molecular replacement* S.G.a.W.W. E.J.Dolson, Editor. 1992, SERC Daresbury Laboratory: Warrington. p. 59-66.
65. Kleywegt, G.J. and T.A. Jones, *In From First Map to Final Model*, R.H.a.D.W. S. Bailey, Editor. 1994, SERC Daresbury Laboratory: Warrington. p. 59-66.
66. Kleywegt, G.J. and T.A. Jones, *xdlMAPMAN and xdlDATAMAN ; Programs for Reformatting, Analysis and Manipulation of Biomacromolecular Electron-Density Maps and Reflection Data Sets*. Acta Cryst. D, 1996. 52(4): p. 826-828.
67. Kleywegt, G.J. and T.A. Jones, *Software for handling macromolecular envelopes*. Acta Cryst. D, 1999. 55(4): p. 941-944.
68. Fiser, A. and A. Sali, *Modeller: generation and refinement of homology-based protein structure models* Meth. Enzymol., 2003. 374: p. 461-491
69. Adams, P.D., et al., *PHENIX: building new software for automated crystallographic structure determination*. Acta Crystallogr D Biol Crystallogr, 2002. 58(11): p. 1948-54.
70. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr D Biol Crystallogr, 2004. 60: p. 2126-32.
71. Laskowski, R.A., et al., *AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR*. J Biomol NMR, 1996. 8(4): p. 477-86.
72. Kleywegt, G.J., *Use of non-crystallographic symmetry in protein structure refinement*. Acta Crystallogr D Biol Crystallogr, 1996. 52(4): p. 842-57.
73. Sanner, M.F., A.J. Olson, and J.C. Spohner, *Reduced surface: an efficient way to compute molecular surfaces*. Biopolymers, 1996. 38(3): p. 305-320.

74. Baker, N.A., et al., *Electrostatics of nanosystems: application to microtubules and the ribosome* Proc. Natl. Acad. Sci. U S A, 2001. 98 p. 10037-10041
75. Pettersen, E.F., et al., *UCSF Chimera - A visualization system for exploratory research and analysis*. Journal of Computational Chemistry, 2004. 25(13): p. 1605-1612.

TABLES

Table 1. Crystallographic analysis of GFLV particles

Virus	GFLV-TD	GFLV-F13
Data collection statistics		
Beamline	ESRF/BM30	SLS/X06DA
Space group (number)	P2 ₁ 3 (198)	P1 (1)
Unit cell lengths a, b, c (Å)	408.0	279.4- 279.5- 293.3
Unit cell angles alpha, beta, gamma (°)	90.0	102.4- 116.4- 108.2
Resolution range (Å)	36 - 2.7	135 - 3.0
Highest resolution shell (Å)	2.77 – 2.7	3.08 - 3.0
No. of unique reflections	563009 (32448)	1214336 (73170)
Completeness (%)	92.0 (72.0)	88.1 (71.7)
Multiplicity	11.0 (3.1)	2.0 (1.9)
R _{merge} (%) [†]	12.5 (68.1)	10.0 (35.0)
$\langle I/\sigma(I) \rangle$	18 (1.9)	9.2 (2.4)
Molecular replacement		
Resolution range (Å)	30 – 15	15 - 6.0
Asymmetric unit content	20-mer	60-mer
Model	EM map	GFLV-TD
Correlation / R-factor (%)	60.4 / 56.4	70.6 / 34.7
Refined atomic structure		
Resolution range (Å)	36 - 2.7	135 - 3.0
R-factor / R-free (%)	19.6 / 21.4	19.1 / 21.0
Number of capsid and solvent atoms	79097 / 556	237051 / –
Rmsd and bonds (Å) and angles (°)	0.009 / 1.30	0.009 / 1.26

$$\dagger R_{\text{merge}} = \frac{\sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle|}{\sum_{hkl} \sum_i I_i(hkl)}.$$

Values in parentheses are for the highest resolution shell

SUPPLEMENTARY MATERIAL

Table S1. Root mean square deviation in C α displacement in the superimposed capsid protein (CP) and the CP domains A, B, and C.

Structural comparison	Superimposed pair		r.m.s.d (Å)	No. of matched residues
GFLV-F13 versus GFLV-TD	CP	CP	0.12	504
GFLV-F13 versus TRSV	CP	CP	1.7	444
	A	A	1.42	143
	B	B	1.30	130
	C	C	1.19	145
GFLV versus Homology model	CP	CP	1.62	429

Table S2. Primers used to produce genes 2C^{CP} genes with mutated residues and characterize GFLV RNA2 progeny.

Name	Direction-position ^a	Sequence ^b
mutDF	F-2918	CCTGCCAGATTGCTTGCTG A TCAAAGTCAGAGAGACATGAGC
mutDR	R-2959	GCTCATGTCTCTCTGACTT T GATCAGCAAGCAATCTGGCAGG
mutAF	F-2918	CCTGCCAGATTGCTTGCTG C TCAAAGTCAGAGAGACATGAGC
mutAR	R-2959	GCTCATGTCTCTCTGACTT T GAGCAGCAAGCAATCTGGCAGG
36	F-2522	GCCACTTCTTTCCCTGGGAAACCAA
18	R-3103	ATCCACCCATACGAAATAGTC
227	F-2862	ATGTGGAAGAGGACGGAAGT
397	R-3722	GGCAAGTGTGTCCAAAGGAC
115	F-1977	CTGTGAGGATTGATAGAAACG

^aPrimer positions are given as the position of the 5' first nucleotide according to the GFLV-F13 RNA2 sequence (NC_003623). F = forward and R = reverse orientation.

^bMutagenic nucleotides are highlighted in grey.

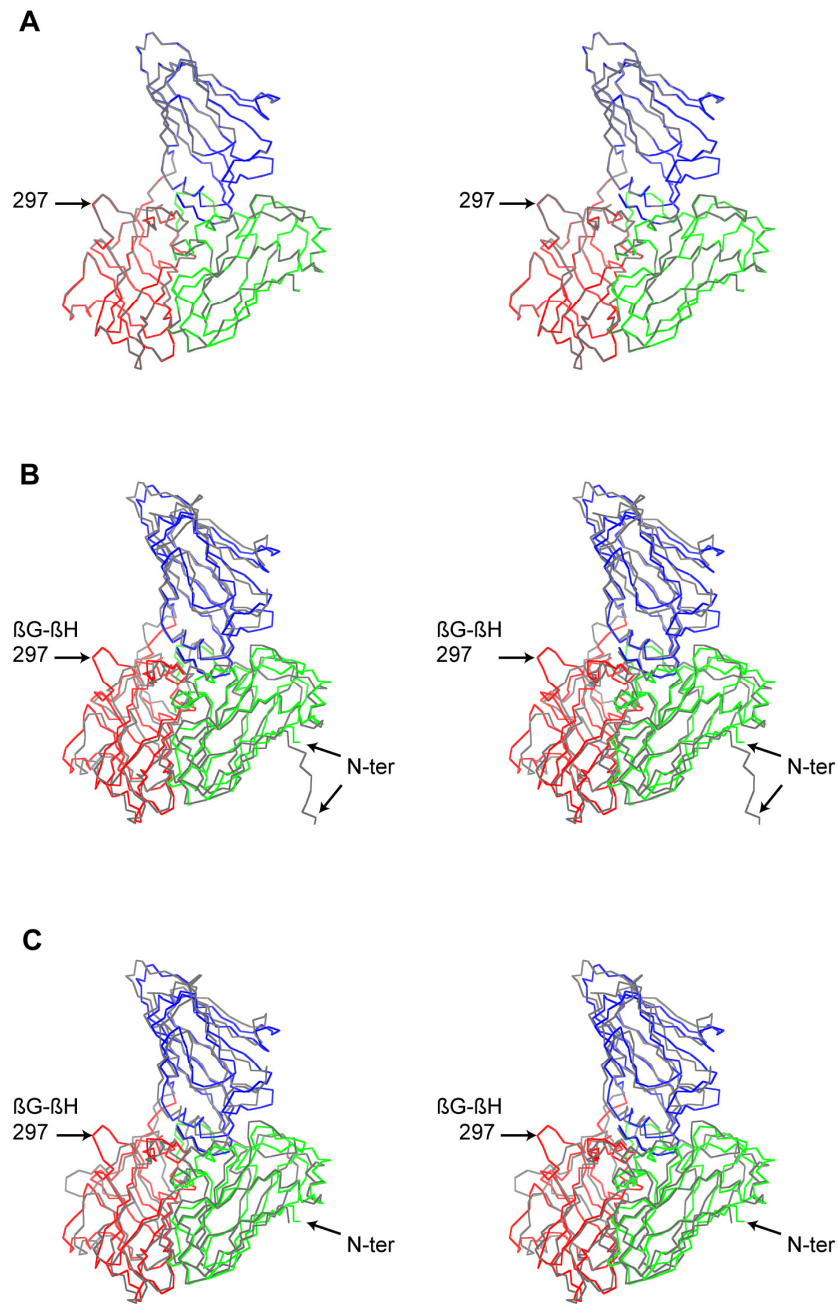


Figure S1. Structural similarity of GFLV and TRSV. This stereoview shows a superposition of the capsid protein ($C\alpha$ trace representation) of GFLV-F13 on capsid proteins of GFLV-TD (A), TRSV (B) and GFLV homology model derived from TRSV crystal structure (C), respectively. The C, B and A domains in GFLV are shown in green red and blue, respectively, and the subunit is viewed from the outside of the capsid. Other structures are depicted in grey. The position of the residue 297 of GFLV is indicated by an arrow. The structures were superimposed using LSQMAN (Kleywegt *et al.*, 1996) and corresponding RMSDs are listed in Table 2.

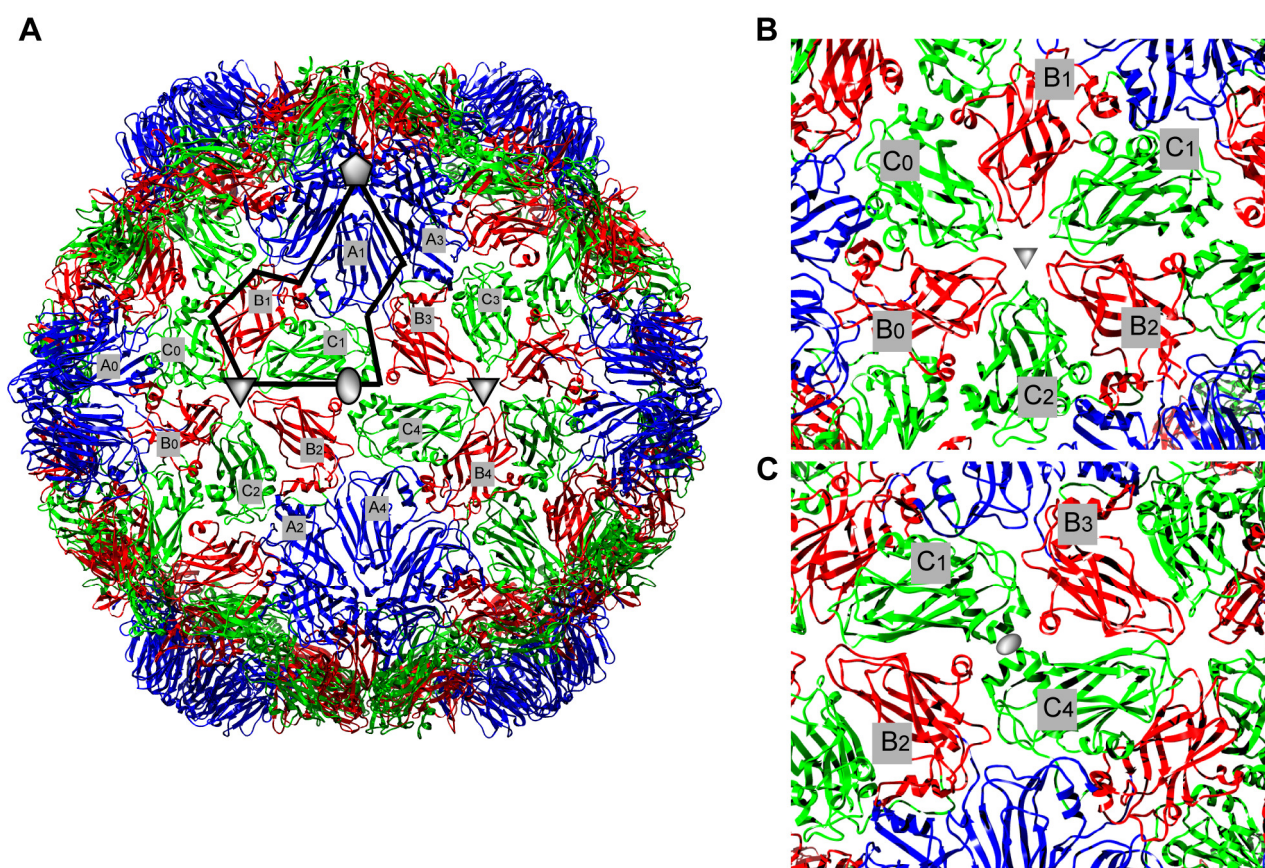


Figure S2. Capsid protein contacts on 3-fold and 2-fold axes. (a) The black line delineates one CP position. The figure indicates the contact between different domains on the 3-fold and the 2-fold axis. The grey pentagon, triangle and oval symbolize the icosahedral 5-fold, 3-fold and 2-fold icosahedral symmetry axes, respectively. Domains of the same CP are labelled with the same number. A domains are exclusively clustered around the 5-fold axis. (B) Six β -barrels domains belonging to 3 different CP interact around the 3-fold axis. (C) Two B and C domains from four CPs interact on the 2-fold axis.

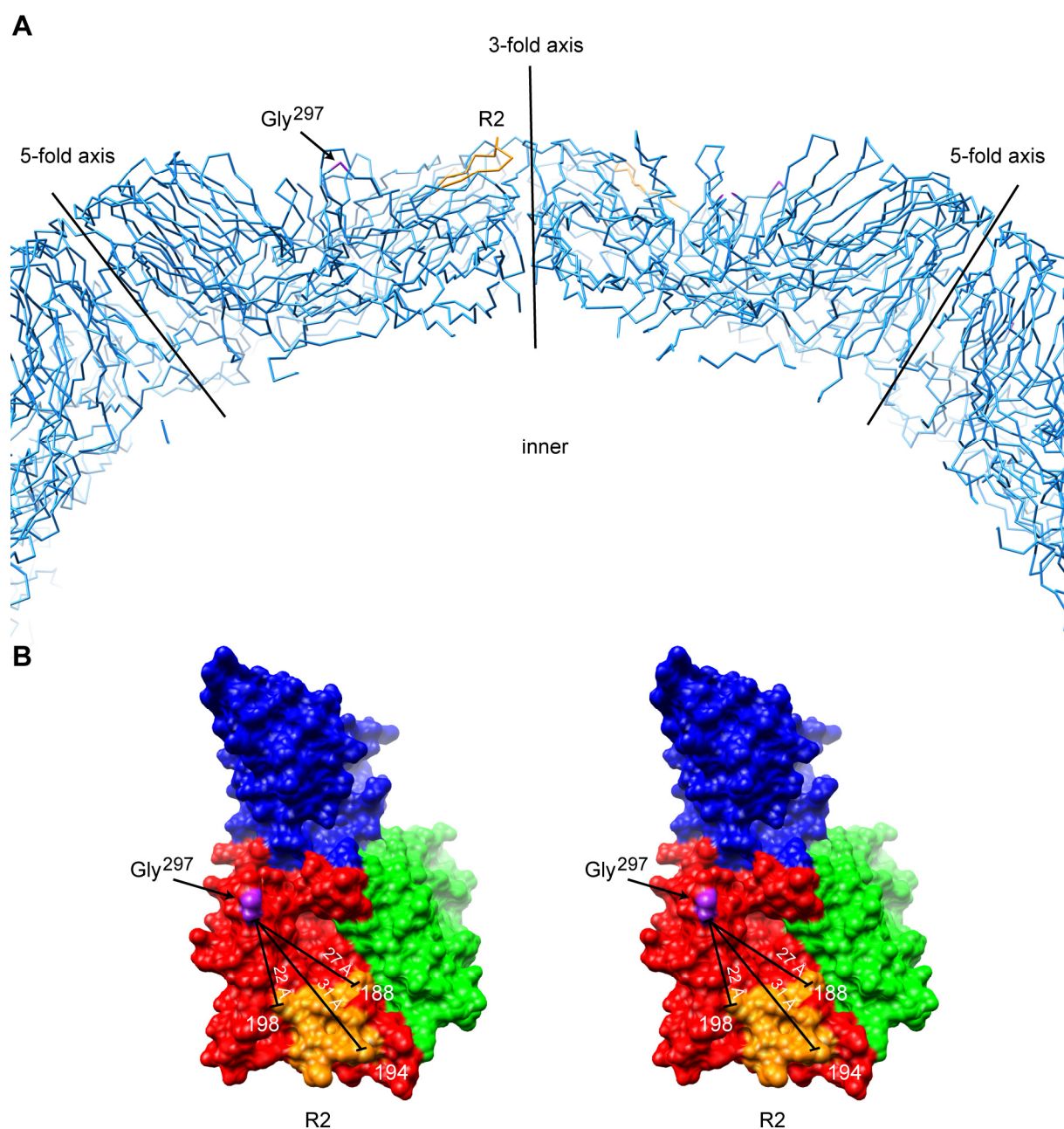


Figure S3. Position of nematode transmission determinants on GFLV capsid surface. (A) This close-up view of the capsid reveals that residue Gly²⁹⁷ (purple) and the region R2 comprising residues 188 to 198 (orange) are surrounding a deep pocket and belong to the same CP domain, as illustrated in (B). In this stereoview, a single CP is seen from the outside of the capsid with C, B, and A domain colored as in figure 2.

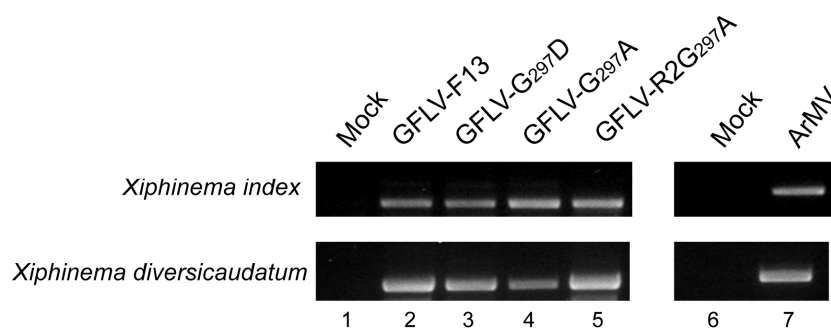


Figure. S4. Virus detection in *Xiphinema spp.* at the end of the acquisition access period (AAP). Randomly selected nematodes analyzed were exposed to source plants either infected with GFLV-F13 (2), GFLV-G₂₉₇D (3), GFLV-G₂₉₇A (4), GFLV-R2G₂₉₇A (5), ArMV (6) or mock inoculated plants (1 and 6). The amplification of specific DNA products confirms that nematodes have ingested all types of viruses during AAP.

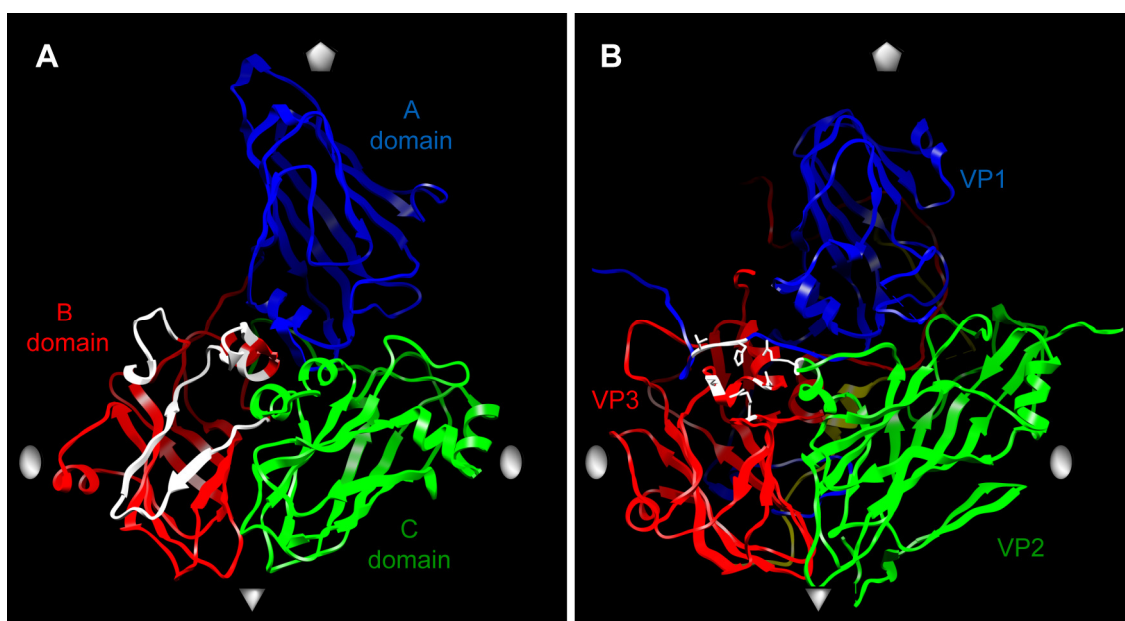


Figure S5. Comparison between the GFLV putative nematode-binding pocket and the FMDV heparin sulphate binding site. (A) The GFLV CP is seen from the outside of the capsid with C, B, and A domain colored as in figure 2, and the putative nematode binding pocket in white. (B) The FMDV viral protein (pdb ID, 1QQP) VP1, VP2, VP3 and VP4, are colored in blue, red, green and yellow, respectively. The residues involved in heparin sulphate binding (Fry *et al.*, 1999) appears in white. The grey pentagon, triangle and oval symbolize the icosahedral 5-fold, 3-fold and 2-fold icosahedral symmetry axes, respectively.

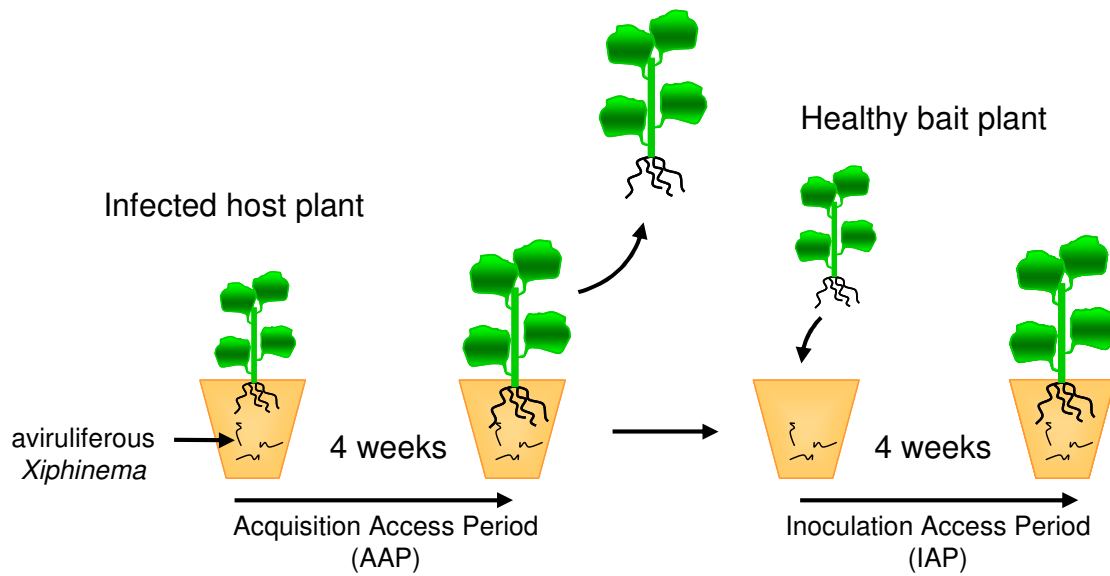


Figure S6. Nematode transmission assays. Prior to the transmission assays, the infectious status of all source plants - *C. quinoa* or *N. benthamiana* - was verified by DAS-ELISA using specific GFLV and ArMV antibodies. 200 aviruliferous nematodes were allowed to feed on the roots of a virus source plant for a four-week acquisition access period (AAP). Then, nematodes were exposed to the roots of healthy bait plants for a four-week inoculation access period (IAP). The successful transmission of viruses by nematodes was verified in the roots of each bait plant by DAS-ELISA using specific GFLV and ArMV antibodies.

5. Compléments d'article

5.1 Structure quaternaire de la capsid

Les surfaces accessibles au solvant (SAS) des résidus de la capsid ont été calculées à partir de la structure cristalline possédant la meilleure résolution, c'est-à-dire celle du GFLV-TD avec le programme MSMS (voir Matériel et Méthode). Les différentes chaînes polypeptidiques sont nommées d'après le schéma qui figure ci-dessous. Sur les 504 résidus de la CP, 208 sont à la surface externe de la capsid et 173 à la surface interne de la capsid (annexes 4). Une large majorité de résidus sont donc accessibles. Dans la capsid, les contacts quaternaires entre les CP sont de différents types en fonction de l'environnement dans lequel se trouve la CP. Trois types d'interactions sont répétées : les contacts pour former les pentamères (D-C et D-E, fig. 5.1), les contacts entre 3 CP au niveau des axes d'ordre 3 (D-S-N), ou encore les contacts entre 2 CP au niveau des axes d'ordre 2 (D-M). Notons qu'il n'y a pas d'interaction entre D-A et D-B. Comme on pouvait le prévoir, la majorité des contacts se fait effectivement entre les sous-unités d'un même pentamère, moins d'interactions se produisent entre les sous-unités regroupées au niveau des axes d'ordre 3, et il y a très peu de contact entre les sous-unités au niveau des axes d'ordre 2 (fig. 5.1). Un nombre quasi-équivalent de résidus sont impliqués dans les contacts de même type, c'est-à-dire entre les sous-unités C/D et D/E ; D/S et D/N.

Contacts	No. résidus	Surface totale (Å ²)
D/C	74	2528,4
D/E	67	2437,7
D/S	31	1078,49
D/N	27	1104,9
D/M	7	228,7

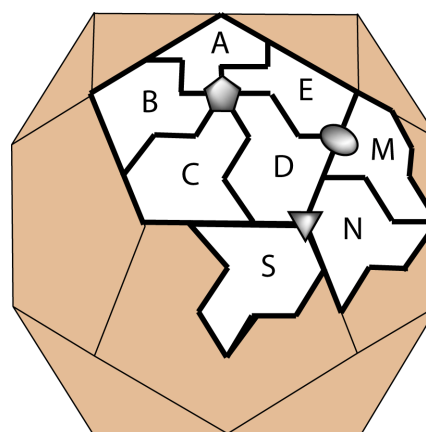


Fig. 5.1. Contacts entre sous-unités. Le schéma indique la nomenclature utilisée dans le fichier pdb de la structure cristalline du GFLV-TD, afin de désigner les différentes chaînes polypeptidiques du premier pentamère et les protéines voisines à D. Les axes d'ordre 5, 3, et 2 sont indiqués respectivement par un pentagone, un triangle et un ovale.

5. 2 Les régions 1 à 5 délimitées d'après la structure du TRSV

Plusieurs régions de la capsidite du GFLV ont été mutées. La structure cristallographique du GFLV confirme leur localisation externe à la capsidite, prédite par le modèle construit par homologie avec la structure cristallographique du TRSV. Cependant, de légers décalages sont tout de même présents entre la localisation prédite des régions et leur position dans la structure cristallographique du GFLV.

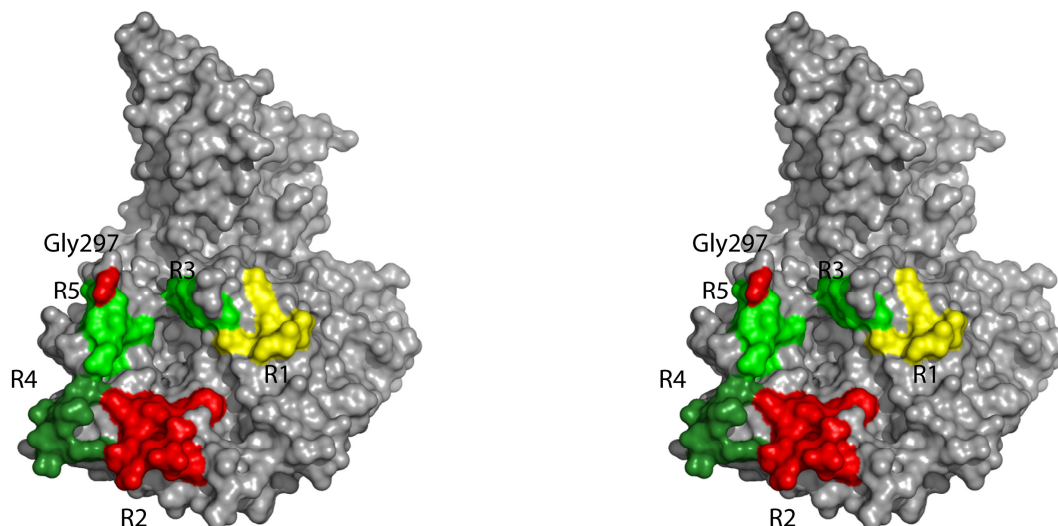
Les cinq résidus de la région R1 et les quatre de la région R3 qui avaient été échangés par leurs équivalents ArMV, sont en surface de la capsidite (fig. 5.2). Contrairement à R1, l'échange de R3 conduit à une chimère défectueuse en encapsidation (G3). Pour restaurer l'encapsidation de G3, et pouvoir étudier l'implication de cette région dans la transmission, il serait intéressant de muter tous les résidus de la région sauf la proline qui contribue à la conformation du brin $\beta C'$.

Tous résidus de la région R2 sont externes à l'exception de la Phe¹⁸⁹ (fig. 5.2B). R2 est situé au niveau de l'axe d'ordre 3 de la capsidite, ce qui explique pourquoi 4 de ses résidus sont impliqués dans des contacts avec une sous-unité voisine (dans l'exemple de la fig. 1, le contact D/S). Néanmoins le changement de ces résidus par ceux de l'ArMV ne modifie pas l'infectivité du virus.

Par rapport aux autres régions mutées, R4 est la région qui participe le plus aux interactions quaternaires des CP. En effet, trois résidus (Thr²⁵⁸, Thr²⁵⁹ et Glu²⁶⁴) et deux résidus (Trp²⁶² et Asn²⁶³) sont respectivement en contact avec la sous-unité voisine du pentamère et une sous-unité présente sur l'axe d'ordre 2 (dans l'exemple de la fig. 1, les contacts D/C et D/S). Même si ces résidus sont externes, la majorité d'entre eux sont donc impliqués dans des contacts entre sous-unités. Précisons que la première Thr²⁵⁸ est faiblement en contact (seulement 1 Å² est protégé par la sous unité C sur les 30 Å² de SAS total du résidu). Cette position à l'interface de plusieurs domaines C et B de la CP pourrait lui conférer un rôle particulier dans l'encapsidation. Pourtant, plusieurs tests d'encapsidation effectués en protoplastes montrent que le mutant G4 est capable de protéger l'ARN. Mais la position de la région dans la structure suggère que le mutant pourrait avoir un problème de conformation ou une dynamique particulière. Il serait donc nécessaire d'approfondir les tests d'encapsidation, par exemple en vérifiant la structure de ce mutant en ME.

La région R5 possède une structure secondaire différente par rapport à celle prédite. Elle comprend deux éléments de structure secondaire et semble donc plus délicate à muter que ne le prédisait le modèle qui n'indiquait aucune structure secondaire. De plus, les deux derniers résidus ne sont pas exposés à la surface de la capsidite. Il est remarquable de constater que pour chaque région, les résidus impliqués dans une structure quaternaire sont globalement mieux conservés entre GFLV, ArMV et TRSV que ceux uniquement en contact avec le solvant.

A



B

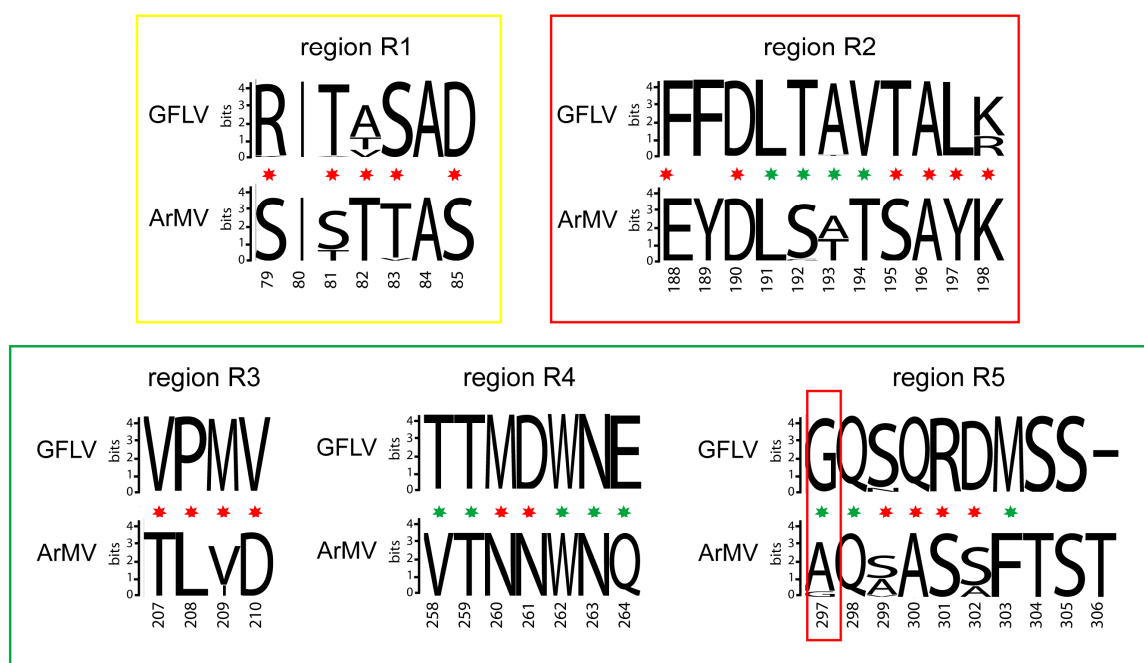


Fig. 5.2. Localisation des régions mutées sur la structure cristallographique de la CP du GFLV-F13. (A) Paire d'image stéréo d'une représentation de surface de la CP, vue depuis son côté externe. La Gly297 et la région R2 (rouge) sont impliquées dans la transmission par *Xiphinema index*. La région R1 (jaune) ne semble pas nécessaire à la transmission. L'implication des régions R3 (vert), R4 (vert foncé) et R5 (vert) dans la transmission n'a pas été étudiée, car la substitution de ces régions par leur équivalent ArMV aboutit à des chimères non infectieux *in planta*. (B) Une représentation LOGO de chaque région montre la conservation des séquences parmi les différents variants du GFLV et de l'ArMV (références des séquences dans le tableau S1 de l'article 2). Les résidus des régions sont soit à l'intérieur de la capsid (pas d'étoile), soit en contact uniquement avec le solvant (étoile rouge), soit en contact à la fois avec le solvant et avec une sous-unité voisine dans la capsid (étoile verte).

De manière générale, chaque région contient des éléments de structures secondaires, à l'exception de R3 (fig. 5.3). Il n'est donc pas possible de faire le lien entre la perte d'encapsidation d'un mutant et la présence de structure secondaire dans la région mutée. L'analyse structurale des

déterminants de la transmission révèle que plusieurs résidus sont impliqués dans la structure quaternaire de la capsid.

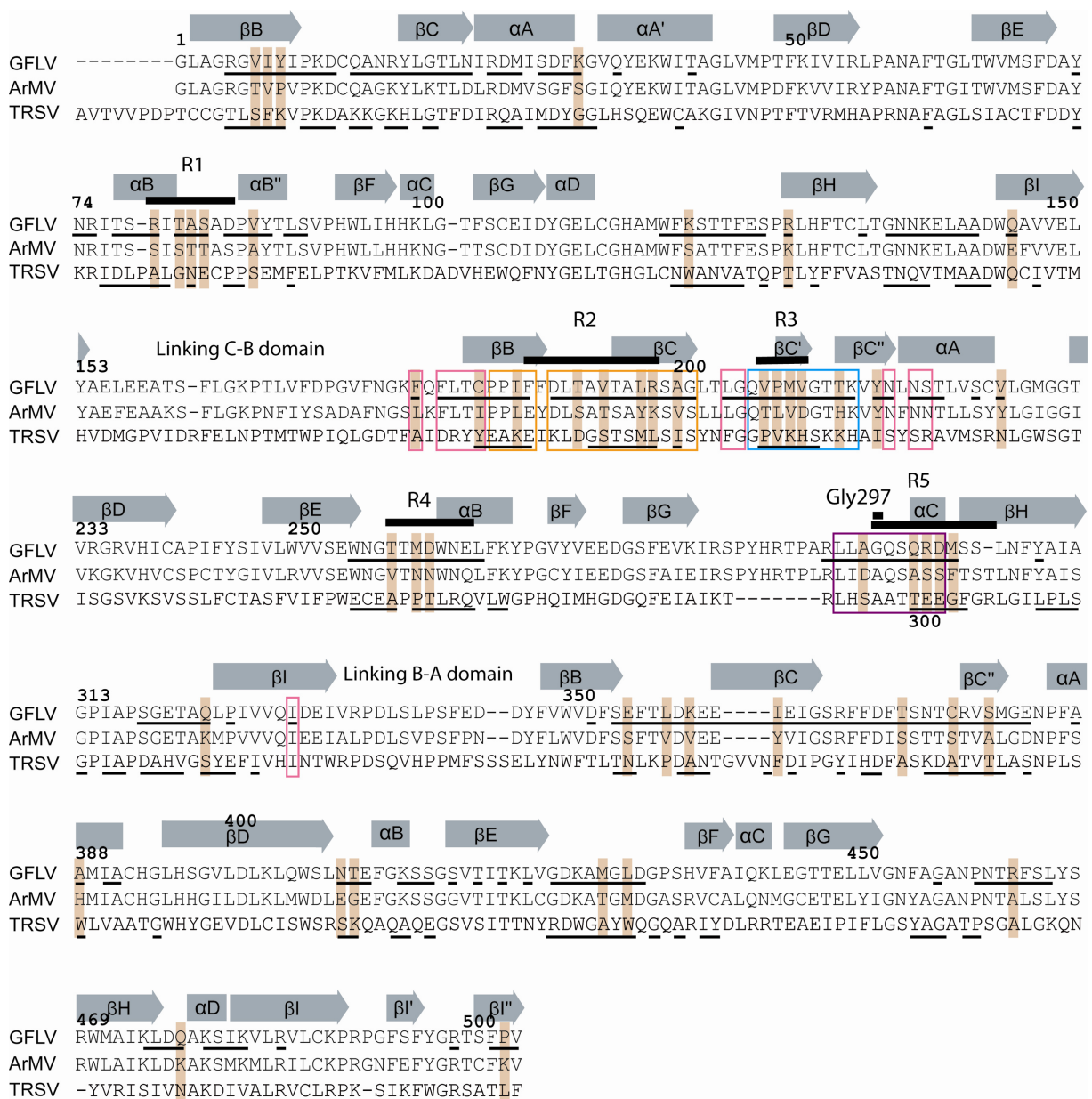


Fig. 5.3. Alignement des séquences protéiques de CP du GFLV, de l'ArMV et du TRSV. Les flèches et les rectangles représentent respectivement les brins β et les hélices α de la structure cristallographique du GFLV. Les résidus de surface du GFLV et du TRSV sont soulignés. Les résidus de surface différents entre les trois espèces virales sont surlignés en brun. Les résidus de surface situés soit au centre, soit sur les bords gauches, droit, ou encore sur le plateau de la cavité sont encadrés respectivement en magenta, violet, cyan et orange.

5. 3 Résidus de surface du GFLV divergeant entre le GFLV, l'ArMV et le TRSV

La figure 5 montre que, malgré une séquence primaire assez divergente entre le GFLV et le TRSV (25 % d'identité), la position des résidus externes à la capsid est relativement bien conservée entre la structure cristallographique des deux virus. Les résidus, à la surface de la capsid du GFLV, différents entre le GFLV, l'ArMV et le TRSV, se regroupent dans plusieurs motifs protéiques (fig. 5.3). Comme nous l'avons déjà noté lors de la comparaison de séquence entre la CP du GFLV et de l'ArMV, le domaine B contient la majorité des régions de surface divergentes (article 2) et ce regroupement est encore plus flagrant lorsque le TRSV est ajouté à l'analyse (fig. 5.3). Plusieurs motifs sont également présents dans le domaine A. De plus, les régions R1, R2, R3, R4, R5 apparaissent clairement comme étant les motifs protéiques contenant le plus de résidus de surface divergents (fig. 5.3).

5. 4 La cavité : site potentiel de liaison à *X. index*

La structure cristallographique du GFLV nous a permis de mettre en évidence une cavité qui sépare les deux régions - Gly297 et R2 - déterminant la transmission du GFLV par *X. index*. Cette cavité possède une forme différente chez le TRSV (fig. 5.4). De plus l'analyse des potentiels de surface montre qu'il s'agit d'une des seules surfaces chargées positivement à l'extérieur de la capsid (article 4). Ce site particulier pourrait donc constituer le domaine de liaison à *X. index*.

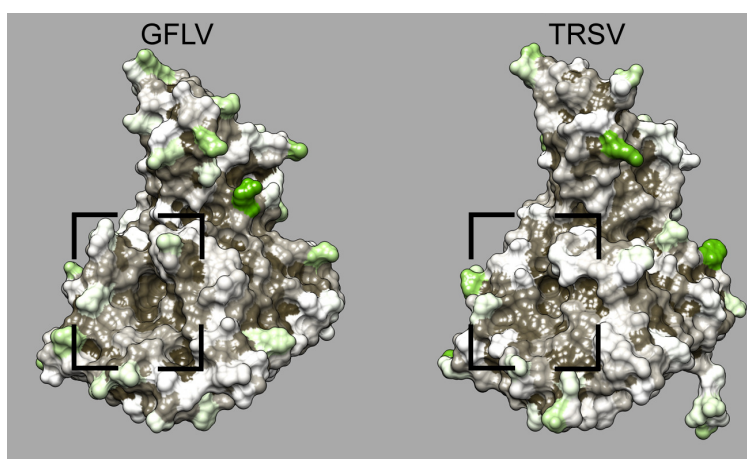


Fig. 5. 4. Différences topologiques de surface entre l'extérieur de la structure cristallographique de la CP du GFLV et celle du TRSV. Représentation de surface avec un code couleur appliquée par rapport à la surface accessible au solvant, allant du brun pour les résidus les moins exposés, au vert pour les plus exposés.

Au niveau de la structure quaternaire, il se situe entre l'axe d'ordre 3 et l'axe d'ordre 5, à l'endroit où se trouve une dépression de la particule virale. Au niveau de la structure tertiaire, il est

formé par le repliement en *jelly-roll* du domaine B. La cavité est asymétrique. Si la CP est orientée en direction de l'axe 5 de la capsid, la Gly²⁹⁷ est localisée sur le bord gauche de la cavité, et la région R2 se localise à l'extrémité d'un plateau (fig. 5.4). Les résidus du domaine B participant à la structure tertiaire de la cavité sont encadrés dans la figure 5.3. Le bord gauche de cette dernière est formé par une partie de la boucle β G- β H (aa 294-302). Le bord droit est formé par les résidus appartenant aux brins β C' et β C'' et à la boucle β C'- β C'' (aa 206-214). Le plateau est formé par les deux brins β B et β C et la boucle β B- β C (aa 185-201). Le centre de la cavité, quant à lui, est constitué de petites alvéoles presque internes à la capsid, mais toujours en contact avec le solvant. La cavité s'enfonce profondément dans la CP, puisque en son centre plusieurs résidus sont à la fois en contact avec l'extérieur et l'intérieur de la capsid (Leu¹⁸², Thr¹⁸³, Leu²¹⁸, Asn²¹⁹, Ile³³⁰).

La figure 5.3 indique les résidus à la surface de la capsid, situés dans la cavité. Parmi les 11 résidus du centre, les 18 résidus des deux protubérances et les 16 résidus du plateau, on retrouve respectivement 2 résidus, 9 résidus et 6 résidus différents entre le GFLV, l'ArMV et le TRSV. La partie centrale de la cavité est donc la plus conservée entre les trois espèces virales (fig. 5.3).

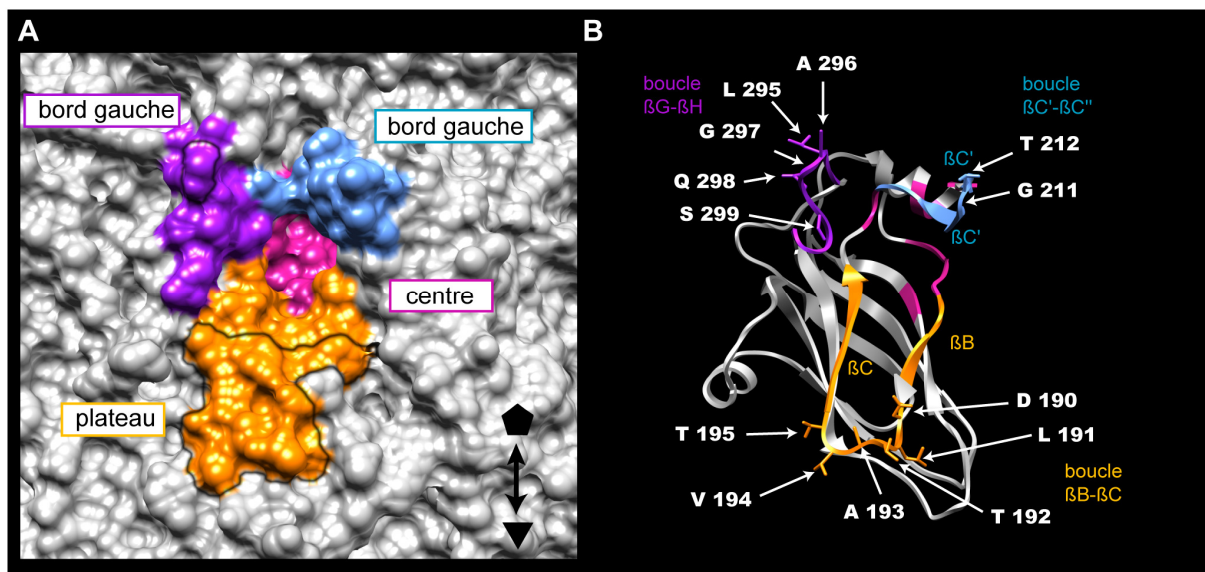


Fig. 5.5. Représentation de surface (A) et représentation *cartoon* (B) de la cavité. Les résidus de surface du bord droit (cyan) et gauche (violet), du plateau (orange) et du centre (magenta) de la cavité sont indiqués dans la figure 5.3. (A) La position de la Gly297 dans le bord droit et de R2 dans le plateau est indiquée par un trait noir. Le croquis en bas à droite de l'image rappelle l'orientation choisie pour la figure. L'axe d'ordre 3 (triangle) se situe en bas de l'image, et l'axe d'ordre 5 (pentagone) en haut de l'image. (B) Seules les chaînes latérales des résidus à l'extrémité de chaque boucle sont indiqués.

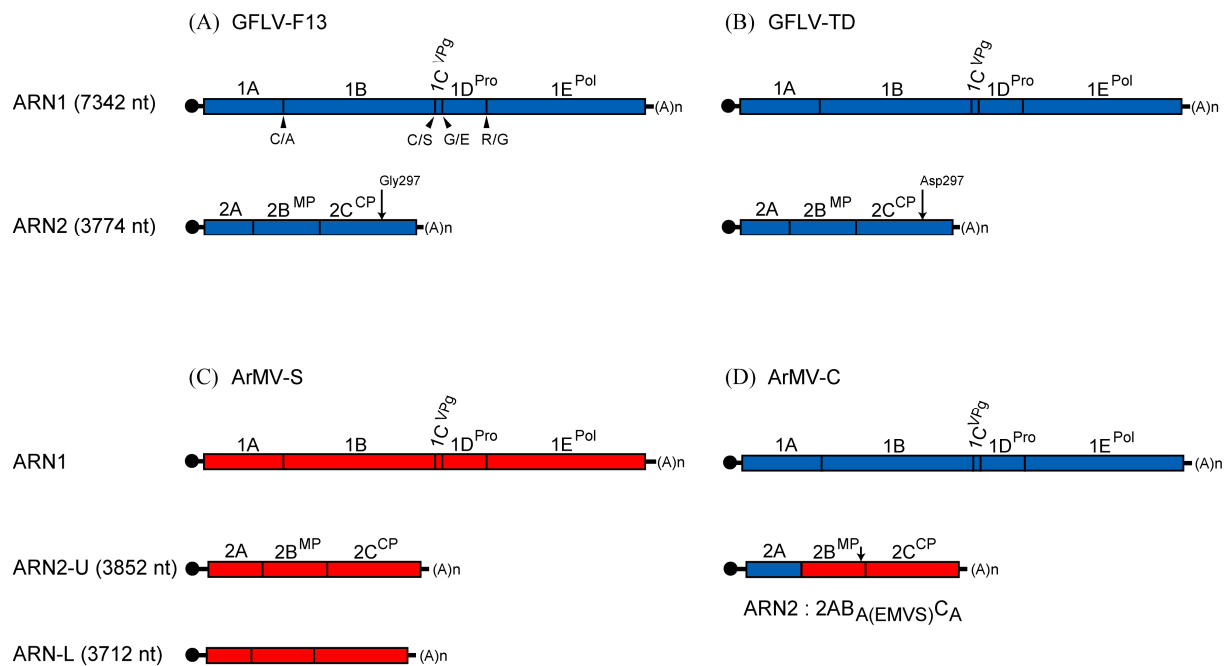


Fig. 6.1. Représentation schématique de l'organisation génétique des virus utilisés dans l'étude structurale. Les rectangles bleus indiquent les séquences codantes d'origine GFLV, ceux en rouges sont d'origine ArMV. (A) GFLV-F13, isolat sauvage de référence (Vuittenez *et al.*, 1964). (B) GFLV-TD, variant naturel du GFLV-F13. Il diffère du GFLV-F13 par la présence d'une mutation ponctuelle affectant le résidu 297 de la CP (Schellenberger *et al.*, article 4). L'acide aminé modifié est indiqué par une flèche noire. (C) ArMV-S, isolat dont le génome est composé d'un ARN1 et de deux ARN2 (Loudes *et al.*, 1995). (D) ArMV-C, chimère GFLV/ArMV dont l'ARN2 a été produit en échangeant les séquences codantes du GFLV-F13 et de l'ARN2-U de l'ArMV-S. La flèche noire indique la présence d'aa d'origine GFLV dans la partie N-terminale de la MP assurant la systémie du virus (Andret-Link *et al.*, 2004b). Chaque ARN est traduit en une polyprotéine, clivée par la protéase virale. Le nom de chaque protéine virale est indiquée au-dessus des séquences codantes correspondantes. La protéine de capsidie conventionnellement nommée 2C^{CP} est également indiquée par l'abréviation CP.

Chapitre 6

Recherche des déterminants viraux de la transmission par comparaison structurale entre le GFLV et l'ArMV

1. Démarche expérimentale

Deux déterminants de la transmission du GFLV ont été identifiés : l'acide aminé Gly²⁹⁷ et la région R2. Les deux déterminants sont en surface de la particule virale, plus précisément sur les bords d'une cavité présente entre l'axe d'ordre 5 et l'axe d'ordre 3. L'évaluation de la transmissibilité des virus chimères GFLV/ArMV ciblant ces deux domaines a abouti à la perte de transmission par *X. index*, mais n'a pas permis de gagner celle par *X. diversicaudatum*. Plusieurs explications à cela : le contexte structural dans lequel se localisent les déterminants de la transmission diffèrent entre les capsides de GFLV et de l'ArMV ; d'autres déterminants viraux interagissant avec les récepteurs des nématodes.

Afin de vérifier ces hypothèses, nous avons entrepris d'identifier les différences structurales existant entre les deux espèces virales. Pour ceci, la structure 3D de l'ArMV devait être déterminée. J'ai donc cherché à obtenir la structure du virus en utilisant les mêmes approches que celles utilisées pour le GFLV dans le chapitre précédent, à savoir la cristallographie aux rayons X et la cryomicroscopie électronique à transmission (cryo-MET).

Au laboratoire, l'isolat ArMV-S avec lequel nous travaillons possède un génome composé de deux ARN2 différents, l'ARN2-U et l'ARN2-L (Loudes *et al.*, 1995)(fig. 6.1). Chaque ARN2 code pour une CP de 505 acides aminés, parmi lesquels neuf résidus sont différents (l'alignement des séquences protéiques est fourni en annexe 5). Pour s'affranchir de toute confusion dans l'interprétation des cartes de densité, due à cette différence de séquences protéiques, nous avons choisi de travailler avec une chimère dont le génome code pour une seule des deux CP. Cette chimère a déjà été utilisée dans une étude présentée dans le chapitre 3 de la thèse (Marmonier *et al.*, 2010). Ce virus est produit à partir de transcrits correspondant à l'ARN1 du GFLV et un ARN2 chimérique, nommé 2AB_{A(EMVS)}C_A, qui possède une MP et une CP de nature ArMV (provenant de la séquence de l'ARN2-U) à l'exception de 3 résidus de la partie C-terminale de la MP (fig. 6.1). Il est transmis spécifiquement par

X. diversicaudatum, avec la même efficacité que celle de l'ArMV sauvage (Marmonier *et al.*, 2010). Dans ce chapitre, nous ferons référence à lui en tant qu'ArMV-C.

Contrairement au GFLV, la cristallisation de l'ArMV-C a échoué dans les différentes conditions testées. Ni l'affinement des conditions de cristallisation du GFLV, ou l'utilisation de plusieurs kits (*screens* commerciaux) de cristallisation, n'ont permis d'obtenir un résultat intermédiaire entre solubilité et précipités irréversibles. Cependant, sachant que les reconstructions 3D à partir d'images de cryo-ME peuvent atteindre des résolutions de plus en plus basses, même inférieures à 5 Å (Chen *et al.*, 2009; Jiang *et al.*, 2008; Yu *et al.*, 2008), cette méthode a été choisie pour poursuivre le travail. La reconstruction 3D de l'ArMV-C a été faite en collaboration avec Patrick Bron (CBS, Montpellier). Mon travail a consisté à produire le virus, à participer aux acquisitions d'images de cryo-MET, et aux premières étapes du traitement d'image. Les détails méthodologiques, ainsi que le principe de reconstruction 3D utilisé, figurent dans la partie Matériel et Méthodes de la thèse.

2. Qualité de la purification de l'ArMV

Les particules virales d'ArMV-C ont été purifiées avec le protocole établi pour le GFLV (Matériel et Méthodes de l'article 3). Elles sont intègres et stables dans le temps, mais le rendement obtenu est faible. Pour le GFLV, le rendement est en moyenne de 2 mg de virus par kg de feuilles infectées, alors que pour l'ArMV-C il est seulement de 0,5 à 1 mg, à une concentration finale de 3mg/ml. Ceci provient probablement de pertes importantes de virus survenant lors des centrifugations à hautes et basses vitesses de la purification. En effet, les analyses de diffusion dynamique de lumière (DLS) et de MET en coloration négative, ont montré qu'une fraction des capsides est présente sous forme d'agrégats (fig. 6.2). Les virions s'oligomérisent, que se soit de manière ordonnée (fig. 6.2) ou non. En DLS, la présence d'agrégats est révélée par la présence de deux pics, dont seul le premier correspond à la détection de particules non-agrégées (fig. 6.2). L'augmentation de la concentration en sel dans la solution ne permet pas d'éviter l'agrégation des particules virales.

Ce phénomène d'agrégation, pourrait expliquer les pertes de virus pendant la purification et la difficulté à le cristalliser.

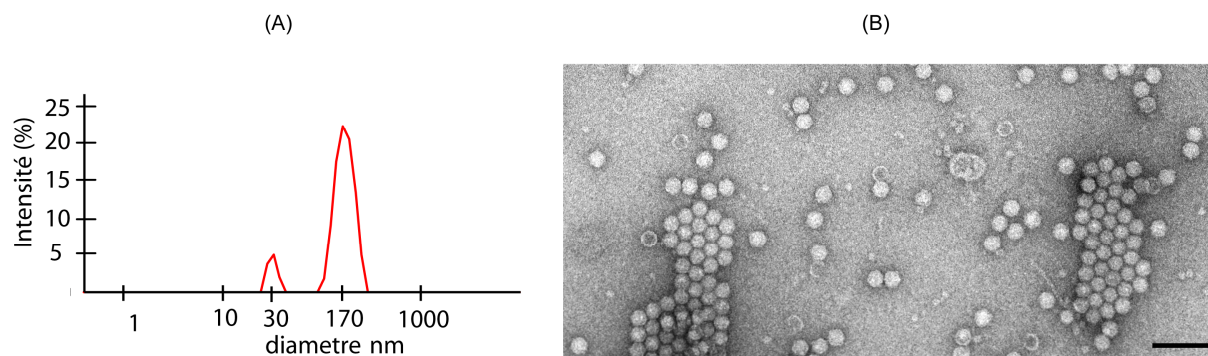


Fig. 6.2. Analyse de la qualité de purification de l'ArMV-C. (A) La diffusion dynamique de lumière montre une polydispersité élevée et la présence de deux pics. (B) La coloration négative en MET montre la tendance des particules virales à interagir les unes avec les autres. La barre d'échelle représente 100 nm.

3. Obtention de la structure de l'ArMV par cryo- MET

Dans un premier temps, les particules virales purifiées ont été congelées en phase vitreuse et observées à l'aide d'un cryo-MET Sphera 200 kV FEI équipé d'une cathode LaB₆ (Equipe Structure et Dynamique des Macromolécules, UMR 6026 CNRS - Université de Rennes I). Les images des particules d'ArMV hydratées-congelées ont été enregistrées en faible dose d'électrons ($15^{\circ}/\text{Å}^2$), à un agrandissement de 28 000 fois, et à des défocalisations comprises entre 1 et 1,5 μm .

A partir des 40 micrographies sélectionnées et numérisées à l'aide d'un scanner Nikon Coolscan 9000 ED, j'ai extrait un total de 4500 particules virales. Les images des capsides ont été normalisées, masquées et corrigées pour la fonction de transfert du microscope selon la méthode dite du « flip des phases ». Le centrage et l'orientation des particules virales a été déterminée par la méthode des transformées de Fourier polaire (PFT). Il s'agit d'une méthode spécifique de reconstruction de particules icosaédriques développée par Baker & Cheng (1996) qui utilise un modèle comme référence pour l'alignement des particules et la détermination de leur orientation d'après les trois angles d'Euler. Le TRSV, filtré à 40 Å , a été utilisé comme modèle initial. Le filtrage des données à haute résolution de ce modèle, permet d'amorcer l'alignement des particules d'ArMV-C, sans que les détails spécifiques du TRSV ne viennent interférer la reconstruction. Après plusieurs cycles itératifs d'alignement et d'attribution des angles d'Euler, nous avons obtenu une reconstruction de l'ArMV-C à une résolution de 13,5 Å , avec 2400 particules virales finales (fig. 6.3). L'indice de corrélation FSC (*Fourier Shell correlation*) appliqué pour déterminer la résolution est de 0,5.

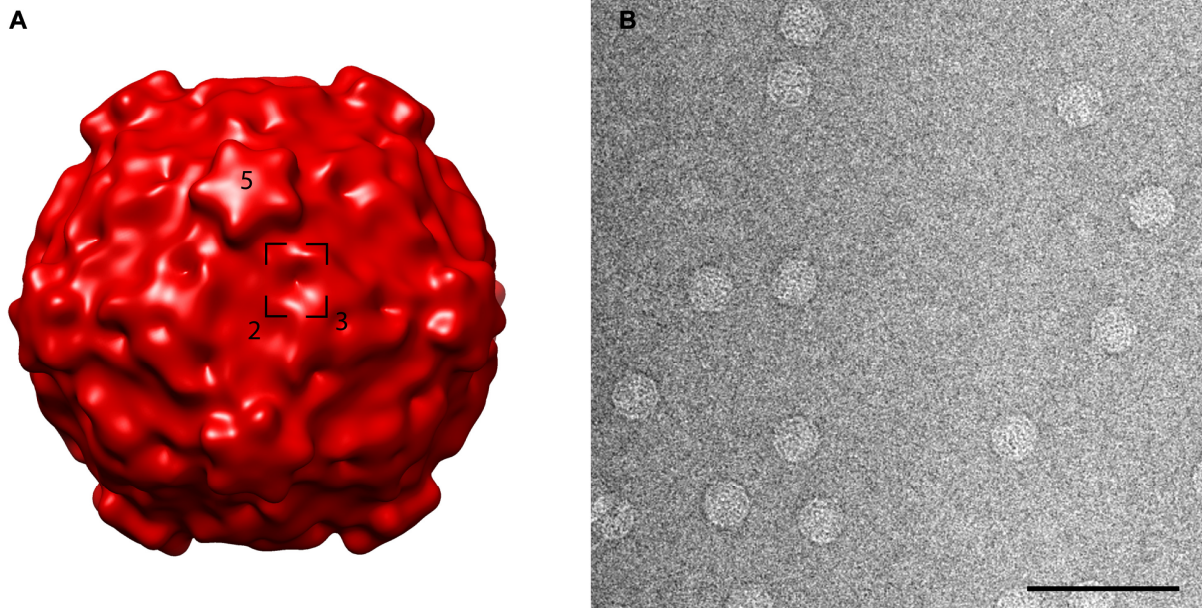


Fig. 6.3. Reconstruction 3D de l'ArMV-C. (A) Une représentation de surface de la reconstruction selon l'axe de symétrie d'ordre 2 de l'icosaèdre. L'encart noir indique la position d'une cavité de surface où se localisent dans le cas du GFLV les déterminants de sa transmission par *X. index*. (B) La micrographie de cryo-MET montre la qualité de la purification et l'homogénéité des particules virales. La barre d'échelle représente 100 nm.

La limite de résolution provient principalement du nombre de particules utilisées pour la reconstruction et le type de microscope utilisé lors de l'acquisition des images. En effet, le faisceau d'électrons produits par émission thermoélectrique, à partir d'une pointe LaB₆, comme celle utilisée pour nos prises d'images, est cohérent, mais bien moins que celui obtenu à partir d'un microscope équipé d'un canon à émission de champs (FEG, *Field emission gun*). La cohérence des électrons permet d'améliorer la fonction d'enveloppe du ME. Cette dernière affecte la fonction de transfert de contraste du microscope (CTF) et atténue le signal transmis à hautes fréquences spatiales.

Dans ce but, des images de virions hydratés-congelés ont donc été acquises à l'aide du cryo-MET Jeol 2010F de l'Institut Pasteur (Paris). Malheureusement, malgré plusieurs sessions d'acquisition, les images se sont avérées de qualité insuffisante pour obtenir une reconstruction 3D des virions à plus haute résolution. Plusieurs obstacles ont été rencontrés : qualité insuffisante de la purification de virus et dérive du porte objet cryo, se traduisant par une perte importante du signal dans une direction. Après étude, il s'est avéré que la dérive du porte-objet cryo était liée à des travaux de construction d'un bâtiment sur le campus de Pasteur. Des nouvelles acquisitions sur un cryo-MET Jeol 2200 FS (Centre de Biochimie Structurale de Montpellier) sont en cours.

3. Comparaison entre la capsidie du GFLV et de l'ArMV

La morphologie de l'ArMV-C est similaire à celle du GFLV. Les plus fortes protubérances du virus sont formées par les domaines A des CP, se rejoignant autour des axes 5 et les dépressions sont retrouvées au niveau des axes 2 (fig. 6.3). On distingue, entre l'axe 5 et l'axe 3, une cavité dont le positionnement dans la particule virale est similaire à celle décrite précédemment pour le GFLV (encart noir de la fig. 6.3).

Pour comparer les deux virus, nous avons utilisé les deux cartes obtenues par cryo-MET, à savoir celle de l'ArMV-C et celle du GFLV-TD à 16,5 Å, dont l'obtention est détaillée dans le chapitre précédant. Compte tenu de la résolution moyenne des deux cartes, nous n'avons pas effectué de carte de différence, mais simplement superposé les deux modèles, à l'aide du logiciel Chimera (Pettersen *et al.*, 2004), en tenant compte du facteur de grossissement respectif de chaque reconstruction.

La superposition indique la présence de trois zones de densités additionnelles à la surface de l'ArMV : l'une au niveau du sommet des pentamères et les deux autres en dessous de ces derniers, au niveau de la cavité (fig. 6.4). Au niveau de l'axe 5, les modifications se traduisent par une rotation et une protrusion plus importante des domaines A de l'ArMV (fig. 6.5). Il semblerait donc que le domaine A, adopte une conformation tertiaire et un positionnement différent dans la capsidie.

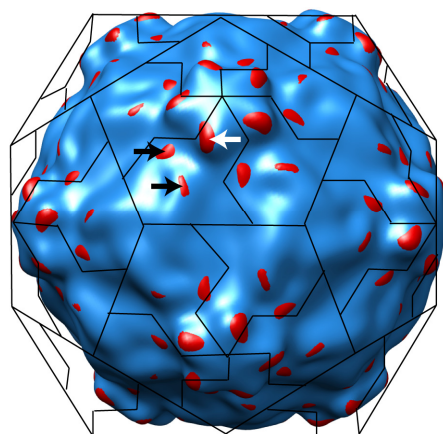


Fig. 6.4. Superposition de la reconstruction 3D du GFLV-TD (cyan) et de l'ArMV-C (rouge). La cage indique l'arrangement des sous-unités dans la capsidie selon la symétrie *pseudo* $T = 3$. Les sous-unités de l'ArMV-C sont plus protubérantes que celles du GFLV au niveau des domaines A (flèche blanche) et en dessous de ces domaines, sur les bords de la cavité (flèches noires).

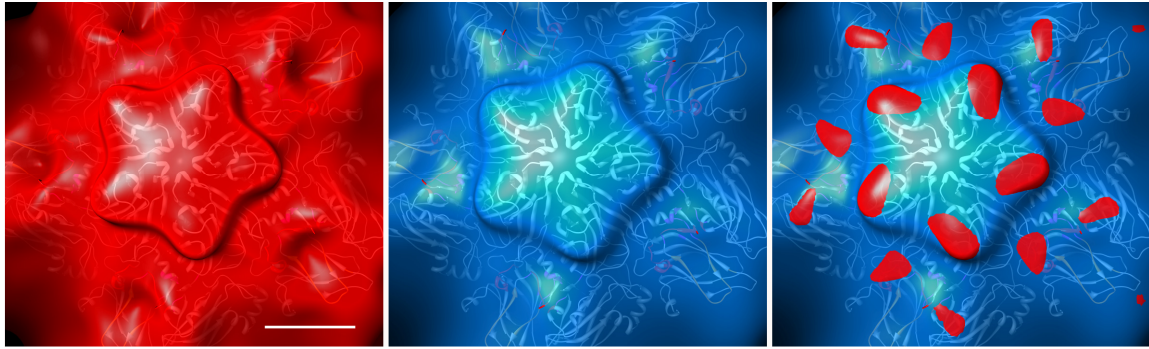
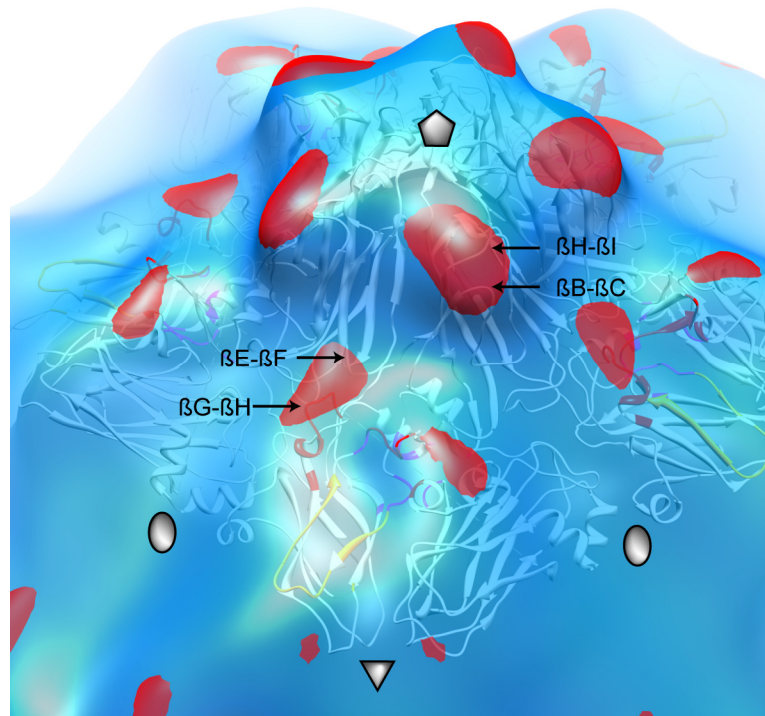


Fig. 6.5. Différences entre les pentamères de l'ArMV-C et du GFLV-TD. Représentation de surface de la reconstruction 3D de l'ArMV-C (gauche), du GFLV-TD (centre) et une superposition des deux cartes (droite). L'axe d'ordre 5 est perpendiculaire au plan de la feuille. La structure cristallographique de 5 sous-unités du GFLV-TD est placée dans la reconstruction 3D. La superposition des deux images montre que l'extrémité des domaines A de l'ArMV-C sont plus vrillés d'après le sens dextrogyne. La barre d'échelle représente 30Å.

Le positionnement de la structure atomique du GFLV dans la carte de densité de l'ArMV, a ensuite permis de localiser plus finement les régions de la CP impliquées dans ces différences.

Au niveau du domaine A, les différences pourraient correspondre aux boucles $\beta\text{H}-\beta\text{I}$ et $\beta\text{B}-\beta\text{C}$ (fig. 6.6). En dessous des domaines A, l'une des différences pourrait correspondre à la boucle $\beta\text{G}-\beta\text{H}$ du domaine B et à la boucle $\beta\text{E}-\beta\text{F}$ du domaine A (fig. 6.6). Il est important de noter que le résidu 297 impliqué dans la transmission du GFLV, est positionné dans la boucle $\beta\text{G}-\beta\text{H}$, laquelle forme le bord gauche de la cavité. La résolution des deux cartes ne permet pas de conclure sur les autres régions de la cavité.

Figure 6.6. Agrandissement de la superposition des reconstructions 3D du GFLV-TD (cyan) et de l'ArMV-C (rouge). La structure cristallographique de 5 sous-unités du GFLV-TD est représentée en *cartoon* (gris). Les régions correspondant au plateau (orange), bord droit et gauche (rouge) et centre (magenta) de la cavité sont indiquées. Les axes de symétries d'ordre 5, 3, et 2, sont représentés respectivement par un pentagone, un triangle et un ovale, pour une UA du virion.



4. Conclusion

La reconstruction 3D de l'ArMV obtenue à 13,5 Å nous a permis de mettre en évidence des différences avec le GFLV au niveau des domaines A de la CP. L'analyse suggère également l'existence d'autres différences, notamment au niveau du bord gauche de la cavité, où se localise un des déterminants de la transmission du GFLV. Aucune différence n'a pu être visualisée au niveau de la région R2, l'autre déterminant de la transmission du GFLV.

A cette résolution, la comparaison structurale reste limitée. Afin de pouvoir confirmer nos hypothèses, il sera nécessaire d'améliorer la résolution de la reconstruction de l'ArMV. Une nouvelle analyse d'image encore en cours, laisse espérer une reconstruction de l'ArMV-C à haute résolution. Ceci permettrait de produire un modèle atomique de l'ArMV, et de le comparer à celui du GFLV.



Conclusion générale et perspectives

Problématique

Au début de la thèse, les données acquises sur les déterminants viraux de la transmission du GFLV étaient limitées. Il avait été établi que le virus utilisait la protéine de capsid (CP) pour assurer sa spécificité de transmission par *X. index*, mais sur les 504 résidus de la protéine, aucune région impliquée dans l'interaction avec son vecteur n'avait été caractérisée. N'ayant aucune information sur des résidus potentiels, l'approche expérimentale, initiée au cours du travail de thèse de P. Andret-Link, s'est basée sur le postulat que ces derniers sont externes à la capsid et différents entre GFLV et ArMV. La sélection des résidus de surface était basée sur un modèle 3D de la capsid, avec toutes les limitations que comporte la modélisation par homologie de séquences et la transposition à partir d'une structure résolue d'un virus apparenté.

Motifs protéiques de la capsid impliqués dans l'encapsidation et l'infection systémique

La capsid est une protéine multifonctionnelle, c'est pourquoi il est difficile d'étudier le mécanisme de transmission et d'isoler des domaines qui sont uniquement impliqués dans cette étape finale d'un cycle viral. Une illustration de cette difficulté est l'étude des régions R3, R4 et R5. Ainsi, les régions R3 et R5 ne peuvent être substituées par leur équivalent ArMV sans empêcher la formation d'une capsid intègre capable de protéger l'ARN génomique viral. Ces conséquences drastiques sur l'encapsidation semblaient surprenantes par rapport au nombre restreint de résidus mutés, et nous avaient conduit à émettre l'hypothèse que ces régions n'étaient pas à leurs positions prédites par le modèle. Cette hypothèse a été confirmée dans le cas de R5. Plusieurs mutations effectuées dans cette région ciblent des résidus participant à la structure secondaire, tertiaire, et même quaternaire de la CP. L'étude des propriétés biologiques des mutants de la CP a également permis de suggérer le rôle d'une région (R4) dans le mouvement du virus de cellule à cellule. La région R4 avait déjà été soupçonnée être nécessaire pour l'interaction CP/MP (Andret-Link, 2005). La structure cristallographique du GFLV obtenue à 2,7 Å, confirme sa localisation en surface de la capsid, ce qui renforce cette hypothèse

Motifs protéiques de la capsid impliqués dans la transmission

La combinaison d'approches biologiques et structurales a permis d'avancer dans la compréhension de la transmission du GFLV. Deux déterminants viraux de la CP ont ainsi été caractérisés.

Dans un premier temps, j'ai montré que la substitution de la région R2 (Phe¹⁸⁸-Lys¹⁹⁸), située au niveau des brins β B et β C du domaine B, par son équivalent ArMV entraîne une perte de transmission

du GFLV par *X. index* et ne permet pas un gain de vection par *X. diversicaudatum*. Cette région est donc nécessaire à la transmission du virus. Tous les 11 résidus de R2 sont exposés à la surface du virion, à l'exception de la Phe¹⁸⁹. Ces résultats suggèrent que la région est directement impliquée dans l'interaction avec le nématode. Dans un deuxième temps, l'implication d'un autre déterminant, le résidu Gly²⁹⁷, situé dans la boucle β G- β H du domaine B, a pu être démontré grâce à la caractérisation moléculaire et fonctionnelle d'un variant naturel du GFLV faiblement transmis par *X. index* (GFLV-TD). Nous avons obtenu la structure cristallographique du variant GFLV-TD et du GFLV sauvage (GFLV-F13) ce qui a permis de les comparer. Cette analyse démontre que le mutant possède une structure identique à celle du GFLV-F13. Seule l'addition à la surface du virus d'une chaîne latérale chargée négativement au niveau du résidu 297, entraîne la perte de transmission du GFLV par son vecteur.

Il s'agit des premiers résultats sur des motifs protéiques de capsidie impliqués dans la vection d'un virus appartenant à la famille des *Secoviridae* et intégrant les données de structures 3D résolues de ces *Nepovirus*.

L'analyse de la structure cristallographique du GFLV indique également que les deux déterminants identifiés sont impliqués dans des structures quaternaires. Même si les contacts CP-CP sont minimales et que la majorité des résidus impliqués dans la transmission sont libres de toute interaction, cette information soulève la possibilité que les mutations peuvent potentiellement modifier la dynamique du virus, en influençant les forces d'interactions entre les sous-unités de la capsidie. Dans le cas du CNV, il a été montré que la capsidie subissait un changement structural lors de son attachement à son vecteur et adoptait un état « gonflé » nécessaire à la transmission du virus (Kakani *et al.*, 2004). Une étude future avec les mutants G2 (virus chimère au niveau de R2) et GFLV-TD pourrait donc consister à analyser l'effet possible de mutations sur la dynamique des particules virales, en étudiant par exemple les conséquences d'un changement de pH sur les changements de conformation de la capsidie ou le processus de décapsidation.

Par ailleurs, la substitution de R2 et du résidu Gly²⁹⁷ dans la CP du GFLV par leurs équivalents ArMV, n'est pas suffisante pour changer la spécificité de vection du GFLV. Pourtant, nous avons démontré dans la première partie de l'étude, que la spécificité de transmission du GFLV et de l'ArMV requiert uniquement la protéine capsidie. Il est donc possible que R2 soit important pour la transmission par *X. index*, mais pas au niveau de la spécificité de vection par *X. diversicaudatum*, qui serait conférée soit par d'autres régions de la capsidie ou par un changement structural de la zone d'interaction virus-nématode. Afin d'identifier les différences structurales entre le GFLV et l'ArMV, la structure de l'ArMV a été obtenue par cryo-MET. La résolution atteinte (13,5 Å) permet de visualiser des différences au niveau des sommets de pentamères, et suggère l'existence d'autres différences structurales, notamment au niveau de la boucle β G- β H du domaine B. Une amélioration de

la résolution de l'ArMV est en cours pour affiner l'analyse des différences structurales et pouvoir ainsi cibler de nouveaux sites potentiels expliquant la spécificité de vection.

Transmission & spécificité de transmission

Une des caractéristiques de la transmission des *Nepovirus* est leur extrême spécificité de vection. Les stratégies développées lors de ce travail de thèse reposent sur ces propriétés, notamment la comparaison entre les systèmes de vection, GFLV-*X. index* et ArMV-*X. diversicaudatum*, afin de caractériser les résidus clés de la spécificité. Cette stratégie a été fructueuse, puisqu'elle a permis d'identifier la région R2. Mais il s'agissait d'un outil de sélection, et cette règle n'était en aucun cas exclusive. En effet, même si nous cherchions des résidus divergents entre des virus transmis par des espèces de nématodes différents, plusieurs résidus identifiés font partie de régions de surface conservées parmi les népovirus, décrits comme le site 1 et 2 par Chandrasekar et Johnson (Chandrasekar and Johnson, 1998). Dans des études ultérieures il serait donc intéressant d'analyser tous les résidus aux environs des déterminants identifiées, qu'ils soit conservés ou non entre le GFLV et l'ArMV. Ceci permettrait de ne plus restreindre l'analyse à la spécificité de la transmission, mais identifier des domaines requis de manière générale dans la transmission du GFLV.

De plus, ce travail met en évidence que le changement de spécificité de vection d'un népovirus semble très difficile à obtenir. La reconnaissance au vecteur englobe certainement des éléments structuraux à « basse résolution », comme par exemple la forme de la cavité ou sa charge électrostatique, et des éléments de « haute résolution », comme nous avons pu le démontrer avec la chaîne latérale de l'Asp²⁹⁷. Afin de savoir si les déterminants viraux de la transmission du GFLV jouent également un rôle dans la transmission de l'ArMV, il serait intéressant d'introduire des mutations dans la boucle β G- β H du domaine B de l'ArMV, et tester si la mutation entraîne la perte de transmission par *X. diversicaudatum*.

La cavité, site potentiel majeur d'interaction avec *X. index*

Dans l'état actuel des connaissances, les résidus intervenant dans la transmission du GFLV se situent sur les bords extérieurs d'une cavité qui pourraient constituer le premier niveau de reconnaissance au vecteur. Si tel est le cas, les résidus les plus protubérants de cette cavité pourraient également être impliqués dans cette étape du cycle viral. Il s'agit des résidus à l'extrémité du plateau (Asp¹⁹⁰ à Val¹⁹⁴), du bord droit (Gly²¹¹ et Thr²¹²) et du bord gauche (Ala²⁹⁶, Gly²⁹⁷, Gln²⁹⁸) de la cavité. L'étude structurale réaffirme donc l'importance d'analyser le rôle de R3 et R5 qui forment les bords droit et gauche de la cavité. Leurs délimitations exactes seraient à redéfinir par rapport aux données structurales acquises. Une question évidente se pose : jusqu'où s'étendent les résidus impliqués dans la vection du GFLV ? Par exemple, est-ce que tous les résidus de la boucle β G- β H sont impliqués dans la vection ? Afin de répondre à cette question, il serait intéressant de substituer les résidus ciblés en acide

aspartique, puisque nous savons que cet échange peut entraîner un effet significatif sur la transmission par *X. index*.

Parmi les quelques études structurales effectuées sur les déterminants de la vection des phytovirus, l'importance d'une cavité à la surface des capsides virales a déjà été rapportée. Dans le cas du CNV, la majorité des résidus impliqués dans la transmission sont localisés au niveau du pseudo-axe 3 (Kakani *et al.*, 2001) et de l'axe d'ordre 3 (Kakani *et al.*, 2004). Or, ces deux régions se situent dans les dépressions de la capside. Également, si nous faisons le parallèle avec le CaMV, l'interaction du facteur assistant la transmission avec la capside a lieu dans les cavités de la particule virale (Plisson *et al.*, 2005).

De manière intéressante, c'est également une cavité (nommé canyon) qui est impliquée dans l'interaction des picornavirus avec leurs récepteurs cellulaires appartenant à la superfamille des immunoglobulines (Rossmann *et al.*, 2002). Dans le cas du *Poliovirus*, le récepteur se lie profondément dans le canyon et l'interaction implique de nombreux résidus (He *et al.*, 2000; Zhang *et al.*, 2008). Dans le cas du *Foot-and-mouth disease virus* (FMDV), le virus se lie à la cellule via deux sites : l'un reconnaissant des intégrines (motif Arg-Gly-Asp (RDG)), l'autre reconnaissant une héparine (Jackson *et al.*, 1996; Mason *et al.*, 1994). Dans ce dernier cas, le site de reconnaissance se situe dans une dépression chargée positivement, localisée à l'interface entre les trois sous-unités VP1, VP2 et VP3 (Fry *et al.*, 1999), ce qui ressemble étonnamment au cas de la cavité du GFLV. Ainsi, malgré la différence apparente entre un phytovirus interagissant de manière non-circulante avec son vecteur et un virus animal infectant une cellule, cette comparaison structurale laisse entrevoir un lien dans l'évolution entre les deux mécanismes de reconnaissance moléculaire.

Développement de stratégie anti-virale

L'identification de plusieurs résidus de la CP impliqués dans la transmission et l'obtention de la structure cristallographique représente une avancée majeure pour développer de nouvelles stratégies anti-virales contre la maladie du court-noué de la vigne. Maintenant qu'un site putatif d'interaction avec le vecteur est identifié, il est envisageable de produire une molécule compétitrice au virus, afin de tester si la production de cette molécule dans la plante serait capable d'empêcher sa rétention dans le nématode.

Recherche du récepteur

Enfin, la thématique des déterminants de la vection du GFLV comprend également les déterminants localisés dans *X. index*. Afin d'avancer dans la compréhension du mécanisme moléculaire de la transmission, il semble maintenant prioritaire d'avancer dans cette voie. La première étape consisterait à caractériser la nature biochimique du ligand du virus au niveau de l'appareil alimentaire du nématode. Le virus interagit-il avec un récepteur de nature protéique, comme cela a été démontré pour le CaMV (Uzest *et al.*, 2007), ou avec un ligand de type polysaccharidique, comme le

suggère le marquage spécifique de ces molécules au niveau des sites de rétention dans les *Xiphinema* (Robertson and Henry, 1986a; Robertson and Henry, 1986b) ? Plusieurs approches sont envisageables : les approches classiques de Far-Western, ou d'interaction *in - situ* qu'il serait possible d'effectuer sur coupes semi-fines de nématodes à l'aide de virus purifiés ; et des approches plus récentes comme la technique biacore ou le criblage de banques d'ADNc. La disponibilité de mutants du GFLV dont la transmission est altérée, pourra être mise à profit dans toutes ces différentes approches.



Matériel et Méthodes

Ce chapitre rassemble les protocoles expérimentaux ne figurant pas dans les paragraphes de «matériels et méthodes» des publications présentées dans les chapitres des résultats. Il concerne donc uniquement le calcul des surfaces accessibles de la structure atomique du GFLV et la reconstruction 3D à partir d'images de cryomicroscopie électronique à transmission.

1. Analyse d'images de cryo-MET

1.1 Principe général

La reconstruction 3D en ME est basée sur principe que le microscope électronique produit une projection 2D de l'objet observé. En combinant l'ensemble des projections 2D de l'objet dans toutes les directions de l'espace, le volume 3D originel peut être reconstitué. La qualité de la reconstruction 3D dépend de la qualité des images utilisées, de leur nombre et de la distribution dans l'espace des différentes orientations. Le but de cette partie méthodologique est d'expliquer les deux étapes clés, nécessaires à la reconstruction 3D : la correction de la fonction de transfert de phase (CTF) et la détermination de l'orientation des particules. Les différentes étapes du procédé sont synthétisées dans la fig. M.1.

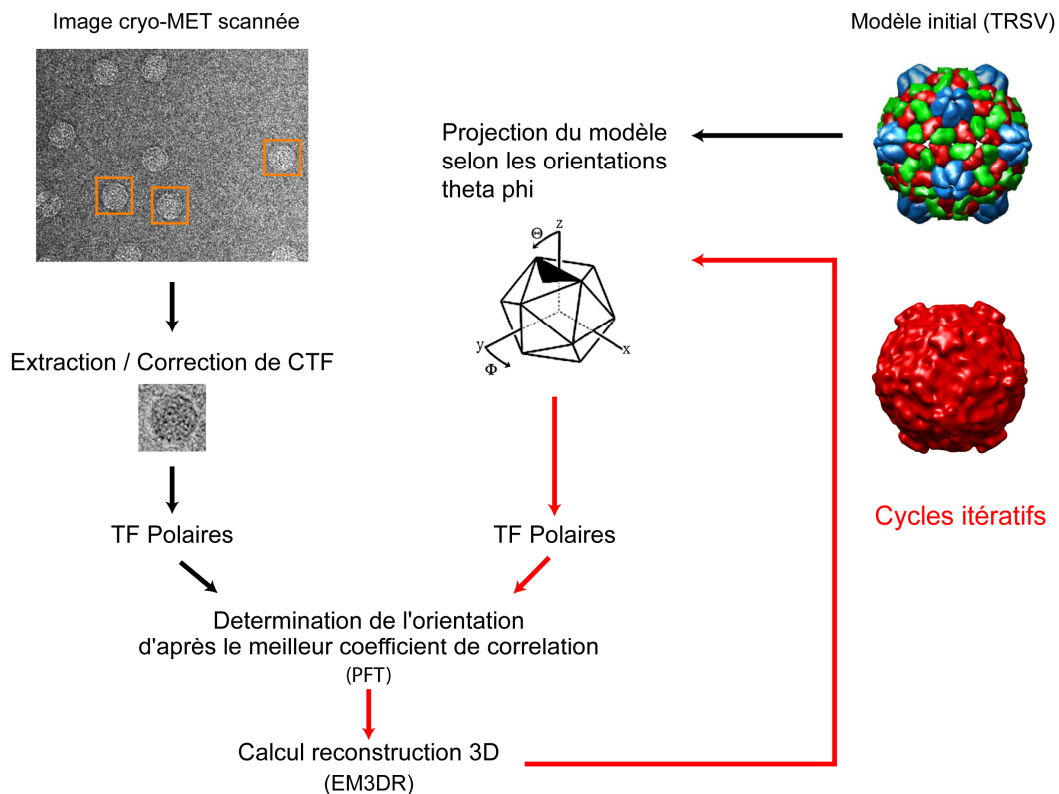


Fig M.1. Méthode de reconstruction 3D de particules virales icosaédriques

1. 1. 1 Détermination de l'orientation des particules virales d'après la méthode PFT

La reconstruction 3D de particules virales icosaédriques se base sur le théorème de la section centrale et des symétries de l'icosaèdre (Crowter *et al.*, 1970). Le théorème de la section centrale postule que la transformée de Fourier (TF) d'une projection d'un objet 3D correspond à une section centrale de la transformée de Fourier de cet objet. Les transformées de Fourier de chacune des projections de l'objet représentent donc différentes sections centrales de la transformée de Fourier de la particule virale. Si l'orientation et l'origine de chaque projection sont déterminées, la combinaison des TF permet de reconstituer la TF de la particule virale. La structure tridimensionnelle, dans l'espace réel, est ensuite calculée par une transformée de Fourier inverse (fig. M.2).

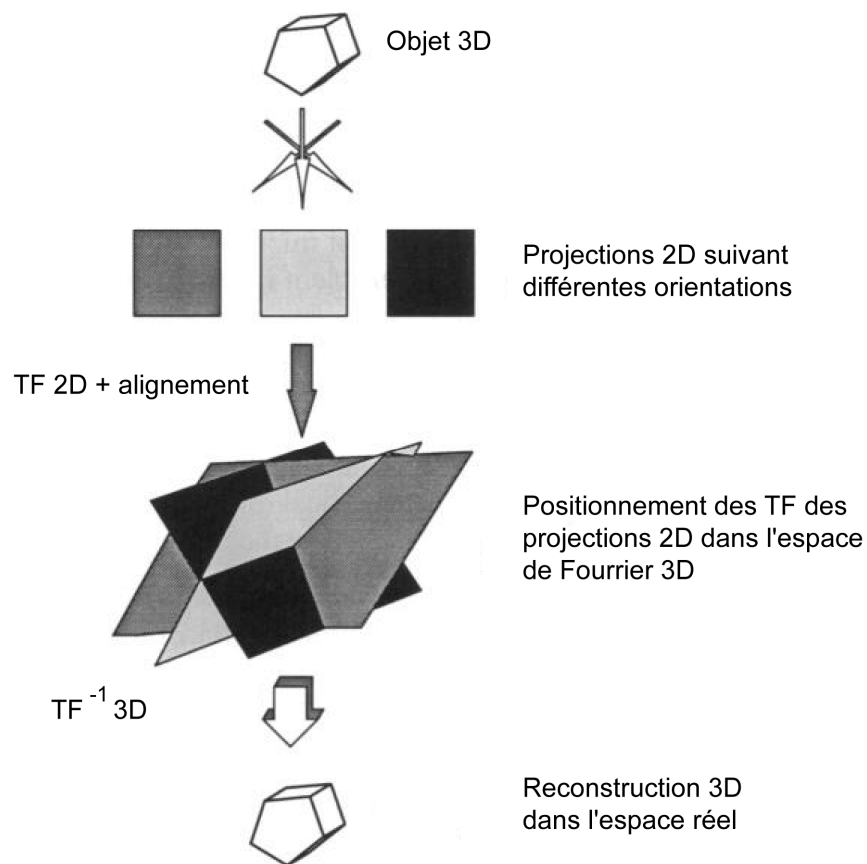
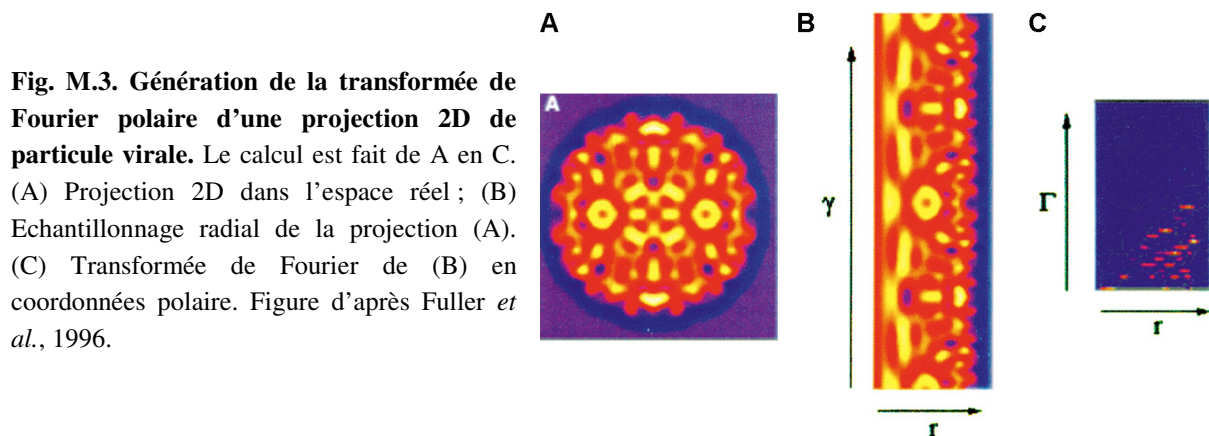


Fig. M.2. Principe de reconstruction 3D d'après le théorème de la section centrale. Figure d'après Chiu et al., 1993.

Dans l'approche de la transformée polaire de Fourier (PFT) développée par Baker & Cheng (Baker and Cheng, 1996), l'image est transformée en coordonnées polaires de Fourier (fig. M.3). Les données tiennent uniquement en compte la densité de la coque protéique icosaédrique. L'origine et

l'orientation des particules virales sont déterminées par comparaison avec les projections théoriques calculées à partir d'un modèle initial. Le programme PFT permet de réaliser une recherche rapide ou affinée de l'orientation des particules en fonction du nombre de projections du modèle utilisé. La haute symétrie facilite la détermination de l'orientation des capsides et diminue le nombre de particules virales nécessaire pour obtenir une résolution donnée (Fuller *et al.*, 1996). L'orientation des particules icosaédriques dans l'espace est définie par convention selon des angles Eulériens $\phi\theta\psi$ défini par Klug et Finch (1968). L'orientation des particules virales est uniquement déterminée dans l'unité asymétrique de la capside.



1. 1. 2 Correction de la CTF

Le microscope produit une image déformée de l'objet observé. Les modifications du signal transmis par le microscope sont modélisées dans la fonction de transfert de Contraste (CTF). Une reconstruction 3D à haute résolution, nécessite de corriger les images par rapport à la contribution de la CTF. Ce qui est possible puisque la transformée de Fourier de l'image (TF(I)) est reliée à la transformée de Fourier de l'objet (TF(O)) par la relation suivante :

$$TF(I) = TF(O) \cdot CTF(S)$$

La formule de la CTF est la suivante :

$$CTF(S) = A \sin [-\gamma (S)] - (1-A) \cos [-\gamma (S)]$$

S correspond à la fréquence spatiale, γ au déphasage introduit par la lentille du microscope et A à la proportion du signal issu du contraste de phase par rapport au contraste d'amplitude. Le déphasage γ dépend de la fréquence spatiale S dans l'image, et sa fonction est donnée par la formule suivante :

$$\gamma(S) = 2\pi/\lambda (Cs(\gamma S)^4/4 - \Delta z(\lambda S)^2/2)$$

Δz est la défocalisation de la lentille de l'objectif, et C_s le coefficient d'aberration sphérique de la lentille.

La CTF dépend donc de différents paramètres, à savoir les caractéristiques du microscope utilisé (C_s), la tension accélératrice du microscope (λ) et la défocalisation (Δf). Cette fonction est une sinusoïde qui varie en fonction de la fréquence spatiale (fig. M.4). Au passage de la CTF en zéro, aucun signal n'est transmis. A chaque passage en zéro, il est inversé. Les basses fréquences, en dessous du premier zéro, ne sont pas altérées par la CTF. Par contre dans les hautes fréquences, les oscillations deviennent de plus en plus nombreuses et rendent l'exploitation des images difficile.

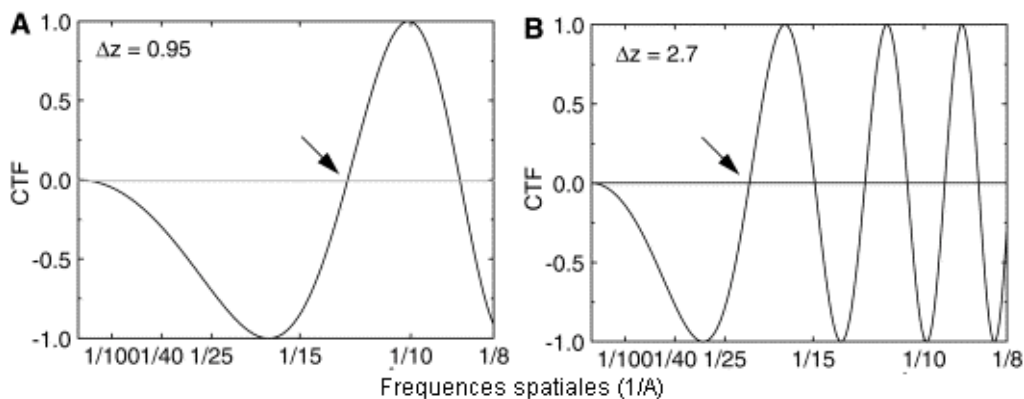


Fig. M.4. Représentation de la CTF. La variation du contraste de phase est représentée en fonction de la fréquence spatiale ($1/\text{\AA}$). Simulation pour une tension accélératrice de 400kV avec une source LaB_6 , $C_s = 4,1\text{mm}$. Le premier zéro de la fonction (flèche) a lieu à $1/12,5 \text{\AA}$ pour une défocalisation de (A) $0,95 \mu\text{m}$ et à $1/21 \text{\AA}$ pour une défocalisation de $2,7 \mu\text{m}$. Cette simulation ne tient pas compte de la fonction d'enveloppe du ME qui atténue exponentiellement l'intensité du signal vers les hautes résolutions. Figure d'après Thuman-Commike & Chiu, 1999.

1. 2 Préparation des échantillons et conditions d'acquisition des images

La cryo-fixation des protéines dans de la glace amorphe permet de limiter les dommages dus aux radiations des électrons, la déshydratation de la protéine et sa déformation induite par les fixateurs et colorants (Dubochet *et al.*, 1983).

La méthode que nous avons utilisée est la suivante :

Quelques microlitres de purification virale sont déposés sur une grille à trou. L'ensemble est monté dans un système de guillotine permettant la congélation ultra-rapide de l'échantillon dans de l'éthane liquide (-183°C), nécessaire à la création de glace vitreuse (Dubochet *et al.*, 1988). Afin d'éliminer l'excès de tampon, la grille est tamponnée avec du papier filtre Whatman (indice 1004) pendant quelques secondes avant la congélation. Le temps d'absorption du tampon est déterminant

afin d'obtenir une glace ni trop épaisse, qui serait opaque aux électrons, ni trop fine, qui entraînerait l'agrégation de l'échantillon sur les bords des trous de la grille. Cette dernière est ensuite transférée dans un porte objet cryo, qui permet de conserver l'échantillon à une température d'environ $-180\text{ }^{\circ}\text{C}$, pendant l'acquisition des images.

Afin de limiter les dommages dus aux électrons sur la structure des protéines, l'acquisition des images se fait à faible dose d'électrons et selon une méthode d'observation permettant de préserver la zone d'intérêt de la grille. Cette dernière est recherchée sur la grille à faible grossissement (mode *search*, 3 500 x). La mise au point, c'est-à-dire la recherche de la focale, est faite à un grossissement plus élevé (150 000 x) au niveau d'une zone proche de la zone à photographier (mode *focus*). L'acquisition de l'image (mode *expo*) est faite à une défocalisation définie par rapport à la focale, avec une dose d'électrons d'environ $15\text{ }^{\circ}/\text{\AA}^2$. Les images sont enregistrées sur des plans films KODAK-SO-163, développés en chambre noire. Les images sont ensuite sélectionnées par inspection visuelle avant la digitalisation.

1. 3 Reconstruction 3D de la structure de l'ArMV-C

Les images ont été enregistrées selon la méthode décrite ci-dessus, à l'aide d'un Sphera 200 kV FEI (FEI, Eindhoven, Netherlands) équipé d'une cathode LaB₆ (Equipe Structure et Dynamique des Macromolécules, UMR 6026 CNRS - Université de Rennes I). Les négatifs sélectionnés ont été scannés à l'aide d'un scanner Nikon Coolscan 9000 ED, avec un pas de 10 μm . Puis, à l'aide du programme BOXER, 4500 particules possédant de l'ARN ont été extraites avec un cadre de 300 pixels de côté. La correction de la CTF a été effectuée avec la suite de programme développé par Conway (Conway and Steven, 1999). Le spectre de puissance des particules a été calculé avec SumPS. Puis, la CTF de chaque spectre a été modélisée avec le logiciel Kaleidagraph afin d'estimer la position des « zéros » de la CTF. CTFZEROS a été utilisé pour estimer la valeur de défocalisation. La correction de la CTF a été effectuée avec CTFMIX. On obtient ainsi des images dites « flippées ».

L'étape suivante consiste à déterminer les centres et l'orientation des particules virales. Pour ceci, les images « flippées » sont comparées aux reprojctions de la structure cristallographique du TRSV, dont la résolution a été filtrée à 40 \AA pour supprimer les détails de la structure. Le programme PFT (Baker and Cheng, 1996) a été utilisé pour la détermination et l'affinement de l'orientation et du centrage. La reconstruction de chaque nouveau modèle est faite avec EM3DR. A chaque cycle itératif, les particules utilisées pour le calcul de la reconstruction sont sélectionnées d'après les deux coefficients de corrélation donné par le programme PFT : PFT-CC et PRJ-CC.

La résolution de la reconstruction est déterminée en calculant 2 reconstructions indépendantes produites à partir de 2 lots de particules séparées de manière aléatoire. Les deux reconstructions sont analysées dans l'espace de Fourier en comparant les séries de coques concentriques de la transformée

3D de chaque reconstruction (FSC, Fourier Shell corrélation) (Saxton and Baumeister, 1982). Les FSC sont calculées avec le programme *bresolve* (Heymann, 2001). La résolution de la reconstruction est fixée à 50 % de corrélation entre les deux reconstructions (FSC = 0,5).

2. Calcul des surfaces accessibles au solvant

La surface accessible au solvant (SAS) des résidus de la protéine de capsid a été calculée avec le programme MSMS (Sanner *et al.*, 1996), à partir de la structure cristallographique du GFLV-TD. Un résidu est considéré comme externe, si une sphère de 1.4 Å – valeur par défaut correspondant au rayon d’une molécule d’eau - peut rentrer en contact avec lui. Si tel est le cas, la valeur SAS sera supérieure à 0 et égale à la distance parcourue par la sphère autour des atomes du résidu. Cette dernière sera donc forcément plus élevée pour un résidu dont la chaîne latérale est longue. Par exemple, même si une glycine et un acide aspartique ont un niveau d’exposition similaire à la surface de la protéine, la valeur SAS sera plus élevée pour l’acide aspartique, parce que ce dernier possède un nombre supérieur d’atomes. Pour cette raison, il est difficile de comparer de manière exhaustive le niveau d’exposition de chaque résidu.

Le calcul a été fait soit en présence soit en absence de protéines voisines de la sous-unité D afin de déduire avec quelle protéine voisine chaque résidu interagit, ainsi que les surfaces totales de contact au niveau des différents environnements. Un monomère (D) complètement entouré de ses voisins (ABCEMNS), est repéré au travers d’une interface graphique est représenté d’après le schéma de la figure 5.1. Les valeurs des différents calculs figures en annexes 4

ANNEXES

1 . Annexe 1. Sources internet utilisées pour illustrer la figure 1.1

Acarien : site ENT - Insect identification :

<http://www.uku.fi/~holopain/ento/Kasvintuotannon-lajisto.htm>

Coleoptère et Thrips: Wikipedia :

http://fr.wikipedia.org/wiki/Wikip%C3%A9dia:Accueil_principal

Punaïse : Photo de J. Wessels (Green mirid adult)

Cicadelle et acariens : Nature photography de Tam Stuart

Delphacide : image de John Jeffry

Aleurode : challenge-service.com

Cochenille : ntlword.com :

http://www.tamstuart.com/PhotoHomePage/photo_galleries_001.htm

Trichodorides et pucerons : site de l'encyclopédie des ravageurs européen, INRA Antibes :

<http://www.inra.fr/hyppz/pa.htm>

Longidorides : source personnelle

Plasmophorides et Olpidium (zoospores) : site internet Microbial word, produit par Jim Deacon :

<http://www.biology.ed.ac.uk/research/groups/jdeacon/microbes/chytrid.htm>

2. Annexe 2. Sources internet et références bibliographiques utilisées dans la table 1. 3

Les structures correspondant aux numéros d'accessions peuvent être retrouvées sur les sites suivant :
Virus Particle Explorer (VIPER): <http://viperdbscripps.edu/> et la Protein Data Bank (PDB):
<http://www.rcsb.org/pdb/home/home.do>

Le site ViralZone est hébergé sur le serveur ExpASy (Expert Protein Analysis System)
http://www.expasy.ch/viralzone/all_by_protein/230.html

Les références bibliographiques sont les suivantes :

- Baker, T.S., Newcomb, W.W., Booy, F.P., Brown, J.C. and Steven, A.C.** (1990) Three-dimensional structures of maturable and abortive capsids of equine herpesvirus 1 from cryoelectron microscopy. *J. Virol.*, 64, 563-573.
- Baker, T.S., Newcomb, W.W., Olson, N.H., Cowser, L.M., Olson, C. and Brown, J.C.** (1991) Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. *Biophys. J.*, 60, 1445-1456.
- Belnap, D.M., Olson, N.H., Cladel, N.M., Newcomb, W.W., Brown, J.C., Kreider, J.W., Christensen, N.D. and Baker, T.S.** (1996) Conserved features in papillomavirus and polyomavirus capsids. *J. Mol. Biol.*, 259, 249-263.
- Booy, F.P., Trus, B.L., Newcomb, W.W., Brown, J.C., Conway, J.F. and Steven, A.C.** (1994) Finding a needle in a haystack: detection of a small protein (the 12-kDa VP26) in a large complex (the 200-MDa capsid of herpes simplex virus). *Proc. Natl. Acad. Sci. U S A*, 91, 5652-5656.
- Butcher, S.J., Aitken, J., Mitchell, J., Gowen, B. and Dargan, D.J.** (1998) Structure of the human cytomegalovirus B capsid by electron cryomicroscopy and image reconstruction. *J Struct Biol*, 124, 70-76.
- Butcher, S.J., Dokland, T., Ojala, P.M., Bamford, D.H. and Fuller, S.D.** (1997) Intermediates in the assembly pathway of the double-stranded RNA virus phi6. *EMBO J.*, 16, 4477-4487.
- Cheng, R.H., Caston, J.R., Wang, G.J., Gu, F., Smith, T.J., Baker, T.S., Bozarth, R.F., Trus, B.L., Cheng, N., Wickner, R.B. and et al.** (1994) Fungal virus capsids, cytoplasmic compartments for the replication of double-stranded RNA, formed as icosahedral shells of asymmetric Gag dimers. *J. Mol. Biol.*, 244, 255-258.
- Chiu, C.Y., Mathias, P., Nemerow, G.R. and Stewart, P.L.** (1999) Structure of adenovirus complexed with its internalization receptor, alphavbeta5 integrin. *J Virol*, 73, 6759-6768.
- Dokland, T., Lindqvist, B.H. and Fuller, S.D.** (1992) Image reconstruction from cryoelectron micrographs reveals the morphopoietic mechanism in the P2-P4 bacteriophage system. *EMBO J.*, 11, 839-846.
- Dokland, T. and Murialdo, H.** (1993) Structural transitions during maturation of bacteriophage lambda capsids. *J. Mol. Biol.*, 233, 682-694.
- Ferlenghi, I., Clarke, M., Ruttan, T., Allison, S.L., Schalich, J., Heinz, F.X., Harrison, S.C., Rey, F.A. and Fuller, S.D.** (2001) Molecular organization of a recombinant subviral particle from tick-borne encephalitis virus. *Mol Cell*, 7, 593-602.
- Katpally, U., Kakani, K., Reade, R., Dryden, K., Rochon, D. and Smith, T.J.** (2007) Structures of T=1 and T=3 particles of cucumber necrosis virus: evidence of internal scaffolding. *J. Mol. Biol.*, 365, 502-512.
- Kendall, A., McDonald, M., Bian, W., Bowles, T., Baumgarten, S.C., Shi, J., Stewart, P.L., Bullitt, E., Gore, D., Irving, T.C., Havens, W.M., Ghabrial, S.A., Wall, J.S. and Stubbs, G.** (2008) Structure of flexible filamentous plant viruses. *J Virol*, 82, 9546-9554.
- Kenney, J.M., von Bonsdorff, C.H., Nassal, M. and Fuller, S.D.** (1995) Evolutionary conservation in the hepatitis B virus core

- structure: comparison of human and duck cores. *Structure*, 3, 1009-1019.
- Khayat, R., Tang, L., Larson, E.T., Lawrence, C.M., Young, M. and Johnson, J.E.** (2005) Structure of an archaeal virus capsid protein reveals a common ancestry to eukaryotic and bacterial viruses. *Proc Natl Acad Sci U S A.*, 102, 18944-18949.
- McKenna, R., Olson, N.H., Chipman, P.R., Baker, T.S., Booth, T.F., Christensen, J., Aasted, B., Fox, J.M., Bloom, M.E., Wolfenbarger, J.B. and Agbandje-McKenna, M.** (1999) Three-dimensional structure of Aleutian mink disease parvovirus: implications for disease pathogenicity. *J. Virol.*, 73, 6882-6891.
- Olson, N.H., Lucken, U., Walker, S.B., Otten, M.T. and Baker, D.** (1995) Cryoelectron microscopy and image reconstruction of spherical viruses with spot scan and FEG technologies. *Proc. Microsc. Microanal.*, 1, 1086-1087.
- Plisson, C., Uzest, M., Drucker, M., Froissart, R., Dumas, C., Conway, J., Thomas, D., Blanc, S. and Bron, P.** (2005) Structure of the mature P3-virus particle complex of cauliflower mosaic virus revealed by cryo-electron microscopy. *J. Mol. Biol.*, 346, 267-277.
- Schoehn, G., Moss, S.R., Nuttall, P.A. and Hewat, E.A.** (1997) Structure of Broadhaven virus by cryoelectron microscopy: correlation of structural and antigenic properties of Broadhaven virus and bluetongue virus outer capsid proteins. *Virology*, 235, 191-200.
- Seitonen, J.J., Susi, P., Lemmetty, A. and Butcher, S.J.** (2008) Structure of the mite-transmitted Blackcurrant reversion nepovirus using electron cryo-microscopy. *Virology*, 378, 162-168.
- Shaw, A.L., Samal, S.K., Subramanian, K. and Prasad, B.V.** (1996) The structure of aquareovirus shows how the different geometries of the two layers of the capsid are reconciled to provide symmetrical interactions and stabilization. *Structure*, 4, 957-967.
- Smith, T.J., Cheng, R.H., Olson, N.H., Peterson, P., Chase, E., Kuhn, R.J. and Baker, T.S.** (1995) Putative receptor binding sites on alphaviruses as visualized by cryoelectron microscopy. *Proc Natl Acad Sci U S A*, 92, 10648-10652.
- Tang, L., Lin, C.-S., Krishna, N.K., Yeager, M., Schneemann, A. and Johnson, J.E.** (2002) Virus-like particles of a fish nodavirus display a capsid subunit domain organization different from that of insect nodaviruses. *J. Virol.*, 76, 6370-6375.
- Thouvenin, E., Laurent, S., Madelaine, M.F., Rasschaert, D., Vautherot, J.F. and Hewat, E.A.** (1997) Bivalent binding of a neutralising antibody to a calicivirus involves the torsional flexibility of the antibody hinge. *J. Mol. Biol.*, 270, 238-246.
- Thuman-Commike, P.A., Greene, B., Jakana, J., Prasad, B.V., King, J., Prevelige, P.E., Jr. and Chiu, W.** (1996) Three-dimensional structure of scaffolding-containing phage p22 procapsids by electron cryo-microscopy. *J. Mol. Biol.*, 260, 85-98.
- Walters, R.W., Agbandje-McKenna, M., Bowman, V.D., Moninger, T.O., Olson, N.H., Seiler, M., Chiorini, J.A., Baker, T.S. and Zabner, J.** (2004) Structure of Adeno-Associated Virus Serotype 5. *J. Virol.*, 78, 3361-3371.
- Xiao, C., Kuznetsov, Y.G., Sun, S., Hafenstein, S.L., Kostyuchenko, V.A., Chipman, P.R., Suzan-Monti, M., Raoult, D., McPherson, A. and Rossmann, M.G.** (2009) Structural Studies of the Giant Mimivirus. *PLoS Biol*, 7, e1000092.
- Yan, X., Chipman, P.R., Castberg, T., Bratbak, G. and Baker, T.S.** (2005) The Marine Algal Virus PpV01 Has an Icosahedral Capsid with T=219 Quasisymmetry. *Microscopy and Microanalysis*, 79, 9236-9243.
- Yan, X., Yu, Z., Zhang, P., Battisti, A.J., Holdaway, H.A., Chipman, P.R., Bajaj, C., Bergoin, M., Rossmann, M.G. and Baker, T.S.** (2009) The capsid proteins of a large, icosahedral dsDNA virus. *J. mol. biol.*, 385, 1287-1299.
- Yeager, M., Berriman, J.A., Baker, T.S. and Bellamy, A.R.** (1994) Three-dimensional structure of the rotavirus haemagglutinin VP4 by cryo-electron microscopy and difference map analysis. *EMBO J.*, 13, 1011-1018.
- Zhang, W., Fisher, B.R., Olson, N.H., Strauss, J.H., Kuhn, R.J. and Baker, T.S.** (2002a) Aura virus structure suggests that the T=4 organization is a fundamental property of viral structural proteins. *J. Virol.*, 76, 7239-7246.
- Zhang, X., Tang, J., Walker, S.B., O'Hara, D., Nibert, M.L., Duncan, R. and Baker, T.S.** (2005) Structure of avian orthoreovirus virion by electron cryomicroscopy and image reconstruction. *Virology*, 343, 25-35.

Zhang, X., Walker, S.B., Nibert, M.L. and Baker, T.S. (2002b) Structure of the human reovirus virion at 9.6Å resolution. *Microscopy and Microanalysis*, 8, 846-847.

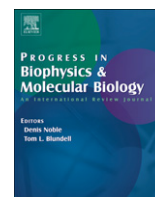
Zhou, Z.H., Chen, D.H., Jakana, J., Rixon, F.J. and Chiu, W. (1999) Visualization of tegument-capsid interactions and DNA in intact herpes simplex virus type 1 virions. *J. Virol*, 73, 3210-3218.

3. Annexe 3. Crystal growth of proteins, nucleic acids, and viruses in gels



Contents lists available at ScienceDirect

Progress in Biophysics and Molecular Biology

journal homepage: www.elsevier.com/locate/pbiomolbio

Crystal growth of proteins, nucleic acids, and viruses in gels

Bernard Lorber^{a,*}, Claude Sauter^a, Anne Théobald-Dietrich^a, Abel Moreno^{b,c},
Pascale Schellenberger^{d,e}, Marie-Claire Robert^f, Bernard Capelle^g,
Sarah Sanglier^h, Noëlle Potierⁱ, Richard Giegé^a

^a Architecture et Réactivité de l'ARN, Université de Strasbourg, CNRS, IBMC, 15 rue René Descartes, 67084 Strasbourg, France

^b Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, CB2 1GA Cambridge, UK

^c Instituto de Química, Universidad Nacional Autónoma de México, México D.F. 04510, México

^d Institut National de la Recherche Agronomique, UMR 1131 Santé de la Vigne et Qualité du Vin, INRA/UDS, Laboratoire de Virologie et Vecton, 28 rue de Herrlisheim, 68021 Colmar, France

^e Institut de Biologie Moléculaire des Plantes, CNRS, UPR2357, 12 rue du Général Zimmer, 67084 Strasbourg, France

^f formerly at the Laboratoire de Minéralogie-Cristallographie de Paris (LMCP), 4 Place Jussieu, 75252 Paris, France

^g UMR 7590, Institut de Minéralogie et de Physique des Milieux Condensés (IMPMC), 140 rue de Lourmel, 75015 Paris, France

^h Laboratoire de Spectrométrie de Masse Bio-organique (LSMBO), IPHC-DSA, Université de Strasbourg, CNRS UMR 7178, 25 rue Becquerel, 67087 Strasbourg, France

ⁱ Institut de Chimie LC3, CNRS UMR 7177, 1 rue Blaise Pascal, 67008 Strasbourg, France

ARTICLE INFO

Article history:

Available online 11 December 2009

Keywords:

Protein
Nucleic acid
RNA
DNA
Virus
Crystallisation
Gel
Agarose
Silica
Lipid phase
Cellulose
Capillary tube
Counter-diffusion

ABSTRACT

Medium-sized single crystals with perfect habits and no defect producing intense and well-resolved diffraction patterns are the dream of every protein crystallographer. Crystals of biological macromolecules possessing these characteristics can be prepared within a medium in which mass transport is restricted to diffusion. Chemical gels (like polysiloxane) and physical gels (such as agarose) provide such an environment and are therefore suitable for the crystallisation of biological macromolecules. Instructions for the preparation of each type of gel are given to urge crystal growers to apply diffusive media for enhancing crystallographic quality of their crystals. Examples of quality enhancement achieved with silica and agarose gels are given. Results obtained with other substances forming gel-like media (such as lipidic phases and cellulose derivatives) are presented. Finally, the use of gels in combination with capillary tubes for counter-diffusion experiments is discussed. Methods and techniques implemented with proteins can also be applied to nucleic acids and nucleoprotein assemblies such as viruses.

© 2009 Elsevier Ltd. All rights reserved.

1. History of crystal growth in gels

For long chemists utilized silica gels to grow single faceted crystals of inorganic compounds reaching a macroscopic size and high optical perfection (for historical reviews see Henisch, 1973, 1988; Wilke, 1988). They added the gel in the crystallisation medium to control the nucleation rate (Henisch et al., 1965), suppress heterogeneous nucleation (e.g. Halberstadt and Henisch, 1968) or immobilise crystal nuclei in the mesh and favour their growth in three dimensions.

The very first trace of an attempt to crystallise a biological macromolecule in a gel dates back to the 1950s. It was at Harvard

University where J. Lewin succeeded in growing well-developed single crystals of a barium derivative of albumin in a rigid gelatin gel (cited in Low and Richards, 1954). In the same Laboratory of Physical Chemistry related to Medicine and Public Health, B.W. Low and F.M. Richards (who invented the optical comparator box called after him) followed this example to prepare in two weeks large single crystals of pure albumin at 10% (m/v) in 5% (v/v) methanol and 1% (m/v) gelatin. They noted that Lewin's observation recalls the use of gels to slow down the mixing of components in the preparation of large single crystals of sparingly soluble compounds, a technique mentioned in Chamot and Mason's Handbook of Chemical Microscopy published in 1947. Their albumin crystals were morphologically and optically identical to those obtained by other methods but showed a marked tendency to twin on the (001) face. In the frame of their investigations they did not exploit these crystals for quantitative measurements of density, composition and

* Corresponding author. Tel.: +33 3 8841 7008; fax: +33 3 8860 2218.
E-mail address: b.lorber@ibmc.u-strasbg.fr (B. Lorber).

unit cell dimensions because they were not sure about their gelatin content (Low and Richards, 1954). Then crystallisation in gelatin seems to have been abandoned either because of the ignorance of the original method or of its advantages. Twenty years later, the gel technique is not mentioned in the reference work on protein crystallography (Blundell and Johnson, 1976).

With the preparation of the first crystallisation experiments of proteins under microgravity in the mid 1980s the use of gels was reborn. In 1987, a short paper published in French by M.-C. Robert, a mineralogist and J. Berthou, a specialist of lysozyme, lays the foundations of the crystallisation method in gels applied to proteins. The convincing results of a simple experiment with lysozyme in tetramethoxysilane gel demonstrate that the approach is applicable. They also announce the real advantages of growing soft crystals in elastic gels which will actually be discovered and confirmed repeatedly over the next twenty years. These include the growth in suspension and in the absence of convection of crystals with well-developed facets which are superb by any standard and possess excellent diffraction properties. Moreover, despite of their fragility these crystals can be extracted without damage from the gel network (Robert and Berthou, 1987). In the absence of gel, similar crystal quality can only be obtained when the level of gravity is reduced. Thus, gels are ideal diffusive media to mimic on Earth crystal nucleation and growth under weightlessness. Agarose gels were introduced immediately after silica gels (Robert and Lefaucheu, 1988). The results obtained with lysozyme and trypsin triggered an avalanche of in-depth investigations of the process of protein crystal growth as well as of the direct and indirect effects of the gel on crystal quality (García-Ruiz, 1991 and references given below).

Unexpectedly, all this did not help popularise enough the gels since, all things considered, they were rarely used to crystallise soluble proteins. At the same time as the gel technique was totally overlooked, advances were made in the difficult field of membrane protein crystallisation. For these delicate macromolecules, crystallisation assays were exclusively limited to proteins which could be dissolved in an appropriate detergent solution. Consequently, many attempts to grow crystals were unsuccessful. After numerous years of effort, a cubic phase of monoolein revealed to be a key to the preparation of well-ordered 3D crystals of bacteriorhodopsin and of other trans-membrane proteins (Rummel et al., 1998). It was also rapidly discovered that the peculiar organization of lipids in gel-like cubic phases does not restrict the use of the latter to the crystallisation of hydrophobic macromolecules. Polar proteins crystallise very well in such media (Landau et al., 1997). Alike what happened to the gels, this convection-less crystallisation method has seriously never been applied to anything else than membrane proteins.

The fact that the last protein structure determined owing to crystals grown in a gel was published in 2006 means that crystallisation in gel has again fallen in disuse. For this reason, we present here in detail the silica and agarose gels employed most frequently to crystallise soluble proteins and icosahedral viruses. Since these media are essentially made of water, they are hydrogels. We explain why macromolecules should preferably be crystallised inside the network of gels which are diffusive media “par excellence”, i.e. where molecules displace solely by diffusion. Further, we insist on the many advantages of crystallisation in gels can have over crystallisation in pure solution. Moreover, we emphasize the potential of gel-like media like lipidic phases for the crystallisation of soluble proteins. We also report results obtained with gels of other chemical compositions that may become useful in future. Further, we discuss the combined use of gels and capillary tubes. The range of applications of gels and of related methods is not limited. It goes from small proteins and nucleic acids to large macromolecular assemblies like nucleoprotein complexes or viruses.

2. Synthesis and structure of gels

2.1. Silica hydrogels

So far polysiloxane hydrogels have rarely been used. There are probably two reasons for this: generally they are not available in biology laboratories and their preparation requires some handling. The polycondensation of silicate ions leads to the formation of a dense network of silicon and oxygen atoms. Silica gels are chemical gel because all inter-atomic bonds are covalent (Fig. 1). So far the silicon-containing compounds (siloxanes) employed to prepare gels destined to protein crystallisation are sodium metasilicate (also called water glass), tetramethyl orthosilicate (TMOS) and tetraethyl orthosilicate (TEOS). These compounds possess a well defined chemical composition. Nevertheless, crystal growers should be aware that the hydrolysis of TMOS and TEOS preceding the condensation of the hydrated silica tetrahedral ions is accompanied by the production of significant amounts of methanol and ethanol, respectively. Hence, these gels must be cured to remove undesired alcohol molecules which may have deleterious effects to biological molecules. Chemical gels are irreversible, elastic and resist to deformation but swell in water.

In the scanning electron microscope (SEM), the flexible polymer network of the silica gel resulting from neutralization of a sodium metasilicate solution with acetic acid has a random pore size distribution from 50 to 150 nm (Moreno et al., 1999a). The distribution is more homogenous in gels prepared at alkaline pH. The pores of TMOS polysiloxane hydrogels have different sizes. Their size and the stiffness of the gel is concentration-dependent; the higher the concentration, the lower the pore size. Their diameter ranges from 150 to 250 nm in 10% (v/v) gels to 50–100 nm in 20% (v/v) gels. Small angle X-ray scattering (SAXS) analyses confirm that the pore size distribution of TEOS gels is also random. Further, condensation involves interactions between particles whose size and fractal dimension depend on the synthesis pathway and on their final location. Three stages of cluster generation have been

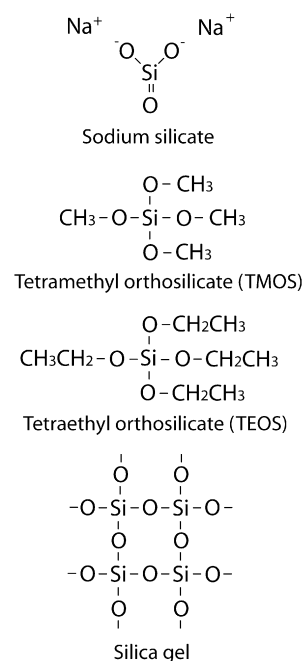


Fig. 1. Silicates and structure of silica gels. (Top) Sodium silicate; (Centre), Tetramethyl orthosilicate (TMOS) and tetraethyl orthosilicate (TEOS); (Bottom) Schematic representation of the silica gel network.

observed in silica gels using SAXS, atomic force microscopy (AFM) and SEM. The structure of these gels is discontinuous from their surface to their inner part (Vinogradova et al., 2003). Table 1 gives a protocol for the preparation of crystallisation assays with silica gel.

2.2. Agarose gels

Agarose is a gelling compound extracted from certain red sea weeds of the Rhodophyceae group (McHugh, 1987). Its long chains are repetitions of the non-charged and hydrophilic disaccharide unit agarobiose (Fig. 2). They are associated in double helices and the latter are grouped in bundles of several thousands. Interactions between chains rely on van der Waals forces or hydrogen bonds of the order of kT . Agarose is a physical gel because the association of its chains is reversible upon heating. The difference between its melting (T_m) and gelling (T_g) temperatures is called hysteresis (Guenet, 1992). At variance with chemical gels, physical gels do not swell but their network breaks apart when a deformation is applied on the rigid fibres.

Agarose was the first sieving medium used in electrophoresis. A cursory glance at manufacturer catalogues reveals a wide variety of commercially available products. Depending on the source, the natural biopolymer contains a variable proportion of β -D-galactose substituted with a methyl or a sulphate group in position 6. A low content of electronegative groups, transparency and absence of colour are indicators of purity. Agaroses distinguish also by the length of their chains and the strength of their gel (usually given for a concentration of 1 or 1.5% m/v). The process of gel formation is influenced by chain length, chemical composition, and degree of derivatisation (Griess et al., 1993). Agarose with shorter chains has a lower melting temperature. The presence of methyl groups increases the melting temperature. Low or ultra-low-gelling temperatures are useful for the crystallisation of biological macromolecules that do not withstand elevated temperatures. The first have a T_m close to 60 °C and a T_g of about 28 °C. For the others T_m is below 50 °C and T_g below 20 °C. Differences between one product and another and variations from one batch to the next make the comparison of the experimental data found in literature very difficult.

The reticular structure of the agarose mesh, its strength and its elasticity vary with concentration. Also, at a given concentration, the size of the pore varies with the nature and concentration of additives (as for example some precipitating agents). For this reason, the mesh of the gel is more reproducible when polymerization occurs in pure water or in the presence of a low concentration of buffering substance. In extreme cases, the polysaccharide

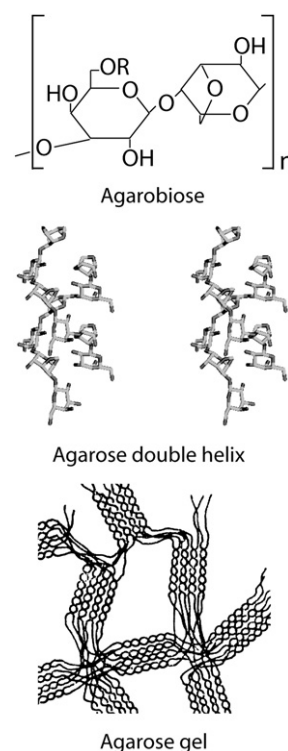


Fig. 2. Elementary unit and structure of agarose gel. (Top) Structure of agarobiose composed of β -D-galactose molecule linked to 3,6 anhydro α -L-galactose via an 1–4 bond. R is hydrogen in the absence of methyl or sulphate group. (Centre) Agarobiose units are linked together via 1–3 bonds. Their number ranges from a few dozens to several thousands. Resulting chains have a mean diameter of 8 nm and an average M_r estimated to 10^5 and adopt a double helical structure. Image created with coordinates from pdb file 1aga using *Pymol* (www.pymol.org). (Bottom) The gel network is an entanglement of chains associated reversibly and exclusively via van der Waals forces or hydrogen bonds (drawing adapted from Arnott et al., 1974). Upon heating agarose the chains separate and adopt a random coil conformation. When the heated solution cools they take the helical conformation and upon crossed linking they form an elastic and transparent gel. Further association of chains and additional crossed links consolidate the network; the gel becomes turbid and rigid. Finally, the gel expulses part of its liquid phase. This phase separation is called syneresis.

chains may be segregated by the chemical additive and polymerization cannot occur (Gonzales-Ramirez et al., 2008). The mean pore diameter decreases with agarose concentration: at low ionic strength (i.e. 10 mM tris-borate-EDTA) it is 520, 500, 350, 300, 230 and 180 nm at 0.7, 1, 2, 3, 4 and 5% (m/v) respectively (Maaloum et al., 1998). It increases with the ionic strength: for a 1% (m/v) gel

Table 1
Preparation of vapour-diffusion crystallisation assays containing silica gel.

Step
1. Dilute a commercial sodium silicate stock solution Na_2SiO_3 with a density of $\rho = 1.39$ g/ml (e.g. Aldrich Cat. No. 338443), with water to prepare a working solution with a density of 1.06 g/ml using the relationship $V_{SS} = (0.06 V_a) / (1.39 - \rho^T)$, where V_{SS} is the volume of stock required to prepare a final volume V_a and ρ^T is the density of water at temperature T . This solution can be kept a few days.
2. Dispense reservoir solutions in the sample wells of crystallisation plates and cover temporarily.
3. Neutralise above silicate solution with a 1 M acetic acid solution to obtain a monosilicic acid (H_4SiO_4) solution. The final solution must be used immediately.
4. Prepare sample drops on glass coverslips, microbridges or microplates by mixing equal volumes of protein solution/reservoir solution/silica solution. The polymerization of the polysiloxane hydrogel occurs in less than 1 h.

Safety note: Siloxane solutions are corrosive. Skin and eyes protection is recommended for handling. Glassware must be thoroughly rinsed with ethanol prior to cleaning with water.

Comments: TMOS and TEOS are not chemically neutral and should never be in direct contact with proteins. In addition, these liquid compounds are not soluble in water. Their gelling process involves a hydrolysis followed by a polycondensation. Experimenters should be aware that the hydrolysis consumes water (vigorous stirring accelerates dissolution and clearing) and releases alcohol (methanol or ethanol, respectively). The molar ratio of water to TMOS is correlated with the size of the pores in the gel matrix (Brinker and Scherrer, 1990). For a TMOS gel of percentage p , the volume of water is $2.5 \times p$ ml and that of methanol $0.9 \times p$ ml. For other silica gel preparation protocols see refs Robert and Lefaucheux (1988), Cudney et al. (1994), and Vidal et al. (1998a,b). The component of the gel can be dispensed using a liquid dispensing apparatus for high-throughput applications (Chayen and Saridakis, 2002).

the mean values are 370 nm (1 mM), 509 nm (10 mM), 966 nm (100 mM) and 1777 nm (1 M). The pore size distribution is broadest at lowest agarose concentrations and at highest ionic strength (Maaloum et al., 1998). In 0.15–0.5% (m/v) gels prepared in 10 mM tris-borate-EDTA pores have diameters ranging from 300 nm to more than 1 μm (Pernodet et al., 1997). Independently of these fluctuations, pore diameters are always compatible with the diffusion of large particles like icosahedral viruses measuring 30–40 nm across. Once the gel is formed the pore size is invariable.

From the practical standpoint, agarose can easily be added to solutions destined to crystallisation. Table 2 gives a protocol for the preparation of the assays. Agarose is easy to handle manually or with an automated pipetting station. Fig. 3 depicts the material required to melt it and maintain a solution at the right temperature. An aluminium block designed to adapt underneath a transparent plastic plate keeps the agarose solution liquid while it is transferred in the crystallisation plate by a Mosquito[®] robot (Fig. 3). Equilibration kinetics derived from refractive index measurements performed at various time intervals, indicate that 0.2% (m/v) low-gelling agarose ($T_g = 28^\circ\text{C}$) does not alter the rate of water evaporation (Fig. 3).

3. Diffusive versus convective mass transfer

Any crystal growth process is accompanied by the formation of solute concentration gradients. Concentration differences generate density gradients. The latter lead to flow in crystallising solutions and the associated differences in refractive index are visible as shadow patterns on Schlieren photographs or as phase shifts on interferograms (see e.g. Chen et al., 1979). They were observed during the growth of macroscopic salt crystals and also around the 0.3–1.7 mm-long (110) faces of lysozyme crystals (Pusey et al., 1988). At the surface of the latter the velocity of convective plumes reaches 50 $\mu\text{m/s}$ at 18 $^\circ\text{C}$. Tiny crystals measuring <20 μm across see their growth rate decline when they are exposed to the direct flow of solution. Comparative studies of inorganic crystal growth revealed that the silica network restricts the transfer of mass and heat between nutrient solution and crystal surface to diffusion (e.g. Rubin, 1969; Gits-Léon et al., 1987). Further, the network of both silica and agarose gels suppresses the convective flow triggered by density gradients in steady solutions as does weightlessness (Robert et al., 1988). Hence, depletions zones around growing

crystals are stable in the gel alike under microgravity (Otalora et al., 2001). It may also keep impurities away from crystal surfaces (see e.g. Chernov, 2003).

The dimensionless Grashof number (Gr) can be used to understand transport processes in solution. It is defined as the quotient of buoyancy over viscosity forces and is represented by,

$$Gr = L^3 \beta \Delta c g \nu^{-2}$$

where L (in cm) is the characteristic length of the system (for example the diameter of a spherical crystallisation drop or of a cylindrical capillary tube), β the solute expansivity (in cm^3/mg), Δc the concentration difference (in mg/cm^3), g the acceleration due to gravity ($9.81 \times 10^2 \text{ cm}^2/\text{s}$ on Earth), and ν the kinematic viscosity (in cm^2/s). When Gr is less than 1, the contribution of convection is so weak that mass transfer occurs essentially by diffusion (see e.g. García-Ruiz et al., 2001a).

Above equation indicates that Gr is proportional to the level of gravity. It can be 10^{-4} – 10^{-9} -fold smaller than on Earth's surface in the microgravity environments existing respectively aboard manned orbiters and satellites. On the other hand, it depends even more upon the characteristic length L of the container since this parameter appears as L^3 . Thus, inside an agarose gel with a pore diameter of 1 μm , Gr is about 10^9 -fold less than in a spherical drop of solution measuring 1 mm across. Consequently, silica and agarose are perfect media to prepare crystals of biological macromolecules. Crystal growth takes place in the solution trapped inside the mesh of the hydrogel. From above theoretical considerations, one may expect to be able to produce on Earth crystals having properties similar to those of crystals that grow under microgravity.

4. Protein crystal growth in gels

4.1. Results from silica gels

Silica gels were seldom used to grow biological crystals since their introduction in the field (Robert and Lefaucheur, 1988). This is surprising because a kit of ready-to-use sodium silicate and acetic acid solutions is commercially available since many years. On the basis of its chemical composition, the silica gel prepared from these chemicals is innocuous. In contrast, the alcohol produced during the polymerization of gels prepared with TEOS or TMOS should be removed before use. Repeated additions and withdrawal of water

Table 2
Preparation of crystallisation assays containing agarose gel and crystal recovery.

Step
1. Prepare a 2% (m/v) stock solution of low-gelling-point agarose ($T_{gel} = 28^\circ\text{C}$). In an Erlenmeyer flask, dissolve the appropriate amount of agarose powder in ultra-pure water and heat it in the microwave at 90–100 $^\circ\text{C}$. Pass the warm solution over a 0.22 μm porosity filter membrane to remove dust and insoluble particles. Dispense the filtrate in 1 ml aliquots and store at 4 $^\circ\text{C}$ or 20 $^\circ\text{C}$.
2. Heat a sample of 2% (m/v) agarose stock solution to 80–90 $^\circ\text{C}$ in a heating block. Once the solution is completely liquid, move it to a block heated at 30–35 $^\circ\text{C}$. Care should be taken to heat agarose suspension long enough (at least 5 min) at a sufficiently high temperature to melt all chains. Otherwise the pore size may not be reproducible. It may be necessary to place the crystallisation plates for a short time at low temperature (e.g. at 5–10 $^\circ\text{C}$) when working with an agarose having a gelling temperature below room temperature.
3. Dispense reservoir solutions in the wells of crystallisation plates, sample wells or depressions of microplates.
4. Prepare crystallisation drops on glass coverslips or polystyrene microbridges by mixing one protein solution with one volume of reservoir solution and add the adequate volume of 2% (m/v) agarose solution to have a final concentration of 0.1–0.4% (m/v). The time required for complete polymerization will depend on agarose concentration.
5. After crystallisation, the gel can be dissected away with microtools under a binocular microscope. A small incision opens the agarose network and liberates the crystal. The enzyme agarase can be used to hydrolyse the polysaccharide chains but it must be free of proteases and/or nucleases. Crystals in a capillary are recovered in a drop of mother liquor by applying a little air pressure on one end. They may also be analysed in situ at room or cryotemperature since they are immobilised in the gel.

Safety note: Skin and eye protection should be worn during handling of hot agarose solutions.

Comments: The gelling temperature of agarose is usually given for a 1, 1.5 or 2% (m/v) solution in water. At a lower concentration the gelling temperature is lower. A liquid agarose solution can be handled by an automated pipetting station (see Fig. 3). Agarose does not form a gel in the presence of some crystallising agents (Gonzales-Ramirez et al., 2008). In this case it can be added first to the macromolecular solution and, after polymerization, the crystallising agent can be diffused into the gel by contact. Proteins such as lectins which bind carbohydrates may interact with agarose chains.

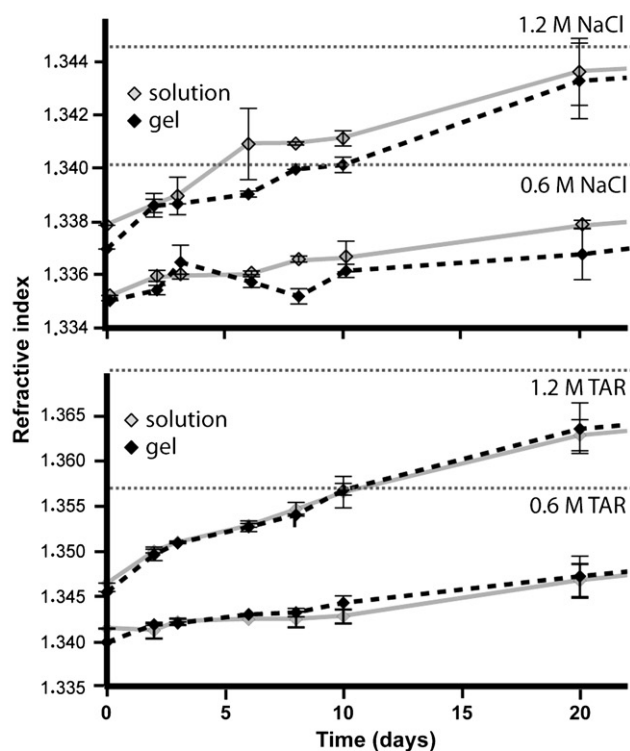
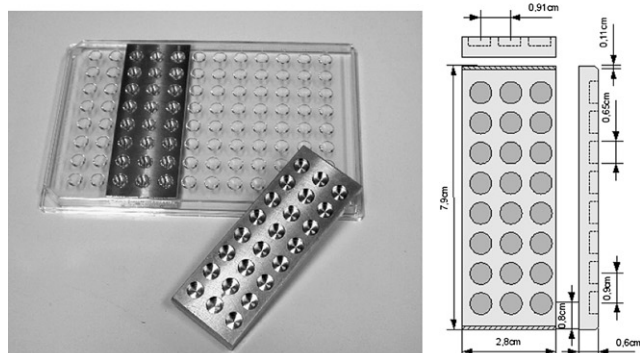
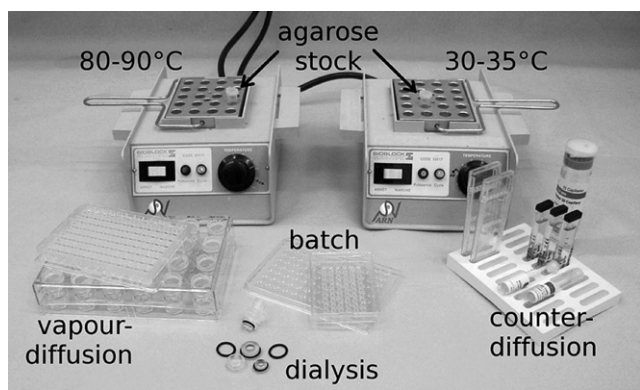


Fig. 3. Handling and properties of agarose gel. (Top) Heating blocks used for melting and maintaining agarose at the right temperature before addition to crystallisation assays. The gel is compatible with all crystallisation methods. (Centre) Aluminium plate fitting underneath a transparent Impact™ (from Greiner BioOne) plastic plate which can be warmed to maintain agarose liquid while it is transferred in crystallisation drops by a Mosquito® pipetting station (www.TTPLabtech.com). (Bottom) Equilibration kinetics for 0.6 and 1.2 M solutions of (upper panel) sodium chloride and (lower panel) sodium tartrate in the absence and in the presence of 0.2% (m/v) low-gelling temperature agarose. Buffers were 0.1 M ADA pH 6.8 and 0.1 M sodium acetate pH 4.5, respectively. Plots represent the variation of the refractive index of crystallisation drops (initial volume 30 μ l) as a function of time. Data are means of four measurements. Horizontal dotted lines correspond to the equilibrium values reached by vapour diffusion in reservoir solutions.

or buffer solution are generally sufficient to reduce significantly the concentration of undesired contaminants. Convection currents are strongly attenuated by the silica gel network and above a critical concentration they are abolished. In the gel, crystals grow at the place where they nucleate (Fig. 4). Within a broad range of concentrations the mechanical properties (e.g. the elasticity) of silica gels are compatible with the growth of crystals with great dimensions (up to several mm on the edge).

The first crystallisations of hen lysozyme and porcine trypsin indicated that these macromolecules diffuse freely through the polysiloxane hydrogel matrix. The crystals are birefringent between crossed polarizers and have the same habit as controls prepared in pure solution but nucleation is less frequent (Robert and Lefaucheux, 1988). Interferential holography images revealed protein concentration gradients around growing crystals (Lefaucheux et al., 1984) (see Fig. 4). Cudney et al. (1994) experimenting with various proteins, one virus and one transfer ribonucleic acid, observed that the presence of the gel decreases the crystal growth rate and leads to fewer nuclei ending in crystals with larger sizes. Crystals were more stable but the habit of some of them was altered. On the other hand, the habit of crystals grown from mixtures of closely related lysozymes from hen and turkey egg-white was much less altered in the presence of TMOS gel (Provost and Robert, 1995).

At high TMOS concentrations surface energy anisotropy is so low that single lysozyme crystals become spherical (García-Ruiz et al., 1998a). The gel network is so strong that a ghost of silica is left after complete dissolution of the protein. In spite of the strength of the gel, the short-range order is maintained; these reinforced crystals diffract X-rays to 1.5 Å, are less sensitive to dehydration and can easily be handled at ambient conditions. The comparison of gel-free and gelled lysozyme solutions by small angle neutron scattering (SANS) over the time intervals of prenucleation and of crystal growth led to the conclusion that a proportion of the protein adsorbs on the gel. The concomitant decrease of the concentration of the soluble protein results in lower supersaturations and nucleation rates. Hence, the silica matrix can be considered as a nucleation inhibitor for small molecules and for lysozyme (Vidal et al., 1998b).

Silica-gel-grown lysozyme crystals produce diffraction patterns with much sharper reflections (extending to at least 1.6 Å resolution) than solution-grown ones. In X-ray topography these crystals distinguish by the presence of single Bragg peaks. The existence of a very good long-range order in the lattice was deduced from rocking curves with full-width at half-maximum (or mosaicity) of only 15 arc seconds and low level of defects visible on topographs (Vidal et al., 1999).

Altogether, these encouraging results should encourage crystal growers to test silica gels with their molecules. Fig. 5 displays photographs of crystals of icosahedral viruses grown in this medium. Their habit is identical to that of crystals prepared in the absence of gel. Crystallisations of concanavalin A, trypsin and C-phycoerythrin in TMOS gel yielded fewer numbers of crystals with augmented volume (Chayen and Saridakis, 2002).

4.2. Clues from agarose gels

Few protein crystal growers have tried this easy-to-use polysaccharide available on the shelves of any molecular biology laboratory. Table 3 lists proteins and nucleic acids crystallized in agarose which had diffraction properties suitable for structure determination. The resolution of the final model is equal or better than 2 Å for 6 amongst the 15 proteins. It is noteworthy that (i) agarose is not only suitable for the crystallisation of proteins but also of nucleic acids, and that (ii) the diffraction limits can be better than 2 Å and even as high as 1.2 Å even when data collection is done at

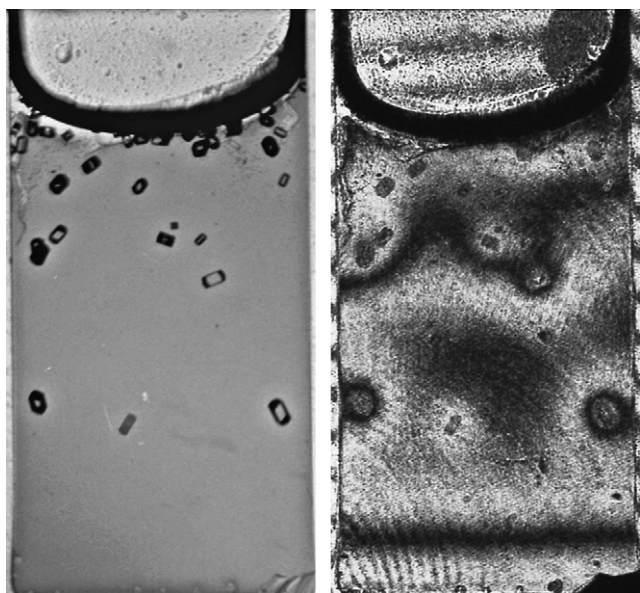


Fig. 4. Growth of hen egg-white lysozyme crystals in silica gel. The protein was at 30 mg/ml in a solution containing 0.4 M sodium chloride, 50 mM sodium acetate and 0.81% (w/v) tetramethyl-polysiloxane. (Left) Crystals visualized in transmitted white light. (Right) Salt gradient imaged by holographic interferometry. The width of the quartz cell is 10 mm. For experimental details, see Lefaucheur et al. (1984). The interference fringes are indicative of the mass transfer and reveal the lower concentration of soluble protein near the crystals.

a temperature between +15 and +20 °C (Table 3). This well illustrate that the presence of the polysaccharide is not deleterious to crystal quality. It is also a strong argument to analyse crystals at a temperature close to that at which they have grown and to which the protein is functional *in vivo*.

Numerous rigorous studies were designed to understand protein crystal growth in agarose. Protein concentration gradients around growing crystals like those seen in silica gel were revealed when optical phenomena were monitored (Lefaucheur et al., 1984). Refractive index variations during lysozyme nucleation in supersaturated solutions in a gel at a concentration of 0.1% (m/v) were interpreted as amorphous precipitate (appearing when protein and crystallising agent are mixed) that does not sediment alike in free solution. Hence, more protein is available for nucleation than in the absence of gel (Robert et al., 1994; Vidal et al., 1996). A study of the evolution of the fringe patterns in Michelson interferometry showed that the protein-depleted zone forms in three steps. First nucleation takes place, then purely diffusive mass transfer occurs, and finally the concentration of soluble protein decreases slowly until crystal growth stops (Vidal, 1997). The prenucleation clusters found by SANS led to the conclusion that the polysaccharide network promotes nucleation and to the hypothesis that lysozyme molecules first associated randomly before they rearrange as ordered nuclei (Vidal et al., 1998a). Agarose is also known to favour the nucleation of many small molecules. It is probably for technical reasons that SAXS analyses did not detect lysozyme aggregates in this gel (Finet et al., 1998).

Actually, the gel behaves as a non-Newtonian liquid when it is present at low concentration. At higher concentrations, it is viscoelastic (García-Ruiz et al., 2001a). Inside the loose network of a 0.15% m/v gel crystals grow as fast as in solution (Lorber and Giegé, 2001). As shown in Fig. 3 vapour diffusion is not slowed down by the gel. Ligands (such as natural substrates, analogs or inhibitors), compounds rich in electrons required for the preparation of heavy atom derivatives, and cryoprotectants like glycerol or

PEG (Gavira et al., 2002) diffuse through the gel without any restriction. A recent kinetic analysis using a temperature-jump technique proposed that the network of the gel inhibits the nucleation of lysozyme and that in turn this results in bigger crystals (Wang and Liu, 2008).

An obvious advantage of the gel network is that crystal nuclei are trapped and cannot move in the solution, an invaluable advantage for photographic (e.g. as reported by Charron et al., 2002) or spectroscopic analyses (Bonneté et al., 1996; Finet et al., 1998; Vidal et al., 1998a,b). The addition of a low concentration of agarose (0.15% m/v) minimises the effects of gravity on fluid dynamics. This is a prerequisite for the implementation of counter-diffusion in glass capillaries (García-Ruiz et al., 1993, 1999; García-Ruiz and Moreno, 1997; Ng et al., 2003). In a similar investigation, 0.05% (m/v) gellan gum (a water-soluble bacterial polysaccharide) sufficed to depress convection and detect wider concentration gradients near the surface of growing lysozyme crystals and lower transport rates than in solution (Hou et al., 2001). On the opposite extreme of the concentration scale, semi-solid gels composed of up to 2% (m/v) agarose are also compatible with protein crystal growth (Sugiyama et al., 2009).

On many occasions agarose was employed to immobilise crystals in solution or inside capillary tubes (for some examples see Fig. 5 and Section 6.3 for details on crystallisation in capillaries). For instance, it was added in vapour diffusion (Provost and Robert, 1991; Bernard et al., 1994), in microdialysis (Thiessen, 1994) and counter-diffusion experiments (García-Ruiz et al., 2002) or in drops kept under oil (Moreno et al., 2002). In another instance it was used in the form of beads (Willaert et al., 2005) or to solidify the reservoir solution (Wong et al., 2009). Also, crystals grown under microgravity in 0.2% (m/v) gel stay at the place where they nucleate. They are less sensitive to mechanical shocks and to temperature fluctuation during their return from space to the laboratory. Their crystallographic quality is superior to that of controls prepared in parallel in the laboratory (Lorber et al., 1999b). The mesh of polysaccharide chains is strong enough to prevent crystal sedimentation (Normand et al., 2000) and its elasticity is no hindrance to the formation of the crystal lattice. Onboard an orbiter the gel attenuated crystal motions like drift and stirring provoked by acceleration peaks (Lorber et al., 2000). In weightlessness nucleation occurs mainly in the bulk of the sample, is more synchronous and 2 times faster than on Earth (Lorber and Giegé, 2001).

Agarose also facilitated the measure of rocking curves on proteins crystal grown in magnetic fields (Lübbert et al., 2004) as well as cryocrystallography analyses (Lopez-Jaramillo et al., 2001; Biertümpfel et al., 2005) and single-wavelength anomalous scattering diffraction (SAD) data collection inside X-ray capillaries (Gavira et al., 2002). It abolishes buoyancy-driven convection in samples subjected to high hydrostatic pressures (Kadri et al., 2002, 2003a,b, 2005) or magnetic fields (Gavira and Garcia-Ruiz, 2009).

The role of agarose in biological crystals has been investigated. SEM images of lysozyme crystals grown in this gel and subsequently frozen and fractured display randomly oriented gel fibers which run right through the crystalline surface (Gavira and Garcia-Ruiz, 2002). The soft material has an ordered structure neither in the core of the crystal nor near its surface and it does not seem to interfere with the diffraction quality of the lattice of biological molecules. On the contrary, crystals containing agarose are less fragile during irradiation either at room temperature (Sauter et al., 2002) or after cryocooling *ex situ* (Zhu et al., 2001; Biertümpfel et al., 2005) or *in situ* (Gavira et al., 2002).

Although some crystals grown in agarose gel have an unusual aspect and seem to be of inferior quality, they may diffract X-rays to

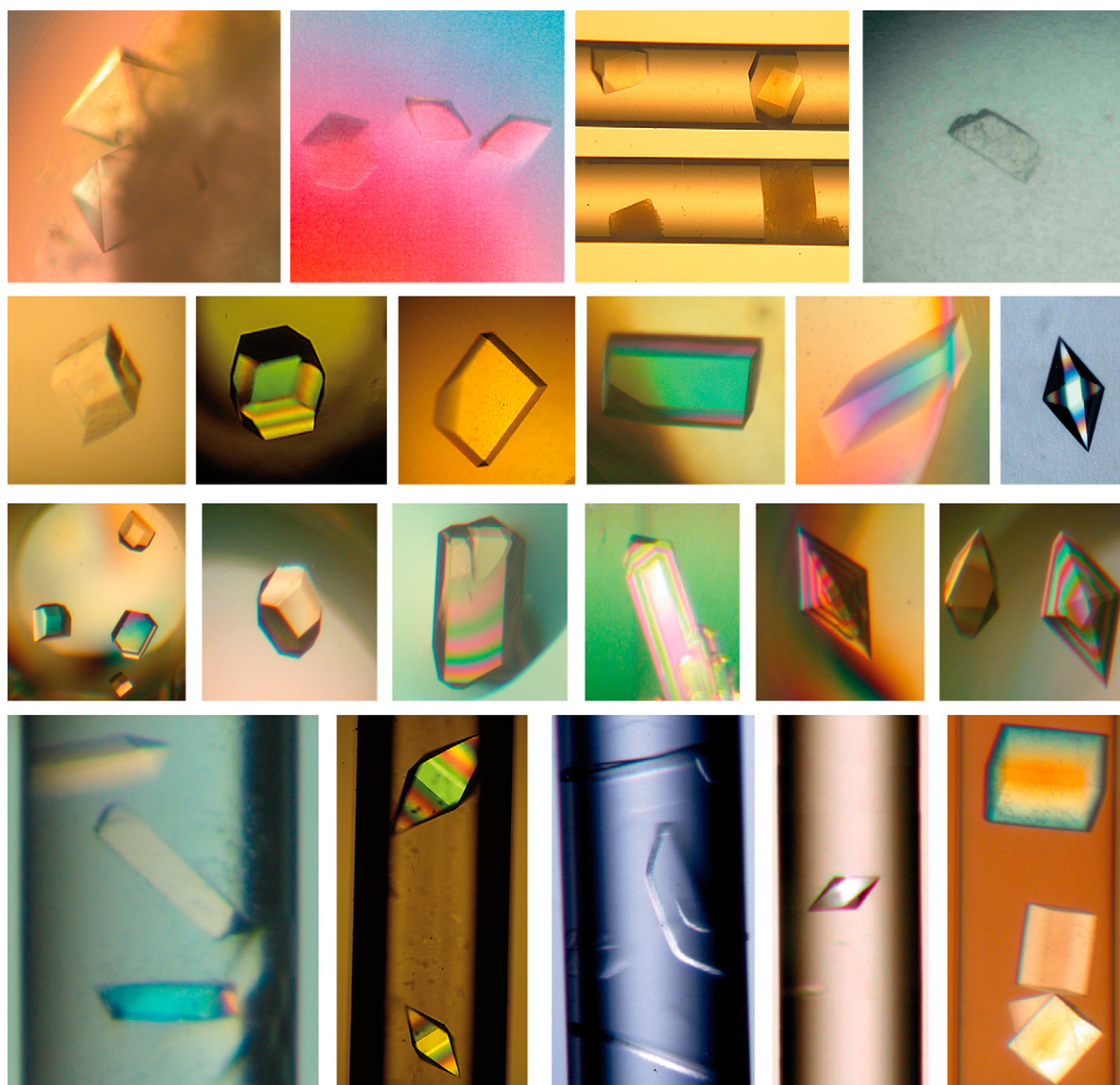


Fig. 5. Virus, protein and nucleoprotein complex crystals grown in gels. (From top to bottom, first row) Viruses in silicate or agarose gel. (From left to right) Brome mosaic virus at 55 mg/ml in 11% (m/v) PEG-8000 and 0.1 M sodium acetate pH 4.5; Turnip yellow mosaic virus at 8.5 mg/ml in 0.6 M ammonium phosphate with 0.1 M 2-(N-morpholino) ethanesulfonic acid pH 3.9 (controls without gel produce crystals with similar habits); Tomato bushy stunt virus at 25 mg/ml grown inside capillaries containing (top) solution without gel and (bottom) 0.2% (m/v) agarose with $T_g = 28^\circ\text{C}$; Grapevine fanleaf virus at 3 mg/ml in 0.2% (m/v) agarose showing a granular structure. All crystals measure 150–300 microns across. (Second row) Proteins in agarose gel. (From left to right) Cubic concanavalin crystal prepared in batch; Tetragonal hen lysozyme crystal prepared by dialysis; Monoclinic plates (in polyethylene glycol) and orthorhombic prisms (in sodium formate) of *Thermus thermophilus* aspartyl-tRNA synthetase-1 (a dimer of M_r 130 000) prepared by vapour diffusion; Orthorhombic crystal of *Deinococcus radiodurans* glutaminyl-tRNA synthetase prepared by batch under paraffin oil in PEG-3350 solution; tetragonal dipyramide of sweet-tasting thaumatin grown by free-interface diffusion. Greatest crystals reach about 600 microns in length. (Third row) Protein crystallisation in novel organic gels. (From left to right) Tetragonal hen lysozyme in sodium chloride solution containing either 2% (m/v) methyl cellulose Metolose® 90SH30000 or 3% (m/v) hydroxypropyl cellulose (M_r 10^5); Hexagonal turkey lysozyme prisms in a solution of the same crystallising agent containing either 2% (m/v) hydroxyethyl cellulose (M_r 1.3×10^6) or 2% (m/v) methyl cellulose Metolose® 90SH15000; Thaumatin dipyramides in sodium tartrate solution containing either 5% (m/v) hydroxyethyl cellulose (M_r 250 000) or 0.5% (m/v) hydroxypropylmethyl cellulose. All crystals were prepared in Cryschem® plates using 10 μl sitting drops equilibrated against 500 μl reservoirs. Hen lysozyme was at 30 mg/ml, turkey lysozyme at 15 mg/ml and thaumatin at 60 mg/ml. The crystals are up to 2 mm long. Metolose® is a registered mark of Shin Etsu Chemical Co., Ltd. (Last row) Protein crystallisation in 0.2% (m/v) agarose gel inside glass capillaries. (From left to right) Prisms of *D. radiodurans* glutaminyl-tRNA synthetase grown in batch; Crystals of T4 endonuclease VII in complex with a DNA cruciform junction grown in the presence of agarose and PEG 6000 (Biertümpfel et al., 2002); Tetragonal thaumatin dipyramides grown by diffusing sodium tartrate through the gel containing the protein; Crystal of Sm1 protein from *Aeropyrum pernix* complexed with a 14-mer RNA in PEG 5000-MME and agarose; Orthorhombic crystals of *T. thermophilus* aspartyl-tRNA synthetase-1 grown by gel acupuncture method (for details see Moreno et al., 2005).

high resolution. Those of the collagen-like peptide (PPG₁₀) prepared on Earth or under microgravity exhibit a high degree of surface roughness (Berisio et al., 2002) and those of the tomato bushy stunt and grapevine fanleaf viruses have a granular structure (Fig. 5). The first diffract X-rays to 1.45 Å (Berisio et al., 2002) and the second (Lorber et al., 1999a) and third (Schellenberger et al., to be published) to at least 3 Å.

5. Crystal quality enhancement

Over the years it has been found that crystallisation in a gel-like agarose may have several benefits. Table 4 lists the advantages which can be expected from the usage of gels during the optimisation of crystallisation conditions. The suppression of convection and the immobilisation of nuclei are favourable to a growth in three

Table 3
Examples of protein and nucleic acid structures derived from crystals grown in agarose gel.

Macromolecule or complex and mass	Crystallising agent ^a	Agarose (% m/v)	Res. (Å)	T (K)	Reference
Protein					
Mitochondrial cytochrom <i>b</i> - <i>c</i> ₁ , 236 000	PEG-4000	0.1–0.4	4.5	n.g.	Yu et al., 1994
T-cell receptor β -chain – superantigen, 40 000	PEG-1000	0.025	3.5	278	Fields et al., 1996
tRNA-guanine transglycosylase, 43 000	PEG-8000	0.2	1.85	288	Romier et al., 1996
Cholera toxin receptor, 14 000	PEG-1000 (PSac)	0.2	2.3	295	Merritt et al., 1997
ι -Asparaginase II, 38 000	MPD	1.0	2.3	295	Kozak and Jaskolski, 2000
Aspartyl-tRNA synthetase-1 (monoclinic), 130 000	PEG-8000	0.1	2.65	100	Zhu et al., 2001
Thaumatococcus, 22 400	Sodium tartrate	0.15	1.2	293	Sauter et al., 2002
Globular head of complement protein C1q, 50 000	PEG-4000 (NDSB)	0.2–0.4	1.9	n.g.	Gaboriaud et al., 2003
3-Dehydroquinase dehydratase, 16 000	Ammonium sulphate	0.1	1.7	100	Maes et al., 2004
Archeal Hjc resolvase, 15 000	PEG-1000	0.3	2.7	100	Biertümpfel et al., 2005
Bacterial metalloprotease FtsH, 51 000	PEG-400	0.1–0.2	2.8	110	Bieniossek et al., 2006
Clam hemoglobin II, 17 000	Ammonium sulphate	0.1	2.0	100	Gavira et al., 2006
Yeast ornithine acetyltransferase, 48 000	PEG-400	0.1	2.8	100	Maes et al., 2006
<i>Sinorhizobium</i> dihydropyrimidinase, 200 000	Sodium formate	0.1	1.85	100	Martinez-Rodriguez et al., 2006
RNA					
Bacterial ribonuclease P RNA, 130 00	MPD	0.05	3.3	100	Kazantsev et al., 2005

Protein structures are listed in chronological order.

Abbreviations: NDSB, non detergent sulfobetaine; n.g.: not given; PSac, pentasaccharide; Res.: resolution; T: temperature.

^a Peculiar additives are indicated between parentheses.

dimensions yielding crystals with fully developed faces and greater volumes (see Fig. 5 for some examples). The polymer chains filling randomly the solvent channels of the crystal make the protein lattice less sensitive to mechanical and osmotic shocks and also to thermal fluctuations or freezing. The quiescent gel medium is an optimal environment for defect-free nucleation and crystal growth. As in the case of small molecules, nucleation is better controlled and impurities have less deleterious effects.

Recent in situ observations performed by laser confocal interference contrast microscopy led to the conclusion that the polysaccharide has almost no effect when the protein solution is pure but it plays the role of a filter when the solution contains large impurities (van Driessche et al., 2008). This is well illustrated by the two following results. (i) In the presence of 0.3% m/v agarose, clear impurity depletion zones are visible around crystals of colourless apoferritin (M_r 450 000) crystallized in a solution contaminated with red-coloured holoferritin dimers (M_r 900 000) (Chernov et al., 2001). (ii) MALDI-TOF and electrospray analyses (performed

according to established procedures, see Potier et al., 2000) detect no significant difference in composition between thaumatin I crystals grown in solution and those grown in 0.2% (m/v) agarose gel when thaumatin I (M_r 22 204) contains traces of thaumatin II (having 98% sequence identity and the same isoelectric point *pI* 8.4 but differing in M_r by only 84 or $\sim 0.4\%$) (Lorber, Potier and Sanglier, unpublished results). There must be a minimum of structural differences between related macromolecules so that the discrimination can occur during the crystallisation process. In other words, the impurity filtering effect operates when structural dissimilarities are strong enough. The same conclusion was drawn from comparative analyses of thaumatin crystals grown in gel either under normal gravity or under microgravity for which chromatography, electrophoresis, N-terminal sequencing and mass spectrometry analyses were also unable to show a difference in macromolecular content (Lorber and Giegé, 2001). In the case of hen and turkey lysozymes which differ by 7 amino acids (M_r 14 313 and 14 209, respectively), by a net charge of 1, and consequently by solubility

Table 4
Some non-exclusive advantages of crystallisation in gel.

Effect of gel ^a	Example of macromolecule, M_r	Reference
3D growth in agarose gel results in crystals with greater volume	tRNA-G transglycosylase, 42 700 Endonuclease-oligoDNA, 49 000	Romier et al., 1996 Biertümpfel et al., 2002
Reinforced crystal lattice in silica gel	Hen lysozyme, 14 500	García-Ruiz et al., 1998a
Protection of crystals in agarose gel during:		
- Handling and irradiation	Thaumatococcus, 22 400	Sauter et al., 2002
- Osmotic stress when soaking with ligands	Aspartyl-tRNA synthetase-1, 130 000	Sauter et al., 2009
- Cryocooling	Aspartyl-tRNA synthetase-1, 130 000 Hjc resolvase, 15 400	Zhu et al., 2001 Biertümpfel et al., 2005
Reproducible diffraction quality in agarose gel	Grapevine fanleaf virus, 6×10^6	Schellenberger et al., to be published
Sharper and more intense reflections when crystals are grown in silica or agarose gels	Hen lysozyme, 14 500 Hen lysozyme, 14 500 Human serum albumin, 66 000	Dong et al., 1999; Lorber et al., 1999a Vidal et al., 1999; Lorber et al., 1999a Miller et al., 1992
Nucleation favored or decreased	Aspartyl-tRNA synthetase-1, 130 000. Dehydroquinase dehydratase, 205 000	Lorber et al., 1999a; Moreno et al., 2005 Maes et al., 2004
Attenuation of impurities effects	Hen lysozyme, 14 500 Several proteins, virus & tRNA Hen lysozyme, 14 500 Turkey lysozyme, 14 500	Robert et al., 1994 Cudney et al., 1994 Vidal et al., 1998a,b Hirschler et al., 1995
Suppression of twinning	Hen lysozyme, 14 500	Provost and Robert, 1995, Hirschler et al., 1995 Hirschler and Fontecilla-Camps, 1996.
Improvement counter-diffusion setup	Alcohol dehydrogenase, 74 000 Insulin, 5740	Sica et al., 1994 Gavira and Garcia-Ruiz, 2002

^a Agarose and polysiloxane concentrations were 0.1–0.5% (m/v) and 0.5–2% (m/v), respectively.

(Provost and Robert, 1995), positive effects were observed. In the case of iron-deprived apoferritin (M_r 476 000) contaminated by dimers of holoferritin containing ~ 4500 iron atoms per particle (M_r 1.44×10^6) (Chernov et al., 2001), the effect was amplified.

Most gel-grown crystals have superior diffraction properties with respect to crystals grown in the absence of gel (see Table 3). Their quality is more reproducible, their diffraction patterns exhibit more intense and sharper reflections. Plot of the average intensity-to-noise I/σ ratio as a function of the resolution for human serum albumin crystals grown in 1.5–0.3% (m/v) agarose gels showed for the first time the real improvement of crystal quality (Miller et al., 1992). Later this was confirmed with thaumatin (Lorber et al., 1999b). Further, the comparison of lysozyme crystals at a resolution of 1.8 Å indicated that laboratory gel-grown ones are close to microgravity-grown ones with regard to their superior diffraction properties (including more intense and sharper reflections) and the number of observable ordered water molecules which is greater than for crystals prepared in solution in the laboratory (Dong et al., 1999).

The enhanced diffraction properties of crystals of aspartyl-tRNA synthetase-1 support the observations made on small- M_r lysozyme, thaumatin and albumin (Moreno et al., 2005). The quality of the crystals of this large size and dimeric enzyme (M_r 130 000) with a multi-domain architecture also benefits from the presence of agarose gel (Fig. 6). Further, the diffraction limit of grapevine fanleaf virus crystals could be extended beyond 6 Å by adding 0.2% (m/v) low-gelling temperature agarose to the crystallisation droplets (Fig. 7).

6. Perspectives for crystallisation in gels

The choice of gels usable for crystallisation is not at all limited to silica and agarose. Two examples of other suitable gel-like media are given below. On the other hand, in addition to silica and agarose gels, convection is also minimal inside capillary tubes. As described below, gels can be used in combination with them.

6.1. Lipidic phases

Mixtures of water and certain lipids organise in complex phases under ambient conditions. These phases were extensively studied as models for the membranes of living organisms (see e.g. Luzzati and Tardieu, 1974). Crystallographic analyses revealed that they are continuous systems made of two compartments (e.g. Mariani et al., 1988; Lindblom and Rilfors, 1989; Luzzati et al., 1993). Later, it was

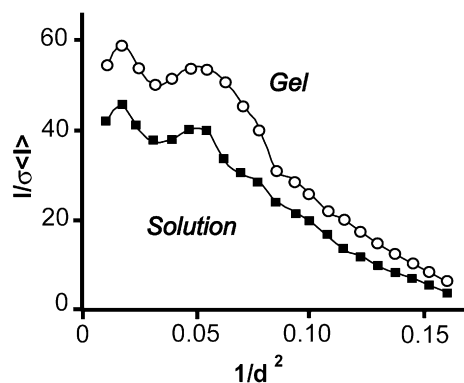


Fig. 6. I/σ vs $1/d^2$ plots for solution- and gel-grown *Thermus thermophilus* aspartyl-tRNA synthetase-1 crystals. The diffraction intensities of orthorhombic crystals with similar dimensions and volumes are compared. Data adapted from Moreno et al. (2005). A similar result had been reported for human serum albumin (Miller et al., 1992) and thaumatin (Lorber et al., 1999b).

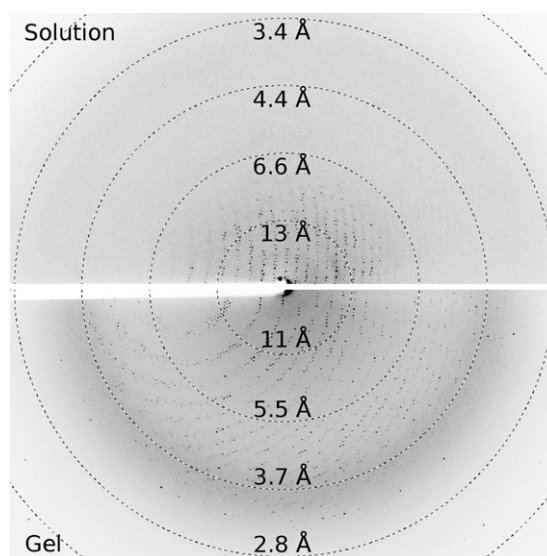


Fig. 7. Gel and cryocooling. Diffraction patterns after cryocooling of grapevine fanleaf virus crystals grown (top half) in the absence and (bottom half) in the presence of 0.2% (m/v) agarose. X-ray diffraction analyses were performed on crystals of comparable volumes. The wavelength was the same ($\lambda = 0.98$ Å) but the crystal-to-detector distance was 250 and 300 mm, respectively.

found that cubic phases formed by the lipids 1-monooleyl-*rac*-glycerol (monoolein, Fig. 8) or 1-monopalmitoleyl-*rac*-glycerol are suitable as matrices for the 3D crystallisation of membrane proteins (Landau and Rosenbusch, 1996; Rummel et al., 1998; for a review of recent developments see Johansson et al., 2009). Indeed, the lipidic component of the mesophase interacts with the hydrophobic domains of membrane proteins and so crystallisation can take place

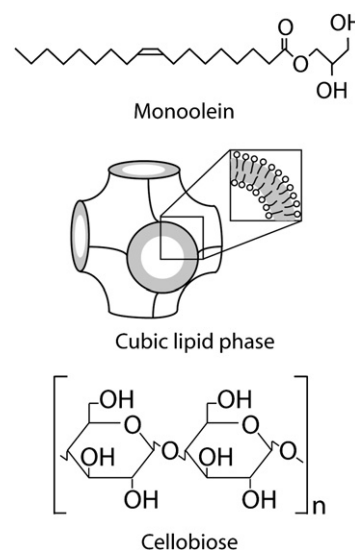


Fig. 8. (Top) Structure of 1-monooleyl-*rac*-glycerol (monoolein). (Centre) Schematic representation of the elementary building unit of a cubic mesophase with a close-up view of the lipid bilayer (Drawing adapted from Lindblom and Rilfors, 1989 and from Landau and Rosenbusch, 1996.) These units do not exist as such in solution. They form a periodical 3D network of nanochannels in which they are located at the nodes of a body-centered cubic lattice (i.e. 1 unit is placed at each of the 8 corners of a cube and one in its centre). The complete system is composed of a continuous hydrophobic phase (composed of the hydrocarbon chains of the lipid, shown in grey colour) and a continuous hydrophilic phase (composed of the polar heads of the lipid shown in white colour) running side-by-side. (Bottom) Cellobiose is made of two D-glucopyranose molecules linked by a β 1–4 bond. In cellulose compounds, the number of cellobiose units varies from a few hundred to several thousands. The hydrogen atoms of free hydroxyl groups can be substituted chemically by methyl, ethyl or propyl groups.

in the absence of detergent (Nollert et al., 1999). The preparation of lipidic mesophases was gradually standardized (see e.g. Caffrey, 2003; Cherezov and Caffrey, 2005). A detailed protocol has been published recently (Caffrey and Cherezov, 2009). Successful crystallisations are listed in the Membrane Protein Data Bank (Raman et al., 2006). Membrane protein crystals can be harvested with the help of microtools (see Caffrey, 2009). They may be recovered from the lipid matrix by enzymatic hydrolysis with *Candida rugosa* lipase (Nollert and Landau, 1998). Solubilization with a detergent like octyl- β -D-glucoside is an alternative.

Interestingly, the aqueous compartment of the cubic phase can accommodate water-soluble proteins. It can be employed as a gel to crystallise non-membrane proteins (Landau et al., 1997; Cherezov and Caffrey, 2003). In this case crystallisation is independent of the type of phase and the packing arrangement of the lipidic compartment. As with agarose the phases are not stable in the presence of all crystallising agents (for results of compatibility tests, see e.g. Landau et al., 1997; Cherezov et al., 2001). From the point of view of the Grashof number, the nanochannels composing the lipidic phases are expected to be even a better diffusive medium than the mesh of agarose gels since their dimensions are smaller (i.e. in the range of the size of protein or not more than 100 nm depending on the degree of swelling).

6.2. Cellulose derivatives

Besides polysiloxane and agarose, a number of chemicals and biochemicals form gels or gel-like media when they are dissolved in water. A priori all those which are not transparent (like carboxyvinyl polymers), not neutral (like polyacrylic and polygalacturonic acids) or free of ions (like alginates and carrageenans) are not suitable for protein crystallisation. This holds for substances of ill-defined composition (such as gums from higher plants or algae, pectins from fruit, chitosans extracted from crustacean carapace, or gelatines from animal skin and bones which may be contaminated by proteins like proteases or nucleases).

Amongst the chemically neutral and biochemically pure ones are cellulose derivatives. Their chains encompassing hundreds to thousands cellobiose units associate in bundles via hydrogen bonds (Fig. 8). An aqueous solution of these fibres (which are either microcrystalline or non-crystalline) behaves like a gel (see e.g. Fortin and Charlet, 1989; Kalinina et al., 2001; Saito et al., 2003; Kondo et al., 2004; Silva et al., 2008). Over a decade ago, carboxymethyl cellulose was already identified as a potential crystallising agent for proteins, nucleic acids and viruses (Patel et al., 1995).

We have used following cellulose derivatives to grow protein crystals at room temperature. For instance, methyl cellulose (at 1% m/v), hydroxyethyl cellulose (with M_r 250 000– 1.3×10^6 at respectively 5% and 1% m/v), hydroxypropyl cellulose (M_r 10^5 – 10^6 at 3–4% and 10% m/v), and hydroxypropyl methyl cellulose (M_r 10^5 at 0.5–1% m/v) dissolved in water form gels which are moderately viscous, transparent and stable over a broad range of temperatures and crystallising agents. These compounds are fully compatible with the crystallisation of soluble model proteins. Tetragonal hen egg-white lysozyme and thaumatin crystals grown at 20 °C in the presence of methyl-, hydroxyl propyl-, hydroxyethyl- and hydroxypropylmethyl cellulose diffract X-ray at least to the same resolution and with the same apparent mosaicity as control crystals (Lorber, unpublished results). The same holds for 2 mm-long hexagonal prisms of turkey lysozyme prepared in hydroxyethyl cellulose and for thaumatin dipyramides grown in 5% m/v hydroxyethyl cellulose (M_r 250 000) or in 10% m/v hydroxypropyl cellulose (M_r 10^6) (Fig. 5). These examples demonstrate that the panel of diffusive media can be extended beyond silica and agarose

gels. Further investigations will tell if these novel gels can replace them and have other hidden properties.

6.3. Gels and capillary tubes

Some crystal growers have used various gels as plugs or dialysis membranes (also sometimes called a physical buffer) to separate the macromolecular solution from the crystallising agent solution (Zepezauer, 1971; Littke and Johns, 1984). Early devices consisting essentially of small-diameter (1 mm or less) glass tubes had the purpose to consume less sample volume. Delicate free-interface diffusion crystallisation experiments (Salemme, 1972) were performed inside very thin tubes. There, capillary forces counter-balance the buoyant forces and suppress any perturbation at the contact area between protein and crystallising agent solutions which would lead to irreversible mixing. Hence, the effects of gravity are minimized, convection is abolished, mass transport restricted to diffusion and concentration gradients around growing crystals are stabilised (for reviews, see García-Ruiz, 2003; Ng et al., 2003). This environment is close to the one attained under microgravity. Further advantages were then expected by combining gel and a capillary geometry.

Three crystallisation methods are represented in Fig. 9. The diameter of the capillary tubes should not exceed 0.3 mm in order to suppress crystal and precipitant sedimentation under gravity. This effect is not negligible when the diameter is superior (see e.g. Moreno et al., 1999b). The first and most simple is batch

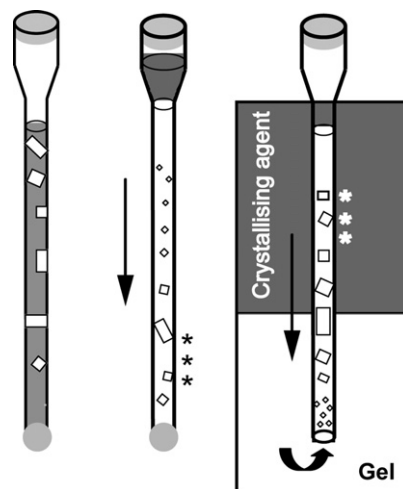


Fig. 9. Crystallisation in transparent capillary tubes. (Left) In batch crystallisation the macromolecule and the crystallising agent are mixed with the gel and transferred inside the capillary. When supersaturation is high enough, the initial mixture may become opaque. At places where crystals grow it will then gradually clear off and a halo will become visible around each crystal. (Centre) Macromolecule (protein or virus) is immobilised in the gel. When the crystallising agent is brought into contact at the top, it diffuses slowly inside the gel. Crystallisation occurs by counter-diffusion. The tomato bushy stunt virus crystals shown in Fig. 6 were prepared in this way in agarose gel inside capillaries. (Right) In the gel acupuncture method (GAME) an open capillary containing the macromolecule (in solution or in gel) is punched into a thick layer of silica or agarose gel. When crystallising agent solution is placed onto the gel, its molecules diffuse into the gel and after some time enter very slowly inside the capillary. As in the centre of the figure, a wave of higher supersaturation is then generated which propagates inside the capillary. This system is reminiscent of a micro-diffusion cell consisting of an X-ray capillary closed on one end by a plug of polyacrylamide gel (for a drawing, see Zepezauer, 1971). To take advantage of capillarity effects, experiments with protein solution are performed in tubes with an inner diameter of 0.3 or less. Greater diameters (up to 1 or 2 mm) may be used with protein solution containing a low concentration of gel. Arrows indicate the direction of the diffusion of crystallising agent molecules. Stars mark the approximate areas where crystals of highest quality are expected to grow.

crystallisation. Capillary tubes were useful to study the effects of physical parameters such as electric (Mirkin et al., 2003) or magnetic fields (Moreno et al., 2007; Gavira and Garcia-Ruiz, 2009; Moreno, Sauter, Giegé and Lorber, unpublished results).

The second method is based on counter-diffusion: the macromolecule mixed with the gel is loaded in the capillary and the crystallising agent brought in contact with the gel diffuses through it to trigger nucleation and crystal growth (García-Ruiz, 2003). In the “gel acupuncture” method (GAME, García-Ruiz and Moreno, 1994, 1997; García-Ruiz et al., 1998b), crystallisation also proceeds very slowly by counter-diffusion. In both cases (Fig. 9), the diffusion of the crystallising agent across the macromolecular solution generates a wave of local supersaturation that displaces along the capillary (García-Ruiz et al., 2001b). The latter creates a self-optimization process (García-Ruiz et al., 2001b) which can be further improved by adding a layer of silicone oil. This fluid slows down the diffusion and thus enhances crystal growth (Moreno et al., 2002). When enough protein is available, crystals can fill completely the tube and take a cylindrical shape (García-Ruiz et al., 1995, Otalora et al., 1996; Moreno et al., 1996). The theory of counter-diffusion was established and the kinetics of the process was simulated (Otalora and García-Ruiz, 1997; García-Ruiz and Moreno, 1997; García-Ruiz et al., 1998b; Carotenuto et al., 2002).

Capillary tubes like X-ray glass capillaries or tubes made of synthetic polymers (Potter et al., 2004) are convenient sample container. A special box was designed to contain several experiments (García-Ruiz et al., 2002). When agarose is present in the protein solution, the crystals can be extracted by displacing it with either air or grease. (A more sophisticated method uses femto-second laser ablation to detach crystals nucleated onto glass walls, see Kashii et al., 2006). The excess of gel is then removed as explained in Table 2. Another advantage of the method is that crystals can be soaked in situ with a cryoprotectant before freezing (Lopez-Jaramillo et al., 2001). Heavy atoms can be diffused into them for structure determination (Gavira et al., 2002) and other small molecules to act on crystal composition (Domínguez-Vera et al., 1996). Thus, crystal quality can be evaluated without manual handling (see e.g. Sugahara et al., 2009).

7. General conclusion

In retrospect the potential of gels has been so far largely underexploited by protein and nucleic acid crystal growers in spite of a body of experimental results which demonstrate that crystallisation in such diffusive media may be more advantageous than crystallisation in pure solution. It is no longer merely our claim that gels may be advantageous for the crystal growth of biological macromolecules. Indeed, the 3D structures of a wide variety of macromolecules derived from crystals prepared in gels are as many pieces of evidence that the technique is applicable not only to proteins but also to nucleic acids (RNA, DNA) as well as nucleoprotein complexes and viruses. Further, the crystals grown in gel may diffract X-rays more strongly and have a less mosaic structure, which in turn facilitates the crystallographic analysis. In addition, gel-grown crystals may be more stable against temperature fluctuations, mechanical shocks, or differences in osmotic pressure occurring upon the soaking with a ligand, inhibitor, cryoprotectant or heavy atom. Their reinforced lattice is more resistant to cryocooling or even to diffraction data collection at room temperature. A supplementary benefit is that all crystallisation methods can accommodate a gel without perturbing the equilibration pathway. For instance, a little amount of agarose can be added in batch, vapour diffusion, dialysis or free-interface diffusion experiments. Recently, agarose was even successfully used inside microfluidic channels (Dhouib et al., 2009). Besides, the combination of gel and

counter-diffusion experiments in capillaries deserves more attention since it offers supplementary possibilities. Indeed, it enables the experimenter to optimize crystal growth under quasi-ideal conditions. We sincerely wish this review will help revive the crystallisation of biological macromolecules in gels and in capillaries and lead to an explosion of the number of well-resolved macromolecular structures.

Acknowledgements

This article is dedicated to the pioneers J. Lewin, B.W. Low, F.M. Richards and J. Berthou. The authors thank So.Bi.Gel (Hendaye, France), SEPPIC (Paris), L'Oréal (Chevilly-Larue, France) and Hampton Research (USA) for generous gifts of gels. They have appreciated discussions with B. Tinland (ICS, Strasbourg) and P. Werlé (Faculté de Pharmacie, Strasbourg). They are grateful to A. Van Dorselaer (LSMBO, Strasbourg) for his interest in crystal content analysis, to G. Demangeat, O. Lemaire (INRA, Colmar) and C. Ritzenthaler (IBMP, Strasbourg) for sharing the GFLV crystallisation project and to F. Stanger and B. Zimmer for their help with vapour-diffusion experiments on gel-containing drops. The authors acknowledge the beam time allocated to the mentioned crystallisation projects by the European Synchrotron Radiation Facility (Grenoble, France) and the Swiss Light Source (Villigen, Switzerland). They thank the French CNES, ESA, CNRS and Université de Strasbourg for financial support. A. Moreno acknowledges the support for a sabbatical year in the University of Cambridge by CONACYT project 91999 and DGAPA-UNAM for supporting travel expenses.

References

- Arnott, S., Fulmer, A., Scott, W.E., Dea, I.C.M., Moorhouse, R., Rees, D.A., 1974. The agarose double helix and its function in agarose gel structure. *J. Mol. Biol.* 90, 269–284.
- Berisio, R., Vitagliano, L., Vergara, A., Sorrentino, G., Mazzarella, L., Zagari, A., 2002. Crystallisation of the collagen-like polypeptide (PPG10) aboard the International Space Station. 2. Comparison of crystal quality by X-ray diffraction. *Acta Crystallogr. D* 58, 1695–1699.
- Bernard, Y., Degoy, S., Lefauchaux, F., Robert, M.-C., 1994. A gel-mediated feeding technique for protein crystals growth from hanging drops. *Acta Crystallogr. D* 50, 504–507.
- Bieniossek, C., Schalch, T., Bumann, M., Meister, M., Meier, R., Baumann, U., 2006. The molecular architecture of the metalloprotease FtsH. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3066–3071.
- Biertümpfel, C., Basquin, J., Suck, D., Sauter, C., 2002. Crystallisation of biological macromolecules using agarose gel. *Acta Crystallogr. D* 58, 1657–1659.
- Biertümpfel, C., Basquin, J., Birkenbihl, R.P., Suck, D., Sauter, C., 2005. Characterization of crystals of the Hjc resolvase from *Archeoglobus fulgidus* grown in gel by counter-diffusion. *Acta Crystallogr. F* 61, 684–687.
- Blundell, T.L., Johnson, L.N., 1976. *Protein Crystallography*. Academic Press New York, pp. 1–565.
- Bonneté, F., Vidal, O., Robert, M.-C., Tardieu, A., 1996. Gel techniques and small angle X-ray scattering to follow protein crystal growth. *J. Crystal Growth* 168, 185–191.
- Brinker, C.J., Scherrer, G.W., 1990. “Sol–Gel Science”: The Physics and Chemistry of Sol–Gel Processing. Academic Press, San Diego, USA.
- Caffrey, M., 2003. Membrane protein crystallisation. *J. Struct. Biol.* 142, 108–132.
- Caffrey, M., 2009. Crystallizing membrane proteins for structure determination: use of lipidic mesophases. *Ann. Rev. Biophys.* 38, 29–51.
- Caffrey, M., Chervov, V., 2009. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* 4, 706–731.
- Carotenuto, L., Piccolo, C., Castagnolo, D., Lappa, M., Tortora, A., García-Ruiz, J.M., 2002. Experimental observations and numerical modelling of diffusion-driven crystallisation processes. *Acta Crystallogr. D* 58, 1628–1632.
- Charron, C., Robert, M.-C., Capelle, B., Kadri, A., Jenner, G., Giegé, R., Lorber, B., 2002. X-ray diffraction properties of protein crystals prepared in agarose gel under hydrostatic pressure. *J. Crystal Growth* 245, 321–333.
- Chayen, N.E., Saridakis, E., 2002. Protein crystallisation for genomics: towards high-throughput optimization techniques. *Acta Crystallogr. D* 58, 921–927.
- Chen, P.S., Schlichta, P.J., Wilcox, W.R., Lefever, R.A., 1979. Convection phenomena during the growth of sodium chlorate crystals from solution. *J. Crystal Growth* 47, 43–60.

- Cherezov, V., Caffrey, M., 2003. Nano-volume plates with excellent optical properties for fast, inexpensive crystallisation screening of membrane proteins. *J. Appl. Crystallogr.* 36, 1372–1377.
- Cherezov, V., Caffrey, M., 2005. A simple and inexpensive nanoliter-volume dispenser for highly viscous materials used in membrane protein crystallisation. *J. Appl. Crystallogr.* 38, 398–400.
- Cherezov, V., Fersi, H., Caffrey, M., 2001. Crystallisation screens: compatibility with the lipidic cubic phase for in meso crystallisation of membrane proteins. *Biophys. J.* 81, 225–242.
- Chernov, A.A., 2003. Protein crystals and their growth. *J. Struct. Biol.* 142, 3–21.
- Chernov, A.A., Garcia-Ruiz, J.M., Thomas, B.R., 2001. Visualization of the impurity depletion zone surrounding apoferritin crystals growing in gel with holoferritin dimer impurity. *J. Crystal Growth* 232, 184–187.
- Cudney, R., Patel, S., McPherson, A., 1994. Crystallisation of macromolecules in silica gels. *Acta Crystallogr. D* 50, 479–483.
- Dhouib, K., Khan-Malek, C., Pflieger, W., Gauthier-Manuel, B., Duffait, R., Thuillier, R., Ferrigno, R., Jacquamet, L., Ohana, J., Ferrer, J.-L., Théobald-Dietrich, A., Giegé, R., Lorber, B., Sauter, C., 2009. Microfluidic chips for the crystallisation of biomacromolecules by counter-diffusion and on-chip crystal X-ray analysis. *Lab-on-the-Chip* 9, 1412–1421.
- Domínguez-Vera, J.M., Rondón, D., Moreno, A., García-Ruiz, J.M., 1996. A preliminary study on the interaction of ferritin single crystals with chelating agents. *J. Crystal Growth* 168, 138–141.
- Dong, J., Boggon, T.J., Chayen, N.E., Raftery, J., Bi, R.-C., Helliwell, J.R., 1999. Bound-solvent structures for microgravity-, ground control-, gel- and microbatch-grown hen egg-white lysozyme crystals at 1.8 Å resolution. *Acta Crystallogr. D* 55, 745–752.
- van Driessche, A.E.S., Otalora, F., Gavira, J.A., Sasaki, G., 2008. Is agarose an impurity or an impurity filter? In situ observation of the joint gel/impurity effect on protein crystals growth kinetics. *Crystal Growth Des.* 8, 3623–3628.
- Fields, B.A., Malchiodi, E.L., Li, H., Ysern, X., Stauffacher, C.V., Schlievert, P.M., Karjalainen, K., Mariuzza, R.A., 1996. Crystal structure of a T-cell receptor β -chain complexed with a superantigen. *Nature* 384, 188–192.
- Finet, S., Bonneté, F., Frouin, J., Provost, K., Tardieu, A., 1998. Lysozyme crystal growth, as observed by small angle X-ray scattering, proceeds without crystallisation intermediates. *Eur. Biophys. J.* 27, 263–271.
- Fortin, S., Charlet, G., 1989. Phase diagram of aqueous solutions of (hydroxypropyl) cellulose. *Macromolecules* 22, 2286–2292.
- Gaboriaud, C., Juanhuix, J., Gruez, A., Lacroix, M., Darnault, C., Pignol, D., Verger, D., Fontecilla-Camps, J.C., Arlaud, G.J., 2003. The crystal structure of the globular head of complement protein C1q provides a basis for its versatile recognition properties. *J. Biol. Chem.* 278, 46974–46982.
- García-Ruiz, J.M., 1991. The uses of crystal growth in gels and other diffusing-reacting systems. *Key Eng. Mater.* 58, 87–106.
- García-Ruiz, J.M., 2003. Counter-diffusion methods for macromolecular crystallisation. *Methods Enzymol.* 368, 130–154.
- García-Ruiz, J.M., Moreno, A., Viedma, A., Coll, M., 1993. Quality of lysozyme crystals grown by the gel acupuncture method. *Mat. Res. Bull.* 28, 541–546.
- García-Ruiz, J.M., Moreno, A., Párraga, A., Coll, M., 1995. Shaped single protein crystals. *Acta Crystallogr. D* 51, 278–281.
- García-Ruiz, J.M., Gavira, J.A., Otalora, F., Guasch, A., Coll, M., 1998a. Reinforced protein crystals. *Mat. Res. Bull.* 33, 1593–1598.
- García-Ruiz, J.M., Moreno, A., Otálora, F., Viedma, C., Rondón, D., Zautscher, F., 1998b. Teaching protein crystallisation by the gel acupuncture method. *J. Chem. Educ.* 75, 442–446.
- García-Ruiz, J.M., Novella, M.L., Otalora, F., 1999. Supersaturation patterns in counter-diffusion crystallisation methods followed by Mach-Zehnder interferometry. *J. Crystal Growth* 196, 703–710.
- García-Ruiz, J.M., Novella, M.L., Moreno, R., Gavira, J.A., 2001a. Agarose as crystallisation media for proteins. I. Transport processes. *J. Crystal Growth* 232, 165–172.
- García-Ruiz, J.M., Otalora, F., Novella, M.L., Gavira, J.A., Sauter, C., Vidal, O., 2001b. A supersaturation wave of protein crystallisation. *J. Crystal Growth* 232, 149–155.
- García-Ruiz, J.M., Gonzales-Ramirez, L.A., Gavira, J.A., Otalora, F., 2002. Granada crystallisation box: a new device for protein crystallisation by counter-diffusion. *Acta Crystallogr. D* 58, 1638–1642.
- García-Ruiz, J.M., Moreno, A., 1994. Investigation on protein crystals grown by the gel acupuncture method. *Acta Crystallogr. D* 50, 484–490.
- García-Ruiz, J.M., Moreno, A., 1997. Growth kinetics of protein single crystals in the gel acupuncture technique. *J. Crystal Growth* 178, 393–401.
- Gavira, J.A., Toh, D., Lopez-Jaramillo, J., Garcia-Ruiz, J.M., Ng, J.D., 2002. Ab initio crystallographic structure determination of insulin from protein to electron density without crystal handling. *Acta Crystallogr. D* 58, 1147–1154.
- Gavira, J.A., De Jesus, W., Camara-Artigas, A., Lopez-Garriga, J., Garcia-Ruiz, J.M., 2006. Capillary crystallisation and molecular replacement solution of haemoglobin II from the clam *Lucina pectinata*. *Acta Crystallogr. F* 62, 196–199.
- Gavira, J.A., Garcia-Ruiz, J.M., 2002. Agarose as crystallisation media for protein: II. Trapping of gel fibers into the crystals. *Acta Crystallogr. D* 58, 1653–1656.
- Gavira, J.A., Garcia-Ruiz, J.M., 2009. Effects of a magnetic field on lysozyme crystal nucleation and growth in a diffusive environment. *Crystal Growth Des.* 9, 2610–2615.
- Gits-Léon, S., Lefaucheur, F., Robert, M.-C., 1987. Mass transport by diffusion in a tetramethoxysilane gel. *J. Crystal Growth* 84, 155–162.
- Gonzales-Ramirez, L.A., Caballero, A.G., Garcia-Ruiz, J.M., 2008. Investigation of the compatibility of gels with precipitating agents and detergent in protein crystallisation experiments. *Crystal Growth Des.* 8, 4291–4296.
- Griess, G.A., Edwards, D.M., Dumais, M., Harris, R.A., Renn, D.W., Serwer, P., 1993. Heterogeneity of the pores of polysaccharide gels: dependence on the molecular weight and derivatization of the polysaccharide. *J. Struct. Biol.* 111, 39–47.
- Guenet, J.M., 1992. Thermoreversible Gelation of Polymers and Biopolymers. Academic Press, London.
- Halberstadt, E.S., Henisch, H.K., 1968. Recent experiments on crystal growth in gels. *J. Crystal Growth* 3–4, 363–366.
- Henisch, H.K., 1973. *Crystal Growth in Gels*, second ed., Pennsylvania State Univ. Press, University Park, Pennsylvania, reprinted in 1996 by Dover Publications, New York.
- Henisch, H.K., 1988. *Crystals in Gels and Liesegang Rings*. Cambridge University Press, Cambridge, pp. 1–197.
- Henisch, H.K., Dennis, J., Hanoka, J.I., 1965. Crystal growth in gels. *J. Phys. Chem. Solids* 26, 493–496.
- Hirschler, J., Charon, M.-H., Fontecilla-Camps, J.C., 1995. The effects of filtration on protein nucleation in different growth media. *Protein Sci.* 4, 2573–2577.
- Hirschler, J., Fontecilla-Camps, J.C., 1996. Contaminant effects on protein crystal morphology in different growth environments. *Acta Crystallogr. D* 52, 806–812.
- Hou, W.B., Kudryavtsev, A.B., Bray, T.L., DeLucas, L.J., Wilson, W.W., 2001. Real time evolution of concentration distribution around tetragonal lysozyme crystal: case study in gel and free solution. *J. Crystal Growth* 232, 265–272.
- Johansson, L.C., Wöhri, A., Katona, G., Engström, S., Neutze, R., 2009. Membrane protein crystallisation from lipidic phases. *Curr. Opin. Struct. Biol.* 19, 372–378.
- Kadri, A., Lorber, B., Jenner, G., Giegé, R., 2002. Effects of pressure on the crystallisation and the solubility of proteins in agarose gel. *J. Crystal Growth* 245, 109–120.
- Kadri, A., Jenner, J., Damak, M., Lorber, B., Giegé, R., 2003a. Crystallogensis studies of proteins in agarose gel – combined effect of high hydrostatic pressure and pH. *J. Crystal Growth* 257, 390–402.
- Kadri, A., Damak, M., Lorber, B., Giegé, R., Jenner, J., 2003b. Pressure versus pH phase diagrams of two lysozymes crystallized in agarose gel. *High Press. Res.* 23, 485–491.
- Kadri, A., Lorber, B., Charron, C., Robert, M.-C., Capelle, B., Damak, M., Jenner, G., Giegé, R., 2005. Crystal quality and differential crystal growth behaviour of three proteins crystallized in gel at high hydrostatic pressure. *Acta Crystallogr. D* 61, 784–788.
- Kalinina, N.A., Boček, A.M., Silinskaya, I.G., Nud'ga, L.A., Petrova, V.A., 2001. Structure of aqueous methyl and propyl methyl cellulose solutions in the initial stage of gelation. *Russ. J. Appl. Chem.* 74, 1568–1572.
- Kashii, M., Fujisawa, R., Adachi, H., Mori, Y., Sasaki, T., Koga, Y., Takano, K., Kanaya, S., Matsumura, H., Inoue, T., Murakami, S., Sugamoto, K., Yoshikawa, H., 2006. Application of femtosecond laser ablation for detaching grown protein crystals from glass capillary tube. *J. Biosci. Bioeng.* 102, 372–374.
- Kazantsev, A.V., Krivenko, A.A., Harrington, D.J., Holbrook, S.R., Adams, P.D., Pace, N.R., 2005. Crystal structure of a bacterial ribonuclease P RNA. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13392–13397.
- Kondo, T., Kasai, W., Brown, R.M., 2004. Formation of nematic ordered cellulose and chitin. *Cellulose* 11, 463–474.
- Kozak, M., Jaskolski, M., 2000. Crystallisation and preliminary crystallographic studies of a new crystal form of *Escherichia coli* L-asparaginase II (Ser58Ala mutant). *Acta Crystallogr. D* 56, 509–511.
- Landau, E.M., Rummel, G., Cowas-Jacob, S.W., Rosenbusch, J.P., 1997. Crystallisation of a polar protein and small molecules from the aqueous compartment of lipidic cubic phases. *J. Phys. Chem. B* 101, 1935–1937.
- Landau, E.M., Rosenbusch, J.P., 1996. Lipidic cubic phases: a novel cubic form for the crystallisation of membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* 93, 14532–14535.
- Lefaucheur, F., Robert, M.-C., Bernard, Y., Gits, S., 1984. Nucleation and growth of some solution grown crystals in gel media: visualisation by interferometric holography. *Crystal Res. Technol.* 19 (12), 1541–1547.
- Lindblom, G., Rilfors, L., 1989. Cubic phases and isotropic structures formed by membrane lipids – possible biological relevance. *Biochim. Biophys. Acta* 988, 221–256.
- Littke, W., Johns, C., 1984. Protein single crystal growth under microgravity. *Science* 225, 203–204.
- Lopez-Jaramillo, F.J., Garcia-Ruiz, J.M., Gavira, J.A., Otalora, F., 2001. Crystallisation and cryocrystallography inside X-ray capillaries. *J. Appl. Crystallogr.* 34, 365–370.
- Lorber, B., Sauter, C., Ng, J.D., Zhu, D.W., Giegé, R., Vidal, O., Robert, M.-C., Capelle, B., 1999a. Characterization of protein and virus crystals by quasi-planar wave X-ray topography: a comparison between crystals grown in solution and in agarose gel. *J. Crystal Growth* 204, 357–368.
- Lorber, B., Sauter, C., Robert, M.-C., Capelle, B., Giegé, R., 1999b. Crystallisation within agarose gel in microgravity improves the quality of thaumatin crystals. *Acta Crystallogr. D* 55, 1491–1494.
- Lorber, B., Ng, J.D., Lautenschlager, P., Giegé, R., 2000. Growth kinetics and motion of thaumatin crystals during USML-2 and LMS microgravity missions and comparison with earth controls. *J. Crystal Growth* 208, 665–677.
- Lorber, B., Giegé, R., 2001. Nucleation and growth of thaumatin crystals within a gel under microgravity on STS-95 mission vs. under Earth's gravity. *J. Crystal Growth* 208, 665–677.
- Low, B.W., Richards, F.M., 1954. Measurement of the density, composition and related unit cell dimensions of some protein crystals. *J. Am. Chem. Soc.* 76, 2511–2518.
- Lübbert, D., Meents, A., Weckert, E., 2004. Accurate rocking-curve measurements on protein crystals grown in a homogeneous magnetic field of 2.4 T. *Acta Crystallogr. D* 60, 987–998.

- Luzzati, V., Varga, R., Mariani, P., Gulik, A., Delacroix, H., 1993. Cubic phases of lipid-containing systems. Elements of a theory and biological connotations. *J. Mol. Biol.* 229, 540–551.
- Luzzati, V., Tardieu, A., 1974. Lipid phases: structure and structural transitions. *Annu. Rev. Phys. Chem.* 25, 79–92.
- Maaloum, M., Pernodet, N., Tinland, B., 1998. Agarose gel structure using atomic force microscopy: gel concentration and ionic strength effects. *Electrophoresis* 19, 1606–1610.
- Maes, D., Gonzales-Ramirez, L.A., Lopez-Jaramillo, J., Yu, B., De Bondt, H., Zegers, I., Afonia, E., Garcia-Ruiz, J.M., Gulnik, S., 2004. Structural study of the type III 3-dehydroquinate dehydratase from *Actinobacillus pleuropneumonia*. *Acta Crystallogr. D* 60, 463–471.
- Maes, D., Crabeel, M., Van de Weerd, C., Martial, J., Peeters, E., Charlier, D., Decannière, K., Vanhee, C., Wyns, L., Zegers, I., 2006. Crystallisation of ornithine acetyltransferase from yeast by counter-diffusion and preliminary X-ray study. *Acta Crystallogr. F* 62, 1294–1297.
- Mariani, P., Luzzati, V., Delacroix, H., 1988. Cubic phases of lipid-containing systems. Structure analysis and biological implications. *J. Mol. Biol.* 204, 165–189.
- Martinez-Rodriguez, S., Gonzales-Ramirez, L.A., Clemente-Jimenes, J.M., Rodriguez-Vico, F., Las Heras-Vazquez, F.J., Gavira, J.A., Garcia-Ruiz, J.M., 2006. Crystallisation and preliminary crystallographic studies of the recombinant dihydropyrimidinase from *Sinorhizobium meliloti* CECT4114. *Acta Crystallogr. F* 62, 1223–1226.
- McHugh, D.J. (Ed.), 1987. Production and Utilization of Products from Commercial Seaweeds. FAO Fisheries Technical Paper 288. Food and Agriculture Organization of the United Nations, Rome.
- Merritt, E.A., Sarfaty, S., Jobling, M.G., Chang, T., Holmes, R.K., Hirst, T.R., Hol, W.G.J., 1997. Structural studies of receptor binding by cholera toxin mutants. *Protein Sci.* 6, 1516–1528.
- Miller, T.Y., He, X.M., Carter, D.C., 1992. A comparison between protein crystals grown with vapor diffusion methods in microgravity and protein crystals using a gel liquid–liquid diffusion ground-based method. *J. Crystal Growth* 122, 306–309.
- Mirkin, N., Frontana-Urbe, B., Rodríguez-Romero, A., Hernández-Santoyo, A., Moreno, A., 2003. The influence of an internal electric field upon protein crystallization using the gel-acupuncture method. *Acta Crystallogr. D* 59, 1533–1538.
- Moreno, A., Rondón, D., García-Ruiz, J.M., 1996. Growth of shaped single crystals of proteins. *J. Crystal Growth* 166, 919–924.
- Moreno, A., Juárez-Martínez, G., Hernández-Pérez, T., Batina, N., Mundo, M., McPherson, A., 1999a. Physical and chemical properties of gels. Applications to protein nucleation control in the gel acupuncture technique. *J. Crystal Growth* 205, 375–381.
- Moreno, A., González-Ramírez, L.A., Hernández Hernández, M.A., Oliver-Salvador, C., Soriano-García, M., Rodríguez-Romero, A., 1999b. Investigation on gravity influence upon protein crystallisation by the gel acupuncture technique. *J. Crystal Growth* 196, 587–594.
- Moreno, A., Saridakis, E., Chayen, N.E., 2002. Combination of oils and gels for enhancing the growth of protein crystals. *J. Appl. Crystallogr.* 35, 140–142.
- Moreno, A., Théobald-Dietrich, A., Lorber, B., Sauter, C., Giegé, R., 2005. Effects of macromolecular impurities and of crystallisation method on the quality of eubacterial aspartyl-tRNA synthetase-1 crystals. *Acta Crystallogr. D* 61, 789–792.
- Moreno, A., Quiroz-García, B., Yokaichiya, F., Stojanoff, V., Rudolph, P., 2007. Protein crystal growth in gels and stationary magnetic fields. *Crystal Res. Technol.* 42, 231–236.
- Ng, J.D., Gavira, J.A., Garcia-Ruiz, J.M., 2003. Protein crystallisation by capillary counter-diffusion for applied crystallography structure determination. *J. Struct. Biol.* 142, 218–231.
- Nollert, P., Royant, A., Pebay-Peroula, E., Laudau, E.M., 1999. Detergent-free membrane protein crystallisation. *FEBS Lett.* 457, 205–208.
- Nollert, P., Landau, E.M., 1998. Enzymic release of crystals from lipidic cubic phases. *Biochem. Soc. Trans.* 26, 709–713.
- Normand, V., Lootens, D.L., Amici, E., Plucknett, K.P., Aymard, P., 2000. New insight into agarose gel mechanical properties. *Biomacromolecules* 1, 730–738.
- Otalora, F., Garcia-Ruiz, J.M., Moreno, A., 1996. Protein crystal quality using rod-shape crystals. *J. Crystal Growth* 168, 93–98.
- Otalora, F., Novella, M.L., Gavira, J.A., Thomas, B.R., Garcia Ruiz, J.M., 2001. Experimental evidence for the stability of the depletion zone around a growing protein crystal under microgravity. *Acta Crystallogr. D* 57, 412–417.
- Otalora, F., Garcia-Ruiz, J.M., 1997. Crystal growth studies in microgravity with the APCF. I. Computer simulation of transport dynamics. *J. Crystal Growth* 182, 141–154.
- Patel, S., Cudney, R., McPherson, A., 1995. Polymeric precipitants for the crystallisation of macromolecules. *Biochem. Biophys. Res. Comm.* 205, 819–828.
- Pernodet, N., Maaloum, M., Tinland, B., 1997. Pore size of agarose gels by atomic force microscopy. *Electrophoresis* 18, 55–58.
- Potier, N., Lamour, V., Poterszman, A., Thierry, J.-C., Moras, D., Van Dorselaer, A., 2000. Characterization of crystal content by ESI-MS and MALDI-MS. *Acta Crystallogr. D* 56, 1583–1590.
- Potter, R., Young, R., Hong, S., Ciszak, E.M., 2004. Use of plastic capillaries for macromolecular crystallisation. *J. Appl. Crystallogr.* 37, 500–501.
- Provost, K., Robert, M.-C., 1991. Application of gel growth to hanging drop technique. *J. Crystal Growth* 110, 258–264.
- Provost, K., Robert, M.-C., 1995. Crystal growth of lysozymes in media contaminated by parent molecules: influence of gelled media. *J. Crystal Growth* 156, 112–120.
- Pusey, M., Witherow, W., Naumann, R., 1988. Preliminary investigations into solutal flow about growing tetragonal lysozyme crystals. *J. Crystal Growth* 90, 105–111.
- Raman, P., Cherezov, V., Caffrey, M., 2006. The membrane protein data bank. *Cell. Mol. Life Sci.* 63, 36–51.
- Robert, M.-C., Berthou, J., 1987. Croissance cristalline du lysozyme de poule en gel de tétraméthoxysilane. *C. R. Acad. Sci. Paris* 305 (Série II), 847–850.
- Robert, M.-C., Lefaucheur, F., Jannot, B., Godefroy, G., Garnier, E., 1988. A comparative study of gel grown and space grown lead hydrogen phosphate crystals. *J. Crystal Growth* 88, 499–510.
- Robert, M.-C., Bernard, Y., Lefaucheur, F., 1994. Study of nucleation-related phenomena in lysozyme solutions. Application to gel growth. *Acta Crystallogr. D* 50, 496–503.
- Robert, M.-C., Lefaucheur, F., 1988. Crystal growth in gels: principles and applications. *J. Crystal Growth* 90, 358–367.
- Romier, C., Figner, R., Reuter, K., Suck, D., 1996. Purification, crystallisation, and preliminary X-ray diffraction studies of tRNA-guanine transglycosylase from *Zymomonas mobilis*. *Proteins: Struct. Funct. Genet.* 24, 516–519.
- Rubin, B., 1969. The growth of single crystals by controlled diffusion in silica gel. *AIChE J.* 15, 206–208.
- Rummel, G., Hardmeyer, A., Widmer, C., Chiu, M.L., Nolert, P., Locher, G.P., Pedruzzi, I., Landau, E.M., Rosenbusch, J.P., 1998. Lipid cubic phases: new matrices for the three-dimensional crystallisation of membrane proteins. *J. Struct. Biol.* 121, 82–91.
- Saito, H., Sakurai, A., Sakakibara, M., Saga, H., 2003. Preparation and properties of transparent cellulose hydrogels. *J. Appl. Polym. Sci.* 90, 3020–3025.
- Salemme, F.R., 1972. A free-interface diffusion technique for the crystallisation of proteins for X-ray crystallography. *Arch. Biochem. Biophys.* 151, 533–539.
- Sauter, C., Lorber, B., Giegé, R., 2002. Towards atomic resolution with crystals grown in gel: the case of thaumatin seen at room temperature. *Proteins: Struct. Funct. Genet.* 48, 146–150.
- Sauter, C., Balg, C., Moreno, A., Dhoub, K., Théobald-Dietrich, A., Chênevert, R., Giegé, R., Lorber, B., 2009. Agarose gel facilitates enzyme crystal soaking with ligand analog. *J. Appl. Crystallogr.* 42, 279–283.
- Schellenberger, P., et al., to be published.
- Sica, F., Demasi, D., Mazzarella, L., Zagari, A., Capasso, S., Pearl, L.H., D'Auria, S., Raia, C.A., Rossi, M., 1994. Elimination of twinning in crystals of *Sulfolobus solfataricus* alcohol dehydrogenase holo-enzyme by growth in agarose gels. *Acta Crystallogr. D* 50, 508–511.
- Silva, S.M.C., Pinto, F.V., Antunes, F.E., Miguel, M.G., Sousa, J.J.S., Pais, A.A.C.C., 2008. Aggregation and gelation in hydroxypropylmethyl cellulose aqueous solutions. *J. Colloid Interface Sci.* 327, 333–340.
- Sugahara, M., Kageyama, Y., Kunishima, N., 2009. A hands-free capillary system for protein crystallography from crystallization to data collection. *J. Appl. Crystallogr.* 42, 129–133.
- Sugiyama, S., Tanabe, K., Hirose, M., Kitatani, T., Hasenaka, H., Takahashi, Y., Adachi, H., Takano, K., Murakami, S., Mori, Y., Inoue, T., Matsumura, H., 2009. Protein crystallization in agarose gel with high strength: developing an automated system for protein crystallographic processes. *Jpn. J. Appl. Phys.* 48, 075502 (5 pages).
- Thiessen, K.J., 1994. The use of two novel methods to grow protein crystals by microdialysis and vapor diffusion in an agarose gel. *Acta Crystallogr. D* 50, 491–495.
- Vidal, O., 1997. Cristallogenèse d'un lysozyme en gel d'agarose et en gel de silice. Thesis, University Pierre et Marie Curie, Paris.
- Vidal, O., Bernard, Y., Robert, M.-C., Lefaucheur, F., 1996. Time evolution of refractive index of gelled and gel-free supersaturated lysozyme solutions. *J. Crystal Growth* 168, 40–43.
- Vidal, O., Robert, M.-C., Boué, F., 1998a. Gel growth of lysozyme crystals studied by small angle neutron scattering: case of agarose as a nucleation promoter. *J. Crystal Growth* 192, 257–270.
- Vidal, O., Robert, M.-C., Boué, F., 1998b. Gel growth of lysozyme crystals studied by small angle neutron scattering: case of silica gel, a nucleation inhibitor. *J. Crystal Growth* 192, 271–281.
- Vidal, O., Robert, M.-C., Arnoux, B., Capelle, B., 1999. Crystalline quality of lysozyme crystals grown in agarose and silica gels studied by X-ray diffraction techniques. *J. Crystal Growth* 196, 559–571.
- Vinogradova, E., Moreno, A., Lara, V.H., Bosch, P., 2003. Multi-fractal imaging and structural investigations of silica hydrogels and aerogels. *Silicon Chem.* 2, 247–254.
- Wang, L., Liu, X.Y., 2008. Kinetic analysis of protein crystal nucleation in gel matrix. *Biophys. J.* 95, 5931–5940.
- Wilke, K.-Th., 1988. *Kristallzüchtung*. Verlag Harri Deutsch, Frankfurt/Main, pp. 938–948.
- Willaert, R., Zegers, I., Wyns, L., Sleutel, M., 2005. Protein crystallization in hydrogel beads. *Acta Crystallogr. D* 61, 1280–1288.
- Wong, T.W., Lee, Y.-H., An, D.-S., Song, H.K., Kim, S.-G., 2009. A simple technique to convert sitting-drop vapour diffusion into hanging-drop vapour diffusion by solidifying the reservoir solution with agarose. *J. Appl. Crystallogr.* 42, 975–976.
- Yu, C.-A., Xia, D., Deisenhofer, J., Yu, L., 1994. Crystallisation of mitochondrial cytochrome *b-c₁* complex from gel or without reduced pressure. *J. Mol. Biol.* 243, 802–805.
- Zeppelauer, M., 1971. Formation of large crystals. *Methods Enzymol.* 22, 253–266.
- Zhu, D.-W., Lorber, B., Sauter, C., Ng, J.D., Benas, P., Le Grimellec, C., Giegé, R., 2001. Growth kinetics, diffraction properties and effect of agarose on the stability of a novel crystal from of *Thermus thermophilus* aspartyl-tRNA synthetase-1. *Acta Crystallogr. D* 57, 552–558.

4. Annexe 4. Surface accessible des résidus de la capsid du GFLV-TD.

La colonne CDESNM, indique la valeur surface accessible au solvant (SAS_{tot}) (en Å²) de chaque résidus de la sous-unité D, lorsque ses sous-unités voisines C, E, S, N et M, sont présentes. Tous les résidus en contact avec le solvant, et externes à la particule virale sont grisés. Les résidus en contact avec d'autres sous-unités de la capsid sont répertoriés dans la colonne D-X, qui signifie en contact avec n'importe quelle sous-unité parmi C, E, M, N, S. Chaque interaction est détaillée dans les colonnes CD, DE, DM, DN, DS, AD, et BD. Les résidus uniquement en contact avec le solvant, sans interaction avec d'autres sous-unités, sont surlignés en gris foncés. Les régions R1 à R5 sont encadrées par un trait noir épais.

GFLV : surface et contacts

Résidus	CDESNM SASStot	D SASStot	CD SASStot	DE SASStot	DM SASStot	DN SASStot	DS SASStot	AD SASStot	BD SASStot	Contact et Surface	Exterieur	Interieur	Contact - Interface						
													D-X	C-D	D-E	D-M	D-N	D-S	
ALA 117	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
MET 118	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
TRP 119	5,00	103,40	103,40	5,00	103,40	103,40	103,40	103,40	103,40	103,40	Oui	ext		Oui		Oui			
PHE 120	10,27	41,86	41,86	10,27	41,86	41,86	41,86	41,86	41,86	41,86	Oui	ext		Oui		Oui			
LYS 121	107,67	171,28	171,28	107,67	171,28	171,28	171,28	171,28	171,28	171,28	Oui	ext		Oui		Oui			
SER 122	59,84	61,29	61,29	59,84	61,29	61,29	61,29	61,29	61,29	61,29	Oui	ext		Oui		Oui			
THR 123	58,01	58,01	58,01	58,01	58,01	58,01	58,01	58,01	58,01	58,01	Oui	ext							
THR 124	26,76	57,00	57,00	26,76	57,00	57,00	57,00	57,00	57,00	57,00	Oui	ext		Oui		Oui			
PHE 125	34,93	41,20	41,20	34,93	41,20	41,20	41,20	41,20	41,20	41,20	Oui	ext		Oui		Oui			
GLU 126	66,64	66,64	66,64	66,64	66,64	66,64	66,64	66,64	66,64	66,64	Oui	ext							
SER 127	9,21	9,21	9,21	9,21	9,21	9,21	9,21	9,21	9,21	9,21	Oui	ext							
PRO 128	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
ARG 129	62,83	62,83	62,83	62,83	62,83	62,83	62,83	62,83	62,83	62,83	Oui	ext							
LEU 130	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
HIS 131	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
PHE 132	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
THR 133	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
CYS 134	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
LEU 135	0,92	0,92	0,92	0,92	0,92	0,92	0,92	0,92	0,92	0,92	Oui	ext							
THR 136	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
GLY 137	16,07	16,07	16,07	16,07	16,07	16,07	16,07	16,07	16,07	16,07	Oui	ext							
ASN 138	4,11	4,11	4,11	4,11	4,11	4,11	4,11	4,11	4,11	4,11	Oui	ext							
ASN 139	16,52	16,52	16,52	16,52	16,52	16,52	16,52	16,52	16,52	16,52	Oui	ext							
LYS 140	107,94	119,54	119,54	119,54	119,54	107,94	119,54	119,54	119,54	119,54	Oui	ext		Oui				Oui	
GLU 141	65,52	92,37	92,37	65,52	92,37	92,37	65,52	92,37	92,37	92,37	Oui	ext		Oui				Oui	
LEU 142	1,12	43,15	43,15	43,15	43,15	43,15	1,12	43,15	43,15	43,15	Oui	ext		Oui				Oui	
ALA 143	21,38	108,58	108,58	108,58	108,58	108,58	21,38	108,58	108,58	108,58	Oui	ext		Oui				Oui	
ALA 144	11,17	63,44	63,44	63,44	63,44	63,44	11,17	63,44	63,44	63,44	Oui	ext		Oui				Oui	
ASP 145	69,70	101,04	101,04	101,04	101,04	101,04	69,70	101,04	101,04	101,04	Oui	ext		Oui				Oui	
TRP 146	0,11	45,91	45,91	45,91	45,91	45,91	0,11	45,91	45,91	45,91	Oui			Oui				Oui	
GLN 147	5,06	79,88	79,88	79,88	79,88	79,88	5,06	79,88	79,88	79,88	Oui	ext		Oui				Oui	
ALA 148	0,00	0,46	0,46	0,46	0,46	0,46	0,00	0,46	0,46	0,46				Oui				Oui	
VAL 149	0,08	33,72	33,72	33,72	33,72	33,72	0,08	33,72	33,72	33,72	Oui			Oui				Oui	
VAL 150	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
GLU 151	0,00	18,74	18,74	18,74	18,74	18,74	0,00	18,74	18,74	18,74	Oui			Oui				Oui	
LEU 152	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
TYR 153	40,78	51,38	51,38	51,38	51,38	51,38	40,78	51,38	51,38	51,38	Oui	int		Oui				Oui	
ALA 154	7,34	7,34	7,34	7,34	7,34	7,34	7,34	7,34	7,34	7,34	Oui	int							
GLU 155	119,62	119,62	119,62	119,62	119,62	119,62	119,62	119,62	119,62	119,62	Oui	int							
LEU 156	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	0,31	Oui	int							
GLU 157	136,43	136,43	136,43	136,43	136,43	136,43	136,43	136,43	136,43	136,43	Oui	int							
GLU 158	62,16	120,01	120,01	62,16	120,01	120,01	62,16	120,01	120,01	120,01	Oui	int		Oui		Oui			
ALA 159	5,53	11,83	11,83	5,53	11,83	11,83	5,53	11,83	11,83	11,83	Oui	int		Oui		Oui			
THR 160	124,16	150,17	150,17	124,16	150,17	150,17	124,16	150,17	150,17	150,17	Oui	int		Oui		Oui			
SER 161	47,08	68,16	68,16	47,08	68,16	68,16	47,08	68,16	68,16	68,16	Oui	int		Oui		Oui			
PHE 162	1,81	44,11	44,11	1,81	44,11	44,11	1,81	44,11	44,11	44,11	Oui	int		Oui		Oui			
LEU 163	31,70	33,92	33,92	31,70	33,92	33,92	33,92	33,92	33,92	33,92	Oui	int		Oui		Oui			
GLY 164	21,93	21,93	21,93	21,93	21,93	21,93	21,93	21,93	21,93	21,93	Oui	int							
LYS 165	145,52	145,52	145,52	145,52	145,52	145,52	145,52	145,52	145,52	145,52	Oui	int							
PRO 166	39,03	39,03	39,03	39,03	39,03	39,03	39,03	39,03	39,03	39,03	Oui	int							
THR 167	37,95	37,95	37,95	37,95	37,95	37,95	37,95	37,95	37,95	37,95	Oui	int							
LEU 168	22,49	22,49	22,49	22,49	22,49	22,49	22,49	22,49	22,49	22,49	Oui	int							
VAL 169	66,18	66,18	66,18	66,18	66,18	66,18	66,18	66,18	66,18	66,18	Oui	int							
PHE 170	9,42	14,53	9,42	14,53	14,53	14,53	9,42	14,53	14,53	14,53	Oui	int		Oui		Oui			
ASP 171	70,78	70,78	70,78	70,78	70,78	70,78	70,78	70,78	70,78	70,78	Oui	int							
PRO 172	8,43	11,85	8,43	11,85	11,85	11,85	8,43	11,85	11,85	11,85	Oui	int		Oui		Oui			
GLY 173	54,59	54,59	54,59	54,59	54,59	54,59	54,59	54,59	54,59	54,59	Oui	int							
VAL 174	40,94	40,94	40,94	40,94	40,94	40,94	40,94	40,94	40,94	40,94	Oui	int							

GFLV : surface et contacts

Résidus	CDESNM SASot	D SASot	CD SASot	DE SASot	DM SASot	DN SASot	DS SASot	AD SASot	BD SASot	Contact et Surface	Exterieur	Interieur	Contact - Interface				
													D-X	C-D	D-E	D-M	D-N
PHE 175	4,48	4,48	4,48	4,48	4,48	4,48	4,48	4,48	4,48	Oui		int					
ASN 176	104,19	104,19	104,19	104,19	104,19	104,19	104,19	104,19	104,19	Oui		int					
GLY 177	1,69	1,69	1,69	1,69	1,69	1,69	1,69	1,69	1,69	Oui		int					
LYS 178	120,53	120,53	120,53	120,53	120,53	120,53	120,53	120,53	120,53	Oui		int					
PHE 179	13,03	13,03	13,03	13,03	13,03	13,03	13,03	13,03	13,03	Oui	ext	int					
GLN 180	114,34	114,34	114,34	114,34	114,34	114,34	114,34	114,34	114,34	Oui		int					
PHE 181	43,47	43,47	43,47	43,47	43,47	43,47	43,47	43,47	43,47	Oui	ext	int					
LEU 182	27,00	27,00	27,00	27,00	27,00	27,00	27,00	27,00	27,00	Oui	ext	int					
THR 183	6,93	6,93	6,93	6,93	6,93	6,93	6,93	6,93	6,93	Oui	ext						
CYS 184	17,74	17,74	17,74	17,74	17,74	17,74	17,74	17,74	17,74	Oui	ext						
PRO 185	56,82	56,82	56,82	56,82	56,82	56,82	56,82	56,82	56,82	Oui	ext						
PRO 186	14,26	14,26	14,26	14,26	14,26	14,26	14,26	14,26	14,26	Oui	ext						
ILE 187	19,97	19,97	19,97	19,97	19,97	19,97	19,97	19,97	19,97	Oui	ext						
PHE 188	64,17	64,17	64,17	64,17	64,17	64,17	64,17	64,17	64,17	Oui	ext						
PHE 189	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
ASP 190	45,13	45,13	45,13	45,13	45,13	45,13	45,13	45,13	45,13	Oui	ext						Oui
LEU 191	5,61	41,08	41,08	41,08	41,08	41,08	41,08	5,61	41,08	Oui	ext		Oui				Oui
THR 192	123,92	141,52	141,52	141,52	141,52	141,52	141,52	123,92	141,52	Oui	ext		Oui				Oui
ALA 193	30,66	32,59	32,59	32,59	32,59	32,59	32,59	30,66	32,59	Oui	ext		Oui				Oui
VAL 194	47,34	145,76	145,76	145,76	145,76	145,76	145,76	47,34	145,76	Oui	ext		Oui				Oui
THR 195	40,30	58,07	58,07	58,07	58,07	58,07	58,07	40,30	58,07	Oui	ext		Oui				Oui
ALA 196	16,97	16,97	16,97	16,97	16,97	16,97	16,97	16,97	16,97	Oui	ext						
LEU 197	46,27	46,27	46,27	46,27	46,27	46,27	46,27	46,27	46,27	Oui	ext						
ARG 198	79,33	79,33	79,33	79,33	79,33	79,33	79,33	79,33	79,33	Oui	ext						
SER 199	49,38	49,38	49,38	49,38	49,38	49,38	49,38	49,38	49,38	Oui	ext						
ALA 200	11,30	11,30	11,30	11,30	11,30	11,30	11,30	11,30	11,30	Oui	ext						
GLY 201	24,31	24,31	24,31	24,31	24,31	24,31	24,31	24,31	24,31	Oui	ext						
LEU 202	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
THR 203	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34	0,34								
LEU 204	21,56	21,56	21,56	21,56	21,56	21,56	21,56	21,56	21,56	Oui	ext						
GLY 205	2,40	2,40	2,40	2,40	2,40	2,40	2,40	2,40	2,40	Oui	ext						
GLN 206	50,79	58,10	50,79	58,10	58,10	58,10	58,10	58,10	58,10	Oui	ext		Oui		Oui		
VAL 207	26,41	26,41	26,41	26,41	26,41	26,41	26,41	26,41	26,41	Oui	ext						
PRO 208	46,07	46,07	46,07	46,07	46,07	46,07	46,07	46,07	46,07	Oui	ext						
MET 209	83,21	83,21	83,21	83,21	83,21	83,21	83,21	83,21	83,21	Oui	ext						
VAL 210	34,15	34,15	34,15	34,15	34,15	34,15	34,15	34,15	34,15	Oui	ext						
GLY 211	72,37	72,37	72,37	72,37	72,37	72,37	72,37	72,37	72,37	Oui	ext						
THR 212	140,31	140,31	140,31	140,31	140,31	140,31	140,31	140,31	140,31	Oui	ext						
THR 213	24,52	24,52	24,52	24,52	24,52	24,52	24,52	24,52	24,52	Oui	ext						
LYS 214	96,04	96,04	96,04	96,04	96,04	96,04	96,04	96,04	96,04	Oui	ext						
VAL 215	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
TYR 216	31,70	31,70	31,70	31,70	31,70	31,70	31,70	31,70	31,70	Oui	ext						
ASN 217	13,80	13,80	13,80	13,80	13,80	13,80	13,80	13,80	13,80	Oui	ext						
LEU 218	4,22	4,22	4,22	4,22	4,22	4,22	4,22	4,22	4,22	Oui		int					
ASN 219	8,30	8,30	8,30	8,30	8,30	8,30	8,30	8,30	8,30	Oui	ext	int					
SER 220	5,41	5,41	5,41	5,41	5,41	5,41	5,41	5,41	5,41	Oui	ext	int					
THR 221	5,68	5,68	5,68	5,68	5,68	5,68	5,68	5,68	5,68	Oui	ext						
LEU 222	13,76	13,76	13,76	13,76	13,76	13,76	13,76	13,76	13,76	Oui		int					
VAL 223	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00			int					
SER 224	6,19	16,71	6,19	16,71	16,71	16,71	16,71	16,71	16,71	Oui	ext		Oui		Oui		
CYS 225	0,19	1,79	0,19	1,79	1,79	1,79	1,79	1,79	1,79	Oui			Oui		Oui		
VAL 226	5,73	6,96	5,73	6,96	6,96	6,96	6,96	6,96	6,96	Oui	ext		Oui		Oui		
LEU 227	0,00	97,70	0,00	97,70	97,70	97,70	97,70	97,70	97,70				Oui		Oui		
GLY 228	0,00	1,41	0,00	1,41	1,41	1,41	1,41	1,41	1,41	Oui			Oui		Oui		
MET 229	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
GLY 230	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
GLY 231	3,69	3,69	3,69	3,69	3,69	3,69	3,69	3,69	3,69	Oui		int					
THR 232	6,05	6,05	6,05	6,05	6,05	6,05	6,05	6,05	6,05	Oui		int					

GFLV : surface et contacts

Résidus	CDESNM SASStot	D SASStot	CD SASStot	DE SASStot	DM SASStot	DN SASStot	DS SASStot	AD SASStot	BD SASStot	Contact et Surface	Exterieur	Interieur	Contact - Interface				
													D-X	C-D	D-E	D-M	D-N
VAL 233	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
ARG 234	79,20	79,20	79,20	79,20	79,20	79,20	79,20	79,20	79,20	Oui		int					
GLY 235	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
ARG 236	102,30	102,30	102,30	102,30	102,30	102,30	102,30	102,30	102,30	Oui		int					
VAL 237	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
HIS 238	28,15	28,15	28,15	28,15	28,15	28,15	28,15	28,15	28,15	Oui		int					
ILE 239	20,68	20,68	20,68	20,68	20,68	20,68	20,68	20,68	20,68	Oui		int					
CYS 240	4,84	4,84	4,84	4,84	4,84	4,84	4,84	4,84	4,84	Oui		int					
ALA 241	11,28	11,28	11,28	11,28	11,28	11,28	11,28	11,28	11,28	Oui		int					
PRO 242	4,05	4,05	4,05	4,05	4,05	4,05	4,05	4,05	4,05	Oui		int					
ILE 243	82,51	108,70	108,70	108,70	108,70	108,70	108,70	82,51	108,70	Oui		int	Oui				Oui
PHE 244	20,41	57,57	57,57	57,57	57,57	57,57	57,57	20,41	57,57	Oui		int	Oui				Oui
TYR 245	2,82	2,82	2,82	2,82	2,82	2,82	2,82	2,82	2,82	Oui		int					
SER 246	7,85	50,65	50,65	50,65	50,65	50,65	50,65	7,85	50,65	Oui		int	Oui				Oui
ILE 247	0,15	3,25	3,25	3,25	3,25	3,25	3,25	0,15	3,25	Oui		int	Oui				Oui
VAL 248	0,00	40,63	40,63	40,63	40,63	40,63	40,63	0,00	40,63	Oui		int	Oui				Oui
LEU 249	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
TRP 250	11,50	11,50	11,50	11,50	11,50	11,50	11,50	11,50	11,50	Oui		int					
VAL 251	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
VAL 252	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
SER 253	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
GLU 254	5,32	5,35	5,32	5,35	5,35	5,35	5,35	5,35	5,35	Oui		int	Oui	Oui			
TRP 255	34,67	60,90	34,67	60,90	60,90	60,90	60,90	60,90	60,90	Oui	ext		Oui	Oui			
ASN 256	89,82	90,31	89,82	90,31	90,31	90,31	90,31	90,31	90,31	Oui	ext		Oui	Oui			
GLY 257	6,26	6,26	6,26	6,26	6,26	6,26	6,26	6,26	6,26	Oui	ext		Oui	Oui			
THR 258	32,20	33,28	32,20	33,28	33,28	33,28	33,28	33,28	33,28	Oui	ext		Oui	Oui			
THR 259	7,48	21,47	7,48	21,47	21,47	21,47	21,47	21,47	21,47	Oui	ext		Oui	Oui			
MET 260	181,96	181,96	181,96	181,96	181,96	181,96	181,96	181,96	181,96	Oui	ext		Oui	Oui			
ASP 261	72,87	72,87	72,87	72,87	72,87	72,87	72,87	72,87	72,87	Oui	ext		Oui	Oui			
TRP 262	88,49	112,49	112,49	112,49	112,49	112,49	112,49	88,49	112,49	Oui	ext		Oui	Oui			Oui
ASN 263	22,19	107,73	107,73	107,73	107,73	107,73	107,73	22,19	107,73	Oui	ext		Oui	Oui			Oui
GLU 264	21,43	70,91	21,43	70,91	70,91	70,91	70,91	21,43	70,91	Oui	ext		Oui	Oui			
LEU 265	2,01	2,01	2,01	2,01	2,01	2,01	2,01	2,01	2,01	Oui	ext						
PHE 266	6,59	72,76	72,76	72,76	72,76	72,76	72,76	6,59	72,76	Oui		int	Oui				Oui
LYS 267	22,98	172,67	75,80	172,67	172,67	172,67	172,67	100,93	172,67	Oui		int	Oui				Oui
TYR 268	8,97	37,78	8,97	37,78	37,78	37,78	37,78	37,78	37,78	Oui		int	Oui	Oui			
PRO 269	33,09	48,05	33,09	48,05	48,05	48,05	48,05	48,05	48,05	Oui		int	Oui	Oui			
GLY 270	16,74	16,74	16,74	16,74	16,74	16,74	16,74	16,74	16,74	Oui		int					
VAL 271	44,67	44,67	44,67	44,67	44,67	44,67	44,67	44,67	44,67	Oui		int					
TYR 272	47,07	123,88	123,88	123,88	123,88	123,88	123,88	47,07	123,88	Oui		int	Oui				Oui
VAL 273	7,97	7,97	7,97	7,97	7,97	7,97	7,97	7,97	7,97	Oui		int					
GLU 274	52,07	99,55	99,55	99,55	99,55	99,55	99,55	52,07	99,55	Oui		int	Oui				Oui
GLU 275	122,13	122,13	122,13	122,13	122,13	122,13	122,13	122,13	122,13	Oui		int					
ASP 276	39,51	39,51	39,51	39,51	39,51	39,51	39,51	39,51	39,51	Oui		int					
GLY 277	36,31	36,31	36,31	36,31	36,31	36,31	36,31	36,31	36,31	Oui		int					
SER 278	83,03	83,03	83,03	83,03	83,03	83,03	83,03	83,03	83,03	Oui		int					
PHE 279	7,70	7,70	7,70	7,70	7,70	7,70	7,70	7,70	7,70	Oui		int					
GLU 280	72,17	72,17	72,17	72,17	72,17	72,17	72,17	72,17	72,17	Oui		int					
VAL 281	1,05	1,05	1,05	1,05	1,05	1,05	1,05	1,05	1,05	Oui		int					
LYS 282	66,37	109,20	66,37	109,20	109,20	109,20	109,20	109,20	109,20	Oui		int	Oui	Oui			
ILE 283	0,00	3,39	0,00	3,39	3,39	3,39	3,39	3,39	3,39	Oui		int	Oui	Oui			
ARG 284	1,04	89,96	1,04	89,96	89,96	89,96	89,96	89,96	89,96	Oui		int	Oui	Oui			
SER 285	0,00	21,89	0,00	21,89	21,89	21,89	21,89	21,89	21,89	Oui		int	Oui	Oui			
PRO 286	6,18	97,47	6,18	97,47	97,47	97,47	97,47	97,47	97,47	Oui		int	Oui	Oui			
TYR 287	29,75	169,40	29,75	169,40	169,40	169,40	169,40	169,40	169,40	Oui		int	Oui	Oui			
HIS 288	12,13	97,48	12,13	97,48	97,48	97,48	97,48	97,48	97,48	Oui		int	Oui	Oui			
ARG 289	31,11	89,91	31,11	89,91	89,91	89,91	89,91	89,91	89,91	Oui		int	Oui	Oui			
THR 290	0,00	0,33	0,00	0,33	0,33	0,33	0,33	0,33	0,33	Oui		int	Oui	Oui			

GFLV : surface et contacts

Résidus	CDESNM SASot	D SASot	CD SASot	DE SASot	DM SASot	DN SASot	DS SASot	AD SASot	BD SASot	Contact et Surface	Exterieur	Interieur	Contact - Interface				
													D-X	C-D	D-E	D-M	D-N
PRO 291	1,19	43,59	1,19	43,59	43,59	43,59	43,59	43,59	43,59	Oui			Oui				
ALA 292	0,00	2,43	0,00	2,43	2,43	2,43	2,43	2,43	2,43				Oui				
ARG 293	16,49	130,03	16,49	130,03	130,03	130,03	130,03	130,03	130,03	Oui	ext		Oui				
LEU 294	12,91	24,47	12,91	24,47	24,47	24,47	24,47	24,47	24,47	Oui	ext		Oui				
LEU 295	17,78	137,89	17,78	137,89	137,89	137,89	137,89	137,89	137,89	Oui	ext		Oui				
ALA 296	35,49	96,40	35,49	96,40	96,40	96,40	96,40	96,40	96,40	Oui	ext		Oui				
ASP 297	151,84	158,75	151,84	158,75	158,75	158,75	158,75	158,75	158,75	Oui	ext		Oui				
GLN 298	44,50	64,72	44,50	64,72	64,72	64,72	64,72	64,72	64,72	Oui	ext		Oui				
SER 299	53,31	53,31	53,31	53,31	53,31	53,31	53,31	53,31	53,31	Oui	ext						
GLN 300	55,94	55,94	55,94	55,94	55,94	55,94	55,94	55,94	55,94	Oui	ext						
ARG 301	130,10	130,10	130,10	130,10	130,10	130,10	130,10	130,10	130,10	Oui	ext						
ASP 302	44,26	44,26	44,26	44,26	44,26	44,26	44,26	44,26	44,26	Oui	ext						
MET 303	14,62	14,86	14,62	14,86	14,86	14,86	14,86	14,86	14,86	Oui	ext		Oui				
SER 304	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
SER 305	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17	0,17								
LEU 306	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
ASN 307	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
PHE 308	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
TYR 309	6,52	6,52	6,52	6,52	6,52	6,52	6,52	6,52	6,52	Oui	ext						
ALA 310	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
ILE 311	0,02	44,69	44,69	44,69	44,69	44,69	44,69	0,02	44,69	Oui			Oui				Oui
ALA 312	0,00	54,70	54,70	54,70	54,70	54,70	54,70	0,00	54,70	Oui			Oui				Oui
GLY 313	0,00	12,94	12,94	12,94	12,94	12,94	12,94	0,00	12,94	Oui			Oui				Oui
PRO 314	0,00	3,98	3,98	3,98	3,98	3,98	3,98	0,00	3,98	Oui			Oui				Oui
ILE 315	0,00	119,81	119,81	119,81	119,81	119,81	119,81	0,00	119,81	Oui			Oui				Oui
ALA 316	0,00	28,74	28,74	28,74	28,74	28,74	28,74	0,00	28,74	Oui			Oui				Oui
PRO 317	30,10	73,63	73,63	73,63	73,63	73,63	30,10	73,63	73,63	Oui		int	Oui				Oui
SER 318	49,56	124,20	124,20	124,20	124,20	124,20	49,56	124,20	124,20	Oui	ext		Oui				Oui
GLY 319	44,25	46,01	46,01	46,01	46,01	46,01	44,25	46,01	46,01	Oui	ext		Oui				Oui
GLU 320	18,83	26,95	26,95	26,95	26,95	26,95	18,83	26,95	26,95	Oui	ext		Oui				Oui
THR 321	112,74	112,74	112,74	112,74	112,74	112,74	112,74	112,74	112,74	Oui	ext						
ALA 322	28,71	28,71	28,71	28,71	28,71	28,71	28,71	28,71	28,71	Oui	ext						
GLN 323	71,95	71,95	71,95	71,95	71,95	71,95	71,95	71,95	71,95	Oui	ext						
LEU 324	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
PRO 325	6,97	6,97	6,97	6,97	6,97	6,97	6,97	6,97	6,97	Oui	ext						
ILE 326	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
VAL 327	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
VAL 328	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
GLN 329	19,09	19,09	19,09	19,09	19,09	19,09	19,09	19,09	19,09	Oui		int					
ILE 330	1,57	1,57	1,57	1,57	1,57	1,57	1,57	1,57	1,57	Oui	ext	int					
ASP 331	19,44	19,44	19,44	19,44	19,44	19,44	19,44	19,44	19,44	Oui		int					
GLU 332	18,66	18,66	18,66	18,66	18,66	18,66	18,66	18,66	18,66	Oui		int					
ILE 333	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15	0,15								
VAL 334	21,87	21,87	21,87	21,87	21,87	21,87	21,87	21,87	21,87	Oui		int					
ARG 335	178,00	178,00	178,00	178,00	178,00	178,00	178,00	178,00	178,00	Oui		int					
PRO 336	67,28	67,28	67,28	67,28	67,28	67,28	67,28	67,28	67,28	Oui		int					
ASP 337	23,88	39,93	23,88	39,93	39,93	39,93	23,88	39,93	39,93	Oui		int	Oui		Oui		
LEU 338	58,12	58,12	58,12	58,12	58,12	58,12	58,12	58,12	58,12	Oui		int					
SER 339	7,26	20,40	7,26	20,40	20,40	20,40	7,26	20,40	20,40	Oui		int	Oui		Oui		
LEU 340	81,33	140,83	81,33	140,83	140,83	140,83	81,33	140,83	140,83	Oui		int	Oui		Oui		
PRO 341	45,27	52,75	45,27	52,75	52,75	52,75	45,27	52,75	52,75	Oui		int	Oui		Oui		
SER 342	0,00	34,75	0,00	34,75	34,75	34,75	0,00	34,75	34,75	Oui		int	Oui		Oui		
PHE 343	0,10	70,42	0,10	70,42	70,42	70,42	0,10	70,42	70,42	Oui		int	Oui		Oui		
GLU 344	44,49	137,99	44,49	137,99	137,99	137,99	44,49	137,99	137,99	Oui		int	Oui		Oui		
ASP 345	94,64	95,25	94,64	95,25	95,25	95,25	94,64	95,25	95,25	Oui		int	Oui		Oui		
ASP 346	29,38	77,43	29,38	77,43	77,43	77,43	29,38	77,43	77,43	Oui		int	Oui		Oui		
TYR 347	32,57	50,18	32,57	50,18	50,18	50,18	32,57	50,18	50,18	Oui		int	Oui		Oui		
PHE 348	0,00	14,20	0,00	14,20	14,20	14,20	0,00	14,20	14,20	Oui		int	Oui		Oui		

GFLV : surface et contacts

Résidus	CDESNM SASStot	D SASStot	CD SASStot	DE SASStot	DM SASStot	DN SASStot	DS SASStot	AD SASStot	BD SASStot	Contact et Surface	Exterieur	Interieur	Contact - Interface					
													D-X	C-D	D-E	D-M	D-N	D-S
VAL 349	5,30	49,68	5,30	49,68	49,68	49,68	49,68	49,68	49,68	Oui		int	Oui					
TRP 350	1,32	1,32	1,32	1,32	1,32	1,32	1,32	1,32	1,32	Oui		int						
VAL 351	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
ASP 352	18,84	18,84	18,84	18,84	18,84	18,84	18,84	18,84	18,84	Oui	ext							
PHE 353	0,52	0,52	0,52	0,52	0,52	0,52	0,52	0,52	0,52	Oui								
SER 354	17,88	17,88	17,88	17,88	17,88	17,88	17,88	17,88	17,88	Oui	ext							
GLU 355	114,20	114,20	114,20	114,20	114,20	114,20	114,20	114,20	114,20	Oui	ext							
PHE 356	15,37	15,37	15,37	15,37	15,37	15,37	15,37	15,37	15,37	Oui	ext							
THR 357	84,67	84,67	84,67	84,67	84,67	84,67	84,67	84,67	84,67	Oui	ext							
LEU 358	98,02	98,02	98,02	98,02	98,02	98,02	98,02	98,02	98,02	Oui	ext							
ASP 359	51,45	51,45	51,45	51,45	51,45	51,45	51,45	51,45	51,45	Oui	ext							
LYS 360	179,03	179,03	179,03	179,03	179,03	179,03	179,03	179,03	179,03	Oui	ext							
GLU 361	9,72	9,72	9,72	9,72	9,72	9,72	9,72	9,72	9,72	Oui	ext							
GLU 362	45,39	45,39	45,39	45,39	45,39	45,39	45,39	45,39	45,39	Oui	ext							
ILE 363	12,09	12,09	12,09	12,09	12,09	12,09	12,09	12,09	12,09	Oui	ext							
GLU 364	91,60	91,60	91,60	91,60	91,60	91,60	91,60	91,60	91,60	Oui	ext							
ILE 365	3,65	3,65	3,65	3,65	3,65	3,65	3,65	3,65	3,65	Oui	ext							
GLY 366	11,70	11,70	11,70	11,70	11,70	11,70	11,70	11,70	11,70	Oui	ext							
SER 367	7,66	7,66	7,66	7,66	7,66	7,66	7,66	7,66	7,66	Oui	ext							
ARG 368	35,93	35,93	35,93	35,93	35,93	35,93	35,93	35,93	35,93	Oui	ext							
PHE 369	6,04	6,04	6,04	6,04	6,04	6,04	6,04	6,04	6,04	Oui	ext							
PHE 370	37,21	37,21	37,21	37,21	37,21	37,21	37,21	37,21	37,21	Oui	ext							
ASP 371	102,83	102,83	102,83	102,83	102,83	102,83	102,83	102,83	102,83	Oui	ext							
PHE 372	10,10	10,10	10,10	10,10	10,10	10,10	10,10	10,10	10,10	Oui	ext							
THR 373	138,00	138,00	138,00	138,00	138,00	138,00	138,00	138,00	138,00	Oui	ext							
SER 374	38,64	38,64	38,64	38,64	38,64	38,64	38,64	38,64	38,64	Oui	ext							
ASN 375	169,02	169,02	169,02	169,02	169,02	169,02	169,02	169,02	169,02	Oui	ext							
THR 376	41,80	41,80	41,80	41,80	41,80	41,80	41,80	41,80	41,80	Oui	ext							
CYS 377	11,61	11,61	11,61	11,61	11,61	11,61	11,61	11,61	11,61	Oui	ext							
ARG 378	149,99	149,99	149,99	149,99	149,99	149,99	149,99	149,99	149,99	Oui	ext							
VAL 379	34,02	34,02	34,02	34,02	34,02	34,02	34,02	34,02	34,02	Oui	ext							
SER 380	46,46	46,46	46,46	46,46	46,46	46,46	46,46	46,46	46,46	Oui	ext							
MET 381	54,71	56,52	54,71	56,52	56,52	56,52	56,52	56,52	56,52	Oui	ext		Oui		Oui			
GLY 382	3,95	12,18	3,95	12,18	12,18	12,18	12,18	12,18	12,18	Oui	ext		Oui		Oui			
GLU 383	68,31	121,52	68,31	121,52	121,52	121,52	121,52	121,52	121,52	Oui	ext		Oui		Oui			
ASN 384	1,27	12,72	1,27	12,72	12,72	12,72	12,72	12,72	12,72	Oui			Oui		Oui			
PRO 385	0,73	10,31	0,73	10,31	10,31	10,31	10,31	10,31	10,31	Oui			Oui		Oui			
PHE 386	0,00	0,71	0,00	0,71	0,71	0,71	0,71	0,71	0,71	Oui			Oui		Oui			
ALA 387	15,23	15,23	15,23	15,23	15,23	15,23	15,23	15,23	15,23	Oui	ext							
ALA 388	5,31	5,31	5,31	5,31	5,31	5,31	5,31	5,31	5,31	Oui	ext							
MET 389	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
ILE 390	4,08	4,08	4,08	4,08	4,08	4,08	4,08	4,08	4,08	Oui	ext							
ALA 391	13,36	13,36	13,36	13,36	13,36	13,36	13,36	13,36	13,36	Oui	ext							
CYS 392	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
HIS 393	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
GLY 394	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	0,07	Oui								
LEU 395	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
HIS 396	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
SER 397	17,43	20,92	17,43	20,92	20,92	20,92	20,92	20,92	20,92	Oui		int	Oui		Oui			
GLY 398	12,64	12,64	12,64	12,64	12,64	12,64	12,64	12,64	12,64	Oui		int						
VAL 399	24,80	32,04	24,80	32,04	32,04	32,04	32,04	32,04	32,04	Oui		int	Oui		Oui			
LEU 400	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	Oui								
ASP 401	29,47	29,47	29,47	29,47	29,47	29,47	29,47	29,47	29,47	Oui		int						
LEU 402	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	0,05	Oui								
LYS 403	77,74	91,89	77,74	91,89	91,89	91,89	91,89	91,89	91,89	Oui		int	Oui		Oui			
LEU 404	0,28	0,28	0,28	0,28	0,28	0,28	0,28	0,28	0,28	Oui		int						
GLN 405	1,74	64,57	1,74	64,57	64,57	64,57	64,57	64,57	64,57	Oui		int	Oui		Oui			
TRP 406	0,00	12,79	0,00	12,79	12,79	12,79	12,79	12,79	12,79	Oui			Oui		Oui			

GFLV : surface et contacts

Résidus	CDESNM SASStot	D SASStot	CD SASStot	DE SASStot	DM SASStot	DN SASStot	DS SASStot	AD SASStot	BD SASStot	Contact et Surface	Exterieur	Interieur	Contact - Interface				
													D-X	C-D	D-E	D-M	D-N
SER 407	0,00	54,41	0,00	54,41	54,41	54,41	54,41	54,41	54,41	Oui			Oui				
LEU 408	0,10	33,64	0,10	33,64	33,64	33,64	33,64	33,64	33,64	Oui			Oui				
ASN 409	82,56	93,67	82,56	93,67	93,67	93,67	93,67	93,67	93,67	Oui	ext		Oui				
THR 410	51,68	56,95	51,68	56,95	56,95	56,95	56,95	56,95	56,95	Oui	ext		Oui				
GLU 411	50,14	114,91	50,14	114,91	114,91	114,91	114,91	114,91	114,91	Oui	ext		Oui				
PHE 412	11,55	122,36	36,78	95,95	122,36	122,36	122,36	113,23	112,06	Oui		int	Oui	Oui	Oui		
GLY 413	10,82	80,13	59,48	31,47	80,13	80,13	80,13	80,13	80,13	Oui		int	Oui	Oui	Oui		
LYS 414	127,61	171,16	171,15	127,62	171,16	171,16	171,16	171,16	171,16	Oui	ext		Oui				Oui
SER 415	0,57	16,47	16,47	0,57	16,47	16,47	16,47	16,47	16,47	Oui	ext		Oui				Oui
SER 416	15,34	63,97	63,97	15,34	63,97	63,97	63,97	63,97	63,97	Oui	ext		Oui				Oui
GLY 417	0,23	26,36	26,36	0,23	26,36	26,36	26,36	26,36	26,36				Oui				Oui
SER 418	2,21	13,36	13,36	2,21	13,36	13,36	13,36	13,36	13,36	Oui	ext		Oui				Oui
VAL 419	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
THR 420	5,98	16,73	16,73	5,98	16,73	16,73	16,73	16,73	16,73	Oui	ext		Oui				Oui
ILE 421	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
THR 422	15,74	17,84	17,84	15,74	17,84	17,84	17,84	17,84	17,84	Oui	ext		Oui				Oui
LYS 423	0,97	8,65	8,65	0,97	8,65	8,65	8,65	8,65	8,65	Oui		int	Oui				Oui
LEU 424	12,43	13,69	13,69	12,43	13,69	13,69	13,69	13,69	13,69	Oui	ext		Oui				Oui
VAL 425	0,00	44,15	44,15	0,00	44,15	44,15	44,15	44,15	44,15	Oui			Oui				Oui
GLY 426	3,86	14,11	14,11	3,86	14,11	14,11	14,11	14,11	14,11	Oui	ext		Oui				Oui
ASP 427	22,95	67,63	67,63	22,95	67,63	67,63	67,63	67,63	67,63	Oui	ext		Oui				Oui
LYS 428	136,36	136,36	136,36	136,36	136,36	136,36	136,36	136,36	136,36	Oui	ext		Oui				Oui
ALA 429	80,21	92,77	92,77	80,21	92,77	92,77	92,77	92,77	92,77	Oui	ext		Oui				Oui
MET 430	100,01	182,75	182,75	100,01	182,75	182,75	182,75	182,75	182,75	Oui	ext		Oui				Oui
GLY 431	14,97	19,33	19,33	14,97	19,33	19,33	19,33	19,33	19,33	Oui	ext		Oui				Oui
LEU 432	15,59	124,03	124,03	15,59	124,03	124,03	124,03	124,03	124,03	Oui	ext		Oui				Oui
ASP 433	30,16	124,82	124,82	30,16	124,82	124,82	124,82	124,82	124,82	Oui	ext		Oui				Oui
GLY 434	0,00	33,82	33,82	0,00	33,82	33,82	33,82	33,82	33,82	Oui			Oui				Oui
PRO 435	12,52	95,37	95,37	12,52	95,37	95,37	95,37	95,37	95,37	Oui		int	Oui				Oui
SER 436	2,30	51,44	51,44	2,30	51,44	51,44	51,44	51,44	51,44	Oui		int	Oui				Oui
HIS 437	30,29	88,99	88,99	30,29	88,99	88,99	88,99	88,99	88,99	Oui		int	Oui				Oui
VAL 438	1,54	107,75	107,75	1,54	107,75	107,75	107,75	107,75	107,75	Oui		int	Oui				Oui
PHE 439	16,17	16,17	16,17	16,17	16,17	16,17	16,17	16,17	16,17	Oui		int	Oui				Oui
ALA 440	0,53	43,01	43,01	0,53	43,01	43,01	43,01	43,01	43,01	Oui		int	Oui				Oui
ILE 441	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
GLN 442	11,87	66,61	65,22	13,26	66,61	66,61	66,61	66,61	66,61	Oui		int	Oui	Oui	Oui		
LYS 443	60,99	155,79	139,74	77,04	155,79	155,79	155,79	155,79	155,79	Oui		int	Oui	Oui	Oui		
LEU 444	3,88	63,99	3,88	63,99	63,99	63,99	63,99	63,99	63,99	Oui		int	Oui	Oui	Oui		
GLU 445	70,72	107,05	70,72	107,05	107,05	107,05	107,05	107,05	107,05	Oui		int	Oui	Oui			
GLY 446	5,68	5,68	5,68	5,68	5,68	5,68	5,68	5,68	5,68	Oui		int					
THR 447	82,15	82,15	82,15	82,15	82,15	82,15	82,15	82,15	82,15	Oui		int					
THR 448	15,11	15,11	15,11	15,11	15,11	15,11	15,11	15,11	15,11	Oui		int					
GLU 449	34,14	69,92	69,92	34,14	69,92	69,92	69,92	69,92	69,92	Oui		int	Oui				Oui
LEU 450	0,00	1,31	1,31	0,00	1,31	1,31	1,31	1,31	1,31	Oui		int	Oui				Oui
LEU 451	1,00	37,89	37,89	1,00	37,89	37,89	37,89	37,89	37,89	Oui		int	Oui				Oui
VAL 452	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
GLY 453	0,39	0,79	0,79	0,39	0,79	0,79	0,79	0,79	0,79	Oui		int	Oui				Oui
ASN 454	0,84	13,76	13,76	0,84	13,76	13,76	13,76	13,76	13,76	Oui		int	Oui				Oui
PHE 455	1,83	147,95	147,95	1,83	147,95	147,95	147,95	147,95	147,95	Oui			Oui				Oui
ALA 456	0,08	25,59	25,59	0,08	25,59	25,59	25,59	25,59	25,59	Oui			Oui				Oui
GLY 457	12,54	26,80	26,80	12,54	26,80	26,80	26,80	26,80	26,80	Oui	ext		Oui				Oui
ALA 458	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
ASN 459	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00								
PRO 460	2,93	2,93	2,93	2,93	2,93	2,93	2,93	2,93	2,93	Oui	ext						
ASN 461	20,52	20,52	20,52	20,52	20,52	20,52	20,52	20,52	20,52	Oui	ext						
THR 462	68,72	68,72	68,72	68,72	68,72	68,72	68,72	68,72	68,72	Oui	ext						
ARG 463	220,73	220,73	220,73	220,73	220,73	220,73	220,73	220,73	220,73	Oui	ext						
PHE 464	45,51	45,51	45,51	45,51	45,51	45,51	45,51	45,51	45,51	Oui	ext						

GFLV : surface et contacts

Résidus	CDESNM SASStot	D SASStot	CD SASStot	DE SASStot	DM SASStot	DN SASStot	DS SASStot	AD SASStot	BD SASStot	Contact et Surface	Exterieur	Interieur	Contact - Interface					
													D-X	C-D	D-E	D-M	D-N	D-S
SER 465	80,99	119,85	119,85	80,99	119,85	119,85	119,85	119,85	119,85	Oui	ext		Oui		Oui			
LEU 466	18,35	52,68	52,68	18,35	52,68	52,68	52,68	52,68	52,68	Oui	ext		Oui		Oui			
TYR 467	1,70	132,74	132,74	1,70	132,74	132,74	132,74	132,74	132,74	Oui			Oui		Oui			
SER 468	3,34	9,40	9,40	3,34	9,40	9,40	9,40	9,40	9,40	Oui		int	Oui		Oui			
ARG 469	3,54	114,30	114,30	3,54	114,30	114,30	114,30	114,30	114,30	Oui		int	Oui		Oui			
TRP 470	24,53	24,57	24,57	24,53	24,57	24,57	24,57	24,57	24,57	Oui			Oui		Oui			
MET 471	0,00	0,54	0,54	0,00	0,54	0,54	0,54	0,54	0,54	Oui			Oui		Oui			
ALA 472	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04	0,04									
ILE 473	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
LYS 474	92,68	102,64	102,64	92,68	102,64	102,64	102,64	102,64	102,64	Oui	ext		Oui		Oui			
LEU 475	2,23	2,23	2,23	2,23	2,23	2,23	2,23	2,23	2,23	Oui	ext		Oui		Oui			
ASP 476	28,29	68,64	68,64	28,29	68,64	68,64	68,64	68,64	68,64	Oui	ext		Oui		Oui			
GLN 477	47,45	58,99	58,99	47,45	58,99	58,99	58,99	58,99	58,99	Oui	ext		Oui		Oui			
ALA 478	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
LYS 479	122,19	122,19	122,19	122,19	122,19	122,19	122,19	122,19	122,19	Oui	ext							
SER 480	9,32	9,32	9,32	9,32	9,32	9,32	9,32	9,32	9,32	Oui	ext							
ILE 481	0,55	0,55	0,55	0,55	0,55	0,55	0,55	0,55	0,55	Oui	ext							
LYS 482	85,56	139,45	85,56	139,45	139,45	139,45	139,45	139,45	139,45	Oui	ext		Oui	Oui				
VAL 483	25,76	43,96	25,76	43,96	43,96	43,96	43,96	43,96	43,96	Oui		int	Oui		Oui			
LEU 484	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
ARG 485	14,99	101,99	14,99	101,99	101,99	101,99	101,99	101,99	101,99	Oui	ext		Oui		Oui			
VAL 486	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
LEU 487	1,52	22,61	1,52	22,61	22,61	22,61	22,61	22,61	22,61	Oui			Oui		Oui			
CYS 488	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
LYS 489	70,93	71,02	70,93	71,02	71,02	71,02	71,02	71,02	71,02	Oui		int	Oui		Oui			
PRO 490	3,52	4,11	3,52	4,11	4,11	4,11	4,11	4,11	4,11	Oui		int	Oui		Oui			
ARG 491	82,74	93,19	93,19	82,74	93,19	93,19	93,19	93,19	93,19	Oui		int	Oui		Oui			
PRO 492	139,35	139,35	139,35	139,35	139,35	139,35	139,35	139,35	139,35	Oui		int	Oui					
GLY 493	43,63	43,63	43,63	43,63	43,63	43,63	43,63	43,63	43,63	Oui		int	Oui					
PHE 494	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19	0,19									
SER 495	13,26	13,26	13,26	13,26	13,26	13,26	13,26	13,26	13,26	Oui		int						
PHE 496	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
TYR 497	16,75	16,75	16,75	16,75	16,75	16,75	16,75	16,75	16,75	Oui		int						
GLY 498	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
ARG 499	12,83	12,83	12,83	12,83	12,83	12,83	12,83	12,83	12,83	Oui	ext							
THR 500	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	0,96	Oui		int						
SER 501	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00									
PHE 502	62,50	62,50	62,50	62,50	62,50	62,50	62,50	62,50	62,50	Oui	ext							
PRO 503	8,56	8,56	8,56	8,56	8,56	8,56	8,56	8,56	8,56	Oui	ext							
VAL 504	101,55	101,55	8,56	101,55	101,55	101,55	101,55	101,55	101,55	Oui	ext							

seuil = 0,01

5. Annexe 5. Alignement des séquences protéiques des CP codées par l'ARN2-U (X81814) et l'ARN2-L (X81815) de l'ArMV-S.

	1					50
X81814	GLAGRGTVPV	PKDCQAGKYL	KTLDLRDMVS	GFSGIQYEKW	ITAGLVMPDF	
X81815	GLAGRDSVQV	PKDCQAGKYL	KTLDLREMVS	GFSGIQYEKW	ITAGLVMPDF	
	51					100
X81814	KVVIRYPANA	FTGITWVMSF	DAYNRITSSI	STTASPAYTL	SVPHWLLHHK	
X81815	KVVIRYPANA	FTGITWVMSF	DAYNRITSSI	STTASPAYTL	SVPHWLLHHK	
	101					150
X81814	NGTTSCDIDY	GELCGHAMWF	SATTFESPCL	HFTCLTGNNK	ELAADWEFVV	
X81815	NGTTSCDLDY	GELCGHAMWF	GATTFESPCL	HFTCLTGNNK	ELAADWEFVV	
	151					200
X81814	ELYAEFEAAK	SFLGKPNFIY	SADAFNGSLK	FLTIPPLEYD	LSATSAYKSV	
X81815	ELYAEFEAAK	SFLGKPNFIY	SADAFNGSLK	FLTIPPLEYD	LSATSAYKSV	
	201					250
X81814	SLLLGQTLVD	GTHKVYNFNN	TLLSYYLIG	GIVKGVHVC	SPCTYGIVLR	
X81815	SLLLGQTLVD	GTHKVYNFNN	TLLSYYLIG	GIVKGVHVC	SPCTYGIVLR	
	251					300
X81814	VVSEWNGVTN	NWNQLFKYPG	CYIEEDGSFA	IEIRSPYHRT	PLRLIDAQSA	
X81815	VVSEWNGVTN	NWNQLFKYPG	CYIEEDGSFA	IEIRSPYHRT	PLRLIDAQSA	
	301					350
X81814	SSFTSTLNFY	AISGPIAPSG	ETAKMPVVVQ	IEEIALPDLS	VPSFPNDYFL	
X81815	SSFTSTLNFY	AISGPIAPSG	ETAKMPVVVQ	IEEIALPGLS	VPSFPNDYFL	
	351					400
X81814	WVDFSSFTVD	VEEYVIGSRF	FDISSTTSTV	ALGDNPF S IIM	IACHGLHHGI	
X81815	WVDFSSFTVD	VEEYVIGSRF	FDISSTTSTV	ALGDNPF A IHM	IACHGLHHGI	
	401					450
X81814	LDLKLMDLE	GEFGKSSGGV	TITKLCGDKA	TGMDGASRVC	ALQNMGCETE	
X81815	LDLKLMDLE	GEFGKSAGGV	TITKLCGDKA	TGMDGASRVC	ALQNMGCETE	
	451					500
X81814	LYIGNYAGAN	PNTALSLSYR	WLAIKLDKAK	SMKMLRILCK	PRGNFEFYGR	
X81815	LYIGNYAGAN	PNTALSLSYR	WLAIKLDKAK	SMKMLRILCK	PRGNFEFYGR	
	501					
X81814	TCFKV					
X81815	TCFKV					

Les acides aminés différents entre les deux protéines sont grisés

Bibliographie

- Abawi, G.S. and Widmer, T.L.** (2000) Impact of soil health management practices on soilborne pathogens, nematodes, and root diseases of vegetable crops. *Appl. Soil Eco.*, **15**, 37-47.
- Acharya, R., Fry, E., Stuart, D., Fox, G., Rowlands, D. and Brown, F.** (1989) The three-dimensional structure of foot-and-mouth disease virus at 2.9 Å resolution. *Nature*, **337**, 709-716.
- Ali, A. and Kobayashi, M.** (2010) Seed transmission of Cucumber mosaic virus in pepper. *J Virol Methods*, **163**, 234-237.
- Andret-Link, P.** (2005) Identification des déterminants viraux responsables de la spécificité de transmission du Grapevine fanleaf virus par son nématode vecteur *Xiphinema index*. In *Thèse en biologie moléculaire et cellulaire*. Strasbourg: Université Louis Pasteur.
- Andret-Link, P. and Fuchs, M.** (2005) Transmission specificity of plant viruses by vectors. *J. Phytopathol.*, **87**, 153-165.
- Andret-Link, P., Laporte, C., Valat, L., Ritzenthaler, C., Demangeat, G., Vigne, E., Laval, V., Pfeiffer, P., Stussi-Garaud, C. and Fuchs, M.** (2004a) *Grapevine fanleaf virus*: Still a major threat to the grapevine industry. *J. Plant Pathol.*, **86**, 183-195.
- Andret-Link, P., Schmitt-Keichinger, C., Demangeat, G., Komar, V. and Fuchs, M.** (2004b) The specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index* is solely determined by the viral coat protein. *Virology*, **320**, 12-22.
- Atreya, C.D., Raceah, B. and Pirone, T.P.** (1990) A point mutation in the coat protein abolishes aphid transmissibility of a potyvirus. *Virology*, **178**, 161-165.
- Atreya, P.L., Atreya, C.D. and Pirone, T.P.** (1991) Amino acid substitutions in the coat protein result in loss of insect transmissibility of a plant virus. *Proc. Natl. Acad. Sci. U S A*, **88**, 7887-7891.
- Atreya, P.L., Lopez-Moya, J.J., Chu, M., Atreya, C.D. and Pirone, T.P.** (1995) Mutational analysis of the coat protein N-terminal amino acids involved in potyvirus transmission by aphids. *J. Gen. Virol.*, **76** (2), 265-270.
- Baker, T.S. and Cheng, R.H.** (1996) A model-based approach for determining orientations of biological macromolecules imaged by cryoelectron microscopy. *J. Struct. Biol.*, **116**, 120-130.
- Baker, T.S., Olson, N.H. and Fuller, S.D.** (1999) Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiol Mol Biol Rev*, **63**, 862-922.
- Ban, N. and McPherson, A.** (1995) The structure of satellite panicum mosaic virus at 1.9 Å resolution. *Nat. Struct. Biol.*, **2**, 882-890.
- Barthels, N., van der Lee, F.M., Klap, J., Goddijn, O.J., Karimi, M., Puzio, P., Grundler, F.M., Ohl, S.A., Lindsey, K., Robertson, L., Robertson, W.M., Van Montagu, M., Gheysen, G. and Sijmons, P.C.** (1997) Regulatory sequences of Arabidopsis drive reporter gene expression in nematode feeding structures. *The Plant cell*, **9**, 2119-2134.

-
- Bawden, F.C. and Pirie, N.W.** (1938) A plant virus preparation in a fully crystalline state. *Nature*, **141**, 513-514.
- Beaty, B.J. and Marquardt, W.C.** (1996) *The Biology of Disease Vectors*.
- Belin, C., Schmitt, C., Demangeat, G., Komar, V., Pinck, L. and Fuchs, M.** (2001) Involvement of RNA2-encoded proteins in the specific transmission of *Grapevine fanleaf virus* by its nematode vector *Xiphinema index*. *Virology*, **291**, 161-171.
- Belin, C., Schmitt, C., Gaire, F., Walter, B., Demangeat, G. and Pinck, L.** (1999) The nine C-terminal residues of the grapevine fanleaf nepovirus movement protein are critical for systemic virus spread. *J. Gen. Virol.*, **80** 1347-1356.
- Belnap, D.M., McDermott, B.M., Jr., Filman, D.J., Cheng, N., Trus, B.L., Zuccola, H.J., Racaniello, V.R., Hogle, J.M. and Steven, A.C.** (2000) Three-dimensional structure of poliovirus receptor bound to poliovirus. *Proc. Natl. Acad. Sci. U S A*, **97**, 73-78.
- Bitterlin, M.W. and D., G.** (1987) Spatial distribution of *X. rivesi* and persistence of Tomato ringspot virus and its vector in soil. *Plant disease*, **87**, 408-411.
- Blanc, S., Ammar, E.D., Garcia-Lampasona, S., Dolja, V.V., Llave, C., Baker, J. and Pirone, T.P.** (1998) Mutations in the potyvirus helper component protein: effects on interactions with virions and aphid stylets. *J. Gen. Virol.*, **79** 3119-3122.
- Blanc, S., Lopez-Moya, J.J., Wang, R., Garcia-Lampasona, S., Thornbury, D.W. and Pirone, T.P.** (1997) A specific interaction between coat protein and helper component correlates with aphid transmission of a potyvirus. *Virology*, **231**, 141-147.
- Bol, J.F.** (2008) Role of capsid proteins. In *Methods in Molecular Biology* (Foster, G.D., Johansen, E., Hong, Y. and Nagy, P.D., eds), pp. 21-31.
- Bouquet, A., L., T. and Chatelet, P.** (2003) Combinaison des méthodes conventionnelles et biotechnologiques dans la sélection de porte-greffe présentant une résistance durable à la transmission de la maladie du court-noué. *Progrès Agricole et Viticole*, **22**, 507-512.
- Bovey, R., Gartel, W., Hewitt, W.B., Martelli, G.P. and Vuittenez, A.** (1990) Soil-borne viruses transmitted by nematodes. In *Virus and virus-like diseases of grapevines* (Bovey R., G.W., Hewitt W. B., Martelli G. P., Vuittenez, A., ed). Lausanne, Switzerland.: Edition payot, pp. 46-50.
- Brown, D.J. and MacFarlane, S.A.** (2001) 'Worms' that transmit viruses. *Biologist (London, England)*, **48**, 35-40.
- Brown, D.J.F.** (1986a) The transmission of two strains of arabis mosaic virus from England by populations of *Xiphinema diversicaudatum* (Nematoda: Dorylaimida) from ten countries. *Revue de Nematologie*, **9**, 83-87.
- Brown, D.J.F.** (1986b) Transmission of virus by the progeny of crosses between *Xiphinema diversicaudatum* from Italy and Scotland. *Revue de Nematologie*, **9**, 71-74.
- Brown, D.J.F., Robertson, W.M. and Trudgill, D.L.** (1995) Transmission of viruses by plant nematodes. *Annu. Rev. Phytopathol.*, **33**, 223-249.
- Brown, D.J.F. and Weischer, B.** (1998) Specificity, exclusivity and complementarity in transmission of plant virus by plant parasitic nematodes. *Fund.Appl. Nematol.* **21**, 1-11.
- Cadman, C.H.** (1963) Biology of soil-borne viruses. *Annu. Rev. Phytopathol.*, **1**, 143-172.

- Campbell, C.L. and Bruehl, G.W.** (1986) Viruses in Vectors: Transovarial Passage and Retention. (H. Ando, T.F., H.H. Storey, ed). St. Paul, Minnesota, USA: APS Press, p. 53.
- Canto, T., López-Moya, J.J., Serra-Yoldi, M.T., Díaz-Ruiz, J.R. and D., L.-A.** (1995) Different Helper Component Mutations Associated with Lack of Aphid Transmissibility in Two Isolates of Potato Virus Y. *Phytopathology*, **85**, 1519-1524.
- Carn, V.M.** (1996) The role of dipterous insects in the mechanical transmission of animal viruses. *Br. Vet. J.*, **152**, 377-393.
- Casjens, S.** (2008) Capsid assembly: Bacterial virus structure and assembly. In *Encyclopedia of virology* (Mahy, B.W.J. and Regenmortel, M.H.V., eds). London, United Kindom: Elsevier/Academic press, pp. 436-444.
- Caspar, D.L.D. and Klug, A.** (1962) Physical principles in the construction of regular viruses. *Cold Spring Harbor symposia on quantitative biology*, **27**, 1-24.
- Chandrasekar, V. and Johnson, J.E.** (1998) The structure of tobacco ringspot virus: a link in the evolution of icosahedral capsids in the picornavirus superfamily. *Structure*, **6**, 157-171.
- Chen, B. and Francki, R.I.** (1990) Cucumovirus transmission by the aphid *Myzus persicae* is determined solely by the viral coat protein. *J. Gen. Virol.*, **71**, 939-944.
- Chen, J.Z., Settembre, E.C., Aoki, S.T., Zhang, X., Bellamy, A.R., Dormitzer, P.R., Harrison, S.C. and Grigorieff, N.** (2009) Molecular interactions in rotavirus assembly and uncoating seen by high-resolution cryo-EM. *Proc. Natl. Acad. Sci. U S A*, **106**, 10644-10648.
- Cheng, Y. and Walze, T.** (2009) The advent of near-atomic resolution in single-particle electron microscopy. In *Annual Review of Biochemistry*, pp. 723-742.
- Conway, J.F. and Steven, A.C.** (1999) Methods for reconstructing density maps of 'single' particles from cryoelectron micrographs to subnanometer resolution. *J Struct Biol*, **128**, 106-118.
- Crick, F.H. and Watson, J.D.** (1956) Structure of small viruses. *Nature*, **177**, 473-475.
- Crowter, R.A., Amos, L.A., Finch, J.T., DeRosier, D.J. and Klug, A.** (1970) Three dimensionnal reconstruction of spherical viruses by Fourier synthesis from electron micrographs. *Nature*, **226**, 421-425.
- Demangeat, G., Esmenjaud, D., Voisin, R., Bidault, J.M., Grenan, S. and Claverie, M.** (2005a) Le court-noué de la vigne : Etat des connaissances sur cette maladie. *Phytoma*, **587**, 38-42.
- Demangeat, G., Komar, V., Cornuet, P., Esmenjaud, D. and Fuchs, M.** (2004) Sensitive and reliable detection of grapevine fanleaf virus in a single *Xiphinema index* nematode vector. *J. Virol. Methods*, **122**, 79-86.
- Demangeat, G., Voisin, R., Minot, J.C., Bosselut, N., Fuchs, M. and Esmenjaud, D.** (2005b) Survival of *Xiphinema index* in Vineyard Soil and Retention of *Grapevine fanleaf virus* Over Extended Time in the Absence of Host Plants. *Phytopathology*, **95**, 1151-1156.
- Dias, H.F.** (1970a) The relationship between cucumber necrosis virus and its vector, *Olpidium cucurbitacearum*. *Virology*, **42**, 204-211.
- Dias, H.F.** (1970b) Transmission of cucumber necrosis virus by *Olpidium cucurbitacearum* Barr & Dias. *Virology*, **40**, 828-839.
- Dias, H.F. and Harrison, B.D.** (1963) The relationship between grapevine fanleaf, grapevine yellow mosaic and arabis mosaic viruses. *Ann. appl. Virol.*, **51**, 97-105.
- Drucker, M., Froissart, R., Hebrard, E., Uzest, M., Ravallec, M., Esperandieu, P., Mani, J.C., Pugniere, M., Roquet, F., Fereres, A. and Blanc, S.** (2002) Intracellular distribution of viral gene products regulates a

complex mechanism of cauliflower mosaic virus acquisition by its aphid vector. *Proc. Natl. Acad. Sci. U S A*, **99**, 2422-2427.

Dubochet, J., Adrian, M., Chang, J.J., Homo, J.C., Lepault, J., McDowell, A.W. and Schultz, P. (1988) Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.*, **21**, 129-228.

Dubochet, J., McDowell, A.W., Menge, B., Schmid, E.N. and Lickfeld, K.G. (1983) Electron microscopy of frozen-hydrated bacteria. *J. Bacteriol.*, **155**, 381-390.

Edwardson, J.R. and Christie, R.G. (1991) Cucumoviruses. *CRC Handbook of Viruses Infecting Legumes*. CRC Press, Boca Raton 293-319.

Esmenjaud, D. (2000) Les nématodes de la vigne. In *Ravageurs de la vigne* (Féret, E. and Stockel, J., eds). Bordeaux, France.

Esmenjaud, D. and Bouquet, A. (2009) Selection and application of resistant germplasm application for grapevine nematodes management. In *Integrated Management of Fruit Crops and Forest Nematodes* (Mukerji, A.C.K.G., ed): Springer, pp. 195-214.

European Commission. (2007) Council Decision of 20 September 2007 concerning the no inclusion of 1,3-dichloropropene in Annex I to Council Directive 91/414/EEC and the withdrawal of authorisations for plant protection products containing that substance (notified under document number C(2007) 4281). Off. J. Eur. Union L249: 11-13.

Fabry, C.M., Rosa-Calatrava, M., Conway, J.F., Zubieta, C., Cusack, S., Ruigrok, R.W. and Schoehn, G. (2005) A quasi-atomic model of human adenovirus type 5 capsid. *EMBO J.*, **24**, 1645-1654.

Fereres, A. and Moreno, A. (2009) Behavioural aspects influencing plant virus transmission by homopteran insects. *Virus Res.*, **141**, 158-168.

Ferlenghi, I., Clarke, M., Ruttan, T., Allison, S.L., Schlich, J., Heinz, F.X., Harrison, S.C., Rey, F.A. and Fuller, S.D. (2001) Molecular organization of a recombinant subviral particle from tick-borne encephalitis virus. *Mol Cell*, **7**, 593-602.

Fragkoudis, R., Attarzadeh-Yazdi, G., Nash, A.A., Fazakerley, J.K. and Kohl, A. (2009) Advances in dissecting mosquito innate immune responses to arbovirus infection. *J. Gen. Virol.*, **90**, 2061-2072.

Franklin, R.E. (1955) Structure of Tobacco Mosaic Virus. *Nature*, **175**, 379-381.

Fry, E.E., Lea, S.M., Jackson, T., Newman, J.W., Ellard, F.M., Blakemore, W.E., Abu-Ghazaleh, R., Samuel, A., King, A.M. and Stuart, D.I. (1999) The structure and function of a foot-and-mouth disease virus-oligosaccharide receptor complex. *EMBO J.*, **18**, 543-554.

Fry, P.R. (1958) The relationship of *Olpidium brassicae* (Wor.) Dang. to the big-vein disease of lettuce. *J. Agric. Res.*, **1**, 301-304.

Fuchs, M. and Gonsalves, D. (2007) Safety of virus-resistant transgenic plants two decades after their introduction: lessons from realistic field risk assessment studies. *Annu. Rev. Phytopathol.*, 173-202.

Fuchs, M., Pinck, M., Serghini, M.A., Ravelonandro, M., Walter, B. and Pinck, L. (1989) The nucleotide sequence of satellite RNA in grapevine fanleaf virus, strain F13. *J. Gen. Virol.*, **70** (4), 955-962.

Fuller, S.D., Butcher, S.J., Cheng, R.H. and Baker, D. (1996) Three-dimensional reconstruction of icosahedral particles-the uncommon line. *J Struct Biol*, **116**, 48-55.

Gaire, F., Schmitt, C., Stussi-Garaud, C., Pinck, L. and Ritzenthaler, C. (1999) Protein 2A of grapevine fanleaf nepovirus is implicated in RNA2 replication and colocalizes to the replication site. *Virology*, **264**, 25-36.

- Gambino, G., Perrone, I., Carra, A., Chitarra, W., Boccacci, P., Marinoni, D.T., Barberis, M., Maghuly, F., Laimer, M. and Gribaudo, I.** (2010) Transgene silencing in grapevines transformed with GFLV resistance genes: Analysis of variable expression of transgene, siRNAs production and cytosine methylation. *Transgenic Res.*, **19**, 17-27.
- Gera, A., Loebenstein, G. and Raccah, B.** (1979) Protein coats of two strains of Cucumber mosaic virus affect transmission by *Aphis gossypii*. *Phytopathology*, **69**, 396-399.
- Gray, S.M. and Banerjee, N.** (1999) Mechanisms of arthropod transmission of plant and animal viruses. *Microbiol Mol Biol Rev*, **63**, 128-148.
- Grimes, J.M., Burroughs, J.N., Gouet, P., Diprose, J.M., Malby, R., Zientara, S., Mertens, P.P. and Stuart, D.I.** (1998) The atomic structure of the bluetongue virus core. *Nature*, **395**, 470-478.
- Grimes, J.M., Jakana, J., Ghosh, M., Basak, A.K., Roy, P., Chiu, W., Stuart, D.I. and Prasad, B.V.** (1997) An atomic model of the outer layer of the bluetongue virus core derived from X-ray crystallography and electron cryomicroscopy. *Structure*, **5**, 885-893.
- Grogan, R.G., Zink, F.W., Hewitt, W.B. and Kimble, K.A.** (1958) The association of *Olpidium* with the big-vein disease of lettuce. *Phytopathology*, **48**, 292-296.
- Hans, F., Fuchs, M. and Pinck, L.** (1992) Replication of grapevine fanleaf virus satellite RNA transcripts in *Chenopodium quinoa* protoplasts. *J. Gen. Virol.*, **73** (10), 2517-2523.
- Harrison, B.D. and Cadman, C.H.** (1959) Role of dagger nematode (*Xiphinema spp.*) in out breaks of plant disease caused by arabis mosaic virus. *Nature*, **184**, 1624-1626.
- Harrison, B.D. and Murant, A.F.** (1977) Nematode transmissibility of pseudo-recombinant isolates of *tomato black ring virus*. *Ann. appl. Virol.*, **22**, 209-212.
- Harrison, B.D., Murant, A.F., Mayo, M.A. and Roberts, I.M.** (1974a) Distribution of determinants for symptom production, host range and nematode transmissibility between the two RNA components of Raspberry ringspot virus. *J. Gen. Virol.*, **22**, 233-247.
- Harrison, B.D., Robertson, W.M. and Taylor, C.E.** (1974b) Specificity of retention and transmission of viruses by nematodes. *J. Nematol.*, **6**, 155-164.
- Harrison, S.C.** (1972) Structure of tomato bushy stunt virus: three-dimensional x-ray diffraction analysis at 30 Angstrom resolution. *Cold Spring Harbor symposia on quantitative biology*, **36**, 495-501.
- Harrison, S.C.** (1991) Viruses. *Curr. Opin. Struct. Biol.*, **1**, 288-295.
- Harrison, S.C. and Jack, A.** (1975) Structure of tomato bushy stunt virus: III. Three-dimensional X-ray diffraction analysis at 16 A resolution. *J. Mol. Biol.*, **97**, 173-191.
- Harrison, S.C., Olson, A.J., Schutt, C.E., Winkler, F.K. and Bricogne, G.** (1978) Tomato bushy stunt virus at 2.9 A resolution. *Nature*, **276**, 368-373.
- He, Y., Bowman, V.D., Mueller, S., Bator, C.M., Bella, J., Peng, X., Baker, T.S., Wimmer, E., Kuhn, R.J. and Rossmann, M.G.** (2000) Interaction of the poliovirus receptor with poliovirus. *Proc. Natl. Acad. Sci. U S A*, **97**, 79-84.
- Hebrard, E., Drucker, M., Leclerc, D., Hohn, T., Uzest, M., Froissart, R., Strub, J.M., Sanglier, S., van Dorselaer, A., Padilla, A., Labesse, G. and Blanc, S.** (2001) Biochemical characterization of the helper component of Cauliflower mosaic virus. *J. Virol.*, **75**, 8538-8546.

-
- Heinz, F.X., Allison, S.L. and Karl Maramorosch, F.A.M.A.J.S.T.J.C.a.T.P.M.** (2003) Flavivirus Structure and Membrane Fusion. In *Adv. Virus Res.*: Academic Press, pp. 63-97.
- Hemmer, C., Vigne, E., Goldschmidt, V., Komar, V., Marmonier, A., Valat, L., Demangeat, D., Vigneron, S., J., M. and Lemaire, O.** (2009) Transgenic rootstock expressing GFLV CP gene in a three years field trial; resistance assessment, impact on GFLV diversity and exchanges between rootstock and scion. . In *16th Meeting of the International Council for the Study of Virus and Virus-like Diseases of the Grapevine (ICVG)*. Dijon, France.
- Hernandez, C., Visser, P.B., Brown, D.J. and Bol, J.F.** (1997) Transmission of tobacco rattle virus isolate PpK20 by its nematode vector requires one of the two non-structural genes in the viral RNA 2. *J. Gen. Virol.*, **78**, 465-467.
- Herrbach, E.** (2004) Arthropod transmission In *Viruses and Virus Disease of Poaceae (Graminae)* (Lapierre, H. and Signoret, A.P., eds). Versailles, France: INRA editions, pp. 114-124.
- Hewitt, W.** (1950) Fanleaf - Another vine disease found in California. *The Bulletin, Department of Agriculture, State of California* **39**, 62-63
- Hewitt, W., Raski, D.J. and Goheen, A.C.** (1958) Nematode vector of soil-borne fanleaf virus of grapevines *Phytopathology*, **48**, 586-595.
- Heymann, J.B.** (2001) Bsoft: image and molecular processing in electron microscopy. *J Struct Biol*, **133**, 156-169.
- Hogenhout, S.A., Ammar el, D., Whitfield, A.E. and Redinbaugh, M.G.** (2008) Insect vector interactions with persistently transmitted viruses. *Annu. Rev. Phytopathol.*, **46**, 327-359.
- Hogle, J.M., Chow, M. and Filman, D.J.** (1985) Three-dimensional structure of poliovirus at 2.9 Å resolution. *Science New York, N.Y.*, **229**, 1358-1365.
- Hoh, F., Uzest, M., Drucker, M., Plisson-Chastang, C., Bron, P., Blanc, S. and Dumas, C.** (2010) Structural insights into the molecular mechanisms of Cauliflower mosaic virus transmission by its insect vector. *J Virol.*
- Huet, H., Gal-On, A., Meir, E., Lecoq, H. and Raccah, B.** (1994) Mutations in the helper component protease gene of zucchini yellow mosaic virus affect its ability to mediate aphid transmissibility. *J. Gen. Virol.*, **75**, 1407-1414.
- Hui, E. and Rochon, D.A.** (2006) Evaluation of the Roles of Specific Regions of the Cucumber Necrosis Virus Coat Protein Arm in Particle Accumulation and Fungus Transmission. *J. Virol.*, **80**, 5968-5975.
- Huss, B., Walter, B. and Fuchs, M.** (1989) Cross-protection between *arabis mosaic virus* and *grapevine fanleaf virus* isolates on *Chenopodium quinoa*. *Ann. appl. Virol.*, **114**, 45-60.
- Husted, K.** (1995) Reason for non-aphid transmissibility in a strain of Kalanchoe mosaic potyvirus. *Virus genes*, **11**, 59-61.
- ICTV International Committee on Taxonomy of Viruses (ICTV) database.
<http://www.ncbi.nlm.nih.gov/ICTVdb/ICTVdB/index.htm>.
- Jackson, T., Ellard, F.M., Ghazaleh, R.A., Brookes, S.M., Blakemore, W.E., Corteyn, A.H., Stuart, D.I., Newman, J.W. and King, A.M.** (1996) Efficient infection of cells in culture by type O foot-and-mouth disease virus requires binding to cell surface heparan sulfate. *J. Virol.*, **70**, 5282-5287.
- Jha, A. and Posnette, A.F.** (1961) Transmission of arabis mosaic virus by the nematode *Xiphinema diversicaudatum* (Micol.). *Virology*, **13**, 119-123.

- Jiang, W., Baker, M.L., Jakana, J., Weigele, P.R., King, J. and Chiu, W.** (2008) Backbone structure of the infectious epsilon15 virus capsid revealed by electron cryomicroscopy. *Nature*, **451**, 1130-1114.
- Johnson, J.E. and Speir, J.A.** (1997) Quasi-equivalent viruses: a paradigm for protein assemblies. *J. Mol. Biol.*, **269**, 665-675.
- Jones, D.R.** (2003) Plant viruses transmitted by whiteflies *Eur. J. Plant. Pathol.*, **109**, 195-219.
- Kakani, K., Reade, R. and Rochon, D.** (2004) Evidence that vector transmission of a plant virus requires conformational change in virus particles. *Journal of molecular biology*, **338**, 507-517.
- Kakani, K., Robbins, M. and Rochon, D.** (2003) Evidence that binding of cucumber necrosis virus to vector zoospores involves recognition of oligosaccharides. *J Virol*, **77**, 3922-3928.
- Kakani, K., Sgro, J.Y. and Rochon, D.** (2001) Identification of specific cucumber necrosis virus coat protein amino acids affecting fungus transmission and zoospore attachment. *J. Virol.*, **75**, 5576-5583.
- Katpally, U., Kakani, K., Reade, R., Dryden, K., Rochon, D. and Smith, T.J.** (2007) Structures of T=1 and T=3 particles of cucumber necrosis virus: evidence of internal scaffolding. *J. Mol. Biol.*, **365**, 502-512.
- Kendall, A., McDonald, M., Bian, W., Bowles, T., Baumgarten, S.C., Shi, J., Stewart, P.L., Bullitt, E., Gore, D., Irving, T.C., Havens, W.M., Ghabrial, S.A., Wall, J.S. and Stubbs, G.** (2008) Structure of flexible filamentous plant viruses. *J Virol*, **82**, 9546-9554.
- Klug, A., Caspar, D.L.D. and Kenneth, M.S.a.M.A.L.** (1961) The Structure of Small Viruses. In *Adv. Virus Res.*: Academic Press, pp. 225-325.
- Knight, R.L., Schultz, K.L.W., Kent, R.J., Venkatesan, M. and Griffin, D.E.** (2009) Role of N-Linked Glycosylation for Sindbis Virus Infection and Replication in Vertebrate and Invertebrate Systems. *J. Virol.*, **83**, 5640-5647.
- Kolatkar, P.R., Bella, J., Olson, N.H., Bator, C.M., Baker, T.S. and Rossmann, M.G.** (1999) Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor. *EMBO J.*, **18**, 6249-6259.
- Komar, V., Vigne, E., Demangeat, G., Lemaire, O. and Fuchs, M.** (2008) Cross-protection as control strategy against Grapevine fanleaf virus in naturally infected vineyards. *Plant Disease*, **92**, 1689-1694.
- Lamberti, F., Taylor, C.E. and Seinhorst, J.W.** (1975) Nematode vector of plant viruses. New York, USA.: Plenum Press.
- Laporte, C., Vetter, G., Loudes, A.M., Robinson, D.G., Hillmer, S., Stussi-Garaud, C. and Ritzenthaler, C.** (2003) Involvement of the secretory pathway and the cytoskeleton in intracellular targeting and tubule assembly of *Grapevine fanleaf virus* movement protein in tobacco BY-2 cells. *The Plant cell*, **15**, 2058-2075.
- Larson, S.B., Day, J., Greenwood, A. and McPherson, A.** (1998) Refined structure of satellite tobacco mosaic virus at 1.8 Å resolution. *J.Mol. Biol.*, **277**, 37-59.
- Le Gall, O., Christian, P., Fauquet, C., King, A., Knowls, N., Nakashima, N., Stanway, G. and Gorbalenya, A.** (2008) *Picornavirales*, a proposed order of positive-sense single-stranded RNA viruses with a pseudo-T = 3 virion architecture. *Arch. Virol.*, **153**, 714-727.
- Leh, V., Jacquot, E., Geldreich, A., Haas, M., Blanc, S., Keller, M. and Yot, P.** (2001) Interaction between the open reading frame III product and the coat protein is required for transmission of cauliflower mosaic virus by aphids. *J Virol*, **75**, 100-106.

-
- Leh, V., Jacquot, E., Geldreich, A., Hermann, T., Leclerc, D., Cerutti, M., Yot, P., Keller, M. and Blanc, S.** (1999) Aphid transmission of cauliflower mosaic virus requires the viral PIII protein. *EMBO J.*, **18**, 7077-7085.
- Lin, T., Chen, Z., Usha, R., Stauffacher, C.V., Dai, J.B., Schmidt, T. and Johnson, J.E.** (1999) The refined crystal structure of cowpea mosaic virus at 2.8 Å resolution. *Virology*, **265**, 20-34.
- Lin, T., Clark, A.J., Chen, Z., Shanks, M., Dai, J.B., Li, Y., Schmidt, T., Oxelfelt, P., Lomonossoff, G.P. and Johnson, J.E.** (2000) Structural fingerprinting: subgrouping of comoviruses by structural studies of red clover mottle virus to 2.4-Å resolution and comparisons with other comoviruses. *J Virol*, **74**, 493-504.
- Lin, T. and Johnson, J.E.** (2003) Structures of picorna-like plant viruses: implications and applications. *Adv. Virus. Res.*, **62**, 167-239.
- Liu, L., Grainger, J., Cañizares, M.C., Angell, S.M. and Lomonossoff, G.P.** (2004) Cowpea mosaic virus RNA-1 acts as an amplicon whose effects can be counteracted by a RNA-2-encoded suppressor of silencing. *Virology*, **323**, 37-48.
- Liu, S., He, X., Park, G., Josefsson, C. and Perry, K.L.** (2002) A conserved capsid protein surface domain of *Cucumber mosaic virus* is essential for efficient aphid vector transmission. *J. Virol.*, **76**, 9756-9762.
- Lopez-Moya, J.J. and A., G.J.** (2008) Potyviruses (Potyviridae). In *Encyclopedia of virology* (Mahy, B.W.J. and Regenmortel, M.H.V., eds). London, United Kingdom: Elsevier/Academic press, pp. 313-322.
- Lopez-Moya, J.J., Wang, R.Y. and Pirone, T.P.** (1999) Context of the coat protein DAG motif affects potyvirus transmissibility by aphids. *J. Gen. Virol.*, **80** 3281-3288.
- Lorber, B., Sauter, C., Theobald-Dietrich, A., Moreno, A., Schellenberger, P., Robert, M.C., Capelle, B., Sanglier, S., Potier, N. and Giege, R.** (2009) Crystal growth of proteins, nucleic acids, and viruses in gels. *Prog. Biophys. Mol. Biol.*, **101**, 13-25.
- Louides, A.M., Ritzenthaler, C., Pinck, M., Serghini, M.A. and Pinck, L.** (1995) The 119 kDa and 124 kDa polyproteins of arabis mosaic nepovirus (isolate S) are encoded by two distinct RNA2 species. *J. Gen. Virol.*, **76** 899-906.
- Luc, M. and Dalmasso, A.** (1975) A "lattice" for the identification of species of *Xiphinema COBB*, 1913. In *Nematode Vectors of Plant Viruses* (Press, L.P., ed). London pp. 53-70.
- Lucas, R.W., Larson, S.B. and McPherson, A.** (2002) The crystallographic structure of brome mosaic virus. *J. Mol. Biol.*, **317**, 95-108.
- Luo, M., Vriend, G., Kamer, G., Minor, I., Arnold, E., Rossmann, M.G., Boege, U., Scraba, D.G., Duke, G.M. and Palmenberg, A.C.** (1987) The atomic structure of Mengo virus at 3.0 Å resolution. *Science (New York, N.Y.)*, **235**, 182-191.
- MacFarlane, S.** (2003) Molecular determinants of the transmission of plant viruses by nematodes. *Molecular plant pathology*, **4**, 211-215.
- MacFarlane, S.A., Wallis, C.V. and Brown, D.J.** (1996) Multiple virus genes involved in the nematode transmission of pea early browning virus. *Virology*, **219**, 417-422.
- Margis, R. and Pinck, L.** (1992) Effects of site-directed mutagenesis on the presumed catalytic triad and substrate-binding pocket of grapevine fanleaf nepovirus 24-kDa proteinase. *Virology*, **190**, 884-888.
- Margis, R., Viry, M., Pinck, M., Bardonnnet, N. and Pinck, L.** (1994) Differential proteolytic activities of precursor and mature forms of the 24K proteinase of grapevine fanleaf nepovirus. *Virology*, **200**, 79-86.

- Margis, R., Viry, M., Pinck, M. and Pinck, L.** (1991) Cloning and in vitro characterization of the grapevine fanleaf virus proteinase cistron. *Virology*, **185**, 779-787.
- Marmonier, A., Schellenberger, P., Esmenjaud, D., Schmitt-Keichinger, C., Ritzenthaler, C., Andret-Link, P., Lemaire, O., Fuchs, M. and Demangeat, G.** (2010) The coat protein determines the specificity of virus transmission by *X. diversicaudatum*. *J. Plant Pathol.*, **92**, 275-279.
- Marsh, M. and Helenius, A.** (1989) Virus entry into animal cells. *Adv. Virus Res.*, **36**, 107-151.
- Martelli, G.P. and Boudon-Padieu, E.** (2006) Directory of infectious diseases of grapevines and viroses and virus-like diseases of the grapevine: bibliographic report 1998-2004. *Bari : CIHEAM-IAMB. Options Méditerranéennes : Série B. Etudes et Recherches*, 55-279.
- Martelli, G.P. and Taylor, C.E.** (1990) Distribution of viruses and their nematode vectors In *Advances in disease vector research* (Harris, K., ed). New York, USA.: Springer-Verlag.
- Mason, P.W., Rieder, E. and Baxt, B.** (1994) RGD sequence of foot-and-mouth disease virus is essential for infecting cells via the natural receptor but can be bypassed by an antibody-dependent enhancement pathway. *Proc. Natl. Acad. Sci. U S A*, **91**, 1932-1936.
- Mayo, M.A. and Robinson, D.G.** (1996) Nepoviruses: molecular biology and replication. In *The plant viruses, Polyhedral virions and bipartite RNA genomes.* (Harrison, B.D. and Murrant, A.F., eds). New York, USA: Plenum Press, pp. 287-288.
- Mc Namara, D.G.** (1980) The survival of *Xiphinema diversicaudatum* in plant-free soil. *Nematologica*, **26**, 170-181.
- McElroy, K.L., Tsetsarkin, K.A., Vanlandingham, D.L. and Higgs, S.** (2006) Role of the yellow fever virus structural protein genes in viral dissemination from the *Aedes aegypti* mosquito midgut. *J. Gen. Virol.*, **87**, 2993-3001.
- McGuire, J.M.** (1973) Retention of Tobacco Ringspot Virus by *Xiphinema americanum*. *Phytopathology*, **63**, 324-326.
- McLean, M.A., Campbell, R.N., Hamilton, R.I. and Rochon, D.M.** (1994) Involvement of the cucumber necrosis virus coat protein in the specificity of fungus transmission by *Olpidium bornovanus*. *Virology*, **204**, 840-842.
- Mello, A.F., Clark, A.J. and Perry, K.L.** (2009) The capsid protein of Cowpea chlorotic mottle virus is a determinant for vector transmission by a beetle. *J. Gen. Viro*, **91** (2), 545-551.
- Moreno, A., Hebrard, E., Uzest, M., Blanc, S. and Fereres, A.** (2005) A single amino acid position in the helper component of cauliflower mosaic virus can change the spectrum of transmitting vector species. *J. Virol.*, **79**, 13587-13593.
- Murant, A.F., Jones, A.T., Martelli, G.P. and Stace-Smith, R.** (1996) Nepoviruses: general properties, diseases and virus identification. . In *The plant viruses. Polyhedral virions and bipartite genomes* (Harrison, B.D. and Murant, A.F., eds). New York and London: Plenum Press, pp. 99-137.
- Natarajan, P., Lander, G.C., Shepherd, C.M., Reddy, V.S., Brooks, C.L., 3rd and Johnson, J.E.** (2005) Exploring icosahedral virus structures with VIPER. *Nat. Rev. Microbiol.*, **3**, 809-817.
- Nault, L.R.** (1997) Arthropod transmission of plant viruses: a new synthesis. *Ann. Entomol. Soc. Am.*, **90**, 521-541.

-
- Navaza, J.** (2008) Combining X-ray and electron-microscopy data to solve crystal structures. *Acta Cryst. D*, **64**, 70-75.
- Ng, J.C. and Falk, B.W.** (2006) Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annu. Rev. Phytopathol.*, **44**, 183-212.
- Ng, J.C., Josefsson, C., Clark, A.J., Franz, A.W. and Perry, K.L.** (2005) Virion stability and aphid vector transmissibility of Cucumber mosaic virus mutants. *Virology*, **332**, 397-405.
- Ng, J.C., Liu, S. and Perry, K.L.** (2000) Cucumber mosaic virus mutants with altered physical properties and defective in aphid vector transmission. *Virology*, **276**, 395-403.
- Olland, A.M., Jana-Valbuena, J., Schiff, L.A., Nibert, M.L. and Harrison, S.C.** (2001) Structure of the reovirus outer capsid and dsRNA-binding protein $\sigma 3$ at 1.8 Å resolution. *EMBO J*, **20**, 979-989.
- Olson, N.H., Kolatkar, P.R., Oliveira, M.A., Cheng, R.H., Greve, J.M., McClelland, A., Baker, T.S. and Rossmann, M.G.** (1993) Structure of a human rhinovirus complexed with its receptor molecule. *Proc. Natl. Acad. Sci. U S A*, **90**, 507-511.
- Peng, Y.H., Kadoury, D., Gal-On, A., Huet, H., Wang, Y. and Raccach, B.** (1998) Mutations in the HC-Pro gene of zucchini yellow mosaic potyvirus: effects on aphid transmission and binding to purified virions. *J. Gen. Virol.*, **79**, 897-904.
- Perry, K.L., Zhang, L. and Palukaitis, P.** (1998) Amino acid changes in the coat protein of cucumber mosaic virus differentially affect transmission by the aphids *Myzus persicae* and *Aphis gossypii*. *Virology*, **242**, 204-210.
- Perry, K.L., Zhang, L., Shintaku, M.H. and Palukaitis, P.** (1994) Mapping determinants in cucumber mosaic virus for transmission by *Aphis gossypii*. *Virology*, **205**, 591-595.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E.** (2004) UCSF Chimera - A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, **25**, 1605-1612.
- Pierro, D.J., Powers, E.L. and Olson, K.E.** (2007) Genetic determinants of Sindbis virus strain TR339 affecting midgut infection in the mosquito *Aedes aegypti*. *J. Gen. Virol.*, **88**, 1545-1554.
- Pierro, D.J., Powers, E.L. and Olson, K.E.** (2008) Genetic determinants of Sindbis virus mosquito infection are associated with a highly conserved alphavirus and flavivirus envelope sequence. *J. Virol.*, **82**, 2966-2974.
- Pinck, L., Fuchs, M., Pinck, M., Ravelonandro, M. and Walter, B.** (1988) A satellite RNA in Grapevine Fanleaf Virus Strain F13. *J. Gen. Virol.*, **69**, 233-239.
- Pinck, M., Reinbolt, J., Loudes, A.M., Le Ret, M. and Pinck, L.** (1991) Primary structure and location of the genome-linked protein (VPg) of grapevine fanleaf nepovirus. *FEBS Lett*, **284**, 117-119.
- Plisson, C., Drucker, M., Blanc, S., German-Retana, S., Le Gall, O., Thomas, D. and Bron, P.** (2003) Structural characterization of HC-Pro, a plant virus multifunctional protein. *J. Biol. Chem.*, **278**, 23753-23761.
- Plisson, C., Uzest, M., Drucker, M., Froissart, R., Dumas, C., Conway, J., Thomas, D., Blanc, S. and Bron, P.** (2005) Structure of the mature P3-virus particle complex of cauliflower mosaic virus revealed by cryo-electron microscopy. *J. Mol. Biol.*, **346**, 267-277.
- Quacquarelli, A., Gallitelli, V., Savino, V. and Martelli, G.P.** (1976) Properties of Grapevine Fanleaf Virus. *J. Gen. Virol.*, **32**, 349-360.

- Raman, S., Lange, O.F., Rossi, P., Tyka, M., Wang, X., Aramini, J., Liu, G., Ramelot, T.A., Eletsy, A., Szyperski, T., Kennedy, M.A., Prestegard, J., Montelione, G.T. and Baker, D.** (2010) NMR structure determination for larger proteins using backbone-only data. *Science New York, N.Y.*, **327**, 1014-1018.
- Ritzenthaler, C. and Hoffmann, C.** (2007) Tubule-guided movement of plant viruses. *Plant Cell Monogr.*, **7**, 64-82.
- Ritzenthaler, C., Laporte, C., Gaire, F., Dunoyer, P., Schmitt, C., Duval, S., Piequet, A., Loudes, A.M., Rohfritsch, O., Stussi-Garaud, C. and Pfeiffer, P.** (2002) Grapevine fanleaf virus replication occurs on endoplasmic reticulum-derived membranes. *J. Virol.*, **76**, 8808-8819.
- Ritzenthaler, C., Viry, M., Pinck, M., Margis, R., Fuchs, M. and Pinck, L.** (1991) Complete nucleotide sequence and genetic organization of grapevine fanleaf nepovirus RNA1. *J. Gen. Virol.*, **72** 2357-2365.
- Robbins, M.A., Reade, R.D. and Rochon, D.M.** (1997) A cucumber necrosis virus variant deficient in fungal transmissibility contains an altered coat protein shell domain. *Virology*, **234**, 138-146.
- Robertson, W.M.** (1987) Possible mechanisms of virus retention in virus vector nematodes *Annual report of the Scottish crop research institute, Invergowrie, 1986*, 127.
- Robertson, W.M. and Henry, C.E.** (1986a) An association of carbohydrates with particles of arabis mosaic virus retained within *Xiphinema diversicaudatum*. *Ann. appl. Virol.*, **109**, 299-305.
- Robertson, W.M. and Henry, C.E.** (1986b) A possible role of carbohydrate in the retention of nematode-transmitted virus *Annual report of the Scottish crop research institute, Invergowrie, 1985*, 113.
- Rochon, D., Kakani, K., Robbins, M. and Reade, R.** (2004) Molecular aspects of plant virus transmission by oospidium and plasmodiophorid vectors. *Annu. Rev. Phytopathol.*, **42**, 211-241.
- Rossmann, M.G., Arnold, E., Erickson, J.W., Frankenberger, E.A., Griffith, J.P., Hecht, H.J., Johnson, J.E., Kamer, G., Luo, M., Mosser, A.G. and et al.** (1985) Structure of a human common cold virus and functional relationship to other picornaviruses. *Nature*, **317**, 145-153.
- Rossmann, M.G., He, Y. and Kuhn, R.J.** (2002) Picornavirus-receptor interactions. *Trends in microbiology*, **10**, 324-331.
- Rossmann, M.G., Morais, M.C., Leiman, P.G. and Zhang, W.** (2005) Combining X-ray crystallography and electron microscopy. *Structure*, **13**, 355-362.
- Ruiz-Ferrer, V., Boskovic, J., Alfonso, C., Rivas, G., Llorca, O., Lopez-Abella, D. and Lopez-Moya, J.J.** (2005) Structural analysis of tobacco etch potyvirus HC-pro oligomers involved in aphid transmission. *J. Virol.*, **79**, 3758-3765.
- Saibil, H.R.** (2000) Macromolecular structure determination by cryo-electron microscopy. *Acta Cryst. D*, **56**, 1215-1222.
- Sanfaçon, H., Wellink, J., Le Gall, O., Karasev, A., van der Vlugt, R. and Wetzel, T.** (2009) Secoviridae: a proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus *Torradovirus*. *Arch. Virol.*, **154**, 899-907.
- Sanner, M.F., Olson, A.J. and Spehner, J.C.** (1996) Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, **38**, 305-320.
- Saxton, W.O. and Baumeister, W.** (1982) The correlation averaging of a regularly arranged bacterial cell envelope protein. *Journal of microscopy*, **127**, 127-138.

-
- Schoehn, G., Moss, S.R., Nuttall, P.A. and Hewat, E.A.** (1997) Structure of Broadhaven virus by cryoelectron microscopy: correlation of structural and antigenic properties of Broadhaven virus and bluetongue virus outer capsid proteins. *Virology*, **235**, 191-200.
- Scott, S.W., Zimmerman, M.T., Jones, A.T. and Le Gall, O.** (2000) Differences between the coat protein amino acid sequences of English and Scottish serotypes of *Raspberry ringspot virus* exposed on the surface of virus particles. *Virus Res.*, **68**, 119-126.
- Serghini, M.A., Fuchs, M., Pinck, M., Reinbolt, J., Walter, B. and Pinck, L.** (1990) RNA2 of grapevine fanleaf virus: sequence analysis and coat protein cistron location. *J. Gen. Virol.*, **71**, 1433-1441.
- Shi, Y., Ryabov, E.V., Van Wezel, R., Li, C., Jin, M., Wang, W., Fan, Z. and Hong, Y.** (2009) Suppression of local RNA silencing is not sufficient to promote cell-to-cell movement of Turnip crinkle virus in *Nicotiana benthamiana*. *Plant Signaling and Behavior*, **4**, 15-22.
- Simons, J.N.** (1976) Aphid Transmission of a nonaphid-transmissible strain of Tobacco etch Virus. *Phytopathology*, **66**, 652-654.
- Smith, T.J., Chase, E., Schmidt, T. and Perry, K.L.** (2000) The structure of cucumber mosaic virus and comparison to cowpea chlorotic mottle virus. *J. Virol.*, **74**, 7578-7586.
- Speir, J.A., Munshi, S., Wang, G., Baker, T.S. and Johnson, J.E.** (1995) Structures of the native and swollen forms of cowpea chlorotic mottle virus determined by X-ray crystallography and cryo-electron microscopy. *Structure*, **3**, 63-78.
- Spence, N.J.** (2001) Virus-vector interaction in plant virus disease transmission and epidemiology. In *Biotic interactions in plant-pathogen associations* (Jeger, M.J. and Spence, N.J., eds). New York, USA: Oxford University Press, pp. 15-26.
- Strauss, J.H., Rumenapf, T., Weir, R.C., Kuhn, R.J. and Strauss, E.G.** (1994) Cellular receptors for alphaviruses. In *Cellular Receptors for Animal Viruses* (Wimmer, E., ed). Plainview, NY: Cold Spring Harbor Laboratory Press, pp. 141-163.
- Syller, J.** (2006) The roles and mechanisms of helper component proteins encoded by potyviruses and caulimoviruses. *Physiol. Mol. Plant. Pathol.*, **67**, 119-130.
- Sylvester, E.S.** (1956) Beet yellow virus transmission by the Green peach aphid. *J. Econ. Entomol.*, **49**, 789-800.
- Sylvester, E.S.** (1988) Virus transmission studies. In *Aphids, their biology, natural enemies and control* (Minks, A.K. and Harrewijn, P., eds). Amsterdam: Elsevier, pp. 69-81.
- Taylor, C.E. and Brown, D.J.F.** (1997) Nematode vectors of plant viruses. Wallingford, England: CAB International.
- Taylor, C.E. and Robertson, W.M.** (1970) Sites of virus retention in the alimentary tract of the nematode vectors, *Xiphinema diversicaudatum* (Micol.) and *X. index* (Thorne and Allen). *Ann. appl. Virol.*, **66**, 375-380.
- Temmink, J.H.M., Campbell, R.N. and Smith, P.R.** (1970) Specificity and site of in vitro acquisition of Tobacco necrosis virus by zoospores of *Olpidium brassicae*. *J. Gen. Virol.*, **9**, 201-213.
- Thomas, C.L., Bayer, E.M., Ritzenthaler, C., Fernandez-Calvino, L. and Maule, A.J.** (2008) Specific targeting of a plasmodesmal protein affecting cell-to-cell communication. *PLoS Biol.*, **6**, e7.
- Uzest, M., Gargani, D., Dombrovsky, A., Cazevielle, C., Cot, D. and Blanc, S.** (2010) The "acrostyle": A newly described anatomical structure in aphid stylets. *Arthropod structure & development*.

- Uzest, M., Gargani, D., Drucker, M., Hebrard, E., Garzo, E., Candresse, T., Fereres, A. and Blanc, S.** (2007) A protein key to plant virus transmission at the tip of the insect vector stylet. *Proc. Natl. Acad. Sci. U S A*, **104**, 17959-17964.
- Valdez, R.B.** (1972) A Micro-container Technique for Studying Virus Transmission by Nematodes. *Plant Pathol.*, **21**, 114-117.
- Valdez, R.B., Mc Namara, D.G., Ormerod, P.J., Pitcher, R.S. and Thresh, J.M.** (1974) Transmission of the hop strain of arabis mosaic virus by *Xiphinema diversicaudatum*. *Ann. appl. Biol.*, **76**, 113-122.
- van den Heuvel, J.F., Hogenhout, S.A. and van der Wilk, F.** (1999) Recognition and receptors in virus transmission by arthropods. *Trends in microbiology*, **7**, 71-76.
- Vasilakis, N. and Weaver, S.C.** (2008) The history and evolution of human dengue emergence. *Adv. Virus Res.*, **72**, 1-76.
- Vassilakos, N., Vellios, E.K., Brown, E.C., Brown, D.J.F. and MacFarlane, S.A.** (2001) Tobravirus 2b protein acts in trans to facilitate transmission by nematodes. *Virology*, **279**, 478-487.
- Vellios, E., Brown, D.J. and MacFarlane, S.A.** (2002) Substitution of a single amino acid in the 2b protein of *Pea early-browning virus* affects nematode transmission. *J. Gen. Virol.*, **83**, 1771-1775.
- Villate, L.** (2008) Origine, variabilité et gestion des populations de *Xiphinema index*, le nématode vecteur du *Grapevine fanleaf virus* (GFLV). Université de Rennes 1, France.
- Villate, L., Esmenjaud, D., M, V.A.N.H., Stoeckel, S. and Plantard, O.** (2010) Genetic signature of amphimixis allows for the detection and fine scale localization of sexual reproduction events in a mainly parthenogenetic nematode. *Molecular ecology*, **19**, 865-873.
- Viry, M., Serghini, M.A., Hans, F., Ritzenthaler, C., Pinck, M. and Pinck, L.** (1993) Biologically active transcripts from cloned cDNA of genomic grapevine fanleaf nepovirus RNAs. *J. Gen. Virol.*, **74** (2), 169-174.
- Vuittenez, M., Munck, M.C. and Kuszala, J.** (1964) Souches de virus à haute agressivité isolées de vignes atteintes de dégénérescence infectieuse. *Etudes Virol. Appliquée*, **5**, 69-78.
- Wake, M.H.** (2003) What is "integrative biology"? *Integrative and Comparative Biology*, **43**, 239-241.
- Walker, M.A., Wolpert, J.A. and Weber, E.** (1994) Field screening of grape rootstock selections for resistance to fanleaf degeneration. *Plant Disease*, **78**, 134-136.
- Walter, B.** (1998) Virus et Viroses de la vigne. *Virologie*, **2**, 435-444.
- Wang, R.Y., Powell, G., Hardie, J. and Pirone, T.P.** (1998) Role of the helper component in vector-specific transmission of potyviruses. *J. Gen. Virol.*, **79** 1519-1524.
- Wang, S. and Gererich, R.C.** (1998) Immunofluorescent Localization of Tobacco Ringspot Nepovirus in the Vector Nematode *Xiphinema americanum*. *Phytopathology*, **88**, 885-889.
- Wang, S., Gergerich, R.C., Wickizer, S.L. and Kim, K.S.** (2002) Localization of transmissible and nontransmissible viruses in the vector nematode *Xiphinema americanum*. *Phytopathology*, **92**, 646-653.
- Wang, X., Bosselut, N., Castagnone, C., Voisin, R., Abad, P. and Esmenjaud, D.** (2003) Multiplex polymerase chain reaction identification of single individuals of the longidorid nematodes *Xiphinema index*, *X. diversicaudatum*, *X. vuittenezi*, and *X. italiae* Using specific primers from ribosomal genes. *Phytopathology*, **93**, 160-166.
- Watson, J.D.** (1954) The structure of tobacco mosaic virus: I. X-ray evidence of a helical arrangement of subunits around the longitudinal axis. *Biochim. Biophys. Acta*, **13**, 10-19.

-
- Watson, M.A. and Robers, F.M.** (1939) A comparative study of the transmission of Hyoscyamus virus 3, potato virus Y and cucumber virus 1 by the vectors *Mysus persicae* (Sulz), *M. circumflexus* (Buckton), and *Macrosiphum gei* (K). *Proc. R. Soc. Lond.* , **127**, 543-577.
- Weaver, S.C. and Barrett, A.D.** (2004) Transmission cycles, host range, evolution and emergence of arboviral disease. *Nat Rev Microbiol*, **2**, 789-801.
- Weaver, S.C. and Reisen, W.K.** (2010) Present and future arboviral threats. *Antiviral Research*, **85**, 328-345.
- Weaver, S.C. and Vasilakis, N.** (2009) Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect Genet Evol*, **9**, 523-540.
- Wikoff, W.R., Tsai, C.J., Wang, G., Baker, T.S. and Johnson, J.E.** (1997) The structure of cucumber mosaic virus: cryoelectron microscopy, X-ray crystallography, and sequence analysis. *Virology*, **232**, 91-97.
- Willits, D., Zhao, X., Olson, N., Baker, T.S., Zlotnick, A., Johnson, J.E., Douglas, T. and Young, M.J.** (2003) Effects of the Cowpea chlorotic mottle bromovirus [beta]-hexamer structure on virion assembly. *Virology*, **306**, 280-288.
- Xiao, C., Bator, C.M., Bowman, V.D., Rieder, E., He, Y., Hebert, B., Bella, J., Baker, T.S., Wimmer, E., Kuhn, R.J. and Rossmann, M.G.** (2001) Interaction of coxsackievirus A21 with its cellular receptor, ICAM-1. *J Virol*, **75**, 2444-2451.
- Yu, X., Jin, L. and Zhou, Z.H.** (2008) 3.88 Å structure of cytoplasmic polyhedrosis virus by cryo-electron microscopy. *Nature*, **453**, 415-419.
- Zhang, P., Mueller, S., Morais, M.C., Bator, C.M., Bowman, V.D., Hafenstein, S., Wimmer, E. and Rossmann, M.G.** (2008) Crystal structure of CD155 and electron microscopic studies of its complexes with polioviruses. *Proc. Natl. Acad. Sci. U S A*, **105**, 18284-18289.
- Zhang, W., Heil, M., Kuhn, R.J. and Baker, T.S.** (2005) Heparin binding sites on Ross River virus revealed by electron cryo-microscopy. *Virology*, **332**, 511-518.
- Ziegler-Graff, V. and Brault, V.** (2008) Role of vector-transmission proteins. *Methods in molecular biology Clifton, N.J.*, **451**, 81-96.