

Université de Montréal

Analyse moléculaire de la séquence leader
du virus de l'artérite équine

par

Ali Kheyar

Département de microbiologie et immunologie

Faculté de médecine

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de
Philosophiæ Doctor (Ph.D.)
en microbiologie et immunologie



Octobre, 2000

© Ali Kheyar, 2000

W
4
N58
2001
No. 112

Page d'identification du jury

Université de Montréal
Faculté des études supérieures

Cette thèse intitulée :

Analyse moléculaire de la séquence leader
du virus de l'artérite équine

présentée par :
Ali Kheyar

a été évaluée par un jury composé des personnes suivantes :

Dr Amer Silim: (Président rapporteur), département de pathologie et microbiologie,
Faculté de médecine vétérinaire, Université de Montréal

Dr Denis Archambault: (Directeur de recherche), département des sciences biologiques et
immunologie, Université du Québec à Montréal

Dre Caroline Alfieri: (Codirectrice de recherche), département de microbiologie et
immunologie, Université de Montréal

Dr Samir Mounir: (Membre du jury), Biochem Thérapeutique, Laval

Dre Estella Cornaglia: (Examinatrice externe), Biovet.

Thèse acceptée le :

SOMMAIRE

Le virus de l'artérite équine (VAE) cause une maladie infectieuse et contagieuse chez les équidés. Il est membre prototype de la famille des *Arteriviridae* du nouvel ordre des *Nidovirales*. C'est un virus enveloppé avec un génome d'ARN de 12,704 nucléotides (nt), monocaténaire, polyadénylé et de polarité positive. Durant l'infection virale, plusieurs ARN sous-génomiques (ARNsg) de tailles différentes avec une extrémité 3' co-terminale et une séquence leader (SL) 5' commune sont générés. Chez le VAE, très peu d'information était disponible quant aux différents rôles que pouvait jouer la SL, excepté son implication dans la synthèse des ARNsg par le biais de son motif pentanucléotidique de jonction situé à son extrémité 3' terminale. Par ailleurs, la séquence quasi complète du VAE a été déterminée en 1991 excepté l'extrémité 5' terminale. Il fut impératif de définir cette séquence manquante avant d'entamer son analyse. Ainsi, quatre souches du VAE ont été utilisées afin de déterminer l'extrémité 5' du génome viral. Pour amplifier cette région, une procédure SLIC ("Single-stranded ligation to single strand cDNA") a été utilisée. Après séquençage, les premiers résultats ont révélé la présence de 17 nt additionnels à l'extrémité 5' du génome viral et l'émergence d'un codon d'initiation ouvrant un cadre de lecture (ORF) intraleader (IL) de 111 nt dans la SL. Étant donné l'existence de variations de virulence entre des souches différentes du VAE indépendantes des variations antigéniques, des études comparatives au niveau de la séquence primaire et des structures secondaires de la SL complète de 11 souches/isolats du VAE ont été réalisées. L'analyse des séquences nucléotidiques primaires a révélé des variations mineures avec un degré élevé d'identité de 92,2 à 98,5%. Quant aux structures secondaires de l'ARN, 5 patrons prédits ont été retenus. Tous les isolats/souches ont présenté 3 boucles en épingle (A, B, et C) conservées et représentées par les mêmes positions de nt malgré quelques substitutions. Une quatrième boucle en épingle D fut observée chez six des isolats. En se basant sur la présence ou l'absence de celle-ci, les isolats du VAE ont été répartis dans deux génogroupes distincts.

L'ORF IL était conservé parmi tous les isolats/souches du VAE. La fonction et l'expression de cet ORF IL ont ensuite été étudiées. Une construction plasmidique dans le vecteur pUC-18 contenant l'ADN complémentaire (ADNc) correspondant à l'ARN 7 du VAE contenant la séquence leader et celle codant pour la nucléoprotéine (N) a été développée. À partir de cette construction parentale, des mutations ciblant quelques nt autour du codon d'initiation de l'ORF IL ont été introduites par PCR pour obtenir des

codons AUG dans un contexte optimal pour l'initiation de la traduction (Kozak), dans un contexte défavorable et un AUG abrogé. Des tests de transcription et de traduction *in vitro* ont été effectués. À l'exception de la construction avec un AUG abrogé, ces tests ont permis de détecter une expression *in vitro* du peptide IL d'un poids moléculaire de 4,3 kDa. Ce fut la première preuve expérimentale chez les nidovirus de l'expression *in vitro* d'un ORF localisé dans une région leader considérée comme non-codante. D'autres analyses à l'aide de ces constructions ont ciblé l'impact de ces mutations sur le niveau d'expression *in vitro* de la protéine N du VAE. Les plasmides avec un AUG IL optimal ou dans le contexte consensus de Kozak montraient des niveaux de traduction similaire au clone parental. Par contre ceux dont le codon AUG a été abrogé ou a été placé dans un contexte défavorable, les niveaux de traduction ont plutôt doublé. Ceci indique que l'ORF IL était probablement impliqué dans l'atténuation de l'expression de la protéine N, suggérant que la synthèse des protéines virales du VAE obéit au mécanisme du "leaky scanning" pour l'initiation de la traduction.

Ces observations ont ensuite mené à des études à une échelle plus complète sur l'impact de cet ORF IL sur le cycle de réplication virale. À cette fin, deux mutants dérivés du clone d'ADNc infectieux du VAE pBRNX1.38 (Glaser et al., 1999) ont été produits; le premier contenait une substitution G à A à la position 16 (AUG IL abrogé), alors que l'autre contenait une substitution T à A à la position 11 (AUG IL optimal). Après transfection des cellules BHK21 avec des transcrits d'ARN dérivés de ces clones mutants, une restauration rapide et spontanée de la séquence au type sauvage a été observée durant les passages viraux subséquents. De plus, l'utilisation des virus mutants dans la technique des plages de lyse a mis en évidence la présence de deux types de plages: un type minuscule et un type large. L'analyse par séquençage a démontré que les plages minuscules portaient le génotype de chacun des AUG IL mutés alors que les plages larges portaient le génotype original du virus sauvage. Ainsi, la restauration des phénotypes des plages de lyse corrélait avec la restauration des génotypes sauvages autour du codon AUG IL. Ces résultats indiquent que les mutations aux positions 11 et 16 étaient responsables de l'atténuation de la réplication et de l'infectivité virale. En conclusion, les bases autour du codon AUG IL du VAE correspondant à la séquence consensus 5' UGUAUGG 3' sont indispensables à la réplication optimale du VAE et que des mutations à son niveau sont soumises à des pressions sélectives pour restaurer et maintenir cette séquence consensus naturelle.

TABLE DES MATIÈRES

SOMMAIRE	iii
TABLE DES MATIÈRES.....	v
LISTE DES TABLEAUX	x
LISTE DES FIGURES	xi
LISTE DES SIGLES ET ABRÉVIATIONS.....	ix
REMERCIEMENTS	xv
DÉDICACE	xvi
AVANT-PROPOS	xviii
CHAPITRE I. INTRODUCTION.....	1
(A) L'ARTÉRITE VIRALE ÉQUINE.....	2
1. Définition	2
2. Historique	2
3. L'artérite virale équine	2
3. 1. Symptômes cliniques.....	2
3. 2. Pathogénèse	4
3. 3. Pathologie.....	4
3. 3. 1. Lésions macroscopiques.....	4
3. 3. 2. Lésions histologiques	4
4. Épidémiologie.....	5
4. 1. Distribution géographique	5
4. 2. Transmission de la maladie	6
4. 3. Persistance virale	6
5. Immunologie.....	8
6. Diagnostic	9
6. 1. Isolement et identification du virus	9
6. 2. Diagnostic sérologique	9
6. 3. Diagnostic basé sur des méthodes moléculaires.....	11
7. Traitement et Prophylaxie	12
7. 1. Traitement.....	12
7. 2. Vaccination.....	12

(B) VIROLOGIE

1. Classification	14
2. Morphologie	16
2.1 Caractéristiques en microscopie électronique	16
3. Composition structurale du virion	19
3.1 Génome et organisation génétique	19
3.2. Protéines structurales virales	20
3.2.1. Protéine de la nucléocapside (N).....	20
3.2.2. Protéine de la membrane (M).....	21
3.2.3. Glycoprotéine G _L	22
3.2.4. Glycoprotéine G _S	23
3.2.5. Protéine E	23
3.3. Cas des ORF3 et ORF4	24
3.4. Protéines virales non structurales	25
3.4.1. Polyprotéines de la réplicase virale	25
4. Propriétés physico-chimiques.....	26
5. Culture cellulaire et caractéristiques biologiques.....	27
6. Réplication virale.....	29
7. Transcription des ARNs viraux.....	31
7.1. Transcription discontinue lors de la synthèse des brins positifs	
d'ARNm sg.....	32
7.2. Transcription discontinue lors de la synthèse des brins négatifs	
d'ARN sg.....	33
(C) SÉQUENCE LEADER	35
1. Définitions et généralités.....	35
2. SL des Nidovirus	36
2.1. SL des Coronavirus	36
2.2. SL des Torovirus	38
2.3. SL des Artérovirus.....	39
2.3.1. VAE.....	39
2.3.1.1. Bref historique :Séquence leader et ARNm	39
2.3.1.2. Étude prospective complémentaire.....	41

2. 3. 1. 3. Données plus récentes	41
2. 3. 2. SL du VFHS	44
2. 3. 3. SL du VLD	46
2. 3. 4. SL du VSRRP.....	47
3. Rôles de la SL.....	51
3. 1. Rôles de la SL chez les coronavirus.....	51
3. 2. Rôles de la SL chez d'autres modèles de virus à ARN.....	51
3. 3. Rôles de la SL chez des modèles de virus à ADN	52
3. 4. Rôles de la SL chez des modèles d'ARNm eucaryotes.....	52
 CHAPITRE I.I. PROBLÉMATIQUE ET OBJECTIFS.....	 54
1. Objectif 1: Détermination de la séquence "leader" complète du VAE.....	55
2. Objectif 2: Étude de la stabilité de la SL.....	55
3. Objectif 3: Étude des rôles de la séquence "leader" dans la réplication.....	
du VAE.....	58
 CHAPITRE III	
Article 1. Sequence determination of the extreme 5' end of the equine arteritis virus leader region. Kheyar, A. , St-Laurent, G., and Archambault, D. 1996.	
Virus Genes 12:291-295.....	59
Abstract	60
Acknowledgments	63
References	63
 CHAPITRE IV	
Article 2. Nucleotide sequence and genetic analysis of the leader region of Canadian, American and European equine arteritis virus isolates.	
Kheyar, A. , St-Laurent, G., Diouri, M., and Archambault, D. 1998.	
Can. J. Vet. Res. 62:224-230.....	65
Abstract	66
Résumé	66
Acknowledgments	71
References	71

CHAPITRE V

Article 3. The intraleader AUG nucleotide sequence context per se is important for equine arteritis virus replication. **Ali Kheyar**, Antoine A. F. de Vries, Peter J. M.

<i>Rottier, and Denis Archambault.</i> Article en préparation pour publication.....	74
ABSTRACT.	75
INTRODUCTION.....	76
MATERIALS AND METHODS.....	77
Cells, virus, and antisera.....	77
RNA extraction and oligonucleotide primers.....	77
Construction of EAV mRNA7 expression plasmids	78
In vitro RNA transcription, translation experiments and radioimmunoprecipitation.....	82
Generation of full-length EAV cDNA clones with mutations in the EAV leader sequence.....	82
RNA transfection of cells and recovery of recombinant viruses.....	84
Serial passage and genetic analysis of recombinant viruses	84
Plaque assays	84
Growth curves	85
RESULTS	85
In vitro expression and significance of the EAV intraleader peptide.....	85
Genetic stability of mutated recombinant viruses	90
Kinetics of recombinant virus growth and of intracellular protein synthesis.....	93
Characterization of plaque formation.....	93
DISCUSSION	98
ACKNOWLEDGEMENTS	101
REFERENCES	102
CHAPITRE V. DISCUSSION GÉNÉRALE.....	108
CHAPITRE VI. CONCLUSION.....	126
CHAPITRE VII. BIBLIOGRAPHIE	131
CHAPITRE VIII. APPENDICE	167

LISTE DES TABLEAUX

CHAPITRE IV. Article 2

TABLE I. Characteristics of equine arteritis virus (EAV) isolates/strains used in this study	67
TABLE II. Nucleic acid sequence identities (%) of the leader sequence of equine arteritis virus (EAV) isolates/strains.....	68

CHAPITRE V. Article 3

TABLE 1. Oligonucleotide primers used in RT-PCR or PCR amplifications.	79
---	----

LISTE DES FIGURES

CHAPITRE I. INTRODUCTION

- Figure 1. Morphologie, organisation génomique et cycle de
réplication du VAE..... 17

CHAPITRE II. Article 1

- Figure 1. Sequence determination of the extreme 5' end of the leader
region of various EAV isolates..... 62
- Figure 2. Complete sequence of the EAV Bucyrus strain leader region 63

CHAPITRE III. Article 2

- Figure 1. Alignment of the nucleotide sequences of the leader region
of 9 arteritis virus isolates and the Arvac vaccine strain
with the ATCC Bucyrus reference strain. 68
- Figure 2. Predicted secondary structure of the leader sequence of eleven EAV
isolates/strains using the FOLD program of
the GCG software package..... 69
- Figure 3. Nucleotide sequences corresponding to the common stem-
loops in the leader sequence of eleven EAV isolates/strains. 70

CHAPITRE IV. Article 3

- Figure 1. Schematic representation of the EAV RNA 7-specific
plasmid and its derivatives 80
- Figure 2. *In vitro* synthesis of the EAV IL peptide..... 86
- Figure 3. Effect of the ILO start codon and its direct surroundings
on the *in vitro* translation of the EAV N protein. 88
- Figure 4. Sequence analysis of the ILO translational start site of
recombinant virus mutants after different cell passages..... 91
- Figure 5. Growth curves of the EAV recombinant viruses..... 94
- Figure 6. Plaque morphology of EAV recombinant viruses..... 96

LISTE DES SIGLES ET ABRÉVIATIONS

AA :	acide aminé
ADAR :	« adenosine deaminase that acts on RNA »
ADN:	acide désoxyribonucléique
ADNc:	ADN complémentaire
AMV:	« avian myeloblastosis virus »
ARN :	acide ribonucléique
ARNm :	ARN messenger
ARN sg :	ARN sous-génomique
ATCC :	« American Type Culture Collection »
AVE :	artérite virale équine
BCV :	« bovine coronavirus »
BEV:	virus Berne
BHK-21 :	cellules de bébé de hamster
BSA :	« bovine serum albumine »
BSC-1 :	cellules de rein de singe
BRV:	virus Breda
CaMV :	« cauliflower mosaic virus »
cDNA:	« complementary DNA »
CITV :	Comité International de Taxonomie des Virus
CPE:	« cytopathic effect »
Da :	daltons
DEAE :	diéthylaminoéthyl
DEPC:	diéthylpyrocarbonate
DI :	défective interférente
DI ₅₀ :	doses infectieuses à 50 % de culture tissulaire
DICT50 :	doses infectieuses infectant 50 % de culture de tissu
DNA:	« deoxyribonucleic acid »
DMEM :	« Dulbecco's modified Eagle's medium »
dsRAD :	« double-stranded RNA adenosine deaminase »
EAV:	« equine arteritis virus »
ECP :	effet cytopathique

EDTA :	« ethylenediaminetetraacetic acid »
ELISA :	« Enzyme linked immunosorbent assay »
EQ :	« equine dermis »
EVA:	« equine viral arteritis »
FBS :	« fetal bovine serum »
GCG:	« Genetic Computer Group »
G _L :	grande («large») glycoprotéine
GP3 :	glycoprotéine membranaire de l'ORF3
G _S :	petite («small»)glycoprotéine
h :	heure
HEPES :	N-(2-hydroxyethyl)piperazine-N'-(2-ethanesulfonic acid)
HmLu :	cellules de poumon de hamster
hpi :	h post-infection
HS :	cellules de tumeur de hamster transformées par le virus de l'hépatite canine
HT-7 :	cellules de tumeur de hamster transformées par le virus de l'hépatite canine
IG:	« Intergenic sequence »
IgG :	immunoglobuline G
IgM :	immunoglobuline M
IL:	intraleader
ILO:	« intraleader ORF »
IRES:	« Internal Ribosomal Entry Site »
IU:	« International Units »
JINET :	cellules de rein de singe <i>Cynomolgus</i>
kb :	kilobases
KLH :	Keyhole Limpet Hemocyanin
kDa :	kiloDaltons
kpb :	kilo-paires de bases
LDV:	« lactate dehydrogenase-elevating virus »
LLC-MK2 :	cellules de rein de singe <i>Rhesus</i>
log ₁₀ :	logarithme décimale
MA104:	lignée de cellules de rein de singe vert africain
MBD :	« metal-binding domain »
MHV :	« mouse hepatitis virus »

min :	minute
MOI:	« multiplicity of infection »
mRNA:	« messenger RNA »
NBL-6 :	cellules de derme équin
nm :	nanomètre
NP-40 :	Nonidet P-40
nt :	nucléotide
ORF :	« open reading frame »
P5	« passage number 5 »
P10	« passage number 10 »
PAGE :	« polyacrylamide gel electrophoresis »
PCP :	« papainlike cysteine protease »
pi :	« post infection »
PM :	poids moléculaire
pns :	protéine non structurale
PAA :	« polyacrylamide »
PBS :	« phosphate-buffered saline »
PBSS :	« phosphate-buffered salt solution »
PPLO::	« pleuropneumonia-like organisms »
Protéine E :	Protéine de l'enveloppe
Protéine M :	Protéine de la membrane
Protéine N :	Protéine de la nucléocapside
PRRSV :	« porcine reproductive and respiratory syndrome virus »
REL :	réticulum endoplasmique lisse
RIPA :	« radioimmunoprecipitation assay »
RK-13 :	cellules de rein de lapin
RNase:	ribonucléase
RT-PCR :	« Reverse Transcription-Polymerase Chain Reaction »
SHFV :	« simian hemorrhagic fever virus »
SL :	séquence leader
RNA:	« ribonucleic acid »
SDS:	« Sodium dodecyl Sulfate »
SP :	« serine-protease » ou « chymotrypsinlike serine protease »

SV40 :	« simian virus 40 »
TAP :	« tobacco acid pyrophosphatase »
TAS :	« transcription-associated sequence »
TCID50 :	« median tissue culture infective dose »
TRS :	« transcription-regulating sequence »
ufc/ml :	unités formant des colonies
ufp/ml :	unités formant des plages par millilitre
UV:	Ultraviolets
VAE :	virus de l'artérite équine
VERO :	cellules de rein de singe vert africain
VFHS :	virus de la fièvre hémorragique simienne
VHM:	virus de l'hépatite murine
VIH1:	virus de l'immunodéficience humaine type 1
VL:	virus Lelystad
VLD :	virus de la lactate déshydrogénase élevée
VSRRP :	virus du syndrome reproducteur et respiratoire du porc

REMERCIEMENTS

Mes remerciements les plus sincères s'adressent;

d'abord à mon directeur de recherche, Dr Denis Archambault, pour m'avoir accueilli dans son équipe de laboratoire, pour sa supervision et son support financier;

en deuxième lieu monsieur Gilles St-Laurent, technicien de laboratoire, pour son aide technique, sa patience et ses diverses discussions enrichissantes;

la Dre Caroline Alfieri d'avoir accepté d'être ma co-directrice et pour ses divers encouragements;

le Dr Yacine Abed pour ses conseils et encouragements divers;

le Dr A. A. de Vries de l'Université Utrecht (Pays-Bas) pour nous avoir fourni son clone infectieux du VAE, ses suggestions et discussions dans la rédaction du dernier article;

le Dr P. P. Rottier pour ses suggestions et discussions dans la rédaction du dernier article l'Université de Montréal, pour m'avoir accordé une bourse d'excellence;

l'ambassade d'Algérie au Canada pour m'avoir supporté avec la bourses d'exemption des frais de scolarité;

mes confrères et mes consoeurs de laboratoire, actuels et précédents, qui ont contribué à créer une ambiance agréable de travail au laboratoire;

mes amis et tous les membres de ma famille pour leur soutien moral sans cesse et leur encouragement au cours de mes études en particulier, mon père qui me disait souvent en ses propres termes : *Adhifetseh Er-Rebi Felekk Ammi Anda Thellet thevvoura N'l-Khir*, et ma grande sœur qui me disait toujours en ses propres termes : *Adhis-Soufegh Er-Rebi Laâthavikk Agh-Thafeth*.

Et enfin je remercie spécialement le personnel de la section de Virologie du Département de Microbiologie et Immunologie de l'Université de Montréal ainsi que le personnel du Département des Sciences Biologiques de l'Université du Québec à Montréal.

Ces travaux de recherche ont été effectués en majeure partie grâce à l'appui financier du Conseil National de Recherche en Sciences Naturelles et en génie.

DÉDICACE

*À ma mère,
mon père
ma soeur
mes frères
et Souâd*

AVANT-PROPOS

Avec l'accord de M. Fernand A. Roberge, vice doyen de la faculté des études supérieures, secteur des sciences de la santé, cette thèse a été rédigée dans un format « par article » et suite aux directives du Guide de présentation publié par la Faculté des études supérieures, édition 1994.

Le corps de la thèse est divisé en sept chapitres. Le chapitre I contient une revue de littérature divisée en trois sous-chapitres, le premier traitant la maladie causée par le virus de l'artérite équine (VAE), le second traitant la virologie et la biologie moléculaire du VAE et le dernier traitant les la séquence leader en général et en particulier chez les nidovirus. Ce premier chapitre servira de source essentiel pour la rédaction d'un article de revue portant sur le sujet.

Le chapitre II traite la problématique et les objectifs tracés pour entamer ce projet.

Le chapitre III contient le premier article dans sa version originale publiée. Il correspond à une partie des matériels et méthodes et résultats répondant au premier objectif tracé pour ce projet.

Kheyar, A., St-Laurent, G., and Archambault, D. 1996. Sequence determination of the extreme 5' end of the equine arteritis virus leader region. 12: 291-295.

Le chapitre VI contient le second article dans sa version originale publiée. Il correspond à une autre partie des matériels et méthodes et résultats répondant au deuxième objectif tracé pour ce projet.

Kheyar, A., St-Laurent, G., Diouri, M., and Archambault, D. 1998. Nucleotide sequence and genetic analysis of the leader region of Canadian, American and European equine arteritis virus isolates. Can. J. Vet. Res. 62:224-230.

Mr Gilles St-Laurent a apporté une contribution dans l'étape de séquençage de la partie 3' terminale de la séquence leader de plusieurs souches du virus de l'artérite équine. Quant au Dr Mounir Diouri, il a contribué dans l'amplification par RT-PCR de l'extrémité 5' de la souche vaccinale ARVAC du VAE.

Le chapitre V contient le troisième article dans sa version prête à être soumise pour publication dans le périodique *Journal of Virology*. Il correspond à une partie des matériels et méthodes et résultats répondant au dernier objectif tracé pour ce projet.

Ali Kheyar, Antoine A. F. de Vries, Peter J. M. Rottier, and Denis Archambault. Mutational analysis of the equine arteritis virus intraleader open reading frame start codon. Article soumis pour publication

Cet article rassemble les résultats de deux principaux travaux, l'un concernant l'expression *in vitro* du peptide intraleader du VAE et son impact sur l'expression de la protéine de la nucléocapside du VAE. L'autre partie concerne l'utilisation d'un clone infectieux de VAE pour évaluer l'impact de mutations ponctuelles introduites autour du codon IL du VAE sur la traduction et la réplication du VAE *in vivo*.

Les Dr Antoine A. F. de Vries, et Peter J. M. Rottier ont contribué au développement du clone infectieux du VAE utilisé dans ce chapitre (Glaser et al., 1999).

Les Drs D. Archambault, A. A. De Vries et P. P. Rottier ont apporté une contribution importante pour la rédaction du dernier article.

Le chapitre VI contient une discussion générale de l'ensemble de ces travaux

Le chapitre VII est le récapitulatif sous la forme d'une conclusion générale de la thèse.

Le chapitre VIII rassemble toutes les références bibliographiques citées dans le corps de la thèse.

Enfin le dernier chapitre IX représente une liste d'autres publications comme contributions supplémentaires dans le cadre du projet de l'artérite équine.

Ma contribution personnelle s'est faite à tous les niveaux de ce projet et inclut une part substantielle des différentes étapes mentionnées dans le corps de la thèse. Il va sans dire que tous ces travaux ont été activement supervisés par le Dr Archambault, qui a participé aussi de façon substantielle à la rédaction de tous les articles.

CHAPITRE I

INTRODUCTION

(A) L'ARTÉRITE VIRALE ÉQUINE

1. Définition

L'artérite virale équine (AVE) est une maladie infectieuse contagieuse affectant exclusivement les équidés (les populations équines et asines). Son agent étiologique est le virus de l'artérite équine (VAE). Celui-ci fut isolé pour la première fois en 1953 par Doll et al. (1957a; 1957b), à partir du tissu pulmonaire de deux fœtus avortés pendant une épizootie de maladie caractérisée par des troubles respiratoires et abortifs dans une ferme de chevaux Standardbred près de Bucyrus, Ohio, États-Unis. Jones et al. (1957) furent les premiers à décrire les lésions vasculaires associées à l'infection expérimentale des chevaux avec le nouveau virus d'où l'appellation de la maladie d'artérite virale équine.

2. Historique :

Depuis le premier isolement du virus (VAE) par Doll et ses collaborateurs (1953), des analyses comparatives des différentes publications parues auparavant et ressemblant à l'AVE ont attribué rétrospectivement des observations cliniques à des infections au VAE (Mumford, 1985). Ainsi, la maladie causée par le VAE fut décrite pour la première fois vers la fin du XIX^{ème} siècle (Pottie 1888). La maladie portait alors depuis longtemps plusieurs synonymes selon la variabilité des signes cliniques apparents (Mumford 1985). On lui attribuait des noms comme l'influenza équin (Pottie 1888), la maladie des yeux roses ou « pink-eye » (Clark 1892), la septicémie aiguë, la cellulite infectieuse, la cellulite épizootique, l'érysipèle à influenza (Bergman 1913) et la fièvre typhoïde du cheval (Brion et al., 1967).

3. L'artérite virale équine

3. 1. Symptômes cliniques

L'apparition et la sévérité de l'infection artériverale dépend de plusieurs facteurs tels que la souche virale, la dose infectieuse, l'âge de l'animal infecté, son sexe, ses conditions

générales et d'environnement, la voie d'infection virale (McCollum et Bryans, 1973). Cependant, la plupart des infections naturelles sont asymptomatiques (Timoney et McCollum, 1993). Les signes cliniques varient entre les individus et entre les différentes manifestations épizootiques. Les symptômes les plus sévères ont été enregistrés durant la première manifestation de VAE reconnue en 1953 (Doll et al., 1953). La période d'incubation varie de 3 à 14 jours. Dans les formes aiguës de la maladie, les symptômes suivants ont souvent été observés: des troubles de l'état général dominés par de la pyrexie avec une température corporelle allant jusqu'à 41°C, de l'anorexie, de la dépression et l'œdème des membres postérieurs rendant la démarche raide. Des symptômes respiratoires rappellent ceux causés par le virus influenza équin avec de la conjonctivite et l'œdème palpébral avec des écoulements lacrymaux, de la rhinite et des écoulements nasaux séreux, parfois de la dyspnée et de la toux. De plus, il y a des signes de dermatites focales localisées sur la tête ou tout le corps, l'œdème diffus dans la région ventrale de l'abdomen en particulier au niveau du prépuce et du scrotum ou au niveau de la glande mammaire, de l'ictère, de la leucopénie et de l'avortement chez les juments (Jones, 1969; Clayton, 1986; Collins et al. 1987; McCollum and Swerczek, 1978). Bien que la plupart des chevaux adultes infectés naturellement guérissent suivant la fin de la virémie, la mortalité des jeunes poulains peut survenir après une phase fatale de pneumonie broncho interstitielle et de nécrose intestinale (Carman et al., 1988; Vaala et al., 1992).

Chez les juments infectées, l'avortement survient de 10 à 33 jours après des infections qui généralement sont sans signes cliniques apparents. Il est souvent associé à la phase terminale de la fièvre ou à la phase initiale de la convalescence (Bryans et al. 1957; Timoney et McCollum, 1993). Les avortons sont expulsés accompagnés des membranes fœtales, sans signes préalables d'avortement menaçant. Dans des infections naturelles et expérimentales, les avortements ont été enregistrés entre 3 et 10 mois après le début de la gestation (Timoney et McCollum, 1993). Les taux d'avortement dans les manifestations épizootiques naturelles varient de 10 à 70 % avec une moyenne estimée à environ 50 % (Cole et al., 1986; Timoney et McCollum, 1993). Il semble que ces taux dépendent de plusieurs facteurs, tels la dose infectieuse ainsi que l'existence de certains isolats du VAE avec un pouvoir abortif distinct (Timoney et McCollum, 1993).

Les signes cliniques de l'infection expérimentale sont plus sévères que dans les infections naturelles. Ils sont parfois mortels. Le taux de mortalité élevé dans les infections expérimentales a souvent été attribué à la virulence de la souche de référence Bucyrus

(Estes et Cheville, 1970; Fukunaga et al., 1981). Il a atteint 50% (Bryans et al. 1957) à 100 % dans une étude d' inoculation intramusculaire de VAE à 12 chevaux adultes (Estes et Cheville, 1970).

3. 2. Pathogenèse

Suite à la pénétration du virus par la voie respiratoire, le site initial de sa multiplication semble prendre place au niveau des macrophages alvéolaires ainsi que des cellules endothéliales. Le deuxième jour, le virus atteint les nœuds lymphatiques bronchiaux d'où commence sa dissémination dans l'organisme via les vaisseaux sanguins soit sous la forme libre extracellulaire ou sous la forme intracellulaire dans les monocytes/macrophages infectés (Plagemann et Moennig, 1992). L'épithélium de certains organes internes (tels les glandes surrénales, le foie, les tubes séminifères, et la thyroïde), le mésothélium et les cellules de la média vasculaire ont tous été identifiés comme des sites secondaires de la réplication virale (Crawford et al., 1973; Timoney et McCollum, 1993).

3. 3. Pathologie

3. 3. 1. Lésions macroscopiques

Dans les formes aiguës suivant l'infection avec la souche Bucyrus, les lésions les plus visibles sont l'œdème, des hémorragies et la congestion éparpillés dans plusieurs tissus. Le stade terminal est caractérisé par des infarctus au niveau des tissus intestinal, pulmonaire et splénique ainsi que des accumulations des fluides dans les cavités pleurale et péritonéale (Prickett et al., 1973).

3. 3. 2. Lésions histologiques

La plupart de ces lésions sont la conséquence de nécroses vasculaires généralisées (Jones et al., 1957; Prickett et al., 1973). Il semble que la dégénérescence des vaisseaux sanguins débute avec des dommages des cellules endothéliales, suivi d'infiltrations des leucocytes au niveau de l'intima et la fragmentation de la couche limitante élastique

interne. L'hyperperméabilité vasculaire entraîne l'inflammation de la média ensuite de l'adventice des vaisseaux sanguins plus petits (Jones, 1969; Estes et Cheville, 1970).

Les lésions vasculaires sont prononcées dans la média des petites artères. Toutefois, les signes cliniques ont été attribués plutôt aux nécroses des veines et des vaisseaux lymphatiques étant donné que la guérison clinique coïncide avec la régénérescence de ces vaisseaux. Habituellement, cette régénérescence vasculaire se complète après 2 semaines. Par contre, la récupération des artères prend environ 2 mois (Jones et al., 1957). Selon Jones et al. (1957), ces différentes lésions vasculaires entraînant des nécroses vasculaires généralisées chez la jument mènent à une dysfonction au niveau des échanges vasculaires entre la mère et le fœtus, ce qui résulte en l'avortement. D'autres cliniciens (Coignoul et Chevilly, 1984) ont suggéré que l'avortement est la conséquence d'une myométrie nécrotique maternelle menant à une compression mécanique des vaisseaux placentaires ou une réduction de la production de la progestérone.

4. Épidémiologie

4. 1. Distribution géographique

Les études séroépidémiologiques ont été menées presque partout dans le monde. L'AVE est une maladie cosmopolite. Elle a été rapportée dans tous les continents et presque dans chaque pays où ces études ont été entreprises (Plagemann et Moennig, 1992). La prévalence est variable entre les différents pays et entre les animaux des différentes fermes et de différents âges. Malgré cette vaste distribution géographique, le VAE cause rarement des manifestations épizootiques. La première manifestation qui fut de fait la plus sévère fut celle déclarée en 1953 à Bucyrus (Doll et al, 1957a). D'autres épizooties ont été rapportées ultérieurement aux Etats-Unis (Timoney, 1984), au Canada (Clayton, 1987; Klavano, 1987), en Autriche (Jaksch, 1973), en Italie (Timoney et McCollum, 1990), en Pologne (Golnik et al., 1986) et en Suisse (Burki, 1970).

Au Canada, la première manifestation épizootique a été rapportée en Alberta en 1986 (Clayton, 1987). Sur 35 chevaux adultes, 10 montrèrent des signes cliniques sévères et sur 26 juments, 11 avortèrent (avec ainsi un taux d'avortement de 42 %). Depuis, d'autres cas cliniques sans manifestation sévère, ont été rapportés en 1987 en Ontario, en Colombie-Britannique et au Québec.

Le premier cas clinique au Québec a été diagnostiqué en 1987 (Vrins et al.). Cependant, très peu de données épidémiologiques y ont été répertoriées sauf celle rapportée par Elazhary et al. (1990). Des échantillons de sérums prélevés en 1985 sur 564 chevaux provenant de 89 écuries sur 97 existant dans le comté d'Hudson et de Saint-Lazare au Québec ont été testés par séroneutralisation. Il s'est avéré que 8,51 % des chevaux étaient séropositifs (Elazhary et al., 1990).

4. 2. Transmission de la maladie

La transmission de la maladie se déroule selon deux modes principaux. Le mode primaire est la voie respiratoire par contact direct avec des aérosols infectieux provenant des sécrétions nasopharyngiennes des chevaux infectés de façon aiguë (McCollum et al., 1970). Néanmoins, les divers produits corporels comme le sang, les matières fécales, les liquides lacrymaux et urinaires de même que les sécrétions vaginales peuvent contribuer à un moindre degré à la propagation horizontale de la maladie (Timoney et McCollum, 1993). L'infection indirecte via les équipements et le personnel est de moindre importance. De plus, l'inexistence de la propagation par l'intermédiaire des insectes et des arthropodes a été démontrée (Bürki et al, 1972).

La deuxième principale voie de transmission du VAE est la voie vénérienne (Timoney et McCollum, 1993). Les étalons infectés de façon aiguë, répandent souvent le virus via leur sperme. Le virus est facilement isolé à partir des écouvillons vaginaux suite au coït ou aux inséminations artificielles (Timoney et al., 1987a). De plus, l'infection devient persistante dans la majorité des étalons infectés. Cependant, même si les étalons sont guéris des symptômes cliniques, ils continuent à disséminer le virus et peuvent infecter de 85 à 100% des juments séronégatives (Timoney et al., 1987b). Il semble que la transmission placentaire du virus ne contribue pas directement dans la dissémination virale. Quand la jument devient infectée tardivement durant sa gestation, elle peut donner naissance à un poulain avec une infection acquise de façon congénitale (Vaala et al., 1992).

4. 3. Persistance virale

Vers la fin du XIX^{ème} siècle, Pottie (1888) et Clark (1892) constatèrent déjà que des étalons apparemment en bonne santé transmettaient la maladie qui ressemblait à l'AVE

à des juments au moment de l'accouplement. Ils postulèrent que l'agent causal demeurait dans le sperme même après la période de convalescence. Cependant, la confirmation virologique de l'existence d'un état de porteur à long terme de VAE chez les étalons n'a été obtenue que vers 1986 par le groupe de Timoney. Le taux moyen de la persistance virale a été estimé à 35 % en général (Timoney et al., 1987a, 1987b), bien qu'un taux aussi élevé que 65% a été rapporté dans un groupe de chevaux infectés expérimentalement (Neu et al., 1988).

Étant donné la diversité des facteurs qui influencent la variabilité des périodes de persistance, trois catégories de groupe de chevaux ont été proposées sur la base de la durée de la sécrétion virale après infection (Timoney et al., 1993). Les porteurs à court terme ou porteurs convalescents sécrètent le virus dans leur sperme pendant 2 à 5 semaines. Les porteurs à moyen terme ou les porteurs intermédiaires sécrètent le virus de 3 à 8 mois. Enfin, les porteurs à long terme ou les porteurs chroniques semblent sécréter le virus durant plus de 10 ans voire presque toute leur vie.

Le VAE a été isolé à partir de plusieurs tissus du tractus reproducteur avec des concentrations relativement plus élevées au niveau des glandes génitales accessoires supérieures et du canal déférent (Neu et al., 1988, Fugunaga et al., 1992). Selon Little et al. (1991), la production du VAE chez les étalons infectés de façon persistante semble être dépendante de la testostérone. Plusieurs observations ont en effet confirmé cette constatation. L'état de porteur persistant n'a jamais été rapporté chez des animaux castrés. Des hongres ayant été traités avec de la testostérone peuvent continuer à produire du virus dans leur semence contrairement aux hongres non traités (Little et al., 1991). De plus, des essais d'induction de la persistance virale chez des poulains (avant la puberté) ont souvent échoué, suggérant la nécessité de la présence de taux élevés de testostérone dans l'organisme animal pour permettre au virus de persister (Holyoak et al., 1993). Enfin la persistance virale n'existe pas chez les juments. Le VAE n'a jamais été isolé du tractus génital de juments séropositives après plus d'un mois d'infection (McCollum et Timoney, 1984). Par ailleurs, des essais d'isolement du virus chez des étalons susceptibles ayant été accouplés aux juments séropositives ont démontré l'inexistence de la transmission de la femelle vers le mâle par cette voie (Bürki et al., 1991; Chirnside, 1992). Le mécanisme par lequel la testostérone peut promouvoir l'établissement de la persistance virale reste encore inconnu bien que certains auteurs ont suggéré l'implication des effets immunosuppresseurs des androgènes empêchant ainsi l'élimination du virus (Little et al., 1991).

5. Immunologie

L'apparition des anticorps neutralisants coïncide avec l'élimination du virus infectieux de la circulation sanguine (Fukunaga et al., 1981). Ces anticorps ainsi que ceux fixant le complément sont détectables une semaine après l'infection naturelle ou expérimentale (Fukunaga et McCollum, 1977; McCollum, 1986). L'immunité humorale contre l'infection artévirale a été bien étudiée par plusieurs groupes. Il semble qu'elle n'est pas suffisante à elle seule pour prévenir la maladie car certaines études ont démontré l'apparition occasionnelle de la maladie chez des animaux possédant des anticorps neutralisants anti-VAE (McCollum, 1986). Curieusement, l'immunité cellulaire anti-AVE n'a pas été l'objet de recherche intensive qui serait complémentaire à instaurer la protection naturelle ou acquise. Le rôle de l'immunité humorale par contre s'est manifesté par l'effet protecteur du colostrum de juments immunisées chez les jeunes poulains (McCollum, 1976).

Les anticorps neutralisants persistent généralement pendant plusieurs années (plus que 7 ans) voire toute la vie de l'animal (Gerber et al., 1978). Certains travaux ont démontré une corrélation entre l'augmentation des titres en anticorps et l'âge de l'animal. Ceci a été expliqué par le fait que des réinfections subséquentes durant la vie de l'animal font maintenir voire augmenter les niveaux des titres et de l'immunité en général (Bürki et al., 1992.).

Le mécanisme de la neutralisation virale semble impliquer la fraction immunoglobuline G (IgG) des sérums immuns tardifs des chevaux, avec une très faible contribution des anticorps IgM qui ne sont présents que précocement après l'infection (Radwan et Burger, 1973a). Cependant, la neutralisation virale n'est pas attribuée exclusivement aux molécules d'IgG. En effet, l'efficacité de la neutralisation virale en outre se voit nettement favorisée et augmentée par la présence du complément ou d'anticorps secondaires anti-IgG (Radwan et Burger, 1973a, Radwan et Burger, 1973b). Deux principaux mécanismes de la neutralisation ont été suggérés. Le premier consiste en l'interférence physique entre le virus et son site d'attachement sur la cellule hôte par des d'anticorps à des concentrations élevées. Le second concerne la virolyse médiée par le complément à de faibles concentrations d'anticorps (Radwan et Burger, 1973b; Radwan et al. 1973, Radwan et Crawford, 1974). Les protéines virales contre lesquelles les anticorps sont produits seront évoquées plus loin.

6. Diagnostic

Il n'y a pas de symptômes pathognomoniques pour l'AVE. Le diagnostic ne peut donc se reposer sur l'observation des signes cliniques, vu que ceux-ci sont tellement variables, et que plusieurs autres maladies infectieuses et non infectieuses lui ressemblent. Ainsi le diagnostic différentiel se fait avec la majorité des infections virales affectant le tractus respiratoire supérieur telles que les infections associées à l'adénovirus équin, à l'herpèsvirus équin, au virus influenza équin, et au rhinovirus équin, ainsi que moindrement d'autres infections telles que l'anémie infectieuse des équidés, la leptospirose, le purpura hémorragique et les troubles de toxicose, de malnutrition ou d'urticaire (Huntington et al., 1990). Le diagnostic de laboratoire reste donc le moyen de détection du VAE. Il repose notamment sur l'isolement et l'identification du virus, la détection de l'antigène viral et/ou sur la sérologie. Cependant le diagnostic histopathologique peut apporter un support supplémentaire. En revanche, les méthodes moléculaires ne font pas encore partie du régime diagnostique de routine.

6. 1. Isolement et identification du virus :

L'isolement du virus se fait à partir de spécimens histologiques frais tels que les lavages et les écouvillons nasopharyngiaux, les écouvillons conjonctivaux et les de leucocytes sanguins. Les échantillons de la semence d'étalons doivent provenir d'éjaculats complets et riches en sperme (Timoney et McCollum, 1993). La propagation virale a été d'abord décrite sur des cultures cellulaires primaires de cheval (McCollum et al., 1973) pour ensuite être adaptées aux différentes lignées cellulaires notamment les cellules de rein de lapin (RK-13) pour l'observation des effets cytopathiques (Timoney et McCollum, 1993).

6. 2. Diagnostic sérologique

La détection des protéines virales par des tests de capture d'antigènes, d'immunodiffusion ou d'immunofluorescence n'a jamais été adoptée comme un outil valable pour le diagnostic de l'AVE à cause de leur faible sensibilité (Crawford et Henson, 1973; Timoney et McCollum, 1993). Cependant les tests de laboratoire les plus utilisés et

qui sont toujours approuvés sont des tests sérologiques comme la séroneutralisation virale dans des microplateaux (Senne et al., 1985), le test de la fixation du complément (Fukunaga et al., 1994; Fukunaga et McCollum, 1977), et à un degré moindre, le test de la réduction des plages (Hyllseth, 1969; McCollum, 1970). Ces différents tests ne permettent pas de distinguer entre les anticorps d'animaux souffrant d'une infection aiguë et ceux provenant d'animaux guéris ou en convalescence. De plus, ils présentent également le défaut de donner parfois des résultats faussement positifs (Fukunaga et al., 1994). Le test de la fixation du complément est rapide et simple. Néanmoins, il est peu sensible et est valable seulement pour le diagnostic d'infections récentes étant donné que les taux d'anticorps IgM baissent rapidement après environ 2 à 3 semaines d'infection (Fukunaga et McCollum, 1977). Par ailleurs, les sommets des titres d'anticorps neutralisants sont généralement atteints après 1 à 3 mois post-infection et peuvent même être détectés à ce niveau après plusieurs années post-infection (McCollum, 1970). La sensibilité de la séroneutralisation peut être nettement améliorée par l'ajout du complément (Hyllseth et Pettersson, 1970).

Actuellement, le test de la séroneutralisation est internationalement approuvé comme étant le test standard pour le séro-diagnostic de l'AVE vu ses meilleures sensibilité et spécificité par rapport aux autres tests (Senne et al., 1985). Les anticorps neutralisants sont dirigés contre la grande glycoprotéine G_L de l'enveloppe qui portent les déterminants neutralisants (Balasurya et al. 1993, Deregt et al. 1994). Toutefois, ce test est coûteux, long à réaliser et requiert la production et le maintien de grande quantité de stocks viraux.

Beaucoup de tentatives de mises au point d'autres tests immunoenzymatiques comme l'«Enzyme linked immunosorbent assay» (ELISA) et le test de l'immunobuvardage (Western blot) ont fait l'objet d'études de plusieurs groupes de laboratoire. Ceux ayant utilisé l'antigène viral total dans le test ELISA ont souvent produit des niveaux inacceptables de faux positifs surtout quand les échantillons sériques proviennent d'animaux vaccinés. Ces derniers développent des anticorps même contre les protéines cellulaires qui restent toujours comme des antigènes contaminant les préparations vaccinales commerciales (Cook et al., 1989). Plus récemment, une équipe japonaise (Kondo et al., 1998) a développé un test ELISA qui semble être plus sensible et spécifique grâce à la qualité hautement purifiée de la préparation virale utilisée.

Par ailleurs, d'autres groupes incluant le nôtre ont plutôt ciblé comme substrats antigéniques des protéines virales structurales recombinantes (protéines de fusion dans des systèmes eucaryotes ou procaryotes) comme la protéine de la nucléocapside N, de la

membrane M, la grande glycoprotéine G_L et/ou la petite glycoprotéine G_S pour les utiliser aussi bien dans l'ELISA que dans le test d'immunobuvardage (Chirnside et al., 1995a, b; Kheyar et al., 1997; Balasuriya et al., 1998). La plupart de ces essais sont encore loin d'être adoptés pour le diagnostic, vu leur faible spécificité qui est compromise (par exemple) par les protéines partenaires de fusion (Chirnside et al., 1995b, c; Kheyar et al., 1997) ou leur faible sensibilité due à leur incompatibilité avec le test de référence de séroneutralisation (Balasuriya et al., 1998)

6. 3. Diagnostic basé sur des méthodes moléculaires

C'est depuis le début des années 1990 que plusieurs tentatives de mise au point de tests de diagnostic basés sur la méthode utilisant la rétrotranscription suivie de l'amplification par la réaction de la polymérase en chaîne (RT-PCR) ont été rapportées. Chirnside et Spaan (1990) ont été les premiers à développer un test qui permettait d'obtenir une sensibilité de détection limite de 600 unités formant des plages (ufp)/ml dans le plasma séminal. Cependant, d'autres chercheurs ont plus tard réussi à augmenter la sensibilité à 2 doses infectieuses infectant 50 % de culture de tissu (DICT50) à partir de surnageant de cultures cellulaires, et 20 à 200 DICT50 à partir du sperme de cheval dépendamment de la région du génome ciblée par les amorces utilisées (St-Laurent et al., 1994). Sekiguchi et al. (1995) ont rapporté des méthodes de RT-PCR et de "nested-PCR", d'hybridation "dot blot", et de digestion avec des endonucléases de restriction permettant de distinguer 7 souches de VAE à partir de surnageants de cultures cellulaires infectées. Ils ont atteint un seuil de sensibilité de 0,5 à 5 unités formant des colonies (ufc/ml). Des taux de sensibilité de 100% et de spécificité de 97%, avec un seuil de détection de 2,5 ufc/ml à partir de plasma séminal d'animaux infectés naturellement ont été obtenus par Gilbert et al. (1997) en utilisant un test de "nested-PCR". Récemment, Ramina et al. (1999) ont développé un test de RT-PCR-ELISA qui était très sensible (de l'ordre de 4 DITC50/ml de surnageant de culture cellulaire). Plus récemment, Starick (1998, 1999) a comparé des méthodes de détection classiques aux méthodes moléculaires. Il a démontré que ces dernières étaient de loin plus efficaces. La méthode de RT-PCR suivie de la "nested-PCR" et d'hybridation "dot bot" semble être la meilleure procédure pour le diagnostic du VAE. Par conséquent, vu la rapidité d'exécution des méthodes moléculaires, il est actuellement recommandé de les intégrer officiellement dans le régime diagnostique routinier (Starick 1998; 1999). Cette

recommandation est encore sous observation probablement dans l'attente d'uniformiser des normes de standardisation.

7. Traitement et Prophylaxie

7. 1. Traitement

Comme dans la majorité des maladies virales, il n'existe pas de traitement spécifique contre l'AVE. Il repose seulement sur le traitement symptomatique particulièrement chez les animaux affectés par une forte fièvre et un œdème scrotal étendu. Dans ces cas, il est recommandé d'administrer des médicaments anti-inflammatoires non stéroïdiens en combinaison avec des substances diurétiques pour réduire les dommages sur la semence de l'animal, tout en permettant un repos sexuel des étalons infectés (Timoney et McCollum, 1993). Par ailleurs, pour éliminer la persistance virale chez les étalons, le seul moyen semble être la castration de l'animal ou le contrôle temporaire des niveaux de testostérone en circulation (Timoney et McCollum, 1993).

7. 2. Vaccination

Présentement, il existe commercialement deux formes de vaccins anti-VAE; un vaccin à virus vivant atténué appelé Arvac (Ford Dodge Laboratories, Fort Dodge, Iowa, USA) et un vaccin à virus chimiquement inactivé par la formaline appelé Artevac (Ford Dodge Laboratories). Le vaccin Arvac a été mis sur le marché en 1985 et est utilisé en Amérique du Nord. Le vaccin Artevac a été introduit en 1993 et est utilisé plutôt dans certains pays d'Europe comme la Grande Bretagne à titre expérimental (Plagemann et Moennig, 1994).

Le phénotype atténué du vaccin vivant a été développé par McCollum et ses collaborateurs (McCollum et al., 1961; 1962a; 1962b), en faisant subir à la souche Bucyrus du VAE 131 passages viraux sur des cellules de rein de cheval, suivis de 111 passages sur des cellules de rein de lapin RK-13, et enfin 24 passages sur des cellules « Equine dermis » NBL-6. Ces passages successifs furent réalisés sans dilutions virales et sans qu'il y eut une réduction notable de l'immunogénicité virale. L'utilité du vaccin vivant atténué a été démontrée à maintes reprises dans des études expérimentales d'infection de défi (challenge)

(Doll et al., 1968; Fukunaga et al., 1982; Harry et McCollum., 1981; McCollum, 1969; McCollum, 1981; 1986).

Il est recommandé en général de vacciner les étalons à environ 3 semaines avant l'accouplement. La vaccination des juments avec le virus atténué est par contre non recommandée notamment durant les deux derniers mois de gestation (Doll et al. 1968). Les poulains doivent recevoir leur vaccin de préférence après 6 semaines d'âge (Timoney et McCollum, 1993). Suite à la vaccination, les anticorps neutralisants se développent durant les 4 à 10 premiers jours induisant une protection clinique contre l'AVE pour une longue durée (plus d'un an). Cependant, la protection n'est jamais totale car la réplication virale et l'infection avec du virus virulent n'est pas compromise (Timoney et al., 1988).

Quant aux vaccins à virus inactivé il existe une forme avec adjuvant correspondant au vaccin Artevac, et une forme sans adjuvant. Dans ce dernier cas, les animaux sont immunisés avec deux injections intramusculaires à 4 semaines d'intervalle. Ceci requiert des doses élevées pour induire une réponse satisfaisante en anticorps neutralisants. De plus, les titres de neutralisation virale sont tellement variables et peu durables que seuls les animaux montrant des titres supérieurs à 1: 80 sont considérés comme étant protégés. Selon les suggestions de Fukunaga et al., (1992) il est recommandé de revacciner avec ce type de virus à chaque 6 mois à un an.

Il est à noter que tous ces vaccins n'arrivent pas à contourner le contrôle de la maladie. De plus, il n'existe pas de moyens de distinguer entre les animaux vaccinés et ceux infectés naturellement. Ceci nécessite à priori d'entreprendre des améliorations dans ce domaine. La vaccination anti-AVE ne procurant pas une protection absolue des animaux, des mesures prophylactiques supplémentaires doivent être entreprises. Des programmes pour prévenir la dissémination de la maladie sont instaurés de façon différente dans plusieurs pays (McKenzie, 1990; Timoney et McCollum, 1988). En général, ils consistent en des mesures de surveillance diagnostique et de vaccination combinées à des mesures hygiéniques et d'isolement des animaux infectés. Le sperme utilisé dans les inséminations artificielles doit faire l'objet de contrôles rigoureux. Les étalons porteurs sains peuvent continuer à être élevés avec des juments ayant déjà été exposées au virus, donc immunisées. Cependant, si ces juments ont été accouplées à ces étalons, elles doivent être gardées en isolement pendant environ trois semaines ou plus pour vérifier si elles n'ont pas contracté la maladie clinique (Timoney et McCollum, 1993).

(B) VIROLOGIE

1. Classification

Comme la recherche en virologie a connu des progrès exponentiels notamment durant les trois dernières décennies, la taxonomie virale semble être la discipline la plus affectée quant à sa stabilité dans le temps. Sur la base de sa composition en acide ribonucléique (ARN), en enveloppe lipidique, et la non contribution d'un insecte vecteur dans sa propagation, le VAE a d'abord été classé dans un groupe nommé *Hemovirus* avec le virus du choléra porcine, et le virus de la diarrhée virale bovine (Bürki, 1966). Par contre, Breese et McCollum (1970) proposèrent de le classer dans le groupe des arbovirus sur la base de sa morphologie générale. D'autres chercheurs, ayant étudié certaines caractéristiques physico-chimiques et morphologiques du VAE, l'ont classifié dans la famille des *Togaviridae*, tantôt dans le genre des *Alphavirus* (Arbovirus groupe A; Hyllseth et al., 1970; Horzinek et al., 1971), tantôt dans le genre des *Flavivirus* (Arbovirus groupe B; Bürki, 1970; Hyllseth, 1973). Un peu plus tard, Horzinek (1973a, 1973b) proposa plutôt de classer le VAE comme un togavirus non-arbovirus au même titre que d'autres membres de ce genre : les pestivirus et les rubivirus, ainsi que le virus de l'élévation de la lactate déshydrogénase, le virus de la fièvre hémorragique simienne et le virus de la fusion cellulaire. Cette proposition fut adoptée en 1977 par le Comité d'Étude sur les Arbovirus (Porterfield et al. 1978), puis en 1984 par le Comité d'Étude sur les *Togavirus*. Le VAE fut ainsi placé dans un nouveau genre *Arterivirus* dans la famille des *Togaviridae* (Westaway et al., 1985). Les virus appartenant à cette famille sont tous sphériques et enveloppés avec un diamètre de 40 à 60 nm et renfermant une nucléocapside icosaédrique de 25 à 35 nm qui à son tour contient une molécule d'ARN simple monocaténaire et de polarité positive de 4 à 6 x 10⁶ daltons (Da) (Plagemann, 1992).

Par ailleurs, les données des analyses moléculaires des années 1980 et 1990 ont montré que l'organisation de son génome ainsi que les stratégies de la réplication du VAE, de la transcription de ses ARN messagers (ARNm) et de la traduction de ses protéines supportent clairement que la famille des *Togaviridae* est la plus inappropriée pour le VAE et les virus apparentés (Cavanagh et al., 1994). Ces stratégies ressemblent étroitement à celle des coronavirus (van Berlo et al., 1982; Spaan et al., 1990; den Boon et al., 1991). C'est pour cela que le Comité International de Taxonomie des Virus (CITV) a émis en

1992 (Cavanagh et Horzinek, 1993) sa recommandation de déclasser les artérovirus de la famille des *Togaviridae*. Les principaux éléments de ressemblance entre les artérovirus, les coronavirus ainsi que les torovirus sont : l'existence d'un groupe d'ARNm sous-génomiques (ARN sg) co-terminaux à l'extrémité 3'; la présence d'une même séquence leader (SL) à l'extrémité 5' de chaque ARN sg; un changement de cadre de lecture du ribosome lors de l'expression des cadres de lecture [ou open reading frame (ORF)] ORF1a et ORF1b pour la synthèse de la polyprotéine donnant naissance aux protéines associées à la réplicase virale, et l'ordre et la fonction de certaines protéines structurales telles que la protéine de la nucléocapside et celle de la membrane (Cavanagh et al., 1994).

Suivant ces considérations, den Boon et al., (1991) ont proposé de classer les artérovirus avec les coronavirus et les torovirus dans une même superfamille des coronavirus. Toutefois, ce terme de superfamille n'est pas adopté dans la taxonomie virale. De plus, l'inclusion du genre *Arterivirus* dans la famille des *Coronaviridae* ne pouvait être adoptée, car, malgré les ressemblances mentionnées ci haut, il y a des différences structurales majeures. En effet, la particule virale des artérovirus est de moitié plus petite que celle des deux autres genres et sa nucléocapside est icosaédrique, alors que celle des coronavirus est hélicoïdale. De plus, les artérovirus ne possèdent pas de projections de surface typiquement proéminentes et leur génome, tout comme la taille du virion, est de loin plus court [12 à 15 kilobases (kb) vs environ 30 kb]. Les alternatives suggérées alors étaient soit l'adoption de trois genres au sein de deux sous-familles *Arterivirinae* (avec un seul genre *Arterivirus*) et *Coronavirinae* (avec deux genres *Coronavirus* et *Torovirus*) elles-mêmes au sein de la famille des *Coronaviridae*, soit l'adoption d'un nouvel ordre et deux familles *Arteriviridae* et *Coronaviridae* (Cavanagh et al., 1994).

C'est seulement durant le X^{ième} Congrès International de Virologie tenu en 1996, que le Comité International de la Taxonomie des Virus (CITV) a approuvé officiellement cette dernière suggestion pour mettre fin à ce long litige taxonomique. Il y a eu rassemblement des quatre membres d'artérovirus dans un seul genre *Arterivirus*, lui-même appartenant à la nouvelle famille *Arteriviridae*. Les quatre membres de ce genre sont le VAE étant le membre prototype de la famille; le virus de l'élévation de la lactate déshydrogénase (VLD) qui affecte l'espèce murine, le virus du syndrome reproducteur et respiratoire du porc (VSRRP) et le virus de la fièvre hémorragique simienne (VFHS) affectant exclusivement les singes. Cette nouvelle famille a été placée avec la famille des *Coronaviridae* (renfermant les deux genres *Coronavirus* et *Torovirus*) dans un nouvel ordre

appelé *Nidovirales* (Cavanagh, 1997). Ce terme provient du latin *nidus* qui signifie nid et fait référence à la principale caractéristique commune de ces différents membres et qui est le set d'ARN sg coterminaux du côté 3' produits durant l'infection cellulaire (de Vries et al., 1997).

2. Morphologie

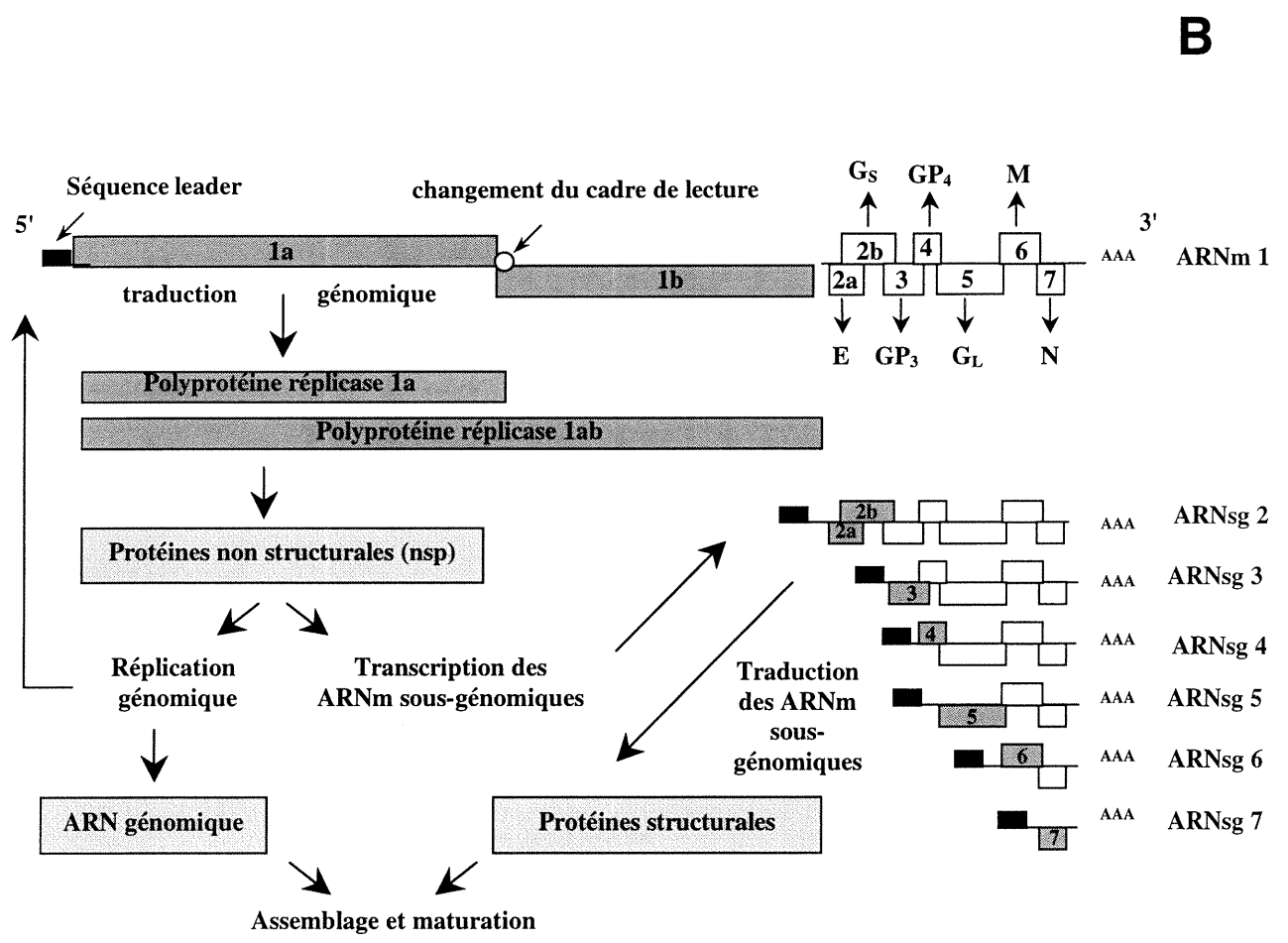
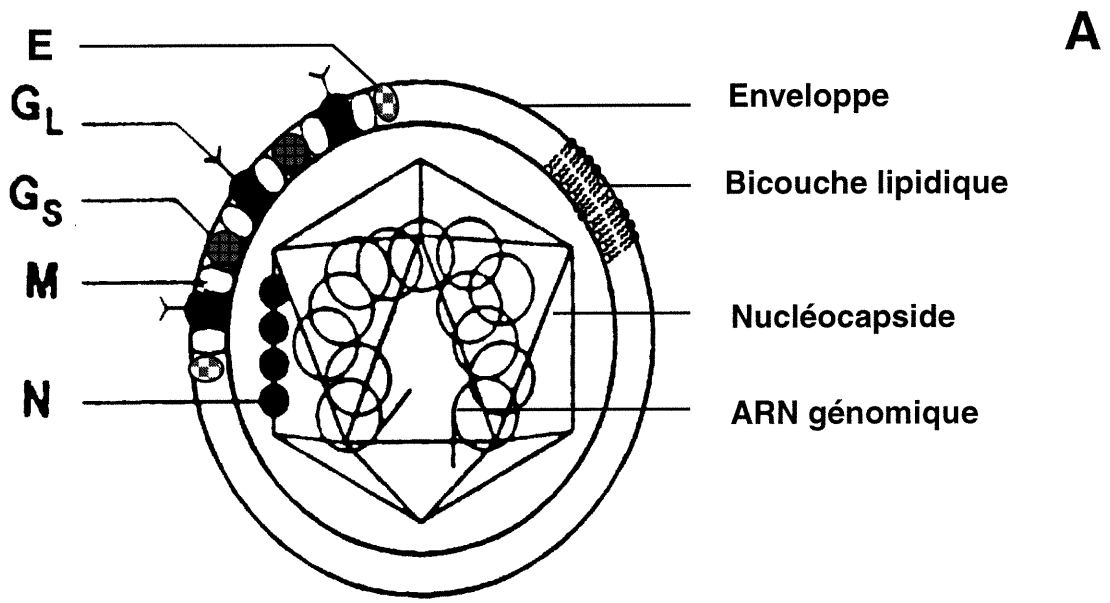
2.1 Caractéristiques en microscopie électronique

Le VAE est une particule de forme sphérique possédant une bicouche lipidique entourant un cœur de forme isométrique qui est la nucléocapside (Fig. 1). La morphologie de la particule virale a été étudiée par différentes méthodes de coloration. Après coloration avec la tungstosilicate de sodium d'une préparation virale purifiée avec un gradient de tartrate de potassium, il fut déterminé que le virus possède un diamètre de 55 ± 10 nm (Magnusson et al., 1970). Par contre, si la particule virale est purifiée dans un gradient de sucrose ou de chlorure de césium, elle possède un diamètre d'environ 50 nm. La particule virale purifiée sur gradient de glycérol ou de sucrose ensuite colorée avec l'acide uranylacétique mesure 60 ± 13 nm. Dans ce cas, la taille de la nucléocapside est de 35 ± 9 nm si la particule est intacte, et de 36 ± 3 nm si la particule est sans enveloppe. Après coloration à l'acide phosphotungstique, le virion intact mesure 67 nm, et la particule nue mesure 37 nm (Horzinek et al., 1971; Maess et al., 1970). De plus, Magnusson et al., (1970) détectèrent sur la surface du virion des projections en forme de la lettre Y de 3 à 5 nm. Par contre, Hyllseth et al. (1973) et Horzinek et al. (1971) montrèrent plutôt sur l'enveloppe virale, des sous-structures en forme de larges anneaux de 12 à 14 nm.

Par ailleurs, les dimensions du virion intracellulaire peuvent également varier par rapport à celle du virion extracellulaire. À partir des observations tirées de différentes cellules infectées, telles que les cellules du derme équin [equine dermis; (EQ); Breese et McCollum, 1970], les cellules de bébé de hamster (BHK-21) (Magnusson et al., 1970), les cellules primaires de rein de cheval (Crawford et Davis, 1970), les cellules endothéliales et les macrophages (Estes et Cheville, 1970), les dimensions moyennes varient de 43 à 60 nm pour le virion entier, et de 25 à 35 nm pour la nucléocapside.

Figure 1. Morphologie, organisation génomique et cycle de réplication du VAE.

A : Représentation schématique de la morphologie du virus de l'artérite équine. Les protéines virales structurales sont représentées à gauche. (Tiré et adapté de de Vries et al., 1997). **B :** Organisation génomique et cycle de réplication du VAE. . Le génome viral comprend la séquence leader (boîte noire) à l'extrémité 5', les ORFs 1a et 1b codant pour la réplicase virale et les ORFs 2a, 2b, 3, 4, 5, 6 et 7 (rectangles gris ou blancs) codant pour les protéines structurales de la nucléocapside (N), de la membrane (M), de la grande glycoprotéine (G_L), de la petite glycoprotéine (G_S), de l'enveloppe (E) ainsi que les glycoprotéines GP3 et GP4 et l'extrémité 3' non codante suivie de la queue poly-A (AAA). Les ARNm sous-génomiques sont indiqués (ARNsg 2 à 7). Les boîtes grises foncées représentent les cistrons fonctionnels avec leur ORF correspondant. Le rôle central des protéines non structurales (nsp) est également montré. (Tiré et adapté de Snijder *et al.* 1999, van Marle *et al.* 1999).



3. Composition structurale du virion

3.1 Génome et organisation génétique

Tout comme les autres membres des nidovirus, le génome du VAE est de nature ribonucléique (Bürki 1965). Cette molécule d'ARN est monocaténaire, de polarité positive et son poids moléculaire a été estimé à 4×10^3 kDa (van der Zeist et Horzinek, 1975). Elle est polyadénylée à l'extrémité 3' (van Berlo et al., 1983) et possède une coiffe à l'extrémité 5' (van Dinten et al., 1997; Glaser et al., 1999). La presque totalité de la séquence nucléotidique pour le VAE fut publiée en 1991 par den Boon et al. (1991). La taille du génome fut alors estimée à 12 704 nucléotides (nt).

Le génome du VAE est structurellement polycistronique avec deux larges cadres de lecture ouverts désignés ORF1a et 1b codant pour les protéines de la réplicase ou la polymérase. Ils occupent trois quarts du génome du côté 5'. En aval de ces ORFs suivent six ORFs (ORFs 2 à 7). Ce qui est caractéristique des nidovirus en général, c'est que ces ORFs ne sont pas exprimés directement par le génome viral mais par des structures d'ARNs messagers viraux ARN sg qui sont traduits durant l'infection virale. Ils codent principalement pour les protéines structurales du virus.

Les ORFs 1a et 1b se chevauchent et ce dernier, l'ORF1b qui est en aval du premier, est exprimé seulement après un changement de cadre de lecture par le ribosome avec un recul d'une base (-1) par rapport au codon stop du premier cadre (den Boon et al., 1991). Cette région de chevauchement est régie par deux signaux promoteurs: un motif de glissement heptanucléotidique (5'-GUUAAAC-3') qui est le site du changement du cadre de lecture proprement dit. Ce signal est suivi en aval d'une structure ribonucléotidique tertiaire appelée pseudo nœud. Il est composé de deux boucles en épingle, la tige de la première est formée par 11 à 12 paires de nt et la tige de la seconde est formée par 6 à 7 paires de nt (den Boon et al., 1991). Ces éléments structuraux sont similaires à ceux déjà décrits auparavant chez les coronavirus (Brierly et al., 1989; Bredenbeck et al., 1990) et les torovirus (Snijder et al., 1990) et font partie d'autres éléments de ressemblance des artérovirus à ces derniers.

Les séquences des autres ORFs sont disposées également de manière à ce que chaque région codante chevauche par un nombre variable de bases, la région avoisinante (den Boon et al., 1991). À l'extrémité 5' se trouve la séquence leader qui est de 211 nt

(Kheyar et al., 1996). Par contre la séquence non codante 3' est de 59 nt sans la queue poly-A (den Boon et al., 1991). Ces séquences aux deux extrémités du génome sont également présentes au niveau des extrémités de tous les ARN sg et sont impliquées dans le processus de la transcription (de Vries et al., 1990).

3.2. Protéines structurales virales

Comme les génomes des virus à ARN de polarité positive, celui du VAE remplit aussi bien le rôle de stockage que le rôle de l'expression de l'information génétique. La reproduction de l'ARN génomique viral est considérée comme un processus combiné de réplication et de transcription. Le terme de réplication est souvent réservé pour la synthèse de l'ARNm de longueur génomique, alors que le terme de transcription est souvent réservé pour la synthèse des ARNm sg viraux (Lai et Cavanagh, 1997; Snijder et Meulenberg 1998). Quant à la traduction des protéines virales, elle est le produit d'expression soit à partir de l'ARNm génomique, soit à partir des ARNm sg. Les produits de traduction de l'ARN génomique sont codés par les ORF1a et 1b et correspondent à des protéines non structurales responsables des processus de réplication et/ou de transcription. Cependant, les produits de traduction des différents ARNm sg sont en règle générale codés par les ORFs situés le plus près de leur extrémité 5' et correspondent aux protéines structurales virales [(dans le cas des ORFs 2a, 2b, 5, 6 et 7) (de Vries et al., 1992; Snijder et al., 1999)] ainsi qu'aux protéines des ORFs 3 et 4 dont le rôle est encore moins bien connu.

3. 2. 1. Protéine de la nucléocapside (N)

La protéine de la nucléocapside N est le produit d'expression de l'ARN 7. Cet ARN sg est composé de la SL, d'une région intergénique de 56 nt (sans le motif de jonction), de l'ORF7 de 333 nt et de la région non codante 3' de 59 nt, ce qui est l'équivalent d'une longueur totale de 659 nt [(sans la queue poly-A) (den Boon et al., 1991); Kheyar et al., article en préparation]. L'ORF7 est celui situé le plus près de l'extrémité 3' du génome viral. La petite protéine N ayant 110 acides aminés est d'un poids moléculaire de 12,3 kDa. Elle est exprimée dans les cellules infectées de façon abondante relativement aux autres protéines virales et constitue 30 à 40 % du contenu en protéines du virion (de Vries et al., 1992). de Vries et al. (1995a) ont observé que la protéine N se trouve sur la particule virale

sous la forme de monomères contrairement à d'autres protéines virales et ne présente aucun site de glycosylation (de Vries et al., 1992). Zeegers et al. (1976) ont démontré que la protéine N est phosphorylée et est très basique. Ceci faciliterait probablement ses interactions avec l'ARN génomique lors du processus de l'assemblage de la nucléocapside. Par des tests sérologiques d'immunobuvadage (Kheyar et al., 1997) et d'immunoprécipitation (Chirnside et al., 1995b) l'immunodominance de la protéine N a été démontrée puisque les sérums de chevaux convalescents contiennent habituellement des titres élevés en anticorps anti-N. Son épitope antigénique dominant a été attribué à la moitié amino terminale de la protéine entre les AA 1 à 69 (Chirnside et al., 1995b). Selon une étude comparative, la séquence codante de l'ORF7 semble être bien conservée entre les différents isolats du VAE (Chirnside et al., 1994). Tout ceci suggère que la protéine N serait un élément viral de prédilection pour le développement d'outil de diagnostic performant (Kheyar et al., 1997).

3. 2. 2. Protéine de la membrane M

La protéine membranaire est exprimée par l'ARN 6. Celui-ci est composé de la SL, d'une région intergénique de 31 nt, de l'ORF6 de 488 nt, d'une séquence silencieuse correspondant à l'ORF7 mais qui n'est pas dans le même cadre de lecture et de la région non codante 3'. La longueur totale de l'ARN 5 est de 1055 nt (sans la queue poly-A) (den Boon et al., 1991). La protéine codée par cet ORF est nommée également la protéine membranaire majeure. C'est la plus conservée des protéines structurales et est composée de 162 AA correspondant à un PM d'environ 16 kDa (de Vries et al., 1992). Elle est non-glycosylée et sa moitié N-terminale semble contenir trois régions trans-membranaires potentielles, ce qui indiquerait une orientation N exo (extra membranaire) – C cyto (intra-cytoplasmique). Certains groupes suggèrent que seulement 10 à 18 AA pourraient être exposés à la surface de la membrane, ce qui expliquerait une immunogénicité relative faible comparativement à la protéine N (Snijder et Meulenberg, 1998). Cependant, son haut degré de conservation parmi les différentes souches virales (Chirside et al., 1994) et son immunogénicité bien démontrée concrètement par notre groupe (Kheyar et al., 1997) ainsi que par d'autres auteurs (Hedges et al., 1998; MacLachlan et al., 1998) font de cette protéine une cible d'intérêt dans le cadre du développement des outils de diagnostic.

La protéine M semble jouer un rôle dans l'assemblage du virion et son bourgeonnement étant donné qu'elle s'accumule dans le réticulum endoplasmique en formant des hétérodimères avec la glycoprotéine majeure G_L (codée par l'ORF5) grâce à des ponts disulfures (de Vries et al., 1995a). Des molécules homodimériques de cette protéine ont également été observées mais seulement dans les cellules infectées (de Vries et al., 1992). Faaberg et al., (1995a) ont démontré, dans le cas du VLD, que la destruction de ces hétérodimères réduit significativement l'infectivité virale, ce qui suggère que ces formations sont importantes dans la liaison du virus au récepteur cellulaire (Balasuriya et al., 2000).

3. 2. 3. Glycoprotéine G_L

La grande glycoprotéine (G_L pour « large glycoprotein») est exprimée par l'ARN sg5. Celui-ci est composé de la SL, d'une région intergénique de 2 nt (ou de 178 nt, car on ne sait encore lequel des motifs de jonction en amont du codon d'initiation de cet ORF5 est utilisé), de l'ORF5 de 767 nt, d'une séquence silencieuse correspondant à l'ORF6 et l'ORF7 mais qui n'est pas dans le même cadre de lecture et de la région non codante 3'. La longueur totale exprimant l'ORF5 est donc de 1 788 nt (ou 1 964 nt) (sans la queue poly-A) (den Boon et al., 1991). La protéine G_L a un PM de 25 à 42 kDa selon que la protéine est traitée avec la glycosylase ou laissée sous sa forme glycosylée (de Vries et al., 1992). Elle possède un site potentiel de N-glycosylation suite à l'addition d'un nombre variable de multiples molécules de lactosamine durant sa maturation (de Vries et al., 1992). Le profil d'hydrophobicité de la protéine G_L est très conservé. Son domaine hydrophobe interne semble traverser l'enveloppe virale une ou trois fois (de Vries et al., 1992). C'est ce qui contient les sites de la N-glycosylation. La glycoprotéine G_L s'accumule avec la protéine M dans le réticulum endoplasmique pour former des hétérodimères. Ces liaisons intermoléculaires sont assurées par des ponts disulfures entre un résidu cystéine de la portion ectodomaine N-terminale de la protéine M et un résidu cystéine de la portion ectodomaine N-terminale de la glycoprotéine G_L .

Les anticorps neutralisant le VAE et provenant des sérums de chevaux infectés sont dirigés essentiellement contre la glycoprotéine G_L . De plus, tous les anticorps monoclonaux neutralisants produits par plusieurs équipes sont spécifiques de cette protéine (Balasuriya et al., 1993; Deregt et al., 1994; Glaser et al., 1995). Le domaine reconnu par ces anticorps est

localisé au niveau des résidus 19 à 115 en AA (Chirnside et al., 1995c, Balasuriya et al., 1997). En effet, ces derniers ont démontré que la délétion de 47 AA entre les positions 66 et 112 ou de 40 AA entre les positions 62 et 101 de l'ectodomaine en question n'affecte pas la capacité d'infection cellulaire par ces mutants échappant à la réponse immunitaire.

3. 2. 4. Glycoprotéine G_S

La petite glycoprotéine de l'enveloppe virale (G_S pour « small glycoprotein ») est exprimée par l'ARN 2 et représente 1 à 2 % de la totalité des protéines structurales virales (de Vries et al., 1992). Celui-ci est composé de la SL, d'une région intergénique de 108 nt, de l'ORF2b de 683 nt, d'une séquence silencieuse correspondant à tous les ORFs 3 à 7 et de la région non codante 3'. La longueur totale de cet ARN sg exprimant la glycoprotéine G_S est donc de 2 280 nt (sans la queue poly-A) (den Boon et al., 1991). Cependant, il est à noter que Snijder et al., (1999) ont démontré récemment un codon d'initiation fonctionnel situé en amont de l'ORF2b et correspondant à l'ORF2a.

La glycoprotéine G_S est codée par une séquence de 227 acides aminés (AA) et possède un PM de 25 kDa avec un site de N-glycosylation. Sa déglycosylation donne un PM de 22 kDa (de Vries et al., 1992). C'est une protéine membranaire avec un complexe de N-glycanes. Elle s'incorpore dans le virion sous la forme dimérique grâce à des ponts disulfures. Par contre, dans les cellules infectées, elle peut se retrouver sous une forme tétramérique et son expression est retenue dans le réticulum endoplasmique (de Vries et al., 1995b). La délétion de la protéine G_S semble abroger l'infectivité virale, ce qui démontre sa nécessité pour la réplication virale (van Dinten et al., 1997).

3. 2. 5. Protéine E

L'ARN sg2 a été démontré comme étant fonctionnellement bicistronique. Il code pour deux ORFs : 2a et 2b. L'ORF2a est plus proximal à l'extrémité 5' et chevauche l'ORF2b avec un cadre de lecture différent. Le codon d'initiation de l'ORF2a est situé à 35 nt en aval de la séquence de jonction qui est la même pour l'ORF2b. L'ORF2a est d'une longueur de 204 nt traduisant un polypeptide de 67 AA et d'un PM d'environ 8 kDa. Cette petite protéine est nommée protéine d'enveloppe E, vu sa ressemblance d'organisation avec les autres nidovirus (Snijder et al., 1999). Elle semble être très hydrophobe et bien stable.

Elle ne forme pas d'oligomères. Contrairement, aux protéines G_L et G_S, elle n'est pas glycosylée. Elle s'associe avec des membranes intracellulaires aussi bien dans les cellules infectées que dans le virion (Snijder et., 1999). Les produits de l'ORF2 (2a et 2b) semblent tous deux importants à la production de progéniture virale infectieuse.

3. 3. Les ORF3 et ORF4

Les produits des ORFs 3 et 4 sont les moins caractérisés des protéines virales du VAE. Aucun rôle ne leur a été associé. Cependant, ils sont plus étudiés chez les autres membres de la famille des *Arteriviridae*. La structure générale des ARN sg de ces produits ressemble à celle des protéines structurales précédentes comme la présence d'une SL d'une séquence intergénique en amont du codon d'initiation du cadre de lecture en question, et d'une séquence silencieuse en aval du cadre de lecture avec la séquence 3' non codante. Cependant, les séquences des ORFs 3 et 4 sont intercalées entre les glycoprotéines G_S et E d'une part, et les protéines G_L, M et N, d'autre part. Il est donc difficile de concevoir l'idée de la présence d'ORFs exprimant des protéines non structurales au milieu d'un segment considéré comme étant la partie du génome exprimant les différentes protéines structurales.

L'ORF3 représente une séquence de 489 nt et celle de l'ORF4 représente une séquence de 456 nt. Ceci correspond à des protéines prédites de 163 AA et 152 AA avec des PM moyens de 18 et 17 kDa, respectivement (den Boon et al., 1991).

Récemment, Hedges et al. (1999a), ont démontré que le produit de l'ORF3 est prédit comme étant une glycoprotéine membranaire (GP3) avec 6 sites potentiels de glycosylation. Leur étude de traduction *in vitro* a révélé que le PM de la forme glycosylée est de 36 à 42 kDa. Cette glycoprotéine semble être associée aux membranes plasmiques et n'est pas libérée sous une forme soluble (Hedges et al., 1999a). Elle se répand dans les membranes des vésicules et est ancrée dans ces membrane via une séquence signal N-terminale, une séquence hydrophobique C' terminale ou avec les deux séquences en même temps. Mais elle ne traverse pas la membrane trois fois comme les autres protéines membranaires. Selon la même équipe, des sérums de chevaux en phase chronique d'infection ou vaccinés contiennent des anticorps spécifiques contre cette protéine GP3, ce qui suggère son immunogénicité.

Quant à l'ORF4, aucune étude n'a caractérisé concrètement son produit. Des prédictions selon sa séquence en AA suggèrent qu'il s'agirait d'une autre glycoprotéine

membranaire mineure (Snijder et Meulenberg, 1998). Toutefois, ces produits des ORF3 et 4 n'ont été détectés ni au niveau des particules virales ni au niveau des cellules infectées (Snijder et Meulenberg, 1998). Les mêmes caractéristiques ont été rapportées chez le VLD sauf que ce sont des protéines solubles associées aux membranes (Faaberg et Plagemann, 1997). Par ailleurs, dans le cas du VSRRP, ces produits ont été identifiés comme étant des glycoprotéines structurales mineures détectées dans la particule virale (van Nieuwstadt et al., 1996). Cela suggère que l'on ne peut écarter la même chose chez le VAE.

3. 4. Protéines virales non structurales

Selon l'organisation génétique de l'ARN viral, les trois quarts du génome viral sont occupés par l'ORF1a et l'ORF1b (den Boon et al., 1991) et le quart restant code pour les protéines d'enveloppe, de membrane et de la nucléocapside virales. (den Boon et al., 1991; de Vries et al., 1992).

3. 4. 1. Polyprotéines de la réplicase virale.

Une des propriétés communes entre les différents membres des nidovirus est la présence d'une série de domaines conservés dans la réplicase virale, l'usage d'un mécanisme de changement de cadre de lecture pour réguler l'expression de la réplicase virale et l'utilisation d'un mécanisme de transcription discontinu pour générer les ARNm sg (de Vries et al., 1997). L'ORF1a code pour un polypeptide de 1 727 AA correspondant à une protéine dont le PM est estimé à 187 kDa. L'ORF1b code pour un plus grand polypeptide de 3 175 AA correspondant à une protéine de 345 kDa (den Boon et al. 1991). L'ORF1b est exprimé lorsque survient un changement de cadre de lecture (de -1 nt par rapport à l'ORF1a) du ribosome vers l'extrémité 3' de l'ORF1a et n'ayant pas terminé la lecture autour du codon stop de cet ORF1a. Ce phénomène semble se produire avec un taux d'efficacité de 15 à 20 % (den Boon et al., 1991).

Les deux polyprotéines de la réplicase virale subissent plusieurs clivages protéolytiques induites par trois protéinases codées par l'ORF1a (Snijder et al., 1992; Snijder et al., 1995; Snijder et al., 1996). Il en résulte la production de 12 protéines non structurales (pns) pns1 à pns12 ainsi que des produits intermédiaires multiples. Huit pns dérivent de la polyprotéine de l'ORF1a (Snijder et al., 1994; van Dinten et al., 1996; van

Dinten et al., 1999; Wassenaar et al., 1997). Quatre pns dérivent de la polyprotéine de l'ORF1b, parmi lesquelles trois pns possèdent les fonctions les plus importantes de la réplication virale. La pns9 porte la fonction de l'ARN polymérase virale dépendante d'ARN (Snijder et Meulenberg, 1998). La pns 10 contient le motif MBD (pour « metal-binding domain ») qui est le site de fixation de quatre ions zinc (Zn^{+2}) essentiels pour la synthèse de l'ARN (van Dinten et al., 2000). La pns11 contient le motif de l'hélicase qui est le site de fixation du complexe nucléoside-triphosphate (van Dinten et al., 2000). Cette pns11 contient aussi un domaine particulier conservé chez tous les membres des nidovirus et dont la fonction est encore inconnue (den Boon et al., 1991).

Les processus de la protéolyse sont assurés par trois différentes protéases virales. La pns1 contient la protéase PCP (« papainlike cystéine protéase ») qui clive entre la pns1 et la pns2. La pns2 peut être générée par autoprotéolyse grâce à une activité cystéine-protéase portée par son extrémité N terminale. La pns4 est une sérine protéase (SP) «chymotrypsinelike serine protéase» qui est responsable du clivage d'un précurseur pns3-8 en différentes sous unités selon deux voies faisant intervenir ou non la pns2. Ceci comprend une voie majeure durant laquelle le site pns4/5 est clivé après association de la pns2 avec le précurseur pns3-8 et une voie mineure durant laquelle le site pns4/5 ne sera clivé qu'après clivage des sites pns6/7 et pns7/8 (Wassenaar et al., 1997). De plus, la pns4 est responsable du clivage des sites de la polyprotéine de l'ORF1b (pns9-12) donnant naissance aux pns9, pns10, pns11 et pns12 (Wassenaar et al., 1997)

À l'exception de la pns1, toutes les pns se localisent au niveau des membranes de la région périnucléaire de la cellule infectée. La synthèse de l'ARN viral s'avère se colocaliser avec ces sous unités de la réplicase, ce qui indique la formation d'un complexe de réplication associé aux membranes. Ce complexe semble s'ancrer dans des structures en forme de vésicules double membranaire par l'intermédiaire des domaines hydrophobes essentiellement portés par la pns2, la pns3 et la pns5 (Pedersen et al., 1999; van Der Meer et al, 1998).

4. Propriétés physico-chimiques

Le VAE peut survivre dans un milieu de culture cellulaire à 4 °C pendant plus de 75 jours, à 37 °C pendant 2 à 3 semaines et à 56 °C pendant 30 minutes (McCollum et al., 1961). À la température de la pièce, le virus peut survivre jusqu'à 6 mois selon Harry et

McCollum (1981). La thermolabilité commence à réduire l'infectivité virale à partir de 60 °C après 30 minutes d'incubation, alors que le virus peut être complètement inactivé en 30 minutes à des températures de 70 à 80 °C (Konishi et al., 1975). Le virus peut être conservé à -70 °C pendant plusieurs années sans perte de son infectivité. À -20 °C et dans les échantillons d'organe, McCollum et al. (1961) ont noté la viabilité du virus pendant au moins 5 ans. La conservation sous la forme lyophilisée est variable (plus d'une année à 4 °C et beaucoup plus longtemps à -20 °C).

Le VAE est sensible au pH bas. À pH 3, les titres infectieux baissent de 1 unité \log_{10} de doses infectieuses à 50 % de culture tissulaire (DI_{50}) en une demi-heure (h) et baissent de 2,25 à 3,25 unités \log_{10} de DI_{50} en 3 h (Konishi et al., 1975). Le VAE est sensible à plusieurs agents chimiques tout comme la plupart des virus enveloppés. Il est rapidement inactivé par l'éther, le chloroforme (Doll et al., 1957a), le déoxycholate de sodium à une concentration de 1 mg/ml (Konishi et al., 1975) et moindrement à la phospholipase C (Hyllseth 1973). Pour provoquer la séparation de l'enveloppe de la nucléocapside virale, la préparation virale est traitée avec divers détergents non ioniques tels que le Nonidet P-40 (0,5 %; Horzinek et al., 1971), le Triton X-100 (Zeegers et al., 1976) et également avec de l'urée à 4M ou de la saponine à 0,05 % (Horzinek et al. 1976).

La densité de flottaison virale est variable selon la méthode de centrifugation sur gradient isopicnique utilisée. Sur des gradients de sucrose, la densité du virus a été estimée à 1,16 g/cm³ (van der Zeijst et Horzinek., 1975), 1,17 g/cm³ (Hyllseth, 1970), et 1,2 g/cm³ (Maess et al., 1970). Quant au coefficient de sédimentation du VAE, les analyses par ultracentrifugation de gradients de sucrose isopicniques ont montré qu'il est autour de 224 ± 8 S (van der Zeijst et al., 1975) alors que le coefficient de sédimentation de la nucléocapside pure est d'environ 158 S (Zeegers et al., 1976).

5. Culture cellulaire et caractéristiques biologiques

Parmi les membres du genre des *Artérvirus* voire même de l'ordre des *Nidovirales*, le VAE est le moins exigeant et le plus facile à manipuler en laboratoire étant donné la variété des lignées cellulaires permissives et pouvant supporter sa réplication dans des conditions simples et standards (Snijder et Meulenberg, 1998). Le VAE infecte plusieurs types de cultures cellulaires primaires et secondaires ayant comme origine le rein de cheval, de lapin, de chat, de hamster, de porc et des cellules de tissu séminal de cheval (Hyllseth,

1969; Konishi et al., 1975; McCollum et al., 1961, 1962a, 1962b; willson et al., 1962). La liste des lignées cellulaires permanentes et susceptibles au VAE est aussi longue. Elle est composée des lignées cellulaires suivantes : des cellules BHK-21 (Hyllseth, 1969); des cellules RK-13 (Radwan et Burger, 1973a ; McCollum 1970); des cellules de rein de singe vert africain (Vero; de Vries et al., 1993; Konishi et al., 1975 ou BSC-1; Crawford et Davis, 1970); des cellules de rein de singe *Cynomolgus* (JINET; Konishi et al., 1975); des cellules de rein de singe *Rhesus* (LLC-MK2, Breese et McCollum, 1970); des cellules de derme équin (NBL-6, Breese et McCollum, 1970); des cellules de poumon de hamster (HmLu; Konishi et al., 1975); et des cellules de tumeur de hamster transformé par le virus de l'hépatite canine (HS et HT-7; Shinagawa et al., 1976). Le temps d'apparition des effets cytopathiques (ECP) varie selon le type de lignée cellulaire utilisé (Konishi et al., 1975).

Il est à noter qu'après transfection de l'ARN génomique, sa réplication peut avoir lieu dans plusieurs types de lignées cellulaires qui ne sont pas permissives et ne peuvent pas être infectées en utilisant des particules virales intactes. Ce phénomène est particulièrement avantageux chez des virus autres que le VAE (Inada et al., 1993; Meulenberg et al., 1998; van Dinten et al., 1997; Snijder et Meulenberg, 1998).

Par observation au microscope photonique ordinaire, l'ECP provoqué par le VAE se caractérise par l'arrondissement cellulaire, la vacuolisation et l'augmentation de la densité optique et de la réfraction (Konishi et al., 1975; McCollum et al. 1962b). Après rétrécissement, pycnose et caryoréxie, la cellule se détache de son substrat de culture. Ceci permet donc l'utilisation des méthodes de quantification standards telles que le titrage viral par dilutions limites ou la techniques des plages à l'aide de la plupart des cellules mentionnées. Toutefois, certaines lignées cellulaires pouvant être plus efficaces que d'autres, le choix des cellules est variable. Par exemple l'emploi des cellules BHK-21, dans la technique des plages, dépend fortement du nombre de passages cellulaires ou de sous cultures subites par ces cellules (observation personnelle; Shinagawa et al., 1976).

L'ajout dans la suspension virale d'un polycation tel que le diéthylaminoethyl (DEAE) - dextran à une concentration finale de 50 µg/ml, semble favoriser l'attachement du virus sur les cellules et améliore donc la production virale (Hyllseth, 1969, van Berlo et al., 1980). Lors de l'infection virale, environ 50 % de l'inoculum s'adsorbe aux cellules après 1 h à 37 °C et presque 100 % le sont après 2 h d'incubation (Hyllseth, 1969, Shinagawa et al., 1976; van Berlo et al., 1980). La courbe de croissance virale a été étudiée et déterminée sur différentes cellules et à différentes températures permissives comme les

cellules VERO à 35 °C (van Berlo et al., 1980), 37 °C (Inoue et al., 1975; Tsukamoto et al., 1982) et 40 °C (van Berlo et al., 1980), des cellules NBL-6 (Inoue et al., 1975) à 37 °C et des cellules BHK-21 à 39 ou 40 °C (van Berlo et al., 1982; 1986b). En général, la progéniture virale commence à être libérée 6 à 8 h post-infection (hpi) et atteint son maximum à 12 à 20 hpi. Dans les meilleures conditions, les titres viraux peuvent atteindre et dépasser les 10^8 pfu/ml de surnageant de culture cellulaire. Sur des cellules BHK-21 et à 40 °C, van Berlo et al. (1986a, 1986b) ont rapporté que la synthèse de l'ARN viral ainsi que celle des protéines virales commencent à 4 hpi et atteignent leur maximum à 6 à 8 hpi.

L'assemblage du VAE s'effectue par bourgeonnement de la nucléocapside dans la lumière du réticulum endoplasmique lisse (REL) et/ou dans l'appareil de Golgi. La libération du virion se fait par exocytose (Magnusson et al., 1970). Wada et al. (1995) ont montré que les protéines de l'enveloppe du VAE s'accumulent dans des membranes intracellulaires alors que leur rôle exact n'était pas encore connu (de Vries et al., 1995b). Cependant, Breese et McCollum, (1970) ont observé une propriété particulière à la réplication des artérovirus soit la formation de vésicules ayant des membranes doubles 3 à 6 hpi. Plus récemment, van Dinten et al. (1996) ont démontré que les protéines de la réplicase virale sont associées aux membranes et sont localisées dans la région périnucléaire où se trouvent surtout ces structures membranaires doubles. Plus tard, Pedersen et al., (1999) ont démontré que ces structures sont induites et dirigées par les pns 2 à 7 de la réplicase indépendamment des autres protéines virales.

Finalement, van der Meer et al. (1998) et Pederson et al. (1999) ont démontré par des tests biochimiques et d'immunofluorescence que d'autres sous unités de la réplicase virale sont ancrées dans des membranes intracytoplasmiques des cellules infectées par le VAE. Des tests de marquage *in situ* de l'ARN viral naissant ont révélé que le site d'accumulation de ces complexes de protéines virales est le même que celui où se déroule la transcription de l'ARN viral.

6. Réplication virale

Par définition, le génome des virus à ARN de polarité positive se charge d'assumer le double rôle d'entreposer et d'exprimer l'information génétique. Ainsi, la reproduction de l'ARN génomique viral est un processus combiné de la réplication du génome et d'expression de l'ARNm. Le cycle de réplication des artérovirus débute avec l'expression

des gènes de la réplicase, c'est-à-dire des ORF1a et 1b, à partir de l'ARN génomique de pleine longueur. En revanche, les produits de la transcription, les ARNm sg, servent à la production des protéines structurales virales (voir Fig. 1).

Peu d'informations sont connues quant aux séquences de l'ARN artéroviral requises pour la réplication génomique. Contrairement aux coronavirus, ce n'est que récemment que le groupe de Snijder et al. (Molenkamp et al., 2000) ont isolé et caractérisé des molécules d'ARN génomique artéroviral déficientes interférentes (DI). Ce sont des molécules d'ARN génomiques tronquées et dans certains cas leurs séquences sont réarrangées. Souvent, elles perdent totalement leur potentiel d'autonomie de réplication, vu que les régions délétées correspondent aux gènes de la réplicase. Leur formation peut être obtenue en effectuant des passages successifs non dilués des stocks de virus. Par contre, leur réplication dépend des enzymes de réplication exprimées par le virus auxiliaire (Mendez et al., 1996; Molenkamp et al., 2000). Ainsi, ils ont délimité quelques séquences requises pour la réplication et la transcription à partir de molécules d'ARN DI et ont par conséquent réduit la molécule d'ARN génomique viral pour déterminer une taille minimale de 3 kb toujours compétente à la réplication. Ces séquences sont éparpillées le long du génome; certaines sont près de l'extrémité 5', d'autres près de l'extrémité 3' et d'autres encore sont situés au milieu du gène de l'ORF1b (Molenkamp et al., 2000). La mise en évidence de molécules d'ARN DI ouvre beaucoup de perspectives pour la caractérisation des séquences virales essentielles aux autres étapes du cycle de biogenèse virale comme la décapsidation et l'encapsidation qui sont encore inconnues.

Pour dupliquer et reproduire des copies d'ARN génomique viral, plusieurs facteurs et cofacteurs viraux et cellulaires sont impliqués. L'ARN génomique viral ou l'ARNm de pleine longueur nécessite vraisemblablement une étape de synthèse de brin négatif d'ARN. Ce processus n'est pas encore bien élucidé, mais on pense qu'il se déroule instantanément ou de façon indépendante, lors de la transcription et la synthèse des différents brins négatifs d'ARN sg y compris le brin négatif de taille génomique (Snijder et Meulenberg, 1998). Ce dernier serait donc la matrice sur laquelle s'associent les différentes sous-unités de la réplicase virale avec l'ARN-polymérase virale ARN-dépendante (produit de la pns9), ainsi que probablement, des protéines cellulaires dont l'identité n'a pas encore été dévoilée. Plusieurs autres groupes travaillant sur les nidovirus ont identifié certaines de ces protéines cellulaires. Hwang et Brinton (1998) en ont identifié quatre au niveau des cellules MA-104 infectées par le VFHS. Ces protéines semblent se fixer sur une séquence de 68 nt à

l'extrémité 3' du brin négatif de molécules d'ARN transcrites *in vitro*, ce qui correspond au brin négatif de la SL ou anti-leader. Ceci serait le point de départ de la synthèse du brin positif de l'ARN génomique (Hwang et Brinton (1998). Curieusement, ces auteurs ont rapporté que ces mêmes protéines dont l'identité exacte n'est pas encore connue se fixent sur des séquences du VAE et du VLD sensiblement au niveau des mêmes régions correspondantes (Hwang et Brinton (1998).

7. Transcription des ARNs viraux

Un des points culminants de rapprochement entre les différents membres des nidovirus est la génération lors du cycle de réplication virale d'une série de plusieurs ARNm sg co-terminaux à partir desquels s'expriment les gènes du quart du génome à l'extrémité 3' et qui codent pour les principales protéines structurales. C'est ce qui constitue l'origine de l'appellation de l'ordre des *Nidovirales* (provenant du mot latin *nidus* signifiant nid). De plus, ces ARNm sg sont également co-terminaux à l'extrémité 5' avec la présence de la SL. Celle-ci dérive de l'extrémité 5' du génome viral et se fusionne au corps des ARNm sg par un mécanisme de transcription discontinue (Snijder et Meulenberg, 1998). Ce mécanisme a été décrit et proposé depuis longtemps chez les coronavirus (Lai et Cavanagh, 1997). Cependant, plusieurs données accumulées durant la dernière décennie avec les artérovirus indiquent que ces derniers adopteraient un mécanisme de transcription similaire aux coronavirus (den Boon et al., 1991, 1995; 1996). Ces résultats ont donc écarté le modèle conventionnel de l'épissage en *cis* comme mode de synthèse des ARNm sg.

Cette existence de la SL fusionnée de façon discontinue à des séquences de taille différentes au sein du génome a toujours suscité l'intérêt d'analyser la région de jonction entre les deux séquences. Curieusement, à proximité de cette jonction, existe un motif pentanucléotidique conservé (5' UCAAC 3') appelé auparavant motif de jonction (de Vries et al., 1990; den Boon et al., 1991) et plus récemment TRS pour « transcription-regulating sequence » (Snijder et Meulenberg, 1998; van Marle et al., 1999).

La détection d'un ensemble de molécules d'ARN sg répliquatives intermédiaires dans des cellules infectées par le VAE a également renforcé le rapprochement du VAE avec les coronavirus (den Boon et al., 1996). Ces molécules correspondent aux copies antisens ou brins négatifs complémentaires des ARNm sg formant ainsi des molécules doubles brins provisoirement. Elles seraient impliquées dans la synthèse des ARNm viraux tel que décrit

auparavant pour les autres nidovirus (Sethna et al., 1989; Sawicki et Sawicki, 1990). Ainsi plusieurs modèles ont été suggéré à propos de cette transcription discontinuée des nidovirus.

7. 1. Transcription discontinuée lors de la synthèse des brins positifs d'ARNm sg.

Dans le premier modèle de transcription où la synthèse des brins positifs d'ARNm sg s'effectue directement, la fusion de la SL avec le corps des ARN sg se déroulerait lors de la transcription elle-même utilisant comme matrice le brin négatif du génome. La connection serait médiée par l'appariement de deux TRS, l'un se trouvant à l'extrémité 3' de la SL et l'autre se trouvant à l'extrémité 5' du corps de l'ARNm sg, respectivement. Cet appariement nécessiterait donc que l'un des deux TRS (c'est-à-dire celui de la SL) soit complémentaire à l'autre. C'est cette idée qui a été à l'origine du modèle appelé « leader-primed transcription » ou la transcription amorcée par la SL (Baric et al. 1983, Spaan et al., 1983). Celle-ci étant considérée ici comme une amorce sens, la complémentarité avec le deuxième segment (anti-sens) se situe au niveau du TRS. Dans ce modèle, l'appariement des deux TRS sens et anti-sens se poursuit par l'extension de l'amorce (SL) pour donner naissance à l'ARNm sg en question.

Cependant, d'où provient cette SL? Est-elle présente sous une forme libre dans les cellules infectées? De quelle façon peut-elle se libérer pour servir d'amorce?

Plusieurs chercheurs ont essayé de répondre à ces questions qui restent toujours paradoxales. Certains qui adoptent ce modèle ont démontré l'existence de SL à l'état libre dans le cytoplasme des cellules infectées par le virus de l'hépatite murine (VHM) (Baric et al., 1985; 1987; Makino et al., 1986; Baker et Lai 1990; Lai, 1986; 1990; Brian et Spaan, 1997). Chez les membres des artérovirus, seul le groupe de Plagemann a exploré cet aspect et n'a pas détecté la présence de SL à l'état libre dans les macrophages infectés par le VLD (Chen et al., 1994).

Par conséquent, l'on admet que la SL est considérée comme une amorce qui provient de deux voies possibles (voir Snijder et Meulenberg, 1998).

a) Soit qu'elle est synthétisée entièrement ou avec des nt supplémentaires en aval du motif TRS à partir du brin négatif du génome, pour être ensuite clivée et libérée dans le cytoplasme de la cellule infectée. Ainsi, elle pourrait se fixer sur les différents sites TRS complémentaires éparpillés le long du brin anti-génomique. Ceci expliquerait aussi les tailles différentes des ARNm sg.

b) Soit qu'elle est synthétisée également à partir de l'extrémité 3' du brin négatif, jusqu'au site TRS où la transcription s'arrête. Ensuite le complexe de transcription effectuerait un saut, par un mécanisme indéterminé, pour se fixer sur un des site TRS le long du brin anti-génomique.

7. 2. Transcription discontinue lors de la synthèse des brins négatifs d'ARN sg.

La détection des molécules d'ARN répliquatives intermédiaires sous forme d'ARN sg anti-sens contenant des séquences anti-leader indique que l'étape de la discontinuité de la transcription surviendrait durant la synthèse des brins négatifs plutôt que des brins positifs. C'est pour cela que les modèles de transcription des nidovirus reste toujours sujets de propositions et de discussions. Le VHM a été le plus étudié sur ce point de vue (van der Most et Spaan, 1995; Sawicki et Sawicki, 1995; Lai et Cavanagh, 1997). Dans ce modèle, l'on pense que la synthèse commence à l'extrémité 3' du génome viral (brin positif). Des molécules d'ARN de brin négatif sont ainsi produites de façon incomplète. Le complexe de transcription s'arrête à chaque fois qu'il rencontre un site TRS (du brin positif), effectue un saut, par un mécanisme indéterminé, pour se fixer sur le site TRS de la SL se trouvant à son extrémité 3'. Ensuite, survient une réinitiation de la transcription qui continue jusqu'à l'extrémité 5'. Ainsi des molécules d'ARN sg de sens négatif serviraient alors de matrice pour la synthèse des ARNm sg disponibles pour l'expression des protéines virales structurales. Quant au brin anti-génomique (génome négatif), le même processus s'effectue, mais sans interruption selon un mécanisme différent. Celui-ci sert de matrice pour la répllication génomique et la synthèse de nouvelles molécules de génome viral qui seront incorporées dans les virions (Lai et Cavanagh, 1997; Snijder et Meulenberg, 1998).

Quant au nombre de TRS se trouvant au long du génome viral, il est différent selon les virus. Chez le VAE, den Boon et al. (1991) en ont dénombré 17, sans compter celui de la SL. Cependant, il semble que seulement six d'entre eux seraient fonctionnels lors de l'étape de la transcription puisque seulement six ARNm sg sont produits. La même situation a été observée chez les autres artérovirus et les nidovirus en général (Lai et Cavanagh, 1997; Snijder et Meulenberg, 1998). Dans ces derniers cas, l'on peut expliquer le phénomène par le fait que les TRS ne sont pas complètement conservés et que certaines différences de transitions, de transversions ou de délétions d'un ou de plusieurs nt ont été rapportées à plusieurs reprises. Il faut, par contre, souligner que dans le cas du VAE en

particulier, les TRS sont parmi les plus conservées le long du génome (den Boon, et al. 1991; 1996). De plus, des TRS constituées de bases variables ont été décrites seulement pour l'ORF3 qui semble en utiliser trois séquences TRS différentes (den Boon, et al., 1996).

Récemment, van Marle et al. (1999) ont rapporté la preuve de l'existence de cet appariement entre la TRS de la SL et celle du brin négatif du génome viral. De plus, ils ont démontré que les nt de la séquence TRS qui reste sur les ARNm sg dérivent du corps de l'ARN génomique et ne dérivent pas de la SL. C'est ce qui a fait que ce groupe supporte l'idée que les ARNm du VAE sont générés par un mécanisme de transcription discontinue durant la synthèse des brins négatifs. Ce mécanisme ressemble à un phénomène de recombinaison appelé « recombinaison d'ARN par choix de copie ». Durant ce phénomène, le complexe transcriptionnel et le transcrit naissant sont libérés d'une matrice primaire, et la transcription est ensuite réinitiée sur une autre matrice secondaire après appariement des TRS de cette dernière matrice et celle de l'extrémité 3' du transcrit naissant (van Marle et al., 1999). Des structures secondaires d'ARN sont évoquées dans ce genre de recombinaison (Chang et al., 1996; Nagy et Simon, 1997) où l'on suppose que la TRS de la SL se situe sur un boucle d'ARN qui présente la TRS elle-même à la matrice sur la quelle s'effectue la réinitiation de la transcription (van Marle et al., 1999).

Par conséquent, le mécanisme par lequel le complexe transcriptionnel sélectionne le bon TRS reste un phénomène inconnu. Des interactions entre des protéines virales et/ou cellulaires ainsi que des séquences d'ARN de la matrice et/ou du brin naissant sont vraisemblablement impliquées (Snijder et Meulenberg, 1998). Par ailleurs, les modèles de transcription décrits ici ont tous été supportés par des travaux expérimentaux concrets. Par conséquent, plusieurs groupes pensent qu'ils ne sont pas tout à fait exclusifs mutuellement mais peuvent plutôt coexister au sein d'une même cellule infectée. Probablement, un choix propre au virus dépendant de certaines conditions particulières déterminerait la voie de transcription que ce dernier doit entreprendre (Lai et Cavanagh, 1997; Brian et Spaan, 1997; Snijder et Meulenberg, 1998).

(C) SÉQUENCE LEADER

1. Définitions et généralités

Tous les ARNm des organismes eucaryotes contiennent deux types de région soit codante et non codante. La région codante débute par un codon d'initiation (principalement AUG) et se termine par un codon de terminaison. Cependant, l'ARNm mature est souvent plus long que la région codante elle-même. Dans une molécule d'ARNm simple monocistronique, des régions supplémentaires sont présentes à chacune de ses extrémités. Ce sont les régions non codantes. La région située du côté 5', précédant le codon d'initiation est appelée la séquence « leader ». La région située en aval par rapport au codon de terminaison est appelée séquence « trailer ». Elles sont également appelées séquences non codantes ou séquences non traduites 5' ou 3' respectivement (Lewin, 1994).

La SL des ARNm eucaryotes est souvent pourvue à son extrémité 5' d'une structure appelée la coiffe ou « cap ». C'est un motif de 7-méthyl guanine introduit après la transcription par une liaison 5'-5' triphosphate à la première base de l'ARNm au niveau de la SL. Ceci donne une structure sous la forme de $7^mG^5,ppp^5,Np$ (où N est considéré comme la première base de l'ARNm). Certains facteurs d'initiation de la traduction se fixent sur cette coiffe pour amorcer le balayage de la chaîne d'ARN par le ribosome (Lewin, 1994) à moins que cet ARN soit dépourvu de coiffe tel que dans le cas de certains virus ou encore quelques ARNm humains et de drosophiles. Ceux-ci sont pourvus d'une structure secondaire particulière appelée site interne d'entrée ribosomal ou IRES (pour « Internal Ribosomal Entry Site ») qui permet au ribosome de s'y fixer pour initier la traduction sans la présence de la coiffe et sans amorcer son balayage à partir de la première base 5' terminale.

Après les divers travaux de Kozak et de beaucoup d'autres chercheurs durant les vingt dernières années, le modèle d'initiation de traduction le plus reconnu est le « Ribosome Scanning Model » qui désigne un mode de balayage de la chaîne d'ARNm par le ribosome (Kozak, 1989). Des facteurs d'initiation reconnaissent et se lient à la structure de la coiffe sur la molécule d'ARN et dirigent la fixation de celle-ci sur le ribosome. Le complexe initiation-ribosome se déplace le long de la molécule d'ARNm dans le sens 5' vers 3' jusqu'à la rencontre du triplet AUG où s'amorce la traduction de la protéine. Cette

initiation de la traduction est souvent l'objet d'une régulation qui se traduit par des variations dans la traduction des protéines.

Chez les nidovirus, étant donné que la transcription signifie la synthèse de molécules d'ARNm à partir de matrices d'ARN (et non à partir de matrices d'ADN comme c'est le cas pour la transcription chez les organismes eucaryotes supérieurs), la SL est donc considérée comme une région essentielle dans la régulation de la transcription. Ceci est dû au fait qu'elle contient des éléments tel que la séquence TRS mentionné auparavant ainsi que d'autres éléments (voir plus bas). La séquence TRS parfois appelée TAS (pour « transcription-associated sequence ») (de Vries et al., 1997) fonctionnerait comme le site d'initiation de la transcription et serait l'équivalent d'un promoteur transcriptionnel (Snijder et Spaan 1995; Hiscox et al., 1995).

2. SL des Nidovirus

Chez les nidovirus, la SL peut être définie de façon plus détaillée à cause de ses particularités quant aux séquences la composant et sa fonction dans la transcription. Il est à noter que selon la définition présentée plus haut, la SL des nidovirus n'est pas considérée comme étant toute la région non codante 5' mais elle fait seulement partie de cette région. Ceci est dû au fait que la région non codante 5' est considérée comme étant toute la séquence s'étendant de la première base 5' jusqu'à la base située juste en amont du codon d'initiation AUG sur la chaîne d'ARNm. Or, chez la plupart des nidovirus, la SL se termine au niveau de la séquence de jonction ou la séquence TRS qui est responsable de sa fusion avec les corps des ARNm sg lors de la transcription (de Vries et al., 1997). Cependant, il existe une séquence supplémentaire (entre la SL et le premier codon d'initiation de l'ARNm génomique ou sous génomique en question) dont le nombre de bases est variable selon la famille, le genre et même l'espèce. (de Vries et al., 1997; Lai et Cavanagh, 1997).

2. 1. SL des Coronavirus

La SL des coronavirus est d'une longueur de 60 à 90 nt dépendamment de l'espèce (Lai, 1982, 1990; Lai et al., 1983, 1984; Spaan 1988). Par exemple, les SL du coronavirus humain, souche 229E, du coronavirus bovin et du virus de la gastroentérite transmissible du porc ont une longueur de 60 nt (Schreiber et al., 1989), 65 nt (Hofmann et al., 1993a) et 90

nt (Hiscox, 1993) respectivement. Néanmoins, en règle générale, les SL de tous les ARNm sont presque identiques au sein d'une même souche notamment chez les coronavirus autres que le VHM où la structure de la séquence TAS est souvent homogène et existe sous forme d'une copie pentanucléotidique constante (Hofmann et al., 1993a). Cependant, il peut y avoir des variations selon différentes souches au sein d'une même espèce tel que chez le coronavirus murin où la SL varie de 72 à 82 nt (Makino et al., 1988). Cette variation peut résulter d'une pluralité du nombre de copies de séquence TAS et d'une hétérogénéité dans l'extrémité 3' de la SL autour de cette séquence TAS. En effet, parfois il peut y avoir 2 à 4 copies de la même séquence TAS (UCUAA) et parfois il peut y avoir des ARNm contenant des copies différentes de la séquence TAS. Ceci indique que la fusion entre la SL et l'ARNm n'est pas bien précise (Makino et al., 1988). La séquence complémentaire de la séquence TAS, située sur le corps de l'ARNm sg, est par contre appelée séquence intérgénique (IG) chez les coronavirus (Lai et Cavanagh, 1997).

Le nombre de copies de la séquence TAS semble jouer un rôle important dans la régulation de la transcription des ARNm (Hofmann et al., 1993a; Lai et Cavanagh, 1997). Quant à la séquence située en aval par rapport à la séquence TAS sur la chaîne de l'ARN génomique, elle est souvent plus longue chez les coronavirus que chez les artérovirus. Sa longueur a été estimée entre 200 et 500 nt chez les coronavirus (Lai et Cavanagh, 1997; de Vries et al., 1997). Il faut cependant rappeler que cette longueur concerne seulement la séquence intermédiaire au niveau de la molécule d'ARN génomique. Ceci est différent dans le cas des ARNm sg et est même variable d'un ARN sg à un autre. Cette variabilité dépend essentiellement de la position de la séquence IG par rapport au codon d'initiation de l'ORF principal exprimé par l'ARNm sg en question.

Depuis plusieurs années, plusieurs travaux sur le rôle de la séquence TAS de la SL, en combinaison ou non avec la séquence IG, ont été réalisés. Dans la plupart des cas, ces travaux ont été élaborés dans le cadre de leur implication dans la transcription des ARNm sg ainsi que sa régulation (Lai et Cavanagh, 1997). La SL est plus étudiée chez les coronavirus que chez les autres membres des nidovirus. Son existence aussi bien au niveau du génome que des ARNm sg a souvent suscité des questions quant à l'origine de son association avec des molécules d'ARNm différentes du même virus. Puisqu'il doit y avoir un processus de fusion entre deux séquences discontinuës (la SL et le corps des ARNm sg) durant et après la transcription, la SL de chaque ARNm peut être échangée de façon libre entre deux virus coinfectant la même cellule (Makino et al. 1986a; 1986b). Baric et al.

(1985) ont détecté la présence de plusieurs ARN de 50 à 90 nt dans le cytoplasme de cellules infectées par le VHM et les ont attribués à des molécules de SL libres de différentes tailles. Ils ont expliqué ceci par le fait que la présence de molécules libres de SL devait passer par un changement quelconque avant leur incorporation dans les molécules d'ARNm sg complètes. Quoiqu'il en soit, la SL et la séquence du corps de l'ARNm sg sont transcrites indépendamment et peuvent par la suite être fusionnées de façon aléatoire (Baric et al., 1985; Makino et al., 1986a, 1986b; Jeong et Makino, 1994; Zhang et al., 1994b).

L'existence de molécules de SL libres dans le cytoplasme et l'implication de la séquence TAS complémentaire à la séquence IG sont les éléments importants à l'origine de la proposition du modèle de la transcription discontinue impliquant la SL comme amorce (Lai 1990; Sawicki et Sawicki, 1995; Lai et Cavanagh, 1997). Néanmoins, il faut noter que la présence de SL à l'état libre dans le cytoplasme n'a été observée que chez les coronavirus. Chez les artérovirus, aucune étude n'a été entreprise à ce sujet sauf chez le VLD où Chen et al. (1994) ont rapporté la présence de brins négatifs d'ARN sg entiers mais non celle de SL à l'état libre.

2. 2. SL des Torovirus

La plupart des études sur le génome des torovirus ont été effectuées sur le virus prototype de cette famille qui est le torovirus équin appelé virus Berne (BEV) et un peu moins avec le torovirus bovin appelé virus Breda (BRV) (Snijder et Horzinek, 1993). Durant la réplication des torovirus, un ensemble de 4 ARNm sg coterminaux du côté 3' sont générés en plus de l'ARN génomique (Snijder et al., 1988). Comparativement au coronavirus et aux artérovirus dont les ARNm contiennent une SL commune à l'extrémité 5', une différence majeure distingue les torovirus chez qui aucune présence de SL commune à l'extrémité 5' des ARNm sg n'a été mise en évidence (Snijder et al., 1990a; 1991b). La région 5' non codante est par contre la plus longue parmi les nidovirus mesurant 824 nt pour le BEV (Snijder et al., 1991b). Puisque la transcription des ARNm sg des coronavirus et des artérovirus se déroule par un mécanisme impliquant la SL par l'intermédiaire d'une séquence TAS ou TRS, on peut admettre que le même mécanisme régirait la transcription des ARNm sg des torovirus. Néanmoins, cette absence d'une SL commune entre les ARNm sg ainsi que l'ARN génomique des torovirus suscite encore plus de questions quant au mécanisme de génération des ARNm toroviraux.

Par analogie aux deux autres familles, Snijder et al. (1990b; 1991b) ont identifié sur le génome du BEV, des séquences intergéniques conservées riches en AU correspondant à la séquence consensus (5' uaUcUUUAGa 3') où les lettres majuscules constituent la partie commune. Cependant, selon leur travaux, il n'y a aucune évidence de l'existence du phénomène de fusion entre des corps d'ARNm et d'une SL commune à ce site. En revanche, ces ARNm de BEV semblent se terminer juste au niveau ou autour de ces séquences intergéniques qui, dans ce cas, peuvent être qualifiées de promoteurs. Il a été proposé qu'un mécanisme impliquant une liaison directe de la polymérase virale sur ces séquences pourrait constituer une alternative au modèle de transcription discontinue décrit chez les coronavirus et les arterivirus (Snijder et Horzinek, 1993).

L'utilisation du terme de SL est donc invraisemblable dans le cas des torovirus. Cependant, la présence d'une séquence intergénique à proximité de l'extrémité 5' terminale de molécules d'ARN déficientes interférentes de BEV peut constituer l'équivalent d'une petite SL de 8 nt (Snijder et al. 1991). Jusqu'à présent, mis à part de ce qui est connu chez le BEV, très peu de travaux ont été effectués sur les mécanismes de transcription chez les torovirus. Des études dans ce créneau seraient donc d'un grand appui pour la compréhension des mécanismes de transcription des *Nidovirales* (de Vries et al., 1997).

Par ailleurs, si l'utilisation de la SL était l'élément essentiel dans la transcription des *Nidovirales* et compte tenu des similarités très étroites entre les coronavirus et les torovirus, on peut suggérer qu'un ancêtre commun aux membres des *Nidovirales* aurait utilisé un mécanisme de transcription commun tout à fait indépendant de la présence de la SL. De plus, celle-ci serait acquise chez les coronavirus et les artérovirus plus tard lors des événements régissant leur divergence (de Vries et al., 1997). Tout cela mène à susciter d'autres questions quant au rôle de la SL dans le cycle de réplication virale autre que la transcription.

2. 3. SL des Artérovirus

2. 3. 1. VAE

2. 3. 1. 1. Bref historique : Séquence leader et ARNm

Les premières études sur l'acide nucléique artéroviral ont démontré que le VAE contenait un brin d'ARN monocaténaire linéaire et infectieux (van Der Zeist et Horzinek 1975). van Berlo et al. (1982) furent par la suite les premiers à effectuer des études sur le mécanisme de réplication du VAE. Ils décrivent la présence, dans les cellules infectées par le VAE, d'un ensemble de 6 ARN viraux polyadénylés avec des poids moléculaires différents. Ils définirent alors 5 ARN sg viraux et un ARN de longueur génomique. Ainsi, ils proposèrent la remise en cause de l'inclusion du VAE dans la famille des *Togaviridae*. Ceci fut reconfirmé plus tard par van Berlo et al. (1986) qui ajoutèrent dans leurs observations que ces différents ARNs contenaient des séquences communes. Par l'usage des techniques d'empreintes digitales (« fingerprinting ») par les oligonucléotides à l'aide de la ribonucléase (Rnase) T1, ils ont montré que ces ARNs formaient un paquet de molécules coterminales grâce à ces séquences communes.

Des études sur la composition nucléotidique de certains segments de l'ARN viral ou des ARN sg ne furent réalisés que vers le début des années 90. En effet, de Vries et al. (1990) ont démontré que durant la réplication virale, les ARNm viraux formaient un set coterminale du côté 3'. De plus, tous ces ARNm sg contenaient une SL commune à l'extrémité 5'. Ils définirent une séquence qu'ils ont appelée motif de jonction (5' UCAAC 3') et qu'ils ont considérée comme étant le site de jonction entre des corps d'ARNm sg et la SL. Ce furent les premières données qui expliquaient que les ARNm sg du VAE étaient composés de la SL, et de séquences du corps génomique non contiguës sur le génome. En combinant les résultats des expériences d'extension d'amorce et d'analyse de la position du site de jonction sur l'ARN6, de Vries et al. (1990) ont calculé la taille de la SL du VAE comme ayant 208 nt sans tenir compte de la séquence de jonction elle-même à son extrémité 3'. L'analyse des séquences d'une librairie d'acides désoxyribonucléiques complémentaires (ADNc) du génome viral amorcés par des oligonucléotides aléatoires et l'utilisation de la technique de l'hybridation avec un oligonucléotide spécifique de la SL ont révélé que l'extrémité 5' comporte la SL qui elle-même contient le motif de jonction ci haut mentionné. Lors de ces travaux (de Vries et al. 1990), en trouvant la SL sur chaque ARNm sg, sa position sur chaque ARNm ne fut pas déterminée. Toutefois, selon eux, la SL des ARNm sg semblait être codée par l'extrémité 5' du génome. Par contre, la séquence exacte de l'extrémité 5' terminale de la SL n'était pas encore déterminée.

À la lumière de ces travaux et en combinaison avec les études d'inactivation des ARNs viraux par les rayonnements ultraviolets, les auteurs conclurent qu'il y avait bien des

ressemblances de structures des ARNm avec ceux des coronavirus mais avec des différences au niveau de leur mécanisme de transcription. La synthèse des ARNs du VAE, selon leur hypothèse, impliquerait la maturation d'un précurseur d'ARN de longueur génomique. Ceci les poussa à proposer un mécanisme d'épissage de l'ARN génomique au niveau de plusieurs sites définis par la séquence de jonction 5' UCAAC 3' (de Vries et al., 1990). Ces observations s'ajoutèrent aux premières décrites plus haut pour remettre en question la classification des artérovirus dans la famille des *Togaviridae*.

2. 3. 1. 2. Étude prospective complémentaire

En 1991, la quasi totalité de la séquence du génome du VAE a été publiée (numéro d'accès de Genbank X59453; den Boon et al., 1991). Dans ce cas, la longueur de la SL a été estimée à 207 nt. Néanmoins, 18 nt à l'extrémité la plus terminale de la SL étaient absentes pour compléter la totalité du génome du VAE (den Boon et al. 1991). Durant ces travaux, il y a eu également la détermination approximative de la taille de ARNm sg et la confirmation de la présence de la SL à l'extrémité 5' de chaque ARN sg. Par contre, ils ont déterminé 18 motifs de jonction le long du génome. Cela voulait dire que chaque ORF décrit plus haut (excepté l'ORF1b) est précédé par un ou plusieurs motifs de jonction avec la SL. Cependant un seul motif de jonction serait utilisé dans la génération d'un seul ARN sg (den Boon et al., 1991). Suite à ces observations, den Boon et al. (1991) rejetèrent l'hypothèse du mécanisme d'épissage alternatif des ARNm du VAE proposé par de Vries al. (1990) et optèrent pour un mécanisme similaire aux coronavirus et torovirus.

2. 3. 1. 3. Données plus récentes

den Boon et al. (1995) ont effectué une étude comparative du VAE avec le coronavirus murin et l'alphavirus Sindbis. La cartographie de la transcription à l'aide des rayons ultraviolets a montré que la majorité des ARNs sg produits au stade tardif de l'infection virale ne dérivait pas d'un seul précurseur d'ARN de la taille du génome. Ainsi cette synthèse des ARN sg n'était pas complètement indépendante de l'ARN génomique durant le cycle de la réplication virale. Ceci suggérait qu'il y aurait un effet au début de l'infection attribué au processing de la polymérase virale. La synthèse de l'ARN génomique, selon cette étude, était indépendante de la synthèse des protéines *de novo* au

stade tardif de l'infection (den Boon et al., 1995). Les auteurs proposèrent alors que si des formes répliquatives d'ARN sg existaient durant l'infection artérovirale comme dans le cas des coronavirus, les observations citées plus haut seraient considérées comme un autre élément de rapprochement entre les deux familles de virus dans le cadre du mécanisme de transcription discontinue.

den Boon et al. (1996) ont analysé plus en détails cet aspect et ont démontré la possibilité de la production de 7 molécules d'ARN de forme répliquative à partir de cellules infectées par le VAE et ce à l'aide d'essais de protection à la RNase. Ainsi, les auteurs suggéraient que les brins positifs d'ARN sg représentaient des transcrits naissant à partir de molécules matrices de polarité négative. Ceci est d'ailleurs appuyé par les mêmes observations obtenues chez les coronavirus (Sawicki et Sawicki 1990; Sethna et al., 1989). Quant à l'existence de molécules d'ARN sg brin négatif sans SL, elle serait un élément à clarifier pour obtenir une meilleure explication au modèle de transcription discontinue (den Boon et al., 1996).

Dans le même cadre d'étude, den Boon et al. (1996) ont effectué des expériences ciblant les motifs de jonction entre la SL et chacun des corps des ARNm sg. L'appariement entre l'extrémité 3' de la SL et la position où prend place la réinitiation de la transcription est l'élément clé dans la transcription discontinue et s'effectue par l'alignement du complexe leader-répliquase le long du promoteur de l'ARN sg. De toute façon, l'appariement est fortement possible à cause de la conservation de la séquence de jonction à des positions en amont de chaque ORF autre que l'ORF1. Selon la séquence génomique rapportée par den Boon et al. (1991), il existe 18 copies de ce motif tout le long du génome viral. Puisque seulement six ARN sg sont produits, cette séquence ne doit pas être le seul facteur déterminant de la fusion des ARNm avec la SL. En amont de chacun des ORF4, 5 et 7, deux motifs de jonction sont présents, et en amont de l'ORF2, il y en a trois. Si l'on considère une base supplémentaire en aval du motif de jonction, l'on observe la prédominance du motif 5' UCAACu 3'. Cependant, le motif de jonction de l'ARN sg6 retient une substitution d'une base du côté 3' soit 5' UCAACc 3'. Quant aux motifs autres que ceux présumés appartenant aux ARN sg (soit 11 motifs) présents le long du génome viral, trois parmi eux ont également la même séquence consensus 5' UCAACu 3'. Pourquoi n'y a-t-il donc pas d'autres ARN sg qui seraient produits et dont le nombre serait égal au nombre de séquences de jonction ? Pourtant malgré la présence de plus d'un motif de

jonction, aucune indication de l'utilisation de plus qu'un motif pour les ARNs 2, 4, 5, et 7 n'a été observée (den Boon et al. 1996).

Dans le cas de l'ARN 3, il y a 3 motifs de jonction en amont de l'ORF3 dont un seul est du type du consensus 5' UCAACu 3'. L'analyse des séquences de plusieurs clones dérivés de l'ARNm 3 a révélé la présence d'une population mélangée de molécules d'ARNm 3 dans les cellules infectées par le VAE (den Boon et al., 1996). Cette hétérogénéité est caractérisée soit par la présence de 2 à 3 nt supplémentaires en amont du motif consensus et marquant la limite de l'appariement de la SL avec cet ARNm 3, soit par la présence de motifs de jonction différents avec délétion d'une base à l'intérieur du motif lui-même (den Boon 1996). L'exemple typique est le manque d'une base C à la position 5 du motif d'un des ARNm 3 mais suivi d'une séquence de 5 nt (5' UACCC 3') immédiatement en aval et qui peut s'apparier avec ceux retrouvés à l'extrémité 3' de la SL.

L'analyse du motif de jonction utilisé pour l'ARNm 7 qui est le plus abondamment produit a révélé des bases supplémentaires dans le site d'appariement. Ces observations suggèrent que l'extension de la complémentarité des séquences entre l'extrémité 3' de la SL du génome et la site de jonction complémentaire (antileader) sur la matrice antisens devrait jouer un rôle important dans le processus de la fusion leader-corps des ARNm sg. De plus, selon den Boon et al. (1996), si cet appariement est impliqué dans la synthèse des ARNm sg, une activité régressive ou nucléasique de la polymérase doit être impérativement une étape précédant l'initiation ou la réinitiation de la transcription. Ceci suggère que l'appariement ne fonctionne que pour aligner initialement les deux séquences. Une fois le complexe de réinitiation est positionné correctement, les autres séquences deviennent superflues. Toutes ces données sont supportées par des observations similaires rapportées chez les coronavirus (Backer et Lai, 1990; Makino et al., 1988; Shieh et al., 1987; van Der Most et al., 1994).

Plus récemment, van Marle et al. (1999) ont clarifié encore plus en détails ces observations dans le cadre du mécanisme de la transcription discontinue. C'est grâce à la génération d'un clone infectieux du VAE (van Dinten et al., 1997) que des travaux plus minutieux ont pu aider à résoudre certains mystères reliés au mécanisme de la transcription des artérovirus et des nidovirus en général. Des expériences de mutagenèse au niveau de la séquence de jonction TRS ont affecté sévèrement l'expression des protéines structurales. Ces mutations ont été dirigées soit au niveau de la séquence TRS de la SL ou sur celle de l'ARNm 7. Ceci confirme aussi que l'appariement entre la séquence TRS de sens positif de

la SL et celle de sens négatif du corps de l'ARN sg brin négatif joue un rôle clé dans le processus de la transcription.

van Marle et al. (1999) ont obtenu des clones d'ADNc correspondant à l'ARNm 7 à partir de produit de RT-PCR utilisant comme gabarit un clone infectieux dont la TRS de la SL génomique est mutée mais pas celle du corps de l'ARNm 7 et vice versa. Curieusement, l'analyse des séquences d'ARNm sg synthétisées a révélé que la TRS du transcrit naissant est dérivée exclusivement de la TRS du corps de l'ARN et pas de celle de la SL. Suite à ces travaux, le mécanisme d'épissage alternatif conventionnel fut complètement rejeté et le mécanisme de transcription discontinue décrit par Sawicki et Sawicki (1995) est donc le plus plausible. Des brins sous gnomiques seraient produits et seraient à leur tour des gabarits pour la synthèse des ARNm sg.

2. 3. 2. SL du VFHS

Sagripanti et ses collaborateurs (1984, 1986) ont été les premiers à caractériser le génome du VFHS. Celui-ci est d'environ 15,7 kb possédant une coiffe de type I sous la forme m⁷G(5')ppp(5')Am et qui est la première de ce type à être caractérisée chez les artérovirus (Sagripanti et al., 1986). Quoique plusieurs publications ont rapporté des séquences de différents segments contigus ou discontigus du génome viral (Zeng et al., 1995; Smith et al. 1997; Godeny et al. 1995; Wang et al. 1998; Hwang et Brinton, 1998; Godeny et al., 1998), il n'y a pas de travail publié où l'on a séquencé la totalité du génome viral. Le plus long fragment de génome séquencé a été rapporté par Smith et al. (1997) qui ont déterminé la séquence de 6,3 kb du côté 3' du génome viral et ont déduit la présence de 9 ORFs chez le VFHS. Zeng et al. (1995) ont identifié, par des tests d'hybridation Northern à l'aide d'une sonde complémentaire à l'extrémité 3' du génome viral, six ARN sg dans les cellules de rein de singe vert africain (MA104) infectées par le VFHS. Le séquençage des extrémité 5' des ARN sg 6 et 7 a révélé la présence d'une SL commune jointe aux corps des ARNm sg par l'intermédiaire d'une séquence de jonction conservée 5' UUAACC 3' ou 5' UCAACC 3' pour chacun des deux ARN sg, respectivement. L'utilisation d'une sonde antileader a mis en évidence l'existence de cette SL au niveau de tous les ARN sg. Cette SL semble avoir une longueur de 208 nt sensiblement égale à celle du VAE (Zeng et al., 1995). Cette longueur a été déterminée suite à l'analyse des séquences de plusieurs clones dont certains dérivait de l'ARN 7 à partir desquels 97 nt de la région 3' ont été déterminés et

d'autres dérivait de la SL génomique à partir desquels 103 nt de la région 5' ont été déterminés. Ceci totalisait donc 200 nt. Cependant, selon les travaux de Zeng et al. (1995), les 2 nt de l'extrémité 5' de la SL manquaient dans leur séquence probablement à cause d'un encombrement stérique provoqué par la structure de coiffe lors de la réaction de RT-PCR. La séquence de jonction à l'extrémité 3' de la SL étant composée de 6 nt, Zeng et al. (1995) ont conclu que la SL entière du VFHS est de 208 nt. Les auteurs ont ensuite comparé la SL du VFHS par rapport aux autres SL des membres des artérovirus connues jusqu'alors en utilisant le programme GAP du système GCG (Genetic Computer Group de l'Université de Wisconsin). Ils ont trouvé un pourcentage de 35,4, 40,1 et 48,6 % d'identité de séquence avec le virus Lelystad (VL), le VAE et le VLD, respectivement. Ils ont conclu que les SL de ces différents membres n'étaient pas conservées d'autant plus qu'ils n'y ont même pas identifié de petites séquences hautement conservées sauf si l'on considère la séquence de jonction dans laquelle les seuls nt conservés parmi les quatre membres sont la triade 5' AAC 3' ou parmi les trois membres (sans le VAE), la séquence 5' AACC 3'.

Les travaux de Godeny et al. (1998) ont confirmé la présence de 9 ORFs le long de l'ARN génomique du VFHS mais leur expression est associée à seulement 8 ARNm sg. Le motif de jonction de la SL avec les corps des ARNm sg appelée TAS, a été redéfini comme ayant la séquence nonanucléotidique consensus 5' UCnUUAACC 3'. Ceci provient du fait qu'ils ont identifié 8 régions distinctes de jonction avec les nt aux positions 7 et 8 étant les plus strictement conservés. Ainsi, la fusion du corps de l'ARNm sg et la SL survient de façon préférentielle à des positions variables à l'intérieur de la séquence TAS selon l'ARNm sg en question. D'après ces auteurs, certaines bases de la séquence nonanucléotidique TAS retrouvées dans les ARNm sg naissants proviennent de la SL génomique et d'autres proviennent du corps de l'ARN génomique (Godeny et al., 1998).

Récemment Hwang et Brinton (1998) ont effectué des travaux ciblant la région interne de la SL et ne touchant pas la séquence de jonction à son extrémité 3'. Pour ces chercheurs, la SL du VFHS est d'une taille de 209 nt. Ils ont réalisé des essais de retardement de la mobilité sur gel pour déterminer des protéines d'extraits cytoplasmiques de cellules MA104 formant des complexes spécifiques avec la séquence antileader du VFHS. Ces protéines au nombre de 4 et dont les poids moléculaires sont de 103, 86, 55 et 36 kDa, ont été visualisées par des essais de liaisons croisées induites par les rayons ultraviolets (« UV-induced cross-linking ») et par des tests d'empreintes de type NorthWestern. Ces protéines ont été confirmées comme étant d'origine cellulaire et non

virale. De plus, la région de liaison de ces protéines a été délimitée à une séquence de 68 nt entre les positions 117 et 184 de l'extrémité 3' du génome brin négatif. Les auteurs ont par la suite démontré que ces mêmes protéines, plus une autre, se liaient avec l'extrémité 3' de la séquence antileader du génome du VLD ainsi que la séquence antileader du VAE (Hwang et Brinton, 1998). Dans le cas du VLD, le site de liaison était sensiblement aux positions similaires que le VFHS. Ces observations ont ainsi révélé que des protéines cellulaires forment des complexes de réplication virale de manière conservée, ce qui suggère un mécanisme de réplication presque identique chez les différents membres des artérovirus. Ce phénomène d'interaction de protéines cellulaires avec des séquences nucléotidiques aussi bien chez les coronavirus que chez d'autres virus à ARN et les ARN eucaryotes a déjà été documenté (Yu et Leibowitz, 1995).

2. 3. 3. SL du VLD

Le génome du VLD est une molécule d'ARN monocaténaire de polarité positive, d'une longueur de 12 à 13 kb (Darnell et Plagemann, 1972; Brinton-Darnell et Plagemann, 1975) avec une séquence poly-A de longueur variable (Contag et al., 1986; Brinton et al., 1986). Cependant, c'est grâce aux travaux de Godeny et al. (1993) que la séquence presque complète du génome viral de la souche VLD-C a été déterminée et estimée à environ 14 222 nt. Elle est composée de 8 ORFs exprimés à partir d'un ensemble de 7 ARNm sg à côté de l'ARN génomique (Kuo et al., 1991). La SL est aussi commune à tous les ARN sg et à l'ARN génomique viral. Le motif de jonction des ARNm 5, 6, et 8 a été déterminé comme étant l'hexanucléotide 5' UAACCA 3', et celui de l'ARNm 7, l'heptanucléotide 5' UAAAACC 3' (Kuo et al., 1992).

Quant à la détermination de la SL complète du virus, Chen et al., (1994) ont utilisé deux approches indépendantes pour la déduire. Dans la première approche, l'ARN génomique viral de la souche VLD-P a été rétro-transcrit pour obtenir un premier brin d'ADNc qui par la suite a été ligué à un oligonucléotide bloqué à son extrémité 3' avant d'être cloné et séquencé. Dans la seconde approche, de l'ARN génomique traité ou non traité avec une enzyme enlevant la structure de la coiffe a été par la suite traité avec la ligase à ARN pour joindre l'extrémité 3' à l'extrémité 5' du génome. Ainsi une réaction de RT-PCR à travers le site de jonction des deux extrémités fut réalisée, et a permis l'obtention de la séquence de la région non codante à l'extrémité 5'. Ces deux procédures

ont abouti à l'indication que la SL du VLD-P est de 156 nt incluant la séquence de jonction 5' UAUAAACC 3'. Ceci correspond à une séquence de 50 nt de moins par rapport aux SL du VAE et du VSRRP (Chen et al., 1994). Cette séquence est apparemment dépourvue d'un quelconque ORF et semble ne pas avoir des identités de séquence significatives avec les séquences correspondantes des virus VAE et VL. Selon Chen et al. (1994), la SL de la souche VLD-P diffère de seulement 5 nt (soit de 3%) de celle de la souche VLD-C déterminée par Godeny et al. (1993). Toutefois, les extrémités 5' et 3' s'étant liguées même dans le cas de l'absence de traitement avec l'enzyme enlevant la coiffe à l'extrémité 5', le génome du VLD-P serait alors dépourvu de la structure de la coiffe (Chen et al., 1994) contrairement à ce qui a été décrit chez le VFHS.

Par ailleurs, le VLD a été le premier artérovirus où l'on a détecté des ARN sg brins négatifs dans les macrophages infectés (Chen et al. 1994). Cela a renforcé l'exploration du mécanisme de transcription qui régit les artérovirus. A priori, des molécules d'ARN de la longueur de l'antigénome ont été détectées par hybridation Northern et des molécules d'ARN sg complémentaires de polarité négative ont été détectées par RT-PCR. De plus la présence de molécules d'ARN sg dans le virion n'a pas été détectée (Chen et al., 1994).

Plus récemment, Li et al. (1999), en se basant sur les différences nucléotidiques entre les souches neuropathogène VLD-C et non neuropathogène VLD-P, ont rapporté l'existence du phénomène de recombinaison tout comme c'est le cas chez les coronavirus. Ils ont considéré que le taux de recombinaison s'élevait à 5%, ce qui donnait naissance à des souches mixtes avec des stabilités variables chez l'hôte naturel.

Quant aux protéines cellulaires susceptibles d'interagir avec la SL du VLD, une étude comparative effectuée par Hwang et Brinton (1998) a démontré l'existence d'interaction avec les mêmes protéines de la cellule hôte et un mécanisme de réplication sensiblement similaire chez tous les membres de la famille des *Arteriviridae*.

2. 3. 4. SL du VSRRP

La maladie causée par le VSRRP a été décrite pour la première fois en 1987 en Amérique du Nord sous le nom de virus de la maladie mystérieuse du porc (Hill, 1990). Elle fut ensuite rapportée en 1990 en Europe (Paton et al., 1991). Elle a par la suite reçu plusieurs appellations telles que le syndrome respiratoire et d'infertilité porcine aux Etats-

Unis et le syndrome reproducteur et respiratoire porcin et le syndrome respiratoire et d'avortement épidémique porcin en Europe (Terpsra et al., 1991; Wensvoort et al., 1992).

C'est grâce aux travaux de Meulenberg et al. (1993a; 1993b) et Conzelmann et al. (1993) que le VSRRP ou le virus de Lelystad (VL) fut classifié dans le groupe des artérovirus. Le génome du VL est un ARN monocaténaire, polyadénylé, composé de 15 088 nt (Meulenberg et al. 1993a). Celui-ci contient 8 ORF dont les produits sont exprimés à partir d'une série de 6 ARNm sg détectés dans les cellules infectées. Ces ARNs sg sont aussi co-terminaux à leur extrémité 3' et contiennent une SL commune dérivée de l'extrémité 5' de l'ARN génomique et jointe à des corps d'ARN génomique du côté 3'. Par contre, la taille du génome ne fut pas définitivement déterminée puisque quelques bases à l'extrémité 5' étaient manquantes (Meulenberg et al., 1993a). Enfin, il a fallu attendre d'autres travaux concernant la production d'un clone d'ADNc infectieux qui exigeait la détermination préalable des extrémités 5' et 3' terminales (Meulenberg et al., 1998). Ainsi, ils ont complété la SL en ajoutant 10 nt supplémentaires par rapport à la séquence du génome publiée en 1993 (Meulenberg et al., 1993). Ceci résultait en une SL de la souche VL de 221 nt débutant avec une base A.

Le motif de jonction de la SL avec les corps des ARN sg peut également différer d'un ARN sg à un autre. Celui du VSRRP situé à proximité de l'ORF1a est similaire à un motif de jonction du VLD. Les ORFs 4 à 7 sont précédés par le motif 5' (U/A)(U/C)AACC 3'. Par contre celui des ARNs 2 et 3 est le motif 5' UAAACC 3' et le motif 5' UUGACC3', respectivement (Meulenberg et al., 1993a). Cependant, il faut noter que ces observations étaient basées sur des déductions de lecture de la séquence génomique. C'est pour cela que ce même groupe a effectué d'autres travaux dans lesquels ils ont déterminé que seul l'ARN 4 est pourvu d'un unique motif invariable puisque 8 clones d'ADNc dérivés de l'ARN 4 ont montré la même séquence de jonction 5' UCAACC 3'. Par contre, les autres ARN sg 2, 3, 5, 6, et 7 ont des séquences avoisinantes hétérogènes, ce qui suggère que le site de fusion de la SL peut être imprécis (Meulenberg et al., 1993b). Meng et al. (1996) ont confirmé cette observation en démontrant que lors de l'infection au VSRRP, il peut y avoir production de 7 ARN sg plutôt que 6. Le septième ARN provenait en effet de ce genre d'imprécision puisque c'est au niveau de l'ARN 3 qu'un autre type de motif de jonction a été utilisé et a donné naissance à cette espèce supplémentaire d'ARN. Ce motif apparaissait chez un isolat viral ayant subi une mutation ponctuelle et ayant rapproché nettement la séquence autour de cette mutation à une séquence de motif de jonction.

Une autre étude effectuée par Faaberg et al. (1998) a également révélé la possibilité de l'utilisation de plus d'un motif de jonction situé en amont de l'ORF7. La détermination des séquences de plusieurs clones d'ADNc dérivés de l'ARN sg7 générés par RT-PCR a permis d'identifier deux espèces de transcrits de cet ARN7 dans les cellules infectées. Un des transcrits contenait une séquence de jonction AUAACC à 123 nt en amont du codon d'initiation de l'ORF7 et l'autre transcrit contenait un motif UAAACC à 9 nt en amont de cet ORF. Ceci a aussi été rapporté par Oleksiewicz et al. (1999) qui ont constaté l'utilisation équitable de plus d'un motif de jonction pour la génération de l'ARN sg7. De plus, Faaberg et al. (1998) ont constaté que la SL de la souche prototype Américaine est différente dans sa longueur de celle de la souche Européenne VL, quoique ces deux souches utilisent la même séquence pour la formation de la jonction leader-corps des ARN sg.

Parallèlement aux travaux de Meulenberg et ses collaborateurs (1993a, 1993b), Conzelmann et al. (1993) ont caractérisé un long segment de 5 kb du côté 3' du génome du VSRRP. Six ORFs sont contenus dans la partie 3' de 3,5 kb. Ils ont rapporté que chaque ARN sg possédait une SL commune dont la taille n'était pas encore déterminée mais qui devait excéder 200 nt. Selon ces auteurs, des motifs de jonction communs pouvant excéder 3 nt n'avaient pas été identifiés. Par contre, les ORFs 2, 5, 6 et 7 étaient précédés du motif AACC qui représentait une partie de la séquence de jonction du VLD (Conzelmann et al., 1993).

Saito et al. (1996) ont également rapporté l'hétérogénéité du motif de jonction au sein d'un segment de 1,7 kb du génome d'une souche virale japonaise du VSRRP. La séquence consensus commune au génome et aux ARN sg a été représentée par le motif UA(A/G)CC. Cette hétérogénéité était détectée plus en amont du motif de jonction. Cela signifie que les nt de la séquence du côté du corps de l'ARN sg étaient identiques mais ils ne l'étaient pas du côté de la SL. Cette observation a été bien clarifiée chez le VAE avec les travaux de van Marle et al., (1999).

Une équipe de chercheurs Danois a récemment déterminé les SL complètes de trois isolats différents du VSRRP soit deux isolats Américain et Européen pathogènes et un type Américain vaccinal (Oleksiewicz et al., 1999). Curieusement, le type Européen était nettement distinct des deux types Américains étant donné que l'analyse de ces deux séquences a révélé seulement 50% d'identité nucléotidique. La SL du premier type avait une longueur de 220 nt alors que celle des deux autres types avait une longueur de 190 nt.

Néanmoins, des régions hautement conservées ont tout de même été identifiées chez les trois isolats. Il s'agit d'un motif de 12 nt débutant à partir de la troisième base et qu'ils ont appelé le « leader start motif » ainsi qu'un segment de 40 nt à l'extrémité 3' de la SL. Ce dernier contenait lui-même trois sous-motifs de 8, 9 et 11 nt invariables. De plus, le motif de 9 nt contient un site CACCC décrit chez d'autres virus tel que le virus Epstein-Barr (Nakagawa et al., 1997) et le virus de la leucémie des cellules T humaines (Okumura et al., 1996) comme étant impliqué dans des interactions avec des facteurs transcriptionnels cellulaires. Ces sites constituent probablement des motifs potentiels pour la fixation de certaines protéines impliquées dans la réplication virale (Oleksiewicz et al., 1999). Ceci pourrait être ajouté aux observations de Hwang et Brinton (1998) concernant les protéines cellulaires se liant à la séquence anti-leader des différents membres des artérovirus dans le but d'expliquer les mécanismes de transcription et de réplication de ces virus. Par ailleurs, Oleksiewicz et al. (1999) ont rapporté que les séquences des différents motifs de jonction utilisés par les souches Américaines ressemblaient à celles de la souche Européenne VL. De plus les séquences non traduites situées entre l'extrémité 3' de la SL et le codon d'initiation de chaque ORF des différents ARN sg avaient des longueurs comprises entre 6 nt pour l'ARN sg 4 et 140 nt pour l'ARN sg 7 (Oleksiewicz et al., 1999).

Plus récemment, Shen et al. (2000) ont déterminé la séquence complète d'une souche vaccinale Nord-Américaine du VSRRP (Prime Pacca PRRS). Elle comprend 15 520 nt incluant la queue poly-A, et contient une insertion dans la région Nsp2 de 36 AA par rapport à d'autres souches Nord-Américaines ou une insertion de 155 AA par rapport à la souche Européenne LV. Ceci rendait cette souche vaccinale la plus longue par rapport aux différentes souches dont la séquence complète a été déterminée. Cependant la comparaison de la séquence nucléotidique à l'extrémité 5' de cette souche SP par rapport aux souches Américaines et Européenne a démontré une conservation considérable entre ces souches. L'analyse des résultats de séquençage en combinaison avec des tests d'extension d'amorce ont confirmé la taille de la SL des deux souches vaccinale (SP) et Nord-Américaine (VR2332) comme étant 190 nt alors que la première nt à l'extrémité 5' était une base A pour la souche vaccinale et une base T pour la souche VR2332. Ces deux souches avaient un pourcentage d'identité très élevé de 98,5 %. Bien que la SL de la souche VL est composée de 221 nt, les auteurs insistent sur le fait que ses premières 12 nt étaient similaires à la souche Américaine VR2332 (Shen et al., 2000).

3. Rôles de la SL

3. 1. Rôles de la SL chez les coronavirus:

Stohlman et al. (1988) ont rapporté à l'aide de tests de NorthWestern blot l'existence d'une interaction entre la protéine N du virus VHM et tous les ARN viraux. Cette interaction se localise en particulier sur un segment de 10 nt entre les bases 56 et 65 de la SL. Selon ces auteurs, ces interactions justifient bien l'importance de la SL dans l'assemblage et probablement dans le processus de la transcription (Stohlman et al., 1988). Hoffman et al. (1993a) ont observé que les 5 nt à l'extrémité 5' de la SL du coronavirus bovin étaient hétérogènes et hypervariables durant les débuts d'infection en culture cellulaire. Par contre, une fois que l'infection persistante était établie, des substitutions à l'extrémité 5' de la SL ont donné naissance à un ORF intraleader (IL) de 33 nt. Ils ont montré par des tests de traduction *in vitro* que cet ORF atténuait la traduction des ORFs situés en aval. Ceci les a mené à suggérer que l'apparition de cet ORF IL durant les stades tardifs d'infection était à l'origine de l'infection persistente par ce virus (Hofmann et al. 1993b). En revanche, un effet inverse à cette observation a été rapporté chez le coronavirus murin VHM. Chen et Baric (1995) ont rapporté des données indiquant que la persistance du VHM était associée à la mutation et l'évolution de la région 5' non traduite. Il s'agissait d'une mutation A-à-G à la position 77 survenant après 55 jours d'infection et résultant en l'émergence d'un ORF de 18 AA en amont de l'ORF1a. Cette mutation semblait augmenter le taux de la traduction des ORFs 1a et 1b situés en aval qui sont indispensables à la répllication et la transcription virales. Tahara et al. (1994) ont examiné la propriété de traduction *in vitro* d'une séquence conservée dans la région leader 5' de l'ARN du coronavirus VHM. Ils ont extrait la SL du VHM et l'ont insérée dans un système contenant un ARN cellulaire. Ils ont alors remarqué une nette augmentation du taux de traduction des ARN contenant cette SL.

3. 2. Rôles de la SL chez d'autres modèles de virus à ARN:

Jang et al. (1989) ont démontré que l'initiation de la synthèse des protéines du virus de l'encéphalomyocardite *in vivo* était médiée par l'entrée interne des ribosomes dans une structure de la région 5' non codante. Rhol et al. (1994) ont identifié des domaines

fonctionnels indépendants dans la région 5' non traduite de l'ARN des entérovirus et qui sont essentiels à la réplication et la traduction. SenGupta et al. (1990) ont démontré que la SL du virus de l'immunodéficience humaine de type 1 (VIH1) s'associait avec une protéine virale accessoire Tat pour la régulation de la traduction. Donze et Spahr (1992) ont démontré le rôle de la SL du virus du sarcome de Rous dans la traduction et l'assemblage de l'ARN viral. Par ailleurs, chez d'autres rétrovirus aviaires comme le virus du sarcome de Rous et le virus de la leucose aviaire, des mutations introduites dans une région de 7 nt faisant partie d'une structure secondaire et où l'appariement d'une des bases est détruit ont causé des troubles d'initiation de la transcription inverse aussi bien *in vivo* qu'*in vitro* (Cobrinik et al., 1991; Aiyar et al. 1992). Quant aux bases appartenant à l'espace de non appariement de ces structures secondaires, le nombre de bases de cette région a été démontré comme ayant un effet plus prononcé dans la transcription inverse que leur substitution elle-même (Aiyar et al., 1994). Chez certains virus des plantes tels que le virus de la mosaïque alfalfa (van der Vossen et al., 1993), le virus du chou-fleur (Futterer et al., 1990) et le virus de la mosaïque du Tabac (Sleat et al., 1988), il a été démontré que la SL jouait aussi des rôles importants dans la réplication et la traduction de l'ARN viral.

3. 3. Rôles de la SL chez des modèles de virus à ADN:

Chez le virus simien 40 (SV40), la SL code pour un petit ORF qui est traduit dans 20 % des cas et dont sa présence réduit de 3 fois la synthèse des protéines virales en aval (Sedman et al., 1989). Chiou et al. (1991) ont attribué un effet d'accumulation des ARNm du virus SV40 à une augmentation de l'efficacité de la polyadénylation médiée par la SL. Dans le cas d'un herpesvirus, le virus Herpes Simplex type 1, la SL est impliquée dans l'augmentation de la transcription et la stabilité des ARNm (Blair et al., 1987).

3. 4. Rôles de la SL chez des modèles d'ARNm eucaryotes:

Sagliocco et al. (1993) ont démontré l'influence de la structure secondaire de la SL sur la fixation du ribosomes sur l'ARNm des levures. Wakiyama et al., (1992) ont rapporté quelques effets *in vitro* de la SL des ARNm sur l'efficacité de la traduction. Celle-ci était nettement affectée par la longueur de la SL ainsi que le nombre de codons d'initiation AUG qui y sont présents. Par ailleurs, les travaux de Kozak ont mis en évidence les

caractéristiques de la séquence 5' non codante dans un modèle de l'ARNm de la globine de lapin (Kozak, 1994). Ainsi, plusieurs règles de l'effet de la SL sur l'ARNm des eucaryotes ont été établies. La longueur de la SL affecte la traduction; plus elle est longue, plus elle favorise l'initiation de la traduction (Kozak, 1988, 1991a, 1991b). De plus, des structures secondaires particulières modulent l'initiation de la traduction par leur forme elle-même, par leur position, ou par le degré leur stabilité (Kozak, 1988, 1990).

CHAPITRE II

PROBLÉMATIQUE ET OBJECTIFS

Ce travail concerne l'analyse moléculaire de la SL du VAE. Ainsi, plusieurs sous objectifs ou objectifs intermédiaires ont été tracés avant de débiter ce projet ou au cours de sa réalisation. Au début du projet, ces objectifs sont justifiés par des données publiées chez d'autres modèles de virus rapportés dans la littérature. Au cours du projet, ils sont jugés impératifs, faisant suite à des résultats précédents pour garder un enchaînement logique et pour arriver au but principal.

1. Objectif 1: Détermination de la séquence "leader" complète du VAE

L'objectif principal de ce projet est d'étudier le rôle biologique de la SL du VAE. Cependant, il est impératif de connaître sa séquence nucléotidique complète avant d'entamer ces analyses. Deux principales publications ont été rapportées au début des années 90 et apportaient les premiers détails de la SL du VAE (de Vries et al., 1990; den Boon et al., 1991). D'après les travaux de de Vries et al. (1990), la taille de la SL du VAE fut estimée à 208 nt. Ceci provenait de la combinaison des résultats des expériences d'extension d'amorce ainsi que de l'analyse de la position du site du motif jonction sur la séquence de l'ARN6. Parmi une série d'une librairie d'ADNc, un clone contenait une partie de la SL du VAE. Cette séquence débutait, selon les estimations, à la 19 ième base de la SL (de Vries et al., 1990). den Boon et al. (1991) ont par la suite répété les expériences d'extension d'amorce pour déterminer la taille de la SL et ce en utilisant d'autres amorces spécifiques d'une région interne de cette séquence. Ces derniers ont ainsi estimé la taille de la SL du VAE à 207 nt plutôt que 208 nt. C'est pour cela que dans leurs travaux (den Boon et al., 1991), un segment d'environ 18 nt de la séquence du génome à l'extrémité 5' était laissé indéfini et marqué avec des lettres N. Ensuite, ce segment se poursuivait avec la base T. Le début de la séquence génomique était ainsi représentée: 5' NNNNNNNNNNNNNN NNNNTGCCATATA...etc 3' (numéro d'accès GenBank X53459). Par conséquent, notre premier objectif était de déterminer l'extrémité 5' du génome du VAE.

2. Objectif 2: Étude de la stabilité de la SL par l'analyse génétique de différentes souches et/ou isolats du VAE de provenance différente et prédiction et analyse de leur structure secondaire.

Des études épidémiologiques antérieures de l'artérite virale équine ont suggéré l'existence de plusieurs souches qui différaient dans leur virulence ainsi que dans leur

pouvoir abortif (Cole et al., 1986, Golnik et al., 1986). Or, depuis longtemps, l'on pensait qu'il n'existait qu'un seul sérotype parmi les différentes souches du VAE (McCollum, 1969). Les études sérologiques déterminées par des tests de séroneutralisation croisée et de fixation du complément ont démontré que les variations étaient mineures (Fukunaga et McCollum, 1977; Fukunaga et al., 1994). Par ailleurs, d'autres études portant sur les différences génétiques en mettant à profit le test des empreintes oligonucléotidiques à la RNase T1 ont également démontré que les variations entre diverses souches du VAE étaient mineures (Murphy et al., 1988, 1992). Plus tard, une étude de comparaison des séquences des gènes codant pour les protéines M et N de plusieurs isolats du VAE a démontré que ces séquences étaient hautement conservées (Chirnside et al., 1994). D'autres publications durant l'accomplissement de ce projet se poursuivaient et indiquaient que d'autres régions montraient aussi une conservation entre plusieurs souches et isolats de provenance différentes. Dans une autre étude comparative, la variation génétique et la relation phylogénétique de l'ORF5 de 3 groupes de souches différentes de laboratoires représentés par la souche prototype originale Bucyrus, la souche vaccinale atténuée (ARVAC, Fort Dodge Laboratories), et un groupe de 18 isolats de champs d'Amérique du Nord et d'Europe ont été déterminées (Balasuryia et al., 1995). Les niveaux d'homologie en séquences nucléotidiques allaient de 85,7 % à 99,7 %, alors que les niveaux d'homologie en séquences d'AA allaient de 89,8 à 99,6 %. Malgré ces petites différences, les auteurs maintenaient l'idée que ces virus étaient étroitement apparentés. Selon leur analyse comparative des séquences d'AA déduites de la protéine G_L, ils ont identifié 3 régions variables pouvant contenir les sites antigéniques de neutralisation. De plus, ceci les a mené à suggérer 4 groupes d'isolats distincts selon les estimations des distances génétiques et de l'arbre phylogénétique établi (Balasuriya et al., 1995). Des analyses sur le gène codant pour l'ORF2 des mêmes groupes de virus ont démontré que cette protéine est plus conservée que la protéine G_L. (Hedges et al., 1996).

D'autres études effectuées dans notre laboratoire, portant sur les variations génétiques des gènes des protéines G_L et G_S ont rapporté des taux similaires d'homologie de séquences obtenues de souches et d'isolats différents. Dans le cas du gène de la protéine G_L, 7 isolats de champ ont été comparés avec la souche de référence Bucyrus. Les pourcentages d'identité nucléotidique avaient des taux de 87,5% à 93,9% alors que les identités en séquences d'AA avaient des taux de 90,2% à 95,1%. Selon les prédictions de cette étude, 2 groupes phylogénétiques distincts ont été retenus (St-Laurent et al., 1997).

Dans le cas du gène de la protéine G_S, les mêmes isolats de champ, ont été comparés avec la souche de référence Bucyrus. Les taux d'homologie de nt entre ces différents isolats étaient de 92,3 à 96,4% et celles des séquences en AA étaient de 93,2 à 95,5%. Selon les prédictions de cette étude, trois groupes phylogénétiques ont été obtenus (Lepage et al., 1996). Enfin, quant aux gènes des ORF 3 et 4, la souche vaccinale Arvac a été ajoutée aux sept derniers isolats de champs pour faire les mêmes études comparatives. Les taux d'homologie de nt entre ces différents isolats étaient de 85,6 à 98,8% et de 90,4 à 98,3% pour les gènes de l'ORF3 et de l'ORF4, respectivement. Les taux d'homologie des séquences en AA étaient de 85,3 à 98,2% et 90,8 à 97,4% respectivement (Archambault et al., 1997). Selon les analyses phylogénétiques basées sur la séquence nucléotidique de l'ORF 3 la souche Européenne Vienne pouvait être classée dans un groupe phylogénétique divergent. Cependant, la séquence de l'ORF 4 semble être moins variable (Archambault et al., 1997).

Les diverses prédictions de ces études n'ont pas été vérifiées avec des tests biologiques. Or, depuis longtemps et jusqu'à présent, les cliniciens ont toujours évoqué l'existence de souches virulentes et de souches avirulentes du VAE. Fukunaga et al. (1993), ont rapporté dans une étude de séroneutralisation en présence du complément qu'un sérum immun préparé contre une souche avirulente réagissait contre la souche virale homologue avirulente mais pas contre des souches virulentes. De plus, depuis la découverte du VAE, celui-ci a été isolé dans plusieurs continents et pays avec des virulences variables. Dans certaines épidémies, la souche virale est virulente avec un pouvoir abortif élevé (Cole et al., 1986; Golnik et al., 1986). Par contre, il existe des cas où la manifestation est sporadique et la souche virale est non-abortive et avirulente (Timoney et McCollum, 1993). Ces variations dans la manifestation clinique peuvent être associées à plusieurs facteurs y compris des différences liées à la souche virale elle-même. En effet, dans le cas du virus de l'encéphalomyélite murine de Theiler, la délétion d'une seule base dans la région 5' non codante d'un clone infectieux d'ADNc provoque l'atténuation de la neurovirulence virale (Pritchard et al., 1992). Bandyopadhyay et al (1993) ont ensuite démontré que l'insertion de seulement 3 bases dans la région 5' non codante du même virus provoque également la réduction de la neurovirulence. Selon ces auteurs, ces effets de virulence étaient dus à la destruction de l'intégrité d'une structure secondaire en boucle d'épingle au niveau de la séquence 5' non codante. Cette structure est particulièrement indispensable et importante pour le déroulement du mécanisme d'initiation de la traduction virale (Pritchard et al.,

1992; Bandyopadhyay et al., 1993). Chez les échovirus de la famille des *Picornaviridae*, des différences de phénotype de virulence ont aussi été associées à des variations génétiques au niveau de la SL de certaines souches virales (Romero et Rotbart, 1995).

Tout cela a conduit à tracer des objectifs d'analyser la stabilité de la SL du VAE en réalisant une étude comparative des différentes souches et/ou isolats du VAE de provenance différente au niveau de leur structure primaire et leur structure secondaire.

3. Objectif 3: Étude du rôle de la séquence "leader" dans la réplication du VAE et l'expression de ses protéines.

Nous avons émis l'hypothèse que toute propriété observée au niveau de la SL ou seulement d'un segment plus ou moins large de la SL serait importante vu la caractéristique de sa présence à l'extrémité de l'ARN génomique et des différents ARNm sg. Ceci laisse présumer l'importance de la SL dans l'expression des gènes de la réplicase et des protéines non structurales au début de l'infection virale et dans l'expression des gènes structuraux. Or, à part l'implication de la SL du VAE dans la transcription des ARNs viraux par l'intermédiaire de son motif de jonction, très peu d'information sur le reste de la SL était disponible. Ceci nous a orienté à effectuer une analyse moléculaire fonctionnelle et une étude du rôle de la SL du VAE.

CHAPITRE III

Article 1

Sequence determination of the extreme 5' end of the equine arteritis virus leader region

Kheyar, A., St-Laurent, G., and Archambault, D. 1996. Virus Genes 12:291-295.

Sequence Determination of the Extreme 5' End of Equine Arteritis Virus Leader Region

ALI KHEYAR, GILLES ST-LAURENT, & DENIS ARCHAMBAULT

Université du Québec à Montréal, Département des Sciences Biologiques, Montréal, Québec, Canada, H3C 3P8

Received February 9, 1996; accepted March 19, 1996

Abstract. The extreme 5' end of the leader sequence of four equine arteritis virus (EAV) strains was obtained by using rapid amplification of cDNA end method (5' RACE), and sequenced. Seventeen more nucleotides were added upstream of the 5' end of the EAV published genomic sequence. A common feature among the analyzed EAV isolates was the presence of an AUG start codon within the added sequence and the appearance of an intraleader open reading frame (ORF) of 111 nucleotides which was predicted to encode a peptide of 37 amino acids. The role of this putative intraleader ORF has yet to be determined.

Key words: equine arteritis virus, 5' terminus, intraleader ORF

Equine arteritis virus (EAV) is the etiologic agent of equine viral arteritis, a debilitating respiratory disease with the most severe form resulting in abortion from pregnant mares (1). Although the virus transmission is primarily via the respiratory route, the virus is also shed in the semen of persistently infected stallions which infect mares at the time of breeding (1). EAV is the prototype species of the Arterivirus group, which includes lactate dehydrogenase-elevating virus (LDV), porcine reproductive and respiratory syndrome virus (PRRSV) and simian hemorrhagic fever virus (SHFV) (2). The EAV genome is a positive, polyadenylated, single stranded RNA of approximately 12.7 kb (3). The replication strategy of EAV resembles that of corona- and toroviruses. During virus replication, a 3' end coterminally nested set of seven virus-specific RNAs is produced with a common leader sequence derived from the 5' end of the EAV genome (4). The leader sequence is joined to each open reading frame (ORF) by a junction sequence motif 5' UCAAC 3' (14). On the basis of primer-extension experiments, the EAV leader sequence has been predicted to be 207 (3) or 208 (4) nucleotides (nt) in length. However, sequence data obtained from cDNA clones derived from a

genomic EAV cDNA library (3) failed to identify the most extreme 5'-terminal nt of the leader sequence.

In this study, the extreme 5' end of the EAV leader sequence of four EAV strains is described, with a prediction of an intraleader open reading frame (ORF) encoding a short peptide of 37 amino acids within the full length leader sequence. To this end, the EAV Bucyrus reference strain (5), and the EAV laboratory 87AR-A1, 86NY-A1 and Vienna strains (6), with low cell passage levels ranging from two to four, were used. The Bucyrus strain was isolated from tissues of a fetus aborted during an abortion endemic episode in standardbred horses (5), while the others were isolated from nasal swab (Vienna) or semen of infected horses (87AR-A1, 86NY-A1) (6). Each virus strain was plaque-purified and propagated for one to two additional passages in rabbit kidney (RK-13) cells (7). After three freeze-thaw cycles, cell culture supernatant was centrifuged at 5000 × g for clarification. Virion RNA was extracted from 300 µl of infected cell culture supernatant by the guanidium isothiocyanate method (8). The RNA sample was resuspended in 20 µl of diethylpyrocarbonate (DEPC)-treated water containing 15 units of hu-

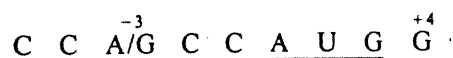
man placental RNase inhibitor (Pharmacia), and kept at -70°C until used.

To determine the 5' end of the EAV leader sequence, the single strand ligation to single-stranded cDNA (SLIC) method (9) was used. For this purpose, the 5'-Amplifinder RACE Kit (Clontech) was used to amplify the 5' distal end of the EAV leader sequence. Briefly, purified EAV RNA was preincubated at 65°C for 5 min and reverse transcribed with avian myeloblastosis virus (AMV) reverse transcriptase (RT) using the cDNA synthesis procedure described in the manufacturer's protocol. The RT reaction was primed with 20 pmol of the antisense oligonucleotide primer PEV-L1 (5'GTGGAGCCGTC CACTTC 3') which is complementary to nt 306 to 323 of the EAV genome (3) (Genebank accession number X 53459). This primer sequence is located downstream from the 3' end of the leader sequence in the ORF 1a genomic region. After RNA hydrolysis, the first single-stranded cDNA (ss-cDNA) was purified using the GENO-BIND™ kit (Clontech). The 3' end blocked AmpliFINDER anchor sequence was ligated to the 3' end of ss-cDNA at room temperature for 18 h with T4 RNA ligase. The single-stranded ligation product was then used in the polymerase chain reaction (PCR) procedure. The oligonucleotide primers used in the PCR consisted of antisense primer PEV-L02 (5'ACCCGTCAAGCCA CAAGATG 3'), which is complementary to an internal sequence (nt 165 to 145) of the EAV leader sequence (3), and the sense AmpliFINDER anchor primer (AFAP). The PCR reaction was then performed using a thermal cycler programmed for 30 successive cycles at conditions recommended by the supplier. The PCR products were then cloned into the *Sma* I cleaved pBluescript II KS+ plasmid vector (Stratagene) and sequenced by the dideoxynucleotide chain-termination method (10).

Sequence analysis of the EAV leader sequence region of 3 clones of the Bucyrus strain revealed the presence of 17 additional nt (5' ACT CGAAGTGTGTATGG 3') absent from the published sequence (3) and located at the 5' end (Fig. 1A). The same additional 17 nt were also obtained from several clones of the 87AR-A1, and Vienna EAV strains (Fig. 1B and C). The results obtained with the 86NY-A1 strain showed a se-

quence similar to the other EAV strains we studied with the exception of two substitutions (G→A) at positions +1 and +7 (Fig. 1D). Sequencing of PCR products from the four virus strains analyzed confirmed that the missing most extreme 5'-terminal nt were of EAV origin, since it was part of the nt sequence of the EAV genome according to the published data (3). With the addition of these 17 nucleotides, the length of the EAV leader sequence would be 206 nt without the junction site motif (Fig. 2). Thus, the EAV leader sequence is identical in length to that reported for PRRSV (11) and is longer than that of LDV (156 nt) (12) or SHFV (202 nt) (13). However, if a cap structure exists, thus an additional G would represent the 207th base, as already suggested (3). In any case, EAV and LDV (14) are, at present, the only two arteriviruses for which the entire leader sequence is determined, in comparison to the PRRSV genome (11) and SHFV genome (13), for which the first two 5' nucleotides of the leader sequence remain to be determined.

An important observation found in this study was the presence of a unique AUG initiator codon located at nucleotide positions +14 to +16. The optimal sequence context for efficient recognition as an internal initiation signal for RNA translation was determined to be



with the -3 and $+4$ nucleotide positions being of primary importance (14). On the basis of the later definition, the initiator codon is designated as strong or weak depending on the presence or absence of these residues at these positions (15). The -3 and $+4$ nucleotide positions of the EAV leader sequence are occupied by U and G, respectively. This suggests that the EAV intraleader AUG might be a weak initiator codon and therefore may provide a suboptimal context for translation initiation (14). Moreover, a UAG terminator codon occurs at positions +125 to +127 of the EAV published genomic sequence (3), thereby showing the presence of an intraleader ORF. This intraleader ORF sequence of 111 nt in length is predicted to encode a short peptide of 37 amino acids (Fig. 2). The estimated molecular mass of this peptide is 4.6 kDa.

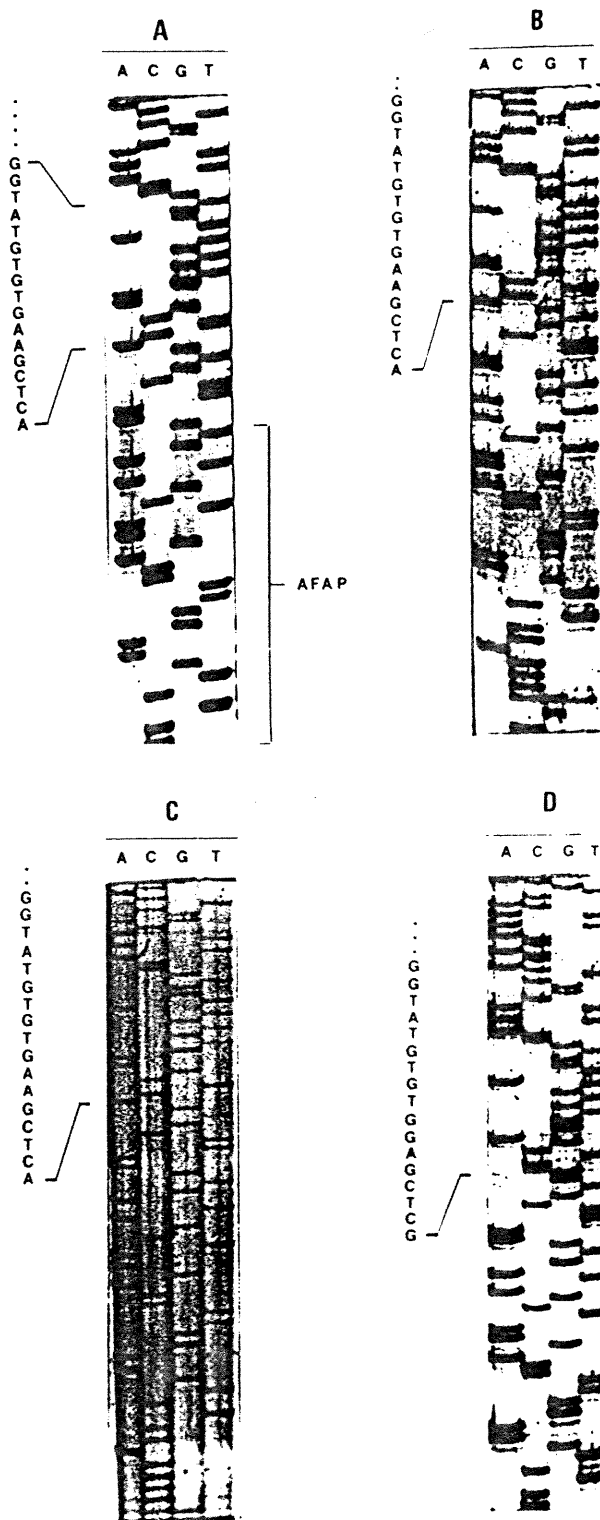


Fig. 1. Sequence determination of the extreme 5' end of the leader region of various EAV isolates. A: Bucyrus reference strain; B: 87AR-A1 strain; C: Vienna strain; D: 86NY-A1 strain. Seventeen additional residues (5' ACTCGAAGTGTGTA TGG 3') are present at the 5' end of the EAV Bucyrus strain genome. The assigned GenBank Database accession numbers are U46944 (Bucyrus), U46945 (86NY-A1), U46946 (87AR-A1), and U46947 (Vienna).

```

      *           *           *           *
1  ACTCGAAGTGTGTATGGTGCCATATACGGCTCACCACCATATACACTGCA
      M V P Y T A H H H I H C K

      *           *           *           *
51 AGAATTACTATTCTTGTGGGCCCCCTCTCGGTAAATCCTAGAGGGCTTTCC
      N Y Y S C G P L S V N P R G L S

      *           *           *           *
101 TCTCGTTATTGCGAGATTGCGTTAGATAACGGCAAGTTCCTTTCTTA
      S R Y C E I R R .

      *           *           *           *
151 CTATCCTATTTTCATCTTGTGGCTTGACGGGTCACTGCCATCGTCGTCGA

201 TCTCTA

```

Fig. 2. Complete sequence of the EAV Bucyrus strain leader region. The 5' end 17 nt added to the previously published sequence (3) are underlined. The intraleader ORF sequence is located between ATG start and TAG stop codons (bolded and underlined). The corresponding putative intraleader ORF-encoded amino acid sequence (bottom line) is also shown.

It is well known that members of both coronaviruses and arteriviruses cause persistent infections in their respective host (16–18). It has been reported that short ORFs within the 5' leader region of some eukaryotic mRNAs attenuate the rate of translation initiation at the downstream ORF (19). In fact, a translation-attenuating intraleader ORF of 33 nt has been described in bovine coronavirus during persistent infection (18). In contrast, a small ORF of 18 amino acids that was observed in the mouse hepatitis virus (which is also a coronavirus) leader sequence during a persistent infection, has been demonstrated to enhance translation of the downstream ORF (20). It is thus possible, by analogy with these virus systems, that the intraleader ORF could be involved in EAV RNA replication and/or translation regulation. However, the presence and role of such an intraleader ORF-encoded peptide in EAV life cycle have yet to be determined. Nevertheless, determination of the complete 5' end sequence provide useful information in construction of an infectious cDNA for a better understanding of the biology of EAV.

Acknowledgments

This work was supported by an operating grant from the National Sciences and Engineering Re-

search Council of Canada to D. Archambault. A. Kheyar is supported by a graduate student fellowship from Université de Montréal. D. Archambault is the holder of a research scholarship from the Fonds de la Recherche en Santé du Québec (FRSQ). We thank Peter J. Timoney and William H. McCollum (Gluck Equine Research Center, Lexington, Kentucky) for providing the EAV laboratory strains. We are also grateful to Amer Silim and Carolina Alfieri for reviewing the manuscript, and to Carole Ville-neuve for secretarial work.

References

1. Timoney P.J. and McCollum W.H., *Vet Clinics North Amer:Equine Practice* 9, 295–309, 1993.
2. *Virology Division News., Arch Virol* 135, 227–237, 1994.
3. Den Boon J.A., Snijder E.J., Chirnside E.D., De Vries A.A.F., Horzinek M.C., and Spaan W.J.M., *J Virol* 65, 2910–2920, 1991.
4. De Vries A.A.F., Chirnside E.D., Bredenbeek P.J., Gravestain L.A., Horzinek M.C., Spaan W.J.M., *Nucleic Acids Res* 18, 3241–3247, 1990.
5. Doll E.R., Knappenberger R.E., and Bryans J.T., *Cornell Vet* 47, 69–75, 1957.
6. Murphy T.W., McCollum W.H., Timoney P.J., Klingeborn B.W., Hyllseth B., Golnik W., and Erasmus B., *Vet Microbiol* 32, 101–115, 1992.
7. Saint-Laurent G., Morin G., and Archambault D., *J Clin Microbiol* 32, 658–665, 1994.
8. Chomzynski P. and Sacchi N., *Anal Biochem* 162, 156–159, 1987.

9. Apte A.N. and Siebert P.D., *BioTechniques* 15, 890-893, 1993.
10. Sanger F., Nicklen S., and Coulson A.R., *Proc Natl Acad Sci (USA)*, 74, 5463-5467, 1977.
11. Conzelmann K.-K., Visser N., Van Woensel P., and Thiel H.J., *Virology* 193, 329-339, 1993.
12. Chen Z., Faaberg K.S., and Plagemann P.G.W., *J Gen Virol* 75, 925-930, 1994.
13. Zeng L., Godeny E.K., Methven S.L., and Brinton M.A., *Virology* 207, 543-548, 1995.
14. Kozak M., *J Mol Biol* 196, 947-950, 1987.
15. Kozak M., *J Cell Biol* 108, 229-241, 1989.
16. Plagemann P.G.W. and Moennig V., *Adv Virus Res* 41, 99-192, 1992.
17. Perlman S., Jakobsen G., Olsen A.L., and Afifi A., *Virology*, 175, 418-426, 1990.
18. Hofmann M.A., Savithra D.S., and Brian D., *Proc Natl Acad Sci (USA)* 90, 11733-11737, 1993.
19. Kozak M., *J Cell Biol* 115, 887-903, 1991.
20. Chen W. and Baric R.S., *J Virol* 69, 7529-7540, 1995.

CHAPITRE IV

Article 2

**Nucleotide sequence and genetic analysis of the leader region of Canadian,
American and European equine arteritis virus isolates.**

*Kheyar, A., St-Laurent, G., Diouri, M., and Archambault, D.
1998. Can. J. Vet. Res. 62:224-230.*

Short Communications

Nucleotide Sequence and Genetic Analysis of the Leader Region of Canadian, American and European Equine Arteritis Virus Isolates

Ali Kheyar, Gilles St-Laurent, Mounir Diouri, and Denis Archambault

ABSTRACT

The extreme 5' end, the entire leader sequence of the Arvac vaccine strain, and 10 equine arteritis virus (EAV) isolates, including the ATCC Bucyrus reference strain and 5 Canadian field isolates, were determined and compared at the primary nucleotide and secondary structure levels. The leader sequence of eight EAV isolates, including the Bucyrus reference strain, and the leader sequence of the Arvac vaccine strain was determined to be 206 nt in length (not including the putative 5' cap structure-associated nucleotide) whereas those of the 86AB-A1 and 86NY-A1 isolates were found to be 205 and 207 nt in length, respectively. The sequence identity of the leader sequences, between the different isolates and the Bucyrus reference strain, ranged from 94.2 to 98.5%. Phylogenetic analysis and estimation of genetic distances, based on the leader nucleic acid sequences, showed that all EAV isolates/strains are likely to represent a large phylogenetically-related group. An AUG start codon found at position 14 in all EAV isolates/strains could initiate an open reading frame (ORF) that could produce a polypeptide of 37 amino acids, except for the 86NY-A1 isolate where the intraleader polypeptide would contain 54 amino acids. Computer-predicted RNA secondary structures were identified in the 11 EAV leader regions analyzed. All EAV isolates/strains showed 3 conserved stem-loops (designated A, B and C). An additional conserved stem-loop (D) was

observed in 7 EAV isolates, including the Bucyrus reference strain. The leader region distal to stem-loop D did not contain conserved sequences or stem-loop structures common to the EAV isolates/strains.

RÉSUMÉ

Les séquences d'acides nucléiques de la région de tête en 5' de l'ARN génomique de la souche vaccinale Arvac et de dix isolats du virus de l'artérite équine (VAE), incluant la souche de référence Bucyrus de l'ATCC et cinq isolats canadiens, furent déterminées et comparées entre elles. La séquence de tête de huit isolats du VAE, incluant les souches Bucyrus et Arvac, était constituée de 206 nucléotides (nt), alors que celles des isolats 86AB-A1 et 86NY-A1 étaient composées de 205 et 207 nt, respectivement. Les niveaux d'identité des bases nucléotidiques entre les différents isolats et la souche Bucyrus ont varié de 94,2 à 98,5 %. Une analyse phylogénétique et une estimation de la distance génétique des séquences d'acides nucléiques a indiqué que tous les isolats et souches de VAE pouvaient être classés dans un même groupe. Un codon d'initiation AUG trouvé à la position 14 et présent chez tous les isolats et souches du VAE pouvait ouvrir un cadre de lecture pouvant potentiellement produire un polypeptide de 37 acides aminés, à l'exception de l'isolat 86NY-A1 dont le polypeptide serait plutôt de 54 acides aminés. Les structures secondaires

au niveau de la séquence de la région de tête de tous les isolats et souches du VAE furent ensuite prédites avec un logiciel approprié. Trois boucles conservées en tête d'épingle (nommées A, B et C) furent observées chez tous les isolats et souches du VAE. Une quatrième boucle conservée et nommée D était aussi présente chez 7 isolats du VAE, incluant la souche de référence Bucyrus. Le reste de la séquence de la région de tête située en aval de la boucle D en termes de structures secondaires et de composition d'acides nucléiques était plus variable.

(Traduit par docteur Serge Messier)

Equine arteritis virus (EAV), the causative agent of equine viral arteritis, is the prototype of the *Arteriviridae* family in the order *Nidovirales*, that also includes lactate dehydrogenase-elevating virus (LDV), porcine reproductive and respiratory syndrome virus (PRRSV) and simian hemorrhagic fever virus (SHFV) (1). EAV was first isolated in Bucyrus, Ohio, from an aborted fetus obtained during an epidemic episode of abortion in horses (2). The symptomatology associated with EAV infection is variable, ranging from subclinical to overt clinical signs (which, when present, generally include pyrexia, conjunctivitis, lacrimation, nasal discharge, and edema) of variable severity, indicating that virulence varies among EAV isolates (2). Abortion is common when pregnant mares become infected (2).

The organization and replication of the EAV genome, a 12.7-kb polyadenylated, single-stranded positive

Département des Sciences Biologiques, Université du Québec à Montréal, C. P. 8888, Succursale centre-ville, Montréal, Québec H3C 3P8.

Correspondence and reprint requests to: Dr. D. Archambault.

Received July 18, 1997.

sense RNA, resemble those of *Coronaviridae* family. EAV messenger RNA involves the production of a 3' co-terminal nested set of 6 subgenomic viral RNAs, each possessing a common leader sequence derived from the 5' end of the viral genome and a 3' poly-A tail (3). The EAV genome contains 8 open reading frames (ORFs), each preceded by the pentanucleotide sequence motif 5'-UCAAC-3', designated by the leader-body junction site (3). ORFs 1a/1b encode for the viral polymerase, and ORF 3- and 4-encoded products are believed to be glycosylated, non-structural proteins (4,5). Four virion structural proteins are produced from ORFs 2, 5, 6, and 7. ORFs 6 and 7 encode an unglycosylated membrane (M) protein of 16 kDa and a 14 kDa nucleocapsid (N) protein, respectively (5). ORFs 2 and 5 encode the glycosylated 25-kDa small (G_S) membrane protein and the heterogeneously-glycosylated 30- to 42-kDa large (G_L) membrane protein (5), respectively. The G_L protein expresses the EAV neutralizing determinants (6-9).

Although it is believed that there is only one EAV serotype (2), minor antigenic differences among various EAV isolates have been reported (10). RNA oligonucleotide fingerprint analysis has indicated genetic variation among various EAV isolates (11). Comparison of M, N, G_L and G_S nucleic acid sequences, whose products induce variable antibody responses (5,9,12), has also demonstrated variation among EAV isolates (13-17). More recently, we also showed variation in the EAV protein-encoding gene ORF 3 of various EAV isolates, whereas the ORF 4 nucleic acid sequences were found to be highly conserved among the EAV isolates (18).

Very little information is available for the EAV leader region. We recently determined the sequence of the terminal (17 nt) 5' leader region of EAV (19). We (19), and other investigators (20), showed that the leader region is 206 nucleotides (nt) in length (not including a putative 5' cap structure-associated nt). We also identified an intraleader ORF of 111 nt in length (19), the role of which has yet to be determined. The goal of this study was to determine and compare the complete leader

TABLE I. Characteristics of the equine arteritis virus (EAV) isolates/strains used in this study

Isolates/Strains	Origin (Year of isolation)	Source	Passage history
Bucyrus ^a	Ohio, USA (1953)	Fetus lung	Horse P14/LLC-MK2, P1 ^b
87AR-A1	Arizona, USA (1987)	Semen	RK, P4/V, P7
86AB-A1	Alberta, Canada (1986)	Fetus	RK, P4/V, P6
86NY-A1	New-York, USA (1986)	Semen	ED, P4/V, P2
84KY-A1	Kentucky, USA (1984)	Nasal swab	RK, P5
Vienna	Austria (1986)	Nasal swab	ED, P1/RK, P2
19933	Guelph, Canada (1992)	Semen	RK, P5
15492	Guelph, Canada (1991)	Semen	RK, P5
11958	Guelph, Canada (1990)	Semen	RK, P5
T1329	Guelph, Canada (1988)	Neonatal lung ^c	PK, P5
Arvac ^d	Fort Dodge Laboratories	—	HK, p131/RK, p111/ ED, p24/V, P2

^a ATCC number: VR-796

^b Cells — ED: equine dermis; HK: primary horse kidney; LLC-MK2: Rhesus monkey kidney; RK: rabbit kidney-13; V: Vero; P: passage number

^c EAV was isolated from 5 d old standardbred foal

^d Vaccine strain of EAV

sequences of the ATCC Bucyrus reference strain of EAV (ATCC number VR-796), the Arvac vaccine strain (2), and 9 EAV field isolates originating from different geographic regions (Canada, $n = 5$, United States, $n = 3$; and Austria, $n = 1$). It took years to isolate and to characterize the genetic variation among these isolates/strains. We also wished to compare the predicted RNA secondary structures in the leader region of these isolates/strains and consider how these structures might have a role in EAV biogenesis.

Each virus isolate/strain used in this study (Table I) was plaque-purified and propagated for 1-2 additional passages in rabbit kidney (RK-13) or Vero cells until 80 to 100% exhibited a cytopathic effect (21). After 3 freeze-thaw cycles, the virus-containing supernatants were collected by centrifugation at 5000 $\times g$. Virion RNA was extracted from 300 μ L of these supernatants by the guanidium isothiocyanate method as previously described (21). RNA pellets were resuspended in 20 μ L of diethylpyrocarbonate (DEPC)-treated water, containing 15 units of human placental RNase inhibitor, and kept at -70°C until used.

The rapid amplification of cDNA ends (RACE) method based on the single strand ligation to single-stranded cDNA (SLIC) was used to obtain sequences of the 5' end of each EAV isolate (19). Briefly, purified EAV genomic RNA was reverse transcribed with avian myeloblastosis virus (AMV) reverse transcriptase

(RT) using the procedure described in the Amplifinder Race Kit (Clontech Laboratories, Palo Alto, California, USA) protocol. The RT reaction was primed with 20 pmol of the antisense oligonucleotide primer PEV-L1 (5'-GTGGAGCCGCTCCACTTC-3') which is complementary to nt 306 to 323 of the EAV genome (4, Genbank Database accession number X53459). This primer sequence is located downstream from the 3' end of the leader sequence in the ORF 1a genomic region. After RNA hydrolysis, the single-stranded cDNA (ss-cDNA) was purified and ligated to the 3' end blocked AmpliFINDER anchor sequence at room temperature for 18 h with T4 RNA ligase. The single stranded ligation product was then amplified in the polymerase chain reaction (PCR) procedure by using the antisense primer PEV-L02 (5'-ACCGTCAAGCCACAAGATG-3'), which is complementary to an internal sequence (nt 165 to 145) of the EAV leader sequence (4), and the sense AmpliFINDER anchor primer (AFAP). A 2nd set of primers was used to obtain, by PCR, overlapping cDNA fragments containing the remainder of the leader sequence at the 3' end. These primers were the sense primer PEV-L0 (5'-TGCCATATACGGCTCACC-3'), which is complementary to nt 1 to 18 according to den Boon et al (4), and the antisense primer PEV-L1 described above. The gel-purified PCR cDNA products for each isolate were cloned into Sma I-cleaved pBluescript II KS+ (Stratagene, La Jolla, California, USA) or PCR II

BUCYRUS	GCTCGAAGTG	<u>TGTATCGTGC</u>	CATATACGGC	TCACCACCAT	ATACACTGCA	50
87AR-A1	A.....G.....G.....	
86AB-A1	-G...GG.....G.....	
86NY-A1G.....G.....	
84KY-A1G.....G.....	
VIENNA	A.....C.....G.....G.....	
19933	T.....G.....G.....	
15492	A.....G.....	
11958G.....G.....	
T1329G.....	
ARVACG.....	
BUCYRUS	AGAATTACTA	TTCTTGTGGG	CCCCTCTCGG	TAAATCCTAG	AGGGCTTTCC	100
87AR-A1	
86AB-A1T.....	
86NY-A1T.....	
84KY-A1T.....	
VIENNAT.....T.....	
19933	
15492G.....	
11958	
T1329T.....C.....	
ARVAC	
BUCYRUS	TCTCGTTATT	GCGAGATTCG	TCGTTAGATA	ACGGCAAGTT	-CCCTTCTT	150
87AR-A1T.....	
86AB-A1	
86NY-A1C.....GT.....C.....	
84KY-A1	
VIENNAT.....T.....C.....TA.....	
19933A.....C.....	
15492T.....	
11958A.....	
T1329A.....	
ARVAC	
BUCYRUS	ACTATCCTAT	TTTCATCTTG	TGGCTTGACG	GGTCACTGCC	ATCGTCTGCG	200
87AR-A1	
86AB-A1	
86NY-A1T.....	
84KY-A1T.....A.....	
VIENNATCT.....A.....	
19933T.....	
15492T.....	
11958	
T1329T.....	
ARVACC.....	
BUCYRUS	ATCTCTA					
87AR-A1					
86AB-A1					
86NY-A1					
84KY-A1					
VIENNA					
19933					
15492					
11958					
T1329					
ARVAC					

Figure 1. Alignment of the nucleotide sequences of the leader region of 9 arteritis virus isolates and the Arvac vaccine strain with the ATCC Bucyrus reference strain. AUG start codon located at positions 14 to 16 is double-underlined. In-frame stop codons (TAG or TGA) are single-underlined.

TABLE II. Nucleic acid sequence identities (%) of the leader sequence of equine arteritis virus (EAV) isolates/strains*

Isolate/Strain	Bucyrus	87AR-A1	86AB-A1	86NY-A1	84KY-A1	Vienna	19933	15492	11958	T1329	Arvac
Bucyrus ^a	100.0	98.1	98.1	98.1	97.1	94.2	97.1	97.6	98.5	97.6	98.5
87AR-A1		100.0	97.6	96.1	95.2	93.2	96.1	97.6	96.6	96.6	98.5
86AB-A1			100.0	97.1	96.1	92.7	96.6	96.1	97.6	95.6	97.6
86NY-A1				100.0	97.1	94.2	97.1	96.6	97.6	97.6	96.6
84KY-A1					100.0	92.2	96.6	95.6	96.6	95.6	95.6
Vienna						100.0	92.2	94.6	92.7	94.6	92.7
19933							100.0	96.6	96.6	96.6	96.6
15492								100.0	96.1	97.1	97.1
11958									100.0	96.1	97.1
T1329										100.0	97.1
Arvac ^c											100.0

* The leader nucleotide sequence data reported for these isolates/strains have been deposited with the Genbank Database under accession numbers AF001259, AF001260, U46945 to U46947 and U65723 to U65728

^a ATCC number: VR-796

^c Vaccine strain of EAV

(Invitrogen, San Diego, California, USA) plasmid vectors. Two or more clones of each PCR product (both strands) were sequenced using the dideoxynucleotide chain termination method as described (16).

As shown in Fig. 1, the RACE method successfully identified the extreme 5' end of the genome of all of the EAV isolates/strains studied and revealed the presence of 17 additional nt for all EAV isolates but one, located upstream of the sequence published by den Boon et al (4). In the case of the 86AB-A1 Canadian isolate, the 1st of these 17 nt was missing. The 17 base sequence in the Arvac vaccine strain and the T1329 Canadian isolate was identical to that of the ATCC Bucyrus reference strain. The identification of these additional nt brings the length of the leader sequence to 206 nt without the junction site motif and the putative 5' cap structure-associated nt for most EAV isolates/strains, except the 86AB-A1 Canadian isolate which is 205 nt. Another exception is the 86NY-A1 isolate, which has a leader region of 207 nt in length due to a C insertion at position 141.

When the nucleotide sequences of the leader region of all EAV isolates/strains were aligned (Fig. 1) and compared to that of the ATCC Bucyrus reference strain by using the Wisconsin Genetics Computer Groups software package (GCG, version 8.1), a total of 29 base positions varied among isolates (14% of the 206 nt positions). Besides a few substitutions that occurred at the extreme 5' end of the leader region (Fig. 1), base substitutions were common at particular positions (for instance, substitutions of A→G at positions 35 and 43, C→T

at position 78, and A→T at position 154). Such base mutations are common in RNA viruses (22).

The variability of the leader sequence of the nine EAV field isolates and the Arvac vaccine strain to the ATCC Bucyrus reference strain is shown in Table II. The percent of identical nucleotide residues ranged from 94.2 (Vienna isolate) to 98.5% (11958 Canadian isolate and the Arvac vaccine strain). When the leader sequences of the EAV field isolates were compared to each other, the 86AB-A1 and 11958 Canadian isolates, the 84KY-A1 and T1329 Canadian isolates, and the 84KY-A1 and 11958 Canadian isolates were the most closely related at 97.6%, while the Vienna and both the 86NY-A1 and 19933 Canadian isolates were the least related at 92.2% identity. Sequences of the North American isolates (5 Canadian and 4 American, including the Bucyrus reference and Arvac vaccine strains), were 95.2 to 98.5% identical when compared to each other, while all were less closely related to the Vienna European isolate where the identity of the sequence ranged from 92.2 (84KY-A1 and 19933 Canadian isolates) to 94.6% (T1329 and 15492 Canadian isolates). When the Arvac vaccine strain was compared to all other EAV isolates, including the Bucyrus reference strain, the lowest level of nucleic acid identity was shown with the Vienna isolate at 92.7%, whereas the highest was observed with the Bucyrus reference strain and the 87AR-A1 isolate at 98.5%. When the EAV isolates/strains were subjected to phylogenetic analysis by using the DNADIST and FITCH programs of the Phylogenic Inference Package (PHYLIP, version 3.5c) (23) on the basis of the nucleic acid sequences of the leader region, all EAV isolates/strains were found to likely represent a large phylogenetically-related group (data not shown).

All EAV isolates/strains analyzed in this study showed the presence of a AUG start codon located at positions 14 to 16 (Fig. 1). Because the primary sequence in the immediate vicinity of an AUG initiation codon is important for initiation of translation (24), this AUG codon is likely to act as a functional start codon because, with the -3 and +4 positions occupied by U and G residue, respectively, it is capa-

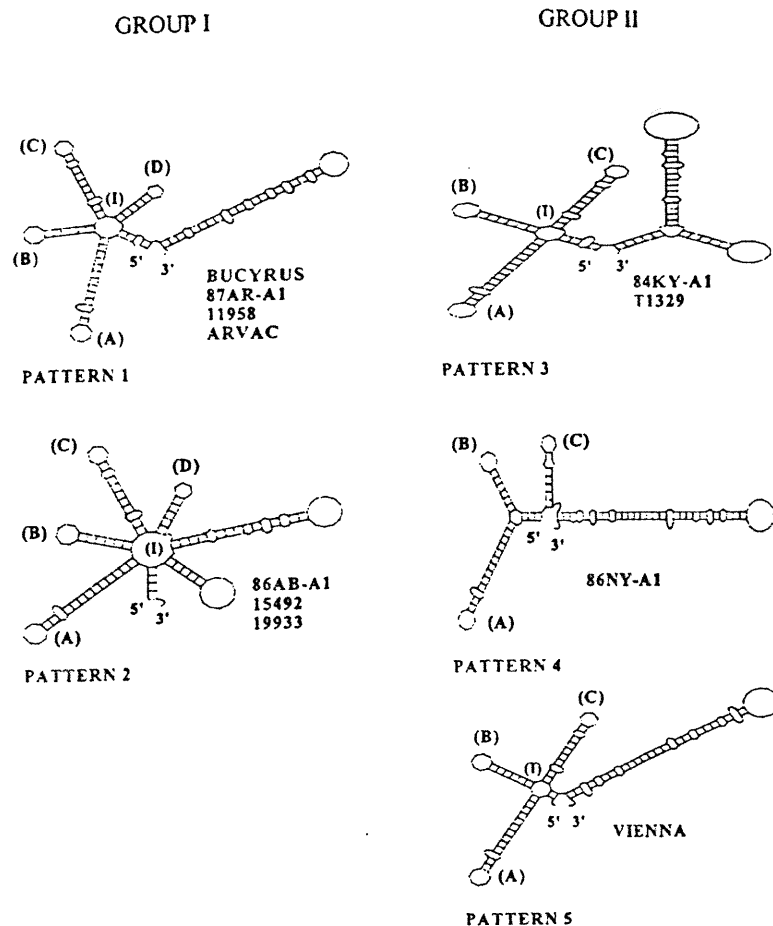


Figure 2. Predicted secondary structure of the leader sequence of eleven EAV isolates/strains using the FOLD program of the GCG software package. The conserved stem-loops (A to D) and the internal loop (I) are designated with upper cases in parenthesis. Five different secondary structure patterns (1 to 5) were obtained. Grouping of viruses (groups I and II) was based on the presence or absence of stem-loop D.

ble, albeit at a suboptimal level, of initiating translation. The presence of an amber UAG stop codon (Bucyrus, Arvac, 87AR-A1, 86AB-A1, 84KY-A1, Vienna, T1329, 11958, and 15492 isolates) or an ochre UAA stop codon (19933 isolate) at positions 125 to 127 in frame with this AUG start codon predicts an open reading frame of 111 nt that could encode a small polypeptide of 37 amino acids. In contrast, the 86NY-A1 isolate contains an opal UGA stop codon at positions 176 to 178, so that its intraleader ORF is 162 nt in length and the predicted polypeptide would include 54 amino acids. When the sequences of the predicted 37 amino acid peptides for all EAV isolates/strains were aligned and compared to that of the Bucyrus refer-

ence strain, only a few amino acid substitutions were observed and, in most cases, were common among the EAV field isolates (for instance, Arg→His at position 8, Met→Ile at position 10, and Ser→Leu at position 22) (data not shown). Finally, no significant homology was found between the 37 amino acid sequences or 54 amino acid sequences (for the 86NY-A1 isolate) and known peptide sequences located in the Genbank Database.

The role of secondary structures in leader sequences in the initiation of translation of eukaryotic mRNAs has been well documented (24). Moreover, in picornaviruses (e.g. poliovirus), mutational analyses demonstrated that neurovirulence of certain

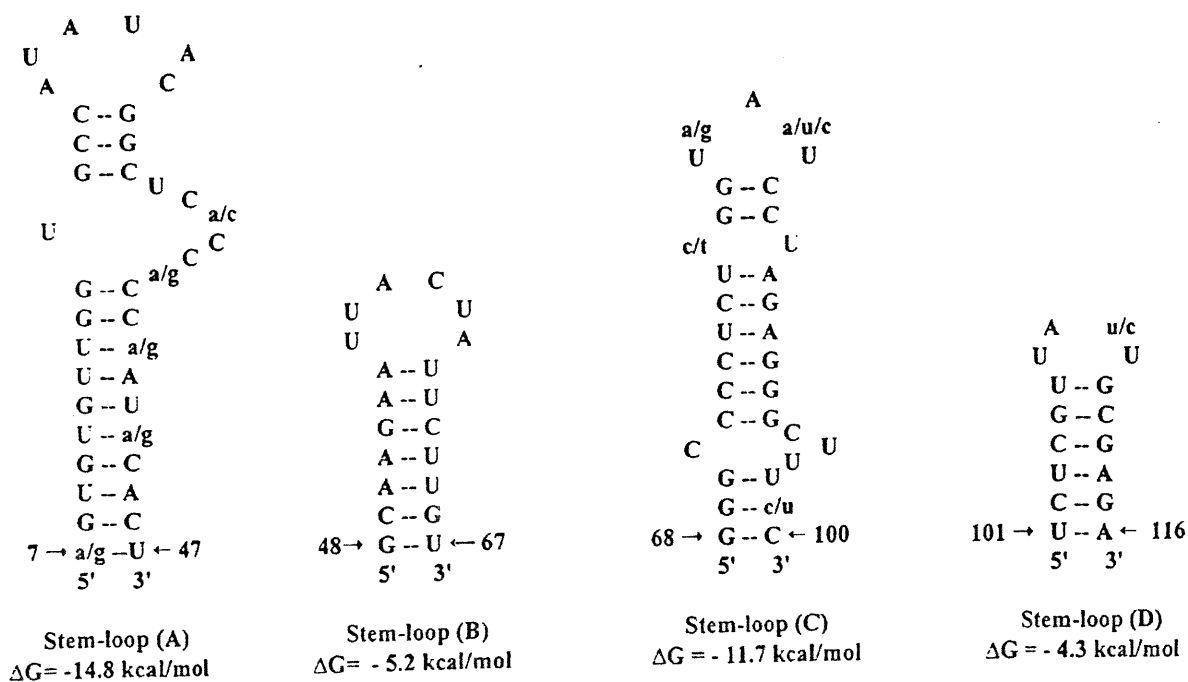


Figure 3. Nucleotide sequences corresponding to the common stem-loops in the leader sequence of eleven EAV isolates/strains. Stem-loops A, B, and C were observed in all predicted patterns. Stem-loop D was observed only in patterns 1 and 2. Lower cases correspond to nucleotide substitutions shown in Fig. 1. The calculated delta G minimum (free energy) values are indicated.

strains was associated with a specific secondary structure in the leader sequence (25). These observations prompted us to determine if putative secondary structures could be present in the EAV leader region. Secondary structures of the leader sequence of all EAV isolates/strains exhibiting minimum-free energies were predicted with the FOLD and SQUIGGLES graphics output programs of the GCG software package. Based on the number and locations of predicted stem-loop structures, 5 different secondary structure patterns were obtained (Fig. 2). Patterns 1 and 2 were represented by 4 (ATCC Bucyrus reference and Arvac vaccine strains, and the 87AR-A1 and 11958 Canadian isolates) and 3 (86AB-A1, 15492 and 19933 Canadian isolates) virus isolates, respectively. Pattern 3 was shared by two EAV isolates (84KY-A1 and T1329 Canadian) while patterns 4 and 5 were represented by 1 isolate each (86NY-A1 and Vienna). Three of the predicted stem-loop structures, which would form in the first half of the leader region at nt 7 to nt 100, designated A, B and C, were

found in all of the analyzed EAV isolates/strains (Fig. 2). Moreover, a 4th stem-loop, designated D, was predicted for the 7 isolates/strains comprising patterns 1 and 2. This suggests that the EAV isolates/strains might segregate into 2 major groups (I and II) based on the presence or absence of stem-loop D.

As shown in Fig. 3, stem-loop A (41 nt in length) forms by base-pairing nucleotides from nt 7 to nt 47, stem-loop B (20 nt in length) uses nt 48 to nt 67, stem-loop C (33 nt in length) uses nt 68 to nt 100 (except for the 86NY-A1 isolate where a shorter stem-loop C is formed from nt 72 to nt 94), and stem-loop D (16 nt in length) uses nt 101 to nt 116. For most isolates, these sets of 3 or 4 stem-loop structures were branched out from an internal loop designated I. Although some nucleic acid substitutions were observed within the sequence of all 4 of these stem-loops, the general form and length of these stem-loops was unaltered regardless of the mutations (Fig. 2 and 3) indicating a high degree of structural conservation for the stem-loop structures

of both groups I and II. Such conservation of the general form of these stem-loop structures indicates that most mutational changes are not tolerated and suggests an important functional role of these structural domains in EAV replication or biogenesis.

The nucleotide sequences and predicted secondary structures downstream from either stem-loop C (from nt 100 for group II) or stem-loop D (from nt 117 for group II) are much more variable than the 5' end. This variability is seen in the sequence following the stop codon at nt 125-127 through nt 154, and is reflected by a variable pattern of additional possible stem-loop structures that would vary from one isolate to another in terms of position and sequence. Variability in this portion of the leader sequence, in marked contrast to the conserved sequence and limited variability in the secondary structures predicted for the 5' end, suggests a lesser role for this region of the genome in regulation of EAV replication. Nevertheless, such variation might also be associated with differences in EAV virulence, as seen elsewhere where variations in

the secondary structure of the leader sequence of echovirus and Sabin poliovirus reflected variation in neurovirulence and the level of virus attenuation in tissue cultures, respectively (26,27).

In summary, the results presented in this study demonstrate a high degree of nucleic acid conservation in the EAV leader region among the analyzed EAV isolates/strains. The results also demonstrate the presence of an intraleader ORF for all analyzed EAV isolates/strains. Preliminary results by using appropriate cDNA clone mutants containing the entire leader region upstream from the EAV ORF 7 nucleic acid sequence indicated a modulating effect of the intraleader ORF nucleic acid sequence on the level of ORF 7 translation product (Kheyar et al. in preparation). This work may be of some interest because, as shown for coronaviruses, an intraleader ORF was reported to modulate persistent infection either by attenuating (in bovine coronavirus, 28) or by enhancing (in murine coronavirus, 29) the translation rate of downstream ORFs. Finally, computer-predicted RNA secondary structure analysis identified 3 conserved stem-loops in the leader region of all EAV analyzed isolates/strains. The role of these secondary structures in EAV biogenesis has yet to be determined.

ACKNOWLEDGMENTS

This work was supported by an operating grant from the National Sciences and Engineering Research Council of Canada to D. Archambault. A. Kheyar has been supported by a graduate student fellowship from Université de Montréal. D. Archambault is the holder of a research scholarship from the Fonds de la Recherche en Santé du Québec (FRSQ). We thank Peter J. Timoney and William H. McCollum (Gluck Equine Research Center, Lexington, Kentucky, USA), and Susy Carman (Veterinary Laboratory Services, Guelph, Ontario) for providing the EAV laboratory strains and field isolates. We are also grateful to John Dahlberg for reviewing the manuscript, and to Carole Villeneuve for secretarial work.

REFERENCES

- DE VRIES AAF, HORZINEK MC, ROTTIER PJM, DE GROOT RJ. The genome organization of the Nidovirales: Similarities and differences between Arteri-, Toro-, and Coronaviruses. *Sem Virol* 1997; 8: 33-37.
- TIMONEY PJ, MCCOLLUM WH. Equine viral arteritis. *Vet Clinics North Amer Equine Pract* 1993; 9: 295-309.
- DE VRIES AAF, CHIRNSIDE ED, BREDEBEEK PJ, GRAVESTEIN LA, HORZINEK MC, SPAAN WJM. All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res* 1990; 18: 3241-3247.
- DEN BOON JA, SNIJDER EJ, CHIRNSIDE ED, DE VRIES AAF, HORZINEK MC, SPAAN WJM. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J Virol* 1991; 65: 2910-2930.
- DE VRIES AAF, CHIRNSIDE ED, HORZINEK MC, ROTTIER PJM. Structural proteins of equine arteritis virus. *J Virol* 1992; 66: 6294-6303.
- BALASURIYA UBR, MACLACHLAN J, ROSSITO PV, DE VRIES AAF, ROTTIER PJM. Identification of a neutralization site in the major envelope glycoprotein (G_L) of equine arteritis virus. *Virology* 1995; 207: 518-527.
- CHIRNSIDE ED, DE VRIES AAF, MUMFORD JA, ROTTIER PJM. Equine arteritis virus-neutralizing antibody in the horse is induced by a determinant on the large glycoprotein G_L . *J Gen Virol* 1995; 76: 1989-1998.
- DEREGT D, DE VRIES AAF, RAASMAN MJB, ELMGREN LD. Monoclonal antibodies to equine arteritis virus protein identify the G_L protein as a target for virus neutralization. *J Gen Virol* 1994; 75: 2439-2444.
- GLASER AM, DE VRIES AAF, DUBOVI EJ. Comparison of equine arteritis virus isolates using neutralizing monoclonal antibodies and identification of sequence changes in G_L associated with neutralization resistance. *J Gen Virol* 1995; 76: 2223-2233.
- FUKUNAGA Y, MATSUMURA T, SUGIURA T, WADA R, IMAGAWA H, KANEMARU T, KAMADA M. Use of the serum neutralization test for equine viral arteritis with different virus strains. *Vet Rec* 1994; 134: 574-576.
- MURPHY TW, MCCOLLUM WH, TIMONEY PJ, KLINGEBORN BW, HYLSETH B, GOLNIK W, ERASMUS B. Genomic variability among globally distributed isolates of equine arteritis virus. *Vet Microbiol* 1992; 32: 101-115.
- KHEYAR A, MARTINS S, ST-LAURENT G, TIMONEY PJ, MCCOLLUM WH, ARCHAMBAULT D. Expression cloning and humoral immune response to the nucleocapsid (N) and membrane (M) proteins of equine arteritis virus. *Clin Diagn Lab Immunol* 1997; 4: 648-652.
- CHIRNSIDE ED, WEARING CM, BINNS MM, MUMFORD JA. Comparison of M and N gene sequences distinguishes variation amongst equine arteritis virus isolates. *J Gen Virol* 1994; 75: 1491-1497.
- BALASURIYA UBR, TIMONEY PJ, MCCOLLUM WH, MACLACHLAN NJ. Phylogenetic analysis of open reading frame 5 of field isolates of equine arteritis virus and identification of conserved and nonconserved regions in the G_L envelope glycoprotein. *Virology* 1995; 214: 690-697.
- HEDGES JF, BALASURIYA UBR, TIMONEY PJ, MCCOLLUM WH, MACLACHLAN NJ. Genetic variation in open reading frame 2 of field isolates and laboratory strains of equine arteritis virus. *Vir Res* 1996; 42: 41-52.
- LEPAGE N, ST-LAURENT G, CARMAN S, ARCHAMBAULT D. Comparison of nucleic and amino acid sequences and phylogenetic analysis of the G_S protein of various equine arteritis virus isolates. *Virus Genes* 1996; 13: 87-91.
- ST-LAURENT G, LEPAGE N, CARMAN S, ARCHAMBAULT D. Genetic and amino acid analysis of the G_L protein of Canadian, American and European equine arteritis virus isolates. *Can J Vet Res* 1997; 61: 72-76.
- ARCHAMBAULT D, LAGANIERE G, CARMAN S, ST-LAURENT G. Comparison of nucleic and amino acid sequences and phylogenetic analysis of open reading frames 3 and 4 of various equine arteritis virus isolates. *Can J Vet Res* 1997; in press.
- KHEYAR A, ST-LAURENT G, ARCHAMBAULT D. Sequence determination of the extreme 5' end of the equine arteritis virus leader region. *Virus Genes* 1996; 12: 291-295.
- VAN DINTEN LC, DEN BOON JA, WASSENAAR ALM, SPAAN WJM, SNIJDER EJ. An infectious arterivirus cDNA clone: Identification of a replicase point mutation that abolishes discontinuous mRNA transcription. *Proc Natl Acad Sci USA* 1997; 94: 991-996.
- ST-LAURENT G, MORIN G, ARCHAMBAULT D. Detection of equine arteritis virus following amplification of structural and nonstructural viral genes by reverse transcription-PCR. *J Clin Microbiol* 1994; 32: 658-665.
- DOMINGO E, MARTINEZ-SALAS E, SOBRINO F, DE LA TORRE JC, PORTELA A, ORTIN J, LOPEZ-GALINDEZ C, PEREZ-BREÑA P, VILLANUEVA N, NAJERA R. The quasispecies (extremely heterogeneous) nature of viral RNA genome populations: biological relevance, a review. *Gene* 1985; 40: 1-8.
- FELSENSTEIN J. PHYLIP (Phylogenetic Inference Package) 3.5c Manual, University of Washington, Department of Genetics SK-50, Seattle, Washington, USA 98195, 1993.
- KOZAK M. An analysis of vertebrate mRNA sequences: intimations of translational control. *J Cell Biol* 1991; 115: 887-903.
- HALLER AA, STEWART SR, SEMLER BL. Attenuation stem-loop lesions in the 5' noncoding region of poliovirus RNA: Neuronal cell-specific translation defects. *J Virol* 1996; 70: 1467-1474.

26. **MACADAM AJ, FERGUSON G, BURLISON J, STONE D, SKUCE R, ALMOND JW, MINOR PD.** Correlation of RNA secondary structure and attenuation of Sabin vaccine strains of poliovirus in tissue culture. *Virology* 1992; 189: 415-422.
27. **ROMERO JR, ROTBART HA.** Sequence analysis of the downstream 5' nontranslated region of seven echovirus with different neurovirulence phenotypes. *J Virol* 1995; 69: 1370-1375.
28. **HOFMANN MA, SENANAYAKE S, BRIAN DA.** A translation attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *Proc Natl Acad Sci USA* 1993; 90:11733-11737.
29. **CHEN W, BARIC RS.** Function of a 5'-end genomic RNA mutation that evolves during persistent mouse hepatitis virus infection in vitro. *J Virol* 1995; 69: 7529-7540.

CHAPITRE V

Article 3

**The intraleader AUG nucleotide sequence context per se
is important for equine arteritis virus replication**

Ali Kheyar, Antoine A. F. de Vries, Peter J. M. Rottier, and Denis Archambault

Article soumis pour la publication

The intraleader AUG nucleotide sequence context per se is important for equine arteritis virus replication

Ali Kheyar¹, Antoine A. F. de Vries^{2#}, Peter J. M. Rottier², and Denis Archambault^{1*}.

¹University of Québec in Montréal, Department of Biological Sciences, P.O. Box 8888, Succursale Centre-Ville, Montréal, Québec, Canada, H3C 3P8; ²Virology Division, Department of Infectious Diseases and Immunology, Veterinary Faculty, Utrecht University, Yalelaan 1, 3584 CL Utrecht, the Netherlands

Running title: EAV intraleader AUG and replication

* Corresponding author and reprint request

Phone: (514) 987-3000, ext. 4622

Fax: (514) 987-4647

E-mail: archambault.denis@uqam.ca

Present address: Gene Therapy Section, Department of Molecular Cell Biology, Leiden University Medical Center, Wassenaarseweg 72, 2333 AL Leiden, the Netherlands

ABSTRACT:

The 5'-terminal leader sequence of the equine arteritis virus (EAV) genome contains an open reading frame (ORF) with an AUG codon in a suboptimal context for initiation of protein synthesis. To investigate the significance of this intraleader ORF (ILO), a plasmid was generated carrying, downstream of a bacteriophage T7 polymerase promoter, a precise DNA copy of the subgenomic mRNA₇. Capped RNA transcribed from this construct was shown to direct in an *in vitro* translation system the synthesis of leader peptide as well as N protein. Plasmids in which the context of the ILO start codon had been mutated to generate either an optimized or a poor sequence for translation initiation yielded leader peptide. However, no peptide was detected when the initiation codon was incapacitated. Translation of the downstream N gene was inversely affected by leader peptide production. N protein synthesis was lowest with the ILO start codon in optimized contexts but doubled when leader peptide expression was blocked. The data are consistent with a leaky scanning mechanism for EAV protein translation. To investigate the role of the leader peptide in EAV replication, we generated two mutant viruses using an infectious EAV cDNA clone. In one of these, the ILO start codon was in an optimal Kozak context for translation initiation while in the other the codon was again incapacitated. Surprisingly, both mutant viruses appeared to be equally viable and to exhibit similar phenotypes in BHK-21 cells. Both the kinetics of replication and the viral yields were reduced in early cell passage relative to that of the wild-type parental virus, as were the plaque sizes of the mutants. Importantly, the mutations introduced into the viruses appear to be rapidly and precisely repaired. Already after one cell passage following cell transfection, a significant fraction of the viruses had regained the wild-type sequence as well as its phenotype. The results suggest that EAV replication is not dependent on the synthesis of the intraleader peptide. They moreover indicate that the similar phenotype of the two mutant viruses does not correlate with the dissimilar levels of ILO expression measured *in vitro*. Consequently, the leader peptide per se does not seem to have any function in the EAV life cycle. Thus, as we elaborate in the Discussion, the precision with which the wild-type sequence around the intraleader AUG is maintained or restored together with other considerations leads us to propose that it is the nucleotide sequence, not its functioning in translation initiation, that is involved in the viral replication cycle

INTRODUCTION

Equine arteritis virus (EAV), the aetiologic agent of equine viral arteritis (EVA), is the prototype member of the *Arteriviridae* family which also includes the lactate dehydrogenase-elevating virus (LDV) of mice, the porcine reproductive and respiratory syndrome virus (PRRSV), and the simian hemorrhagic fever virus (SHFV). Together with the *Coronaviridae*, the arteriviruses form the recently established order *Nidovirales* (4). The EAV genome is a single-stranded, 5'-capped and 3'-polyadenylated mRNA molecule with a length of 12,704 nucleotides (nt). During EAV replication, a 3'-coterminal nested set of seven virus-specific RNAs (RNAs 1 through 7) is produced. Each open reading frame (ORF) in the EAV genome is preceded by the pentanucleotide sequence motif 5'-UCAAC-3', also called the leader-body junction site (6, 9) or the transcription regulating sequence (TRS) (8, 44), which is involved in the formation of six subgenomic RNAs with a common leader sequence of 206 nt derived from the extreme 5' terminus of the viral genome.

The nidovirus leader sequence is believed to have a pivotal role in the transcription process. In the discontinuous minus strand transcription model, the body TRS in the nascent minus strand transcript base pairs with the leader TRS of the plus strand genome, thereby allowing the addition of the antileader sequence (32). Alternatively, in the leader-primed-discontinuous transcription model, the leader sequence acts as a primer during plus strand synthesis by base pairing of its 3'-TRS with a complementary TRS on the minus genome strand (26). The complementary body TRS motifs of the minus genomic RNA strand are then fused to the leader TRS for the generation of the viral sg mRNAs of various lengths (8, 43).

We have previously determined the extreme 5' end of the leader sequence of EAV and reported the presence of an AUG codon at genome positions 14 to 16 that could open an intraleader ORF (ILO) of 111 nt in length (19). This ILO was predicted to encode a peptide of 37 amino acids and appeared to be conserved among different EAV isolates (21) which suggested that it may be of functional significance for the virus. Further analysis showed that the ILO start codon was present in a suboptimal context to initiate protein translation, which might explain the efficient translation of the ORFs located downstream of the leader sequence on each EAV mRNA. Because short ORFs within the 5' leader region of other viral mRNAs have been reported to down-regulate the rate of translation of downstream ORFs (5, 16, 24), we argued that the EAV ILO may similarly affect the synthesis of the EAV proteins or have another regulatory function in the life cycle of the virus.

In this study, we performed *in vitro* transcription and translation experiments to study the expression of the ILO and its effect on the synthesis of EAV N protein. We also studied by using two mutant infectious cDNA clones the effect of mutations around the ILO AUG codon on the EAV replication.

MATERIALS AND METHODS

Cells, virus and antisera

African green monkey kidney (Vero) and baby hamster kidney (BHK-21) cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS) and antibiotics (100 IU/ml of penicillin, 100 µg/ml of streptomycin, and 60 µg/ml of anti-pleuropneumonia-like organisms (PPLO) agent. The EAV reference Bucyrus strain (ATCC number VR-796) was propagated in cells at a multiplicity of infection (MOI) of 1 (41). Rabbit antisera directed against bacterial fusion products of the maltose binding protein and the EAV M or N protein have been described (20). Antisera against whole EAV and against the leader protein were raised in laboratory rabbits following standard immunization procedures (13) (three antigen inoculations three weeks apart using the Freund's adjuvants with the final bleed 14 days after the last inoculation) using as immunizing antigens a pelleted stock of the Canadian EAV isolate T1329 [10^9 median tissue culture infective dose (TCID₅₀) per inoculation] propagated in cell culture (27) and a synthetic peptide (500 µg per inoculation) corresponding to the entire predicted amino acid sequence of the leader protein (19), respectively. The titer of the antiserum raised against the leader peptide 50 000 as determined in a Western blot assay as antigen substrate was the synthetic peptide.

RNA extraction and oligonucleotide primers

Total RNA was extracted from 5×10^6 EAV-infected cells with a commercial kit (TRIzol reagent; Gibco/BRL, Gaithersburg, Md.) as previously described (2). The final RNA pellet was dissolved in 30 µl of diethylpyrocarbonate (DEPC)-treated water and stored at -80°C until used.

The oligonucleotide primers used in this study are listed in Table 1. They were designed on the basis of the genomic sequence of the Bucyrus reference strain of EAV as determined by den Boon et al. (9) (GenBank accession number X53459) and Kheyar et al., (21) (GenBank accession number AF001259).

Construction of EAV mRNA7 expression plasmids

Recombinant DNA procedures were adopted from Sambrook et al. (36) or performed according to the instructions supplied with specific reagents. Three μ l of total RNA isolated from the EAV-infected cells at 10 h post infection (pi) were used as template in a reverse transcription (RT) reaction catalyzed by the avian myeloblastosis virus (AMV) reverse transcriptase (Stratagene, La Jolla, Calif.). As primer for cDNA synthesis, an oligonucleotide (Peav3') complementary to the extreme 3' end of the EAV genomic and sg viral RNAs was used. The cDNA was then amplified by the polymerase chain reaction (RT-PCR) using *Pfu* DNA polymerase (Stratagene) and a primer pair consisting of the antisense oligonucleotide Peav3' and the sense oligonucleotide P(wt)atg which corresponds to the extreme 5' end of the EAV leader sequence (Table 1). By doing so, we generated a DNA copy of EAV mRNA7 downstream of the bacteriophage T7 RNA polymerase promoter which was introduced by the sense primer. After agarose gel purification, the RNA 7-specific DNA fragment was cut with *Bam*HI and *Sal*I and ligated into *Bam*HI- and *Sal*I-digested plasmid vector pUC18 (InVitrogen, San Diego, Calif.). The resulting plasmid, designated F7wtATG, was subjected to nucleotide sequence analysis (19) by using a T7 DNA polymerase sequencing kit (Amersham Pharmacia Biotech, Piscataway, N.J.) to verify the correctness of the cloned fragment. As shown in Fig. 1, this construct would theoretically give rise to a RNA transcript of 674 nt in length containing the entire EAV mRNA7 including a 15-A tail.

The construct F7wtATG was then used as a template to introduce point mutations in order to modify the sequence context of the ILO initiation codon. PCRs were carried out under the same conditions as described above except that oligonucleotide P(wt)atg was replaced with primer P(-)atg, P(nf)atg, P(opt)atg, or P(koz)atg (Table 1). As a consequence, the generated PCR products lacked an ILO start codon, or contained an IL initiation codon in a highly unfavorable, improved, or optimal context for protein translation initiation (24, 25). The mutated PCR products were then inserted into pUC18 to generate the F7(-)ATG, F7(nf)ATG, F7(opt)ATG, and F7(koz)ATG constructs (Fig. 1). All these constructs were subjected to sequence analysis to confirm the introduced mutations.

TABLE 1. Oligonucleotide primers.

Name	Sequence	Polarity	Position ^a
P(wt)atg	5' AAGCTTGGATCCTAATAACGACTCACTATA GCTCGAAGTGTGTATGGTGC 3'	(+)	-29 to 20
P(-)atg	5' AAGCTTGGATCCTAATAACGACTCACTATA GCTCGAAGTGTGTATTGTGC 3'	(+)	-29 to 20
P(opt)atg	5' AAGCTTGGATCCTAATAACGACTCACTATA GCTCGAAGT GAGT ATGGTGC 3'	(+)	-29 to 20
P(mf)atg	5' AAGCTTGGATCCTAATAACGACTCACTATA GCTCGAAGTGTGTATGTGC 3'	(+)	-29 to 20
P(koz)atg	5' AAGCTTGGATCCTAATAACGACTCACTATA GCTCGAAGCC ACC ATGGTGC 3'	(+)	-29 to 20
Peav3'	5' ATCGAATTTCGTCGACTTTTTTTTTTTT TTTGGTTCCTGGGTGGC 3'	(-)	12704 to 12691 ^b
PT7(wt)atg	5' caagcggcgcgggtacctaat TAACGACTCACTATA GCTCGAAGTGTGTATGGTGC 3'	(+)	-34 to 20
PT7(ko)atg	5' caagcggcgcgggtacctaat TAACGACTCACTATA GCTCGAAGTGTGTATTGTGC 3'	(+)	-34 to 20
PT7(opt)atg	5' caagcggcgcgggtacctaat TAACGACTCACTATA GCTCGAAGTGA GTATGGTGC 3'	(+)	-34 to 20
P962	5' AGGACCAGTCTGC TT CAGAG 3'	(-)	981 to 962
Peav5'	5' AAGCTTGGATCCGCTCGAAGT G 3'	(+)	-12 to 10
P182	5' ACC CGT CAA GCC ACA AGA TG 3'	(-)	182-163

^a According to Genbank accession number X53459 adjusted after the addition of 17 nt at the 5' end of the EAV genome (Kheyar, et al. 1996).

^b Exclusive of the oligo (T) trail. These positions do not include the poly T sequence.

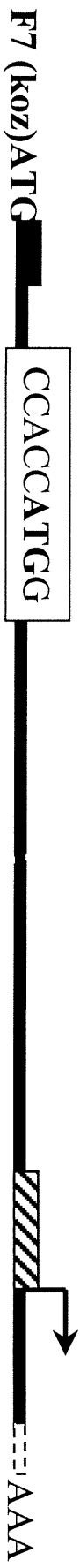
Italicised capitals refer to recognition sites for the restriction enzymes *Hind*III, *Bam*HI, *Eco*RI, *Sal*I, *Nor*I and *Kpn*I; Bolded capitals represent T7 RNA polymerase promoter sequences; normal capitals refer to EAV genomic sequences; underlined capitals correspond to the start codon of the ILO ORF with surrounding nucleotides; lower case letters refer to pBR322 specific nucleotides.

Figure 1 : Schematic representation of the EAV mRNA7-specific plasmid and its derivatives. The parental cDNA clone was generated by RT-PCR and inserted into pUC18 as described in Materials and Methods. Indicated are, going from 5' to 3': the T7 RNA polymerase promoter sequence (T7, black box), the leader sequence (LS, large black line) containing the ILO start codon with the desired point mutations as indicated (capitals in open box), the intergenic sequence (IS) between the leader sequence, the nucleocapsid start codon (ATG[N]), and the poly(A) tail.

Parental Clone *T7* *LS* *IS ATG (N)*



Mutant clones



In vitro RNA transcription, translation experiments and radioimmunoprecipitation

Plasmid DNA of the aforementioned constructs was prepared by the alkaline lysis method, linearized with *SalI*, and purified. In vitro transcription experiments were performed with 2 µg template DNA in a 50 µl reaction volume for 2 h at 37°C by using a T7 transcription kit (Promega, Madison, Wisc.). Cap analogue of type 0 (7mG(5')ppp(5')G) at a concentration of 5mM was incorporated during the transcription reaction as recommended in the manufacturer's instructions. The yield and integrity of the RNAs was analyzed by denaturing agarose gel electrophoresis (36), and the concentration of each RNA was estimated by densitometric scanning.

In vitro translation was performed with 2 µg of each template RNA in a 50-µl translation mixture by using a reticulocyte lysate system (Promega) After 90 min of incubation at 30°C, 5 µl of each sample was loaded onto a sodium dodecyl sulfate (SDS)-15% polyacrylamide (PAA) gel. After electrophoresis, the gels were fixed, impregnated with a scintillant (En³hance™, Dupont, Mississauga, Ont.) for 30 min, washed, dried, and exposed for 6 h using an X-OMAT AR-5 film (Eastman Kodak Company, Rochester, N.Y.). Fluorograms were subsequently analyzed by densitometry to measure the effects of each mutation on N synthesis.

The specificity of the in vitro translated products was analysed by radioimmunoprecipitation assay (RIPA) (1) by using the appropriate rabbit sera. The RIPA samples were analyzed in SDS-15% PAA gels. The gels were then processed as described above and subjected to densitometrical analysis for protein quantitation.

Generation of full-length EAV cDNA clones with mutations in the EAV leader sequence

The construction of the full-length EAV cDNA pBRNX1.038 has been described previously (12). In this pBR322-based construct, the genomic sequence of the Utrecht variant of the Bucyrus strain of EAV is preceded by a bacteriophage SP6 RNA polymerase promoter. To generate three new EAV cDNA clones, a series of PCRs was performed as described above by using plasmid pBRNX1.038 as a template, the ORF 1a-specific oligonucleotide P962 as an antisense primer, and the leader-specific oligonucleotides PT7(wt)atg, PT7(ko)atg, or PT7(opt)atg as alternative sense primers (Table 1). The resulting PCR products were digested with *NotI* and *ApoI* and stored as frag-

ments of 897 nt in length until used. Another aliquot of plasmid pBRNX1.038 was digested with the unique restriction enzymes *NotI* and *EcoRV* which generates fragments of 12,672 and 4,282 nt, respectively. The small fragment, which contained the SP6 promoter sequence and the wild-type ILO start codon, was purified from the gel and digested with *ApoI* to obtain a 891-nt *NotI-ApoI* fragment and a 3391-nt *ApoI-EcoRV* fragment. The latter DNA fragment was then gel-purified and stored until further use. It is noteworthy that there was a 6-nt difference in length between the *NotI-ApoI* fragments generated by PCR and the *NotI-ApoI* fragment obtained by digestion of plasmid pBRNX1.038. This difference is due to the insertion of a *KpnI* site between the *NotI* site and the T7 promoter sequence in primers PT7(wt)atg, PT7(ko)atg and PT7(opt)atg (Table 1). The additional *KpnI* site allowed us to distinguish, by restriction enzyme analysis, the original SP6 promoter-containing infectious cDNA clone pBRNX1.038 from the newly generated T7 promoter-containing full-length EAV cDNA clones.

Each of the *NotI*- and *ApoI*-digested PCR products was dephosphorylated by using the calf intestinal alkaline phosphatase (Pharmacia Biotech) and ligated to the 3391-nt *ApoI-EcoRV* fragment of plasmid pBRNX1.038 to generate *NotI-EcoRV* fragments of 4,288 bp in length. These fragments were gel-purified and ligated to the 12,672-nt *EcoRV-NotI* fragment of pBRNX1.038 to generate the T7NX(wt)ATG wild-type full-length cDNA clone and the two mutant full-length cDNA clones T7NX(ko)ATG and T7NX(opt)ATG. Each of these constructs was used to transform *E. coli* UltraMax DH5 α (Gibco/BRL) cells. Finally, the resulting plasmids were sequenced in the leader region by using primer P182.

It is acknowledged that these clones might have contained unintended mutations elsewhere in the viral genome introduced during genetic manipulation. Thus, to ensure that any phenotypic difference between wild-type EAV and the mutant viruses was indeed the result of the point mutations introduced in the vicinity of the ILO start codon, we restored the wild-type genotype in the mutant full-length EAV cDNA clones. For this purpose, oligonucleotide PT7(wt)atg containing the wild-type ILO initiation codon was used as the sense primer and oligonucleotide P962 as the antisense primer (Table 1). To reconstruct full length cDNA clones that would include these newly generated PCR fragments, we used the same procedure as described above and derived all remaining sequences from the plasmids that were subjected to restoration, i.e. T7NX(ko)ATG and T7NX(opt)ATG. The resulting constructs were designated T7NX(ko-to-wt)ATG and T7NX(opt-to-wt)ATG to indicate restoration of the ILO translational start site.

RNA transfection of cells and recovery of recombinant viruses

Capped run-off RNA transcripts from the *XhoI* linearized EAV cDNA clones T7NX(wt)ATG, T7NX(ko)ATG, T7NX(opt)ATG, T7NX(ko-to-wt)ATG, and T7NX(opt-to-wt)ATG were prepared by using MEGAscript transcription kit (Ambion, Austin, Tex.). The integrity of the RNA transcripts was verified by denaturing agarose gel electrophoresis prior to cell transfection.

RNA transcripts were transfected into subconfluent monolayers of BHK-21 cells by electroporation. Electroporations were carried out at room temperature or on ice-water with two consecutive pulses of 300 V, low R ohms (4 000 Ω), 60 μ F, by using the Cell-PoratorTM electroporation system (Gibco/BRL). The cells were then seeded in 25-cm² tissue culture flasks until observation of typical EAV cytopathic effect (CPE). Recombinant viruses were harvested essentially as previously described (41).

Serial passage and genetic analysis of recombinant viruses

To study the stability of the introduced mutations, the recombinant viruses recEAV(ko)AUG and recEAV(opt)AUG were serially propagated on BHK-21 cells until the 15th passage and collected at each passage when 80 to 90% of the cells exhibited a cytopathic effect (41). Virus-containing supernatants from transfected cells (passage P0) and from viral passages 5, 10 and 15 were analyzed for the presence of the introduced point mutations by RT-PCR sequencing using the T7 sequencing kit (Amersham Pharmacia Biotech). Primers used in the PCR were the sense Peav5' primer (corresponding to the first 10 nt at the 5' end of the EAV leader sequence) and the antisense P182 primer (Table 1). The PCR products were then cloned into the *SmaI*-digested pBlueScript KS⁺ (Stratagene) vector. Two clones derived from each viral passage level were finally sequenced.

Plaque assay

Plaque assay was performed with virus samples collected at different cell passages in order to compare the morphology and size of plaques of the viruses recEAV(ko)AUG and recEAV(opt)AUG with those of recombinant wild-type EAV. The virus samples were diluted to such an extent that one 10-cm² cell culture dish contained 10 to 100 plaques. The plaque assay was carried out upon standard procedures (Hierholzer and Killington, 1996).

Growth curves

BHK-21 cell monolayers in 6-well tissue culture plates were infected with each of the recombinant viruses (one well per recombinant virus) at an MOI of 0.5. At various times pi, the cell cultures were subjected to three cycles of freezing and thawing. The cell lysates were then clarified by centrifugation for 15 min at 5,000 x g and the viral titer (TCID₅₀ per milliliter) of the supernatants was determined as previously described (41).

RESULTS

In vitro expression and significance of the EAV intraleader peptide

Computer analysis (18) predicted that the IL peptide contained at least one linear epitope. Accordingly, we were able to successfully generate a rabbit antiserum directed against a synthetic version of the protein. However, with this serum, we failed to detect the IL-peptide in EAV-infected cells either by immunofluorescence, or by Western blot, or by RIPA (data not shown). We next generated a plasmid (F7wtATG) containing the EAV mRNA7 sequence downstream of a bacteriophage T7 RNA polymerase promoter. In vitro transcripts derived from F7wtATG were translated in vitro and the translation products were directly analyzed by SDS-PAGE. As shown in Fig. 2A (lane 1), polypeptide species of 14 and 4 kDa were detected corresponding to the N protein and the IL peptide, respectively. This result demonstrates that EAV mRNA7 is functionally bicistronic. Because the ILO start codon present in this wild-type construct is in the suboptimal context for translation initiation (24), we decided to introduce point mutations at the 5' end of the ILO to improve or abrogate its translation. As shown in Fig. 2A (lanes 2 and 3), the IL peptides from mutant plasmids F7(koz)ATG and F7(opt)ATG (in which the context of the ILO initiation codon was improved) were detected at levels comparable to that of the wild-type construct. As expected, lower amounts of the IL peptide were obtained with the construct F7(nf)ATG which had the ILO AUG codon placed in an unfavorable context for initiation of translation (Fig. 2A, lane 5). No translation product was observed from the mutant clone F7(-)ATG in which the ILO initiation codon was abrogated (Fig. 2A, lane 4). To confirm that the 4 kDa translation product is encoded by the ILO, RIPAs were performed using the rabbit anti-IL peptide serum or the corresponding preimmune serum (Fig. 2B). This experiment proved that the 4 kDa protein is encoded by the ILO and corroborated the previous finding that the amount of IL peptide synthesized depends on the context of the ILO start codon.

Figure 2: In vitro synthesis of the EAV IL peptide. **A:** Direct visualization of [³⁵S] methionine-labeled in vitro translation products in tricine-buffered SDS-10% PAA gels. The transcripts used for in vitro translation were derived from F7(wt)ATG (lane 1), F7(koz)ATG (lane 2), F7(opt)ATG (lane 3), F7(-)ATG (lane 4) and F7(nf)ATG (lane 5). Lanes 6 and 7 show the result of in vitro translation reactions without RNA (negative control) and with a transcript encoding the luciferase (Luc) protein (positive control), respectively. The positions of the IL peptide (ILP) and the nucleocapsid protein (NP) are indicated. **B:** RIPA of the EAV IL peptide synthesized in vitro using a rabbit antiserum specific for the IL peptide (+) or the corresponding preimmune serum (-). [¹⁴C]-labeled molecular (Mr) weight standards (kDa) (Amersham Pharmacia Biotech) are indicated at the left.

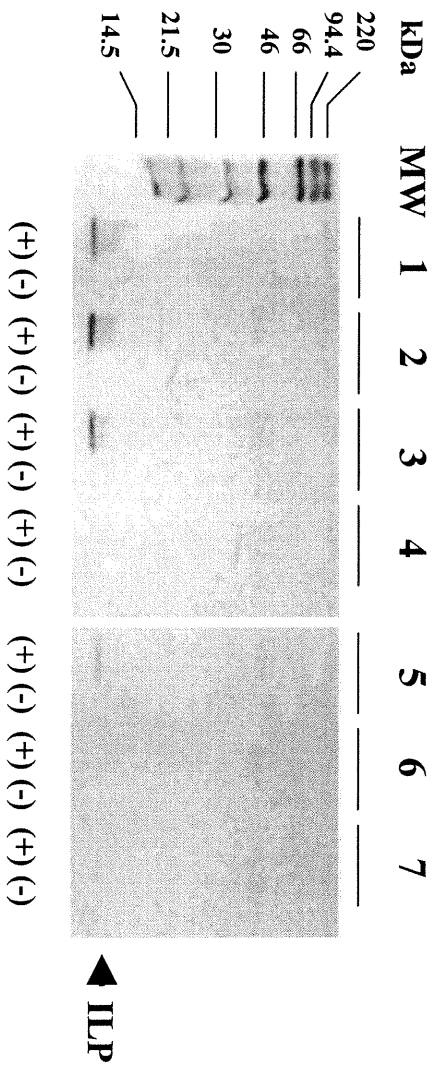
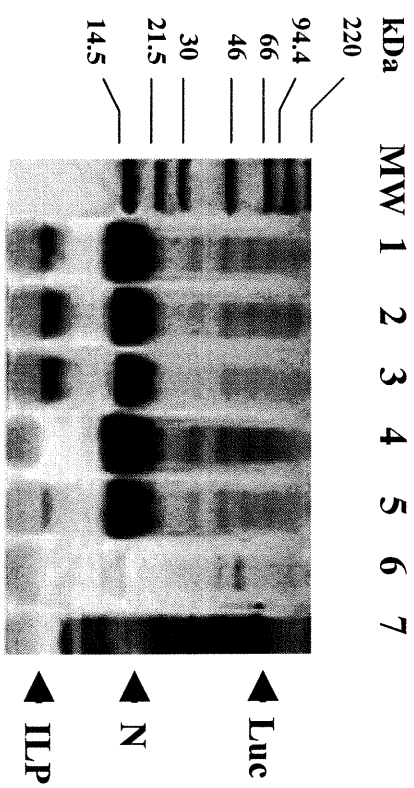
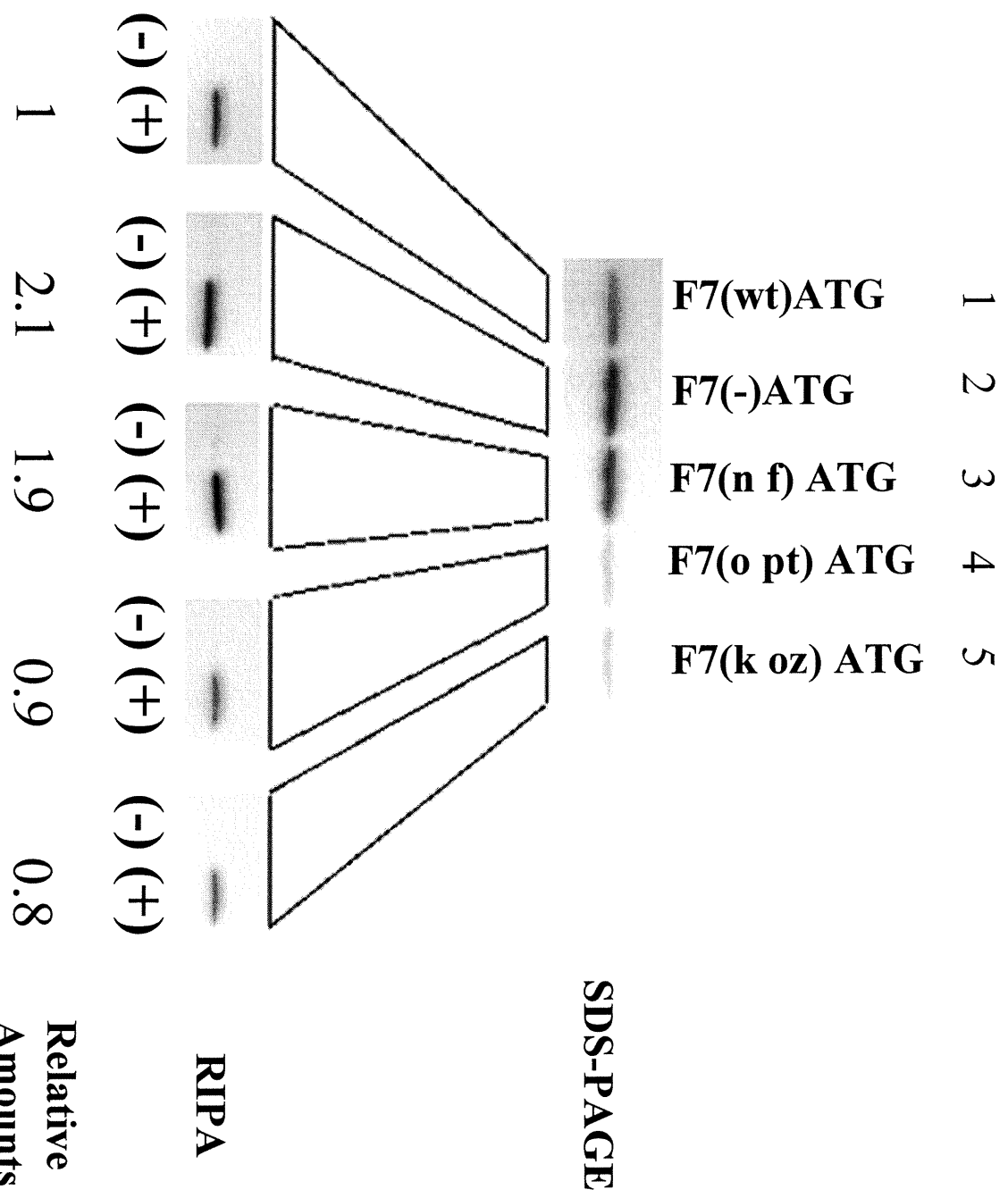


Figure 3: Effect of the ILO start codon and its direct surroundings on the in vitro translation of the EAV N protein. In vitro translation products of the mRNA7 expression plasmids F7(wt)ATG, F7(-)ATG, F7(nf)ATG, F7(opt)ATG, and F7(koz)ATG (lanes 1 to 5, respectively) were analyzed directly by SDS-PAGE and after immunoprecipitation (RIPA) using a rabbit antiserum raised against the EAV N protein (+) or the corresponding preimmune serum (-). The mean relative amounts of N protein were calculated from both experiments and for each construct and normalized to the parental construct (F7(wt)ATG) as indicated at the bottom



We next analyzed the effect of the mutations around the ILO start codon on the production of the N protein. As shown in Fig. 3, a approximately twofold increase (as determined by densitometric analysis) in the amount of N protein was observed with the mutant constructs F7(-)ATG (lane 2), or F7(nf)ATG (lane 3) relative to that of the wild-type plasmid F7(wt)ATG (lane 1). In contrast, slightly diminished N protein levels (15 to 20% less) were obtained with in vitro transcripts from the mutant clones F7(opt)ATG and F7(koz)ATG (lanes 4 and 5, respectively). Based on these results, we concluded that translation of the ILO has a negative effect on N protein synthesis. These results are suggestive of a leaky scanning mechanism (24, 25) for translation of the ORF immediately downstream of the EAV leader sequence in each viral (sg) mRNA.

Genetic analysis of mutated recombinant viruses

Following transfection of cells with RNA transcripts derived mutant full-length EAV cDNA clones, sequence analysis of each of the recombinant viruses was carried out and confirmed the presence of the respective mutations

In order to determine the genetic stability of the mutations introduced in the EAV leader region, the mutant viruses (recEAV[ko]AUG and recEAV[opt]AUG), as well as virus derived from the wild-type EAV cDNA clone (recEAV[wt]AUG), were each passaged fifteen times. EAV RNA from supernatants of transfected (P0) cells and at passages P5, P10 and P15 was then isolated and used for RT-PCR to amplify the 5' 182 nt of the viral genomes by using the high-fidelity *Pfu* polymerase. The PCR products were then sequenced. Surprisingly, both mutants appeared to revert to the wild-type sequence. Already within five passages, the knock-out mutation had been restored while this took some ten to fifteen passages for the virus carrying the optimized ILO start codon (Fig. 4). These results indicated that both recombinant virus mutants are subject to selective genetic pressure under which viruses acquiring the wild-type sequence around the ILO start codon have a selective advantage. Consistently, the leader sequence of the wild-type virus appeared to be genetically stable (data not shown).

Figure 4 : Sequence analysis of the ILO translational start site of recombinant virus mutants after different cell passages. Total RNA extracted from BHK-21 cells was used to amplify the EAV leader region by RT-PCR, and the leader-specific amplification products were cloned and sequenced. Sequences were determined for two independent plasmid clones (clone # 1 and clone # 2) of both the recombinant viruses recEAV(ko)AUG and recEAV(opt)AUG following cell transfection (P0) , five (P5), ten (P10) and fifteen (P15) cell passages. The nucleotide sequences surrounding the ILO start codon are indicated next to each gel picture. The experimentally mutated nucleotides and revertants hereof are shown in large capitals and their position in the sequence gels is indicated with an arrow.

	recEAV(ko)AUG		recEAV(opt)AUG	
	Clone #1	Clone #2	Clone #1	Clone #2
P 0	<p>ACGT 3'</p> <p>← TTAGT 5'</p>	<p>ACGT 3'</p> <p>← TTAGT 5'</p>	<p>3' .</p> <p>GGTATGA →</p> <p>ACGT</p> <p>5' .</p>	<p>3' .</p> <p>GGTATGA →</p> <p>ACGT</p> <p>5' .</p>
P 5	<p>3' .</p> <p>GGTATGT ←</p> <p>5' .</p>	<p>3' .</p> <p>GGTATGT ←</p> <p>5' .</p>	<p>. GGTATGA →</p> <p>5' .</p>	<p>. GGTATGA →</p> <p>5' .</p>
P 10	<p>3' .</p> <p>GGTATGT ←</p> <p>5' .</p>	<p>3' .</p> <p>GGTATGT ←</p> <p>5' .</p>	<p>. GGTATGT →</p> <p>5' .</p>	<p>3' .</p> <p>GGTATGA →</p> <p>5' .</p>
P 15	<p>3' .</p> <p>GGTATGT ←</p> <p>5' .</p>	<p>3' .</p> <p>GGTATGT ←</p> <p>5' .</p>	<p>. GGTATGT →</p> <p>5' .</p>	<p>. GGTATGT →</p> <p>5' .</p>

Kinetics of recombinant virus growth and of intracellular protein synthesis

We studied the growth properties of the viruses derived from the transfected cells. As indicated by their growth curves (Fig. 5), both virus mutants derived from the transfected cells (P0) replicated less efficiently than the recombinant wild-type virus. This resulted in more than tenfold lower viral titers as measured during the period of 14 to 24/26 h pi and to lower N and M protein synthesis levels at 12, 16, and 18 h pi (data not shown). In contrast, the growth curves of viruses in which the wild-type leader sequence had been restored by genetic manipulation (recEAV[ko-to-wt]AUG and recEAV[opt-to-wt]AUG) or by repeated passage (recEAV[ko]AUG and recEAV[opt]AUG) were similar to that of the recombinant wild-type virus.

Characterization of plaque formation

To further characterize the recombinant mutant viruses, we assessed the size and morphology of their plaques induced in Vero cells after different passages. As shown in Fig. 6, the recombinant wild-type virus (recEAV[wt]ATG) yielded a homogeneous population of large plaques with an average diameter of 5 mm. The size and morphology of these plaques were similar to those of the recombinant viruses recEAV(ko-to-wt)AUG and recEAV(opt-to-wt)AUG and to those of the Bucyrus strain of EAV which were used as controls. In contrast, plaques of variable size were obtained for the mutant viruses recEAV(ko)AUG and recEAV(opt)AUG after a single cell passage. It is important to note that most of these plaques (more than 70%) had an average diameter of 1.5 mm. However, this heterogeneous plaque phenotype was no longer observed after cell passage 15. At this stage, the plaque phenotype had become identical to that of the recombinant viruses with a wild-type leader sequence.

The differences observed in plaque morphology and size prompted us to analyze further the genotype of the virus found in each type of plaque. To this end, total cellular RNA was extracted from individual plaques and used to amplify the leader region by RT-PCR as described above. The PCR products were sequenced. The results showed that the small plaques were produced by the mutant viruses whereas the medium-size and large plaques contained viruses in which the ILO translation start site had reverted to the wild-type sequence. These results indicate that restoration of the wild-type genotype occurred already early in the life cycle of the mutant viruses and correlate with the fact that plaques of both sizes were observed with both virus mutants following cell transfection.

Figure 5: Growth curves of the EAV recombinant viruses. **A:** Comparison of growth kinetics of the virus in which the ILO start codon was knocked out (recEAV[ko]AUG) with those of the recombinant wild-type virus (recEAV[wt]AUG) and recombinant viruses in which the wild-type leader sequence was restored by genetic manipulation (recEAV[ko-to-wt]AUG) or by repeated passages (recEAV[ko]AUG-15P). **B:** Comparison of growth kinetics of the virus in which the context of the ILO start codon was improved (recEAV[opt]AUG) with those of the recombinant wild-type virus (recEAV[wt]AUG) and recombinant viruses in which the wild-type leader sequence was restored by genetic manipulation (recEAV[opt-to-wt]AUG) or by repeated passages (recEAV[opt]AUG-15P).

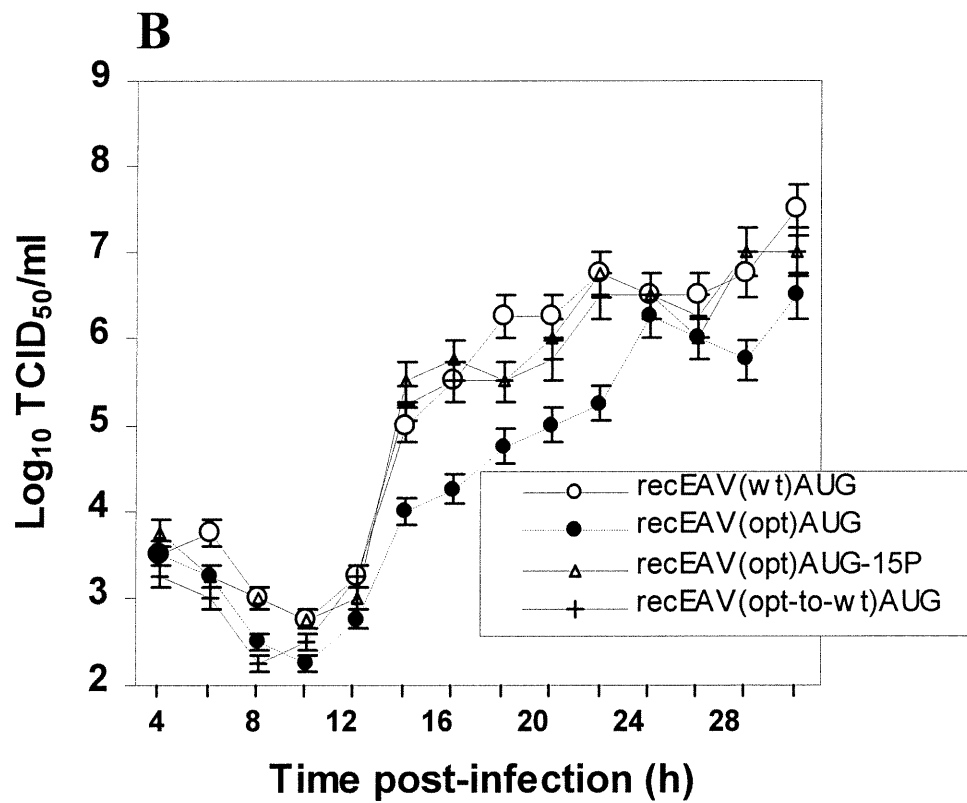
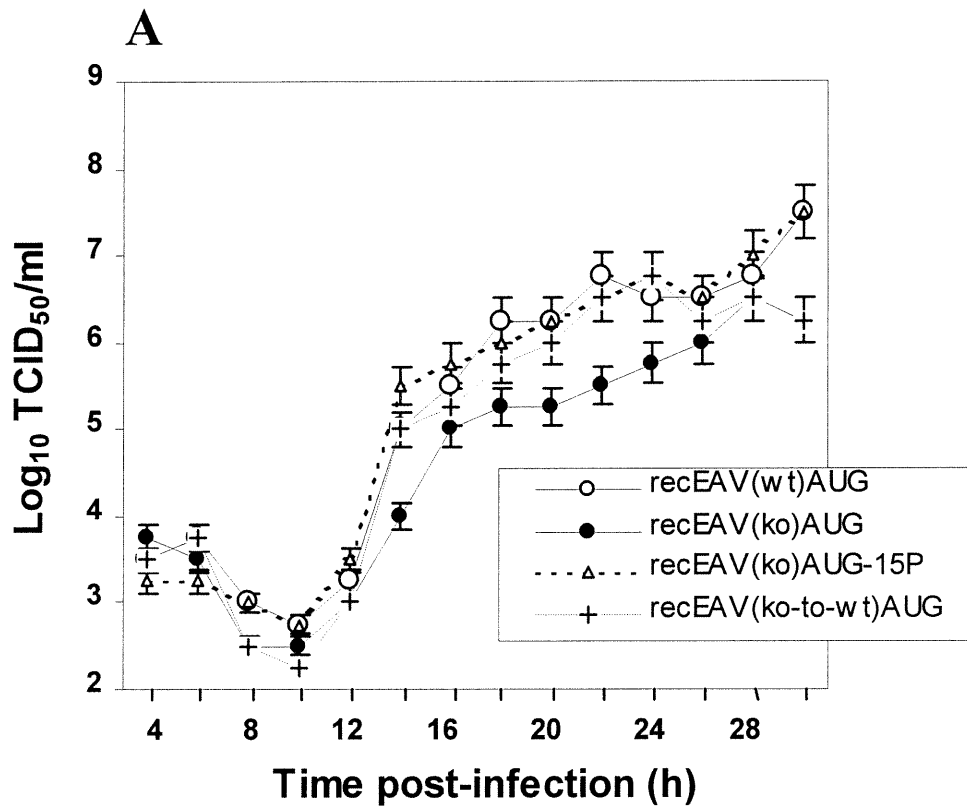
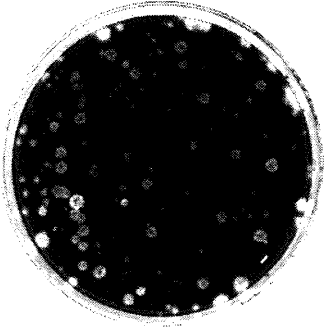
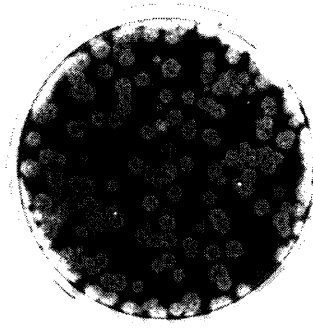


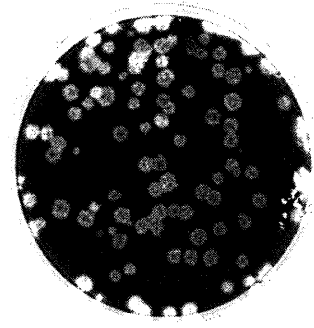
Figure 6: Plaque morphology of EAV recombinant viruses. Monolayers of Vero cells were either mock-infected (negative control) or infected with the indicated recombinant viruses or the Bucyrus strain of EAV (positive control) and covered with an agar overlay. At 4 days pi, the overlay was removed and the monolayers were fixed and stained with 0.9 % violet crystal in 25 % formaldehyde and 5 % ethanol. The recombinant viruses used for the plaque assay were those recovered after cell transfection or had been used after the fifteenth cell passage (P15).



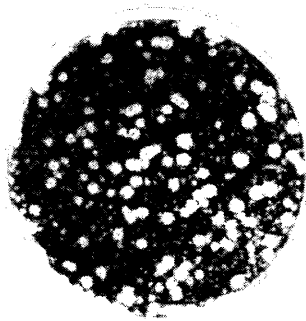
recEAV(ko)AUG



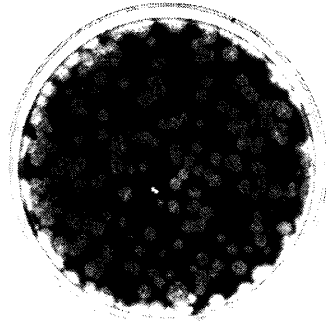
recEAV(ko)AUG-15P



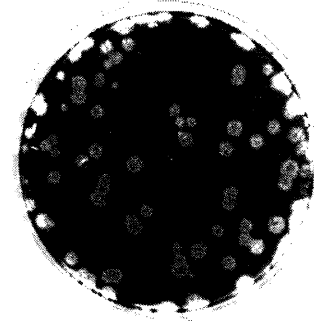
recEAV(ko-wt)AUG



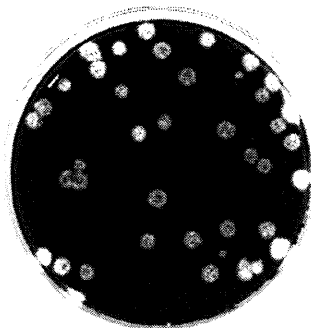
recEAV(opt)AUG



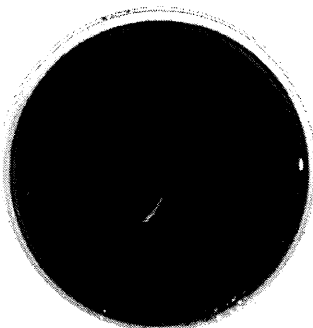
recEAV(opt)AUG-15P



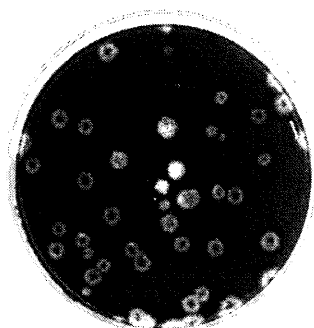
recEAV(opt-wt)AUG



recEAV(wt)AUG



mock



Bucyrus

DISCUSSION

Leader sequences, defined as nontranslated sequences of mRNA, upstream of the initiation codon, occur in mRNAs of many different viruses. These sequences have been demonstrated to control various functions in the viral life cycle. In papillomaviruses, for instance, an agnoprotein translated from an upstream ORF has been reported to regulate the levels of processes as different as transcription, mRNA processing and translation, viral assembly, and maturation or release of mature viruses (14, 17, 32, 34, 35). In Rous Sarcoma virus, the first and third ORFs upstream of the gag gene in the viral leader RNA have been shown to be involved in translational regulation and viral RNA packaging (10). In the cauliflower mosaic virus (CaMV), translation of leader peptides is necessary for reinitiation at downstream ORFs by a shunt ribosomal mechanism (11). Moreover, other important steps in the CaMV infection cycle are also regulated or influenced by elements located within the leader sequence (33). Finally, in several RNA viruses, leader sequences have, either at the primary or at the secondary structure level, been reported to play an important role in virus virulence or pathogenicity. Notable examples are the Sindbis (23, 28) and Venezuelan equine encephalomyelitis viruses (22) and the Sabin strain of poliovirus (42).

As we previously reported, the 5' leader region of the EAV genome contains an ILO of 111 nt that was predicted to produce an IL peptide of 37 aa (19). Because this intraleader ORF is present in various EAV isolates, a putative role in the EAV life cycle was suggested (21). Here, the expression of the EAV IL peptide was detected in an *in vitro* cell-free translation system programmed by a plasmid containing a cDNA copy of mRNA7. Although small ILOs with undetermined function have been observed in the genomes of PRRSV and SHFV, two other members of the *Arteriviridae* family (29, 44), there was no evidence of an expressed leader peptide.

As a consequence of its particular replication strategy, all genomic and subgenomic viral mRNAs in EAV-infected cells carry the ILO. We were therefore quite surprised that, despite this abundance, we were unable to detect the IL peptide in infected cells, neither by RIPA nor by Western blotting or immunofluorescence assay. This negative result can be explained in a number of ways. First, the IL peptide might simply not be produced during virus replication or be synthesized in very low amounts. Second, the folding of the IL peptide in a reticulocyte lysate may be different from that in virus-infected cells and may

hence interfere with its detection. Third, it is also possible that the IL peptide is rapidly degraded in cells. This would also be consistent with our inability to detect the protein by expression from the plasmid carrying the ILO in a wild-type AUG context using the vaccinia virus vTF7-3 expression system (data not shown).

Bi- and multicistronic mRNAs occur - though rarely - in normal eucaryotic cells (for a review, see 31) but appear more frequently in viral mRNAs. Expression of the different genes is regulated by particular sequence conditions. Unless driven by an internal ribosome entry site (IRES), the expression of a downstream ORF is generally inefficient. For instance, a 61 amino acids leader ORF present in a bicistronic mRNA species of SV40, which is translated by a leaky scanning mechanism, has been shown to reduce the synthesis of the downstream protein threefold (39). Similarly, the second of three short upstream ORFs in the leader of the human cytomegalovirus gp48 gene also inhibits downstream translation (3). Remarkably, this inhibition depended on the particular amino acid sequence specified by this leader ORF. In addition, the authors demonstrated that expression of the cistron downstream from this leader ORF depended on ribosomes bypassing its AUG codon by a leaky scanning mechanism (3). Accordingly, the reduced expression of N protein that we observed in vitro when the EAV ILO initiation codon was functional would also suggest a leaky scanning mechanism. The significance, however, of such an attenuating effect of the ILO on the expression of all EAV genes is as yet hard to understand. In this connection, it is interesting to note that in the related coronaviruses, which are normally devoid of an intraleader ORF, short ILO's can sometimes emerge as a result of single base substitutions (5, 16). Their appearance in the genomes of bovine coronavirus (BCV) and mouse hepatitis virus (MHV) has been associated with virus persistence in vitro and in vivo, respectively (5, 16). In the case of BCV, the presence of an ILO of 33 nt in length was shown to attenuate translation of downstream ORFs, whereas in MHV, the ILO (54 nt) appeared to enhance translation of downstream ORFs in vitro. Although both ILOs had a significant impact on the expression of downstream ORFs, their translation products have never been detected.

To investigate the role of the ILO in the life cycle of EAV, we introduced mutations in the leader region of an infectious cDNA clone of the virus. The effects of these mutations were then analyzed by comparing the phenotypic properties of the viruses derived from the mutant full-length EAV cDNA clones with those of a recombinant wild-type virus. Interestingly, a mutant EAV clone in which the expression of the ILO had been blocked,

was able to produce infectious virions, consistent with a recent note (30) where EAV infectivity was also reported to be maintained after disruption of the intraleader AUG initiation codon. Obviously, the leader peptide per se is not essential for the EAV replication cycle. However, whereas in the other study the authors did not observe any phenotypic difference between mutant and wild-type virus (30), our results showed that modulating the ILO translation efficiency, either positively or negatively, by mutation of the AUG codon or of its context sequence, significantly affected viral characteristics. Mutations that were confirmed in vitro to either block or enhance the expression of the ILO both resulted in viruses that exhibited delayed growth kinetics, reduced virus yields and smaller plaque size.

The mutations that we introduced into the EAV genome appeared to be rapidly restored during serial passages of the viruses. Already following cell transfection, a significant fraction of both mutant viruses had reverted to the wild-type phenotype and this effect became progressively complete following subsequent rounds of infection. Phenotypic restoration indeed correlated with genetic correction of the mutations as was demonstrated by sequencing of the relevant genomic region. The wild-type phenotype was also regained by experimentally restoring the mutations ruling out the possibility that mutations elsewhere in the genome were responsible for the small-plaque phenotype and reduced growth.

Taking all our data together, we cannot draw any conclusion about the function of the intraleader peptide in the EAV replication cycle. Two different interpretations are possible: either the peptide has no function at all or, though not essential, it does play some subtle role. In the latter case, it would seem that a certain, delicate level of the peptide is required for optimal replication. The amount of peptide synthesized would be regulated precisely by the particular sequence context of the initiating AUG. This context is understandably suboptimal as the efficient expression of the ILO would, through its negative effect on downstream ORFs, be at the unfavourable expense of the synthesis of all other viral proteins. The rapid repair of the mutations that we introduced in or around the AUG sequence would also be consistent with the requirement for a dosed translation of the ILO. What function the translation product would serve remains as yet unknown. In the alternative interpretation, the IL peptide per se is irrelevant. In that case, it must be the nucleotide sequence at and around the AUG that is of importance. In general, 5'- and 3'-terminal domains of RNA virus genomes play essential roles in replication and

transcription. Their sequences are important for recognition by viral and host cell proteins that mediate these processes. In nidovirus replication, the 3'-terminal sequences of the minus-strand RNAs, i.e. the sequences complementary to the leader, direct the synthesis of genomic and subgenomic RNAs. Hwang and Brinton (1998) reported that a sequence of 68 nt within the antileader sequence of Simian Hemorrhagic Fever Virus was the specific site of fixation of four cellular proteins. The same cellular proteins were found to bind to similar regions in the antileader of the other members of arteriviruses. They also predicted the secondary structures corresponding to the 68 nt sequence. In EAV, we have predicted and compared the secondary structures of several EAV isolates (Kheyar et al., 1998). The ILO AUG codon base-pairs with a three-nt sequence at positions 38 to 40 in an upstream stem-loop structure. By using the *mfold version 3.1* software of Zuker et al (1999) to predict the secondary structure of the Bucyrus strain of EAV, the G to T mutation leads to the formation of a small internal loop and a small bulge downstream and in front of the mutation site, respectively. However, the T to A mutation leads to the formation of a small internal loop at the mutation site.

It is not known whether the secondary structure in the vicinity of the ILO AUG might have a pivotal role in the viral RNA-cellular protein interactions but it is thus quite plausible that strong selective pressures exist to maintain these sequences. The precision with which the G-to-T mutation in the IL AUG and the T-to-A mutation at position -3 were repaired supports this view as does the apparent stability of the reported A-to-G mutation in the AUG (30). The latter mutation had an identical effect (i.e. complete inhibition) on the expression of the ILO as had our G-to-T mutation. Yet, in contrast to our results, viral replication efficiency nor plaque size were significantly affected. We therefore favour the idea that the IL peptide has no functional significance but that it is the precise nucleotide sequence that is essential for the life cycle of the virus.

ACKNOWLEDGEMENTS

This work was supported by an operating grant from the National Sciences and Engineering Research Council of Canada to D. Archambault. A. Kheyar was supported by a graduate student fellowship from Université de Montréal. D. Archambault is the holder of a research senior scholarship from the Fond de la Recherche en Santé du Québec.

REFERENCES

1. **Anderson, D. J., and G. Blobel.** 1983. Immunoprecipitation of proteins from cell-free translations. *Methods Enzymol.* **96**:111-120.
2. **Archambault, D., G. Laganière, and G. St-Laurent.** 1998. Genetic variation and phylogenetic analysis of open reading frames 3 and 4 of various equine arteritis virus isolates. *Adv. Exp. Med. Biol.* **440**:813-819.
3. **Cao, J., and A. P. Geballe.** 1995. Translational inhibition by a human cytomegalovirus upstream open reading frame despite inefficient utilization of its AUG cod. *J. Virol.* **69**:1030-1036.
4. **Cavanagh, D.** 1997. Nidovirales: a new order comprising *Coronaviridae* and *Arteriviridae*. *Arch. Virol.* **142**:629-633.
5. **Chen, W., and R. S. Baric.** 1995. Function of a 5'-end genomic RNA mutation that evolves during persistent mouse hepatitis virus infection in vitro. *J. Virol.* **69**:7529-40.
6. **de Vries, A. A. F., E. D. Chirnside, P. J. Bredenbeek, L. A. Gravestien, M. C. Horzinek, and W. J. M. Spaan.** 1990. All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res.* **18**:3241-3247.
7. **de Vries, A. A. F., E. D. Chirnside, M. C. Horzinek, and P. J. M. Rottier.** 1992. Structural proteins of equine arteritis virus. *J. Virol.* **66**:6294-6303.
8. **den Boon, J. A., M. F. Kleijnen, W. J. M. Spaan, and E. J. Snijder.** 1996. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J. Virol.* **70**:4291-4298.
9. **den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. F. de Vries, M. C. Horzinek, and W. J. M. Spaan.** 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J. Virol.* **65**:2910-2930.

10. **Donzé, O., P. Damay, and P. F. Spahr.** 1995. The first and third uORFs in RSV leader RNA are efficiently translated: implications for translational regulation and viral RNA packaging. *Nucleic Acids Res.* **23**:861-868.
11. **Fütterer, J., Z. Kiss-Laszlo, and T. Hohn.** 1993. Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* **73**:789-802.
12. **Glaser, A. L., A. A. F. de Vries, M. J. B. Raamsman, M. C. Horzinek, and P. J. M. Rottier.** 1999. An infectious cDNA clone of equine arteritis virus: a tool of future fundamental studies and vaccine development. *In* Wernery U., Wade J. F., Mumford J. A., and Kaaden O.- R. (ed), *Proceedings of the Eighth International Conference on Equine Infectious Diseases*, Dubai 1998, pp.166-176. R & W Publications, Ltd., Newmarket, England.
13. **Harlow, E., and D. Lane.** 1988. *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.
14. **Hay, N., and Y. Aloni.** 1985. Attenuation of late simian virus 40 mRNA synthesis is enhanced by the agnoprotein and is temporally regulated in isolated nuclear systems. *Mol. Cell Biol.* **5**:327-1334.
15. **Hedges J. F., U. B. Balasuriya, and N. J. MacLachlan.** 1999. The open reading frame 3 of equine arteritis virus encodes an immunogenic glycosylated, integral membrane protein. *Virology* **264**:92-98.
16. **Hierholzer, J. and R. Killington.** 1996. Quantitation of virus; plaque assay. *In* B. W. Mahy and H. O. Kangro (ed). *Virology Methods Manual*. pp. 38-39. Academic Press, London.
16. **Hofmann, M. A., S. D. Senanayake, and D. A. Brian.** 1993. A translation-attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *Proc. Natl. Acad. Sci. USA* **90**:11733-11737.

17. **Hou-Jong, M.-H. J., S. H. Larsen, and A. Roman.** 1987. Role of the agnoprotein in regulation of simian virus 40 replication and maturation pathways. *J. Virol.* **61**:937-939.
18. **Jameson, B. A., and H. Wolf.** 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. *Comput. Appl. Biosci.* **4**:181-186.
19. **Kheyar, A., G. St-Laurent, and D. Archambault.** 1996. Sequence determination of the extreme 5' end of the equine arteritis virus leader region. *Virus Genes* **12**:291-295.
20. **Kheyar, A., S. Martin, G. St-Laurent, P. J. Timoney, W. H. McCollum, and D. Archambault.** 1997. Expression cloning and humoral immune response to the nucleocapsid and membrane proteins of equine arteritis virus. *Clin. Diagn. Lab. Immunol.* **4**:648-652.
21. **Kheyar, A., G. St-Laurent, M Diouri, and D. Archambault.** 1998. Nucleotide sequence and genetic analysis of the leader region of Canadian, American and European equine arteritis virus isolates. *Can. J. Vet. Res.* **62**:224-230.
22. **Kinney, R. M., G.-J. Chang, K. R. Tsuchiya, J. M. Sneider, J. T. Roehrig, T. M. Woodward, and D. W. Trent.** 1993. Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5'-noncoding region and the E2 envelope glycoprotein. *J. Virol.* **67**:1269-1277
23. **Kobiler, D., C. M. Rice, C. Brodie, A. Shahar, J. Dubuisson, M. Halevy, and S. Lustig.** 1999. A single nucleotide change in the 5' noncoding region of Sindbis virus confers neurovirulence in rats. *J. Virol.* **73**:10440-10446.
24. **Kozak, M.** 1991. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* **115**:887-903.
25. **Kozak, M.** 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene* **234**:187-208.

26. **Lai, M. M., and D. Cavanagh.** 1997. The molecular biology of coronaviruses. *Adv. Virus Res.* **48**:1-100.
27. **Lepage, N., G. St-Laurent, S. Carman, and D. Archambault.** 1996. Comparison of nucleic and amino acid sequences and phylogenetic analysis of the G_S protein of various equine arteritis virus isolates. *Virus Genes* **13**:87-91.
28. **McKnight, K. L., D. A. Simpson, S.-C. Lin, T. A. Knott, G. M. Polo, D. F. Pence, D. B. Johannsen, H. W. Heidner, N. L. Davis, and R. E. Johnston.** 1996. Deduced consensus sequence of Sindbis virus strain AR339: mutations contained in laboratory strains which affect cell culture and in vivo phenotypes. *J. Virol.* **70**:1981-1989.
29. **Meulenberg, J. J. M., J. N. A. Bos-de Ruijter, R. van de Graaf, G. Wensvoort, and R. J. M. Moormann.** 1998. Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. *J. Virol.* **72**:380-387.
30. **Molenkamp, R., H. van Tol, B. C. Rozier, Y. van Der Meer, W. J. Spaan, and E. J. Snijder.** 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J Gen Virol.* **81**:2491-2496.
31. **Morris, D. R., and A. P. Geballe.** 2000. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* **20**:8635-8642.
32. **Ng, S. C., J. E. Mertz, S. Sanden Will, and M. Bina.** 1985. Simian virus 40 maturation in cells harboring mutants deleted in the agnogene. *J. Biol. Chem.* **260**:1127-1132.
33. **Pooggin, M. M., T. Hohn, and J. Futterer.** 1998. Forced evolution reveals the importance of short open reading frame A and secondary structure in the cauliflower mosaic virus 35S RNA leader. *J. Virol.* **72**: 4157-4169.

34. **Resnick, J., and T. Shenk.** 1986. Simian virus 40 agnoprotein facilitates normal nuclear location of the major capsid polypeptide and cell-to-cell spread of virus. *J. Virol.* **60**:1098-1106.
35. **Rinaldo, C. H., T. Traavik, and A. Hey.** 1998. The agnogene of the Human Polyomavirus BK Is Expressed. *J. Virol.* **72**:6233-6236.
36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. *Molecular cloning: a laboratory manual*. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
37. **Sawicki, S. G, and D. L. Sawicki.** 1998. A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* **440**:215-219.
38. **Schägger, H., and G. von Jagow.** 1987. Tris-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166**:368-379.
39. **Sedman, S. A., G. W. Gelembiuk, and J. E. Mertz.** 1990. Translation initiation at a downstream AUG occurs with increased efficiency when the upstream AUG is located very close to the 5' cap. *J. Virol.* **64**:453-457.
40. **Snijder, E. J., H. van Tol, K. W. Pedersen, M. J. B. Raamsman, and A. A. F. de Vries.** 1999. Identification of a novel structural protein of arteriviruses. *J. Virol.* **73**:6335-6345.
41. **St-Laurent, G., G. Morin, and D. Archambault.** 1994. Detection of equine arteritis virus following amplification of structural and nonstructural viral genes by reverse transcription PCR. *J. Clin. Microbiol.* **32**:658-665.
42. **Stewart, S. R., and B. L. Semler.** 1998. RNA structure adjacent to the attenuation determinant in the 5'-non-coding region influences poliovirus viability. *Nucleic Acids Res.* **26**:5318-5326.

43. **van Marle, G., J. C. Dobbe, A. P. Gulyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder.** 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. USA* **96**:12056-12061.

44. **Zeng, L., E. K. Godeny, S. L. Methven, and M. A. Brinton.** 1995. Analysis of simian hemorrhagic fever virus (SHFV) subgenomic RNAs, junction sequences, and 5' leader. *Virology* **207**:543-548.

- 45 **Zuker, M., D.H. Mathews, and D.H. Turner.** 1999. Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide In J. Barciszewski and B. F. C. Clark, eds. *RNA Biochemistry and Biotechnology*. Pp. 11-43. Kluwer Academic Publishers, NATO ASI Series.

CHAPITRE VI

DISCUSSION GÉNÉRALE

Durant l'accomplissement de ce projet, le VAE ainsi que les virus membres du genre *Arterivirus* auquel il appartient, ont subi des modifications d'ordre taxonomique. Malgré les publications qui sont apparues depuis le début des années 90 et qui préparaient ces virus à être déclassés de leur ancienne famille les *Togaviridae*, la reclassification officielle dans le nouvel ordre des *Nidovirales* avec ses deux familles des *Arteriviridae* et *Coronaviridae* n'a été adoptée que vers 1996 dans une réunion spéciale (CITV) (Cavanagh, 1997). L'établissement enfin, de cette nouvelle famille se doit aux propriétés moléculaires et biologiques uniques de ses quatre membres. Leur affiliation avec la famille des *Coronaviridae* reconnaît l'existence d'une relation ancestrale à l'échelle de l'organisation et la réplication génomique. La mise en évidence des séquences des différents membres de virus a été le facteur essentiel dans cette reclassification. Par ailleurs, d'autres facteurs liés à la nature du génome favorisent l'avancement rapide des études de ces virus. Parmi ces facteurs se retrouve le fait que l'ARN de ces membres est de polarité positive donc se comportant d'emblée comme infectieux. Par contre, la taille du génome de ces membres (entre 13 et 15 kb) leur donne un avantage par rapport aux autres membres des *Coronaviridae* dont la taille est de presque le double. Cela s'est manifesté par le développement récent de clones d'ADNc infectieux du VAE (van Dinten et al., 1997; Glaser et al., 1999) et du VSRRP (Meulenberg et al., 1998), ce qui certainement contribuera à l'analyse plus sophistiquée du cycle biologique et de la réplication des artérovirus. Ainsi, l'on peut s'attendre à ce que les coronavirologistes puiseront désormais leur source d'inspiration des artérovirologistes contrairement à ce que l'on a constaté depuis longtemps à travers l'évolution des publications concernant ces deux familles.

L'établissement d'un clone d'ADNc infectieux nécessite la connaissance de la totalité de la séquence du génome en question. Or, les séquences les plus difficiles à déterminer dans une longue molécule d'ARN sont souvent les extrémités 5' et 3' des ARN non polyadénylés. Notre projet a contribué au début de notre étude, de façon significative à la mise en évidence de l'extrémité 5' de la SL du VAE, étape indispensable à l'établissement des clones infectieux rapportées par van Dinten et al. (1997) et Glaser et al., (1999).

Les observations de de Vries et al. (1990) et den Boon et al. (1991) concernant la taille de la SL du VAE qui était estimée à 208 ou 207 nt respectivement a été confirmée et vérifiée avec notre étude. Nous avons en effet démontré que la taille exacte de la SL sans la séquence de jonction était de 206 nt pour la souche prototype Bucyrus du VAE (Fig. 2 de

l'article 1). Cela correspondait à l'ajout de 17 nt en amont de la première base rapportée dans la séquence publiée par den Boon et al. (1991). Ces différences dans la taille de la SL peuvent être attribuées à la base 5' proximale. Quoique la possibilité de la reconnaissance de la structure de la coiffe par la polymérase lors de l'amplification ou de la réaction d'extension d'amorce ne doit pas être complètement écartée (Hirzmann et al. (1993), elle peut expliquer la différence d'une base entre les travaux de de Vries (1990), de den Boon (1991) ou des nôtres (Kheyar et al., 1996). Cependant, nous avons authentifié au départ notre observation par le séquençage de la SL de 2 à 3 clones différents de chacune de quatre souches différentes de VAE : Bucyrus, Vienne et les isolats de champ 87AR-A1 et 86NY-A1. Malgré cela, la première base d'une longue chaîne d'ARN semble être plus incertaine pour plusieurs raisons comme la possibilité qu'elle corresponde à la structure de la coiffe ou encore des erreurs de reconnaissance et d'incorporation de bases par des polymérases au début du processus de la transcription. En effet, lors de l'établissement du premier clone infectieux du VAE, van Dinten et al. (1997) ont également rapporté l'existence de 17 bases à l'extrémité 5', sauf que la première base était hétérogène sur plusieurs clones. puisque quatre clones portaient à la première position une base G et deux avaient plutôt une base A. Les auteurs ont considéré la base G comme premier nt du clone d'ADNc pleine longueur du VAE. De même, selon Glaser et al. (1999) et de Vries (communication personnelle), lors de l'établissement du second clone infectieux, ils ont trouvé que sur 7 plasmides recombinants portant la SL complète du VAE, trois contenaient les 17 bases additionnelles et débutaient avec la base G. Par contre, deux clones manquaient la première base, un autre manquait les deux premières et le dernier contenait deux bases supplémentaires en amont. Meulenberg et al. (1998), lors de l'établissement de leur clone infectieux du VSRRP, avaient également constaté cette hétérogénéité des premières bases du génome. Sur 26 clones portant la SL, 22 étaient identiques, alors que les quatre autres manquaient une à trois bases au début de la séquence.

Lors de la détermination de la SL complète de 11 souches/isolats différents de VAE, trois souches avaient la base A à la première position (87AR-A1, Vienne et 15492) et un seul isolat (86AB-A1) avait la délétion de la première base. Les 7 souches/isolats restants débutaient par la base G avec la séquence contenant les 17 nt supplémentaires (Kheyar et al., 1998). Il faut souligner que la souche Bucyrus utilisée durant la première étude avait été l'objet de plusieurs passages en culture cellulaire avant l'isolement de son ARN utilisé comme gabarit dans les réactions de transcription inverse (Kheyar et al., 1996). Par contre,

la souche Bucyrus utilisée dans la seconde étude correspondait à un passage peu élevé de la souche provenant directement de l'ATCC (Kheyar et al., 1998). Ainsi la première base de la première souche analysée consistait en la base A alors que celle de la souche Bucyrus de l'ATCC était un G.

La connaissance de la première base de l'ARNm est d'une grande importance dans l'établissement d'un clone infectieux car elle peut influencer l'initiation de la transcription. Pour obtenir des rendements élevés de transcription, la plupart des chercheurs utilisent le promoteur de la polymérase du bactériophage T7 en amont de la séquence à transcrire. Or, ce système trouve son efficacité plus élevée lorsque la première base située en aval de la séquence du promoteur T7 est une base G tel que rapporté par Milligan et Uhlenbeck (1989) qui ont évalué les meilleures conditions de l'utilisation de ce système.

Lors de la mise en évidence des 17 nt à l'extrémité 5' de la SL, quelques substitutions ont été observées avec la souche NY86-A1 aux positions 1 et 7 (Fig. 1 de l'article 1) ainsi que d'autres substitutions en aval des 17 nt et avec les autres souches par rapport à la souche de référence (résultats non montrés dans l'article 1). Étant donné la position de la SL à la tête du génome viral ainsi que son mouvement, sa présence active et son implication certaine lors de la transcription, on se demandait du degré de sa stabilité parmi plusieurs souches de provenance géographique différente et de son expression. De plus, les études comparatives de séquences effectuées dans notre laboratoire sur différents segments et différentes souches n'avaient pas démontré des variations significatives ou une association de variations séquentielles avec des différences antigéniques ou de virulence. C'est pour cela que nous avons entrepris notre travail sur la détermination de la SL complète de la souche Bucyrus du VAE et ensuite sa comparaison entre 11 souches et isolats dont nous disposions dans notre laboratoire (Tableau I de l'article 2).

L'analyse de la SL complète à l'échelle primaire, sans le motif de jonction (Fig. 1 de l'article 2), a révélé que la plupart des souches avaient la même taille de 206 nt excepté deux isolats dont la première base est délétée (205 nt) (isolat 86AB-A1) ou une base supplémentaire est insérée (207 nt) (isolat 86NY-A1). Les substitutions entre les différentes souches et isolats sont considérées comme des variations mineures puisqu'elles sont éparpillées le long de la SL et ne représentaient pas des motifs bien particuliers pouvant susciter de l'intérêt. De plus, les pourcentages d'identité entre elles étaient élevés, se situant entre 92,2 à 98,5 % (Tableau. II de l'article 2). Fait intéressant cependant, la souche Vienne de VAE qui était la seule souche européenne à être utilisée dans cette étude comparative

était celle qui présentait le plus faible pourcentage d'identité par rapport aux autres souches Nord-Américaines. L'analyse d'autres souches virales d'origine encore plus diversifiée et de divers continents apporterait plus de détails quant à cette observation. Par contre, avec les pourcentages d'homologie apportés dans ce travail, l'on ne peut pas avancer l'hypothèse d'une nette divergence de souche de provenance géographique différente. De plus, la plupart des travaux de comparaison de séquences autres que la SL ont mentionné des pourcentages d'homologie inférieurs à 90% qui étaient considérés comme étant mineurs (Chirnside et al.1994; Hedges et al., 1996; Lepage et al., 1996.; St-Laurent et al. 1997). Finalement, lorsque les différentes séquences furent soumises à une analyse phylogénétique assistée par ordinateur, les souches et isolats étudiés représentaient un seul groupe relié phylogénétiquement excepté une légère divergence de la souche Vienne (résultats non montrés).

Quelques mutations ont été observées comme étant éparpillées et touchant les mêmes positions (7, 35, 43, 78 et 154) chez les différentes souches et isolats. Vingt neuf substitutions totales entre toutes les souches et isolats représentaient environ 14 % de la SL entière (Fig. 1 de l'art. 2). Malgré cela les variations au niveau des structures nucléotidiques primaires étaient mineures. Il était donc impératif de pousser nos analyses à l'échelle des structures secondaires. Jusqu'au début de nos travaux, il n'y avait aucune preuve que l'initiation de la traduction des protéines du VAE requérait la présence de la coiffe et mettait en jeu un système du balayage par le ribosome (« ribosome scanning model » de Kozak, 1989c). Or, le système de l'initiation interne de traduction du ribosome par l'intermédiaire de la structure « IRES » chez les picornavirus et les flavivirus (Jang et al., 1990; Tsukiyama-Kohara et al., 1992) est plus dépendant de la structure secondaire de la SL. La seule indication pouvant laisser présumer que le génome du VAE pourrait bien être pourvu de la structure de la coiffe était son existence chez un membre apparenté aux *Arteriviridae*; le VFHS (Sagripanti et al., 1986). Néanmoins, quel que soit le mode d'initiation de la traduction, la structure secondaire de la SL pouvait jouer un rôle direct ou indirect dans cette fonction. Il a été même rapporté que la neurovirulence de certaines souches de poliovirus est associée à des structures secondaires de la SL (Haller et al., 1996).

Les structures secondaires des souches et isolats du VAE ont été déterminées à l'aide d'un ensemble de programmes informatiques du système GCG utilisés par la plupart des chercheurs faisant de telles études (voir article 2). D'après ces analyses, cinq patrons

particuliers de structures secondaires ont été obtenus pour l'ensemble des 11 souches et isolats. Les patrons 1 à 5 regroupent 4, 3, 2, 1 et 1 souches/isolats, respectivement. Tous ces patrons ont en commun la présence de trois boucles d'épingle (A, B et C) constituées par les mêmes positions de nt. Cependant, ces patrons ont été séparés en deux groupes principaux I (patrons 1 et 2) et II (patrons 3 à 5) selon la présence ou l'absence d'une quatrième boucle d'épingle (D), respectivement. Les bases constituant ces structures ont été déterminées. Bien que certaines positions correspondant à des sites de mutation, la forme de la structure n'était nullement affectée pour les 4 boucles d'épingle A à D (Fig. 3 de l'article 2) suggérant une stabilité de ces boucles. Le fait que la souche Vienne ait formé un patron distinct (patron 5) confirme l'observation de sa divergence à l'échelle de la structure primaire mentionnée plus haut.

Le degré de stabilité de ces structures secondaires n'a pas été vérifié par des tests biochimiques faisant appel à diverses ribonucléases. Il faut admettre que la majorité des chercheurs qui ont étudié des structures secondaires par des tests biologiques ont témoigné de la fidélité moyenne des programmes informatisés utilisés (Konings et Gutell, 1995). La seule indication sur laquelle on peut prédire un certain degré de stabilité des structures secondaires est la valeur ΔG qui exprime l'énergie minimale libre responsable de l'appariement de certaines bases appartenant à un même brin d'ARN. Ces valeurs sont calculées à l'aide de formules intégrées dans les programmes utilisés pour déterminer ces structures secondaires et tiennent compte de la somme de l'énergie libre de chacune des bases de la structure en question. Lors des travaux de Hwang et Brinton (1998), les structures secondaires de la séquence antileader du VFSH, du VLD (souche C) et du VAE ont été comparées sur le plan de leur interactions avec des protéines cellulaires. Curieusement, bien que la SL du VAE étudiée par ce groupe représentait le brin négatif de la SL de la souche Utrecht retrouvé dans GenBank, certains points de similitude ont été notés par rapport à la structure secondaire prédite dans notre étude pour la souche de référence Bucyrus (Fig. 2 et 3 de l'article 2). Premièrement, cinq boucles d'épingle forment principalement la structure secondaire complète. Deuxièmement, les boucles sont sensiblement formées par les mêmes positions de bases alors que très peu de différences sont observées au niveau du nombre de bases appariées formant la tige de chaque boucle d'épingle.

Récemment, Stadejek et al. (1999) ont effectué l'étude comparative impliquant un grand nombre de souches/isolats de VAE obtenus à partir d'une vaste distribution

géographique et isolés à travers une période de 40 ans. Le nombre total de virus analysés était de 67 avec 22 virus dont les séquences provenaient de GenBank. Les parties étudiées étaient des fragments des gènes de la glycoprotéine G_L et de la réplicase. Les arbres phylogénétiques obtenus avec le gène de la G_L ont révélé 2 grands groupes (I et II) et 5 sous-groupes. Le premier grand groupe consistait principalement en des virus provenant d'Amérique du Nord, alors que le second consistait en ceux provenant essentiellement d'Europe. Les résultats de cette étude ont confirmé la valeur du gène de la G_L dans l'estimation de la diversité génétique pour le VAE, particulièrement au niveau des sites de neutralisation rapportés par Balasuriya et al. (1997). Ces derniers avaient identifié quatre sites distincts de neutralisation. Des substitutions en AA à leur niveau pouvaient altérer le phénotype de neutralisation de certaines souches de VAE (Balasuriya et al., 1997). Ceci a été reconfirmé par Hedges et al. (1999b) qui ont même trouvé que durant l'infection persistante au VAE, des pressions sélectives s'exerçaient sur la région variable V1 du gène G_L ainsi que sur le gène de la GP3. Ce phénomène pourrait donc être à l'origine de l'émergence de nouveaux variants viraux avec quelques propriétés de neutralisation différentes (Hedges et al. 1999b). L'analyse comparative des séquences des six ORFs 3' terminaux (ORF2 à 7) de plusieurs virus isolés à partir du sperme de deux chevaux infectés durant l'épizootie de 1984 dans l'état du Kentucky a été réalisée pour en étudier l'évolution génétique des nouveaux virus isolés plus tard à des périodes différentes (Hedges et al. 1999b). Il a été démontré que des variants viraux de la population de VAE apparaissaient dans le tractus reproducteur avec un taux de variation estimé à 1 % par année et que l'hétérogénéité des quasi-espèces augmentait avec de longues périodes d'infection persistante. Cette évolution pourrait influencer des propriétés phénotypiques telles que la neutralisation et probablement la virulence (Hedges et al., 1999b; Patton et al., 1999).

La variation et l'évolution génétique peuvent être aussi expliquées par le phénomène de la recombinaison, un mécanisme bien connu chez les coronavirus (Chang et al., 1996; Brian et al., 1997) et chez un membre artéroviral, le VLD (Li et al., 1999). Ces quelques variations mineures rapportées dans le gène N avaient d'ailleurs été associées à un mécanisme possible de recombinaison chez le VAE (Chirnside et al., 1994).

Les structures secondaires de la SL du VAE devraient participer dans ces mouvements évolutifs à l'intérieur de l'espèce et leur caractérisation plus poussée devrait nous guider sur des éléments de première importance dans le cycle de la réplication du virus. Hwang et Brinton (1998) ont mis en évidence une séquence de 68 nt (positions 117-

184) au sein de la séquence antileader du VFHS permettant la fixation de protéines cellulaires. Ils ont également mis en évidence des régions similaires chez les autres membres artéroviraux. Dans leur rapport, ils ont même prédit que les structures secondaires étaient formées de six boucles d'épingle et ont identifié les boucles d'épingle 4 et 5 comme correspondant aux sites de fixation mentionnés plus haut (Hwang et Brinton, 1998). Les quatre protéines des cellules MA104 ayant des PM de 103, 86, 55, et 36 kDa étaient les mêmes à se fixer sur les séquences antileader du VFHS, du VAE et du VLD. Cependant, une cinquième protéine cellulaire additionnelle de 31 kDa a été observée chez le VAE.

Il faut souligner que les résultats de Hwang et Brinton, (1998) ont été attribués aux séquences primaires dont les structures secondaires ont été par la suite prédites. En réalité, le rôle de ces structures secondaires, indépendamment des structures primaires, n'a pas été étudié et des analyses conformationnelles plus minutieuses seraient le seul moyen décisif à ce sujet. Par ailleurs, ces données concernant la séquence antileader des artérovirus ont témoigné d'un rôle important de la SL dans la réplication ou la transcription, ce qui n'est pas nécessairement le cas pour la traduction qui n'implique normalement que de l'ARN brin positif. Dans notre travail, des études plus poussées sur les structures secondaires en effectuant des délétions des boucles d'épingle avec des combinaisons d'altérations de la boucle, de la tige, des deux ou de plusieurs ensembles, pourrait constituer un objectif intermédiaire et être une suite logique aux études de structure secondaire de la SL du VAE.

L'évolution des résultats obtenus au cours de la réalisation de ce projet nous a permis de poursuivre les analyses sur la SL du VAE à l'échelle primaire. Ceci était rendu impératif par le fait que l'addition des 17 nt à l'extrémité 5' du génome avait révélé la présence d'un codon d'initiation AUG (aux positions 14 à 16) ouvrant un cadre de lecture dont le codon de terminaison était situé à l'intérieur de la SL (aux positions 125 à 127). Cet ORF de 111 nt pouvait coder pour un peptide de 37 AA d'un poids approximatif de 4,6 kDa (Fig. 2 de l'article 1). Ce codon d'initiation est considéré comme étant du type faible puisque les bases avoisinantes lui confèrent théoriquement un contexte sous optimal pour l'initiation de la traduction (Kozak, 1987a). La traduction des protéines pourrait ainsi obéir au système du « leaky scanning » établi par Kozak (1989c, 1991b). Cela ajoute donc un élément additionnel dans l'organisation des structures primaires de tous les ARNs du VAE. Avec la mise en évidence de ce nouvel ORF IL, chaque protéine du VAE provient de l'expression du deuxième cistron plutôt que du premier situé à proximité de l'extrémité 5' de l'ARN sg en question.

L'existence et le rôle de pareils codons d'initiation et ORFs à l'intérieur de la SL, ou de la région 5' non codante ont largement été observés et analysés chez d'autres virus à ARN dotés du mécanisme d'initiation interne de la traduction par le biais de l'IRES tels les picornavirus (Pelletier et al., 1988; Slobodskaya et al., 1996), le virus de l'hépatite A (Tesar et al., 1992) ainsi que d'autres virus adoptant le système traditionnel de balayage à partir de l'extrémité 5' tel le virus du sarcome de Rous (Hackett et al., 1986, Moustakas et al. 1993a, 1993b, Donze et al., 1995). Chez les virus apparentés, deux publications rapportaient l'existence de petits ORFs IL chez le coronavirus murin (Hofmann et al., 1993b) et le coronavirus bovin (Chen et Baric, 1995). Cependant, leurs codons d'initiation n'existaient pas au début d'infection mais apparaissaient après mutation spontanée et après l'établissement de l'infection persistante. Dans notre cas, toutes les souches et isolats analysés avaient montré que quel que soit leur passage, la présence du codon d'initiation IL et du peptide IL (excepté la souche NY86-A1 qui avait un ORF IL plus long) témoignait d'une grande stabilité de l'ORF IL du VAE (Fig.1 de l'article 2). Dans le cas des autres membres des artérovirus, les SL du VSRRP (Meulenberg et al., 1998) et du VFHS (Zeng et al., 1995) montrent aussi la présence de petits ORFs (aux positions 1 ou 4 à 38 équivalant à 13 ou 12 AA et aux positions 35 à 73 équivalant à 13 AA, respectivement) mais aucune étude analytique ne leur a été consacrée. D'où se dégageait l'importance d'entamer des études sur l'expression de l'ORF IL du VAE et de son rôle spécifique ou général dans le cycle de la réplication virale.

À l'aide des techniques de transcription et de traduction *in vitro*, l'expression du peptide IL du VAE dans un système acellulaire fut détectée (Fig. 2 de l'article 3). Cela représente la première démonstration et preuve expérimentale de la possibilité d'expression de peptide de la SL chez les nidovirus. En effet, aucune expression des ORFs IL émergeant lors d'infection persistantes chez les coronavirus (Hofmann et al, 1993b; Chen et Baric, 1995) n'avait été détectée ni *in vitro* ni *in vivo*.

Toutes nos tentatives de détection d'une expression du peptide IL *in vivo* dans les cellules infectées par le VAE ont été vaines. Par contre, ceci ne signifie pas définitivement sa non expression naturelle à notre égard. Plusieurs facteurs pourraient influencer cette hypothèse: (i) la traduction de quantités très infimes indétectables par les méthodes actuelles et classiques utilisées; (ii) l'absence d'éléments cellulaires et/ou viraux actifs dans le système de traduction *in vitro* acellulaire pourrait permettre l'expression de ce peptide, alors que leur présence dans des cellules infectées pourrait s'y lier et empêcher sa détection;

(iii) il n'est pas impossible qu'une faible quantité de peptide ayant une courte demi-vie soit traduite pour être ensuite rapidement dégradée après avoir rempli sa fonction requise. Ce dernier cas a été d'ailleurs observé chez le virus SV40 chez lequel un peptide IL de 61 AA et ayant la fonction d'une protéine de fixation sur l'ADN s'accumulait tard durant l'infection et avait une demi-vie d'environ 2 heures (Jay et al., 1981; Mertz et al., 1983).

L'expression *in vitro* du peptide IL du VAE a été démontrée par l'introduction dans le plasmide recombinant de mutations ponctuelles autour du codon d'initiation lui conférant des contextes différents. Ainsi, les contextes sous optimal (wild-type), optimal, et non favorable (à un degré plus faible) ont permis cette expression. Seule la construction dans laquelle le codon IL a été abrogé n'a pas permis la détection du peptide (Fig. 2 de l'article 3).

L'expression du peptide IL a ensuite été confirmée par les tests de radioimmunoprécipitation (Fig. 2 de l'article 3). Lors de la fabrication d'un antisérum de lapin anti-peptide IL de VAE, trois formes antigéniques synthétiques ont été préparées soit une forme conjuguée peptide(37 AA)-BSA (peptide-albumine sérique bovine), une forme conjuguée peptide-KLH (peptide-Keyhole Limpet Hemocyanin) et une forme non conjuguée (Fig. 2 de l'article 1). Des titres élevés allant jusqu'à $10^{4,5}$ aux tests d'immunoblots avaient été obtenus même avec l'antisérum anti-peptide non conjugué (résultats non montrés), de sorte que ce dernier fut utilisé pour les expériences de radioimmunoprécipitation.

L'expression *in vitro* du peptide intraleader ayant été prouvée dans des constructions contenant l'ADNc provenant de l'ARN sg 7 renfermant la séquence de la protéine de la nucléocapside, nous pouvions par conséquent conclure que nous étions en présence d'un système bicistronique dans lequel les deux cistrons sont fonctionnels (du moins *in vitro*). Le mécanisme régissant leur traduction ne pouvait être autre que le « leaky scanning » décrit par Kozak (1989c, 1991b). Ce mécanisme a aussi été décrit chez plusieurs autres modèles de virus (Kozak 1991b) et plus récemment chez le coronavirus bovin chez lequel une protéine nommée I de 23 kDa a été démontrée comme étant exprimée *in vitro* par le même ARNm exprimant la protéine de la nucléocapside de 49 kDa (Senanayake et al., 1992; Senanayake et Brian 1997).

Ces constatations ont été confirmées par les mêmes expériences de traduction *in vitro* avec les mêmes constructions mentionnées en analysant les taux d'expression de la protéine de la nucléocapside. Ces taux ont doublé lorsque le codon IL a été abrogé ou dans

un contexte non favorable mais ont baissé lorsque le contexte a été optimisé (Fig. 3 de l'article 3). Par conséquent, il faut conclure que le codon IL wild-type qui est dans un contexte sous optimal exerce un effet d'atténuation de la synthèse de la nucléocapside *in vitro*. Ceci peut donc être extrapolé pour les autres protéines de VAE puisqu'elles sont exprimées par des ARNm ayant une organisation génétique similaire et possédant la même SL située en amont de chaque ORF.

Plusieurs recherches ont été menées pour établir les règles régissant l'expression des ARNm bi- ou polycistroniques.

(1) La présence de la coiffe est indispensable à l'extrémité 5' des ARNm sauf pour ceux pourvus de la structure de l'IRES (Kozak, 1991c).

L'existence de la coiffe du type 0 [(m⁷(5')pppG(5')G] chez le VAE n'a été rapportée que récemment avec l'établissement des clones infectieux (van Dinten et al., 1997; Glaser et al., 1999). Par contre, lorsque nous avons entamé nos recherches sur la détermination de l'extrémité 5' du génome du VAE, nous avons constaté l'existence de la coiffe étant donné que nous avons réussi à obtenir une ligation des extrémités 5' et 3' du génome viral à condition que celui-ci soit traité avec une enzyme (la TAP; tobacco acid pyrophosphatase) enlevant la coiffe (résultats non montrés).

(2) La présence du premier codon d'initiation dans un contexte sous-optimal favorise le « leaky scanning » (Kozak, 1989b).

(3) Plus le premier ORF du côté 5' est court, plus il y a possibilité de réinitiation au second codon d'initiation et expression du second ORF (Kozak, 1987b; Sedman et Mertz, 1988; Abastado et al., 1989). Sedman et al. (1990) ont rapporté que le peptide LP1 de l'ARNm 16S du virus SV40 avait 61 AA alors que Hackett et al. (1986) ont démontré que le peptide IL du rétrovirus aviaire (virus du sarcome de Rous) était exprimé *in vivo* malgré sa petite taille de seulement 7 AA.

(4) Dans ce dernier cas, plus la séquence entre les deux cistrons est longue (c'est à dire la distance entre le codon de terminaison du premier ORF et le codon d'initiation du second ORF), plus l'efficacité de réinitiation augmente (Kozak, 1987b). La distance minimale au-dessus de laquelle elle n'a plus d'effet a été établie à environ 79 bases (Kozak, 1987b). Or, dans notre cas, la distance séparant le codon de terminaison IL et le codon d'initiation de l'ORF7 est de 140 bases.

(5) La présence d'une structure secondaire stable en aval du codon d'initiation retarde la sous unité 40 S du ribosome durant son mouvement de balayage et augmente sa

reconnaissance du codon non favorable ou sous optimal (Kozak, 1990). La stabilité de la structure secondaire a été limitée à une énergie libre $\Delta G \Sigma -19$ kCal/J (Kozak, 1990).

(6) Cependant, la présence en amont d'une telle structure secondaire entre la coiffe et le codon d'initiation exerce l'effet inverse et rend la reconnaissance de ce codon difficile (Kozak, 1989a).

(7) Plus la distance entre le premier codon d'initiation et la coiffe est égale ou supérieure à 12 nt, plus il y a augmentation de l'efficacité d'initiation à son niveau, notamment en dépassant les 20 nt (Kozak, 1991a; 1991c). Dans notre cas, le codon d'initiation IL est situé à la 14^{ième} base.

Ainsi, les règles de Kozak semblent être parfaitement en accord avec notre démonstration de l'expression du peptide IL du VAE.

Si l'effet de l'ORF IL est d'atténuer les niveaux d'expression de la protéine de la nucléocapside *in vitro*, quel en serait l'effet *in vivo*? De plus, si cette protéine est exprimée à partir d'un transcrit d'ARNm contenant la SL tout comme les autres ARN sg du VAE, quel serait l'effet de cet ORF IL sur l'expression des autres protéines structurales et non structurales du VAE?

Pour répondre à ces questions, le même genre de mutations ciblant le codon d'initiation IL doit être introduit sur la SL de chaque ARN génomique et sg viraux. La meilleure façon est de profiter de la caractéristique de la provenance de la SL des ARN sg à partir de l'ARN génomique viral. Ainsi, l'utilisation du clone d'ADNc infectieux est l'outil de choix pour effectuer ces études. Celui-ci nous a été fourni gracieusement par l'équipe de laboratoire de P. J. M. Rottier de l'Institut de Virologie de l'Université Utrecht, Pays Bas. Ce clone d'ADNc infectieux a été le résultat de l'assemblage de trois fragments larges provenant de produits d'amplification de RT-PCR et de leur insertion dans le vecteur pBR322. La progéniture virale de ce clone infectieux a été authentifiée après avoir été comparée avec le virus parental sur le plan de la morphologie des plages d'effet cytopathique produites, des caractéristiques de croissance à différentes températures et des profils de synthèse des protéines et des ARN viraux (Glaser et al., 1999).

Étant donné que l'effet du codon d'initiation avec un contexte non favorable était similaire à celui du codon IL abrogé, et que l'effet du codon AUG IL avec un contexte optimal était semblable à celui du codon AUG IL avec un contexte de Kozak sur l'expression de la protéine N dans les expériences de traduction *in vitro*, seuls deux types de ces mutations ponctuelles ont été retenues et introduites sur la SL du clone infectieux :

un clone contenant un AUG IL abrogé (ko) et un clone avec un AUG IL dont le contexte d'initiation de la traduction est optimal. De plus, puisque les rendements des produits de transcription *in vitro* à partir de la version originale du clone infectieux contenant la séquence du promoteur SP6 étaient relativement faibles lors des expériences préliminaires, ce promoteur fut substitué par celui de la polymérase T7 lors de la mutagenèse du codon AUG IL par PCR. Cela veut dire que le clone infectieux contenant le codon AUG IL wild-type a été également manipulé pour contenir le promoteur de la T7 polymérase et pour servir de témoin de clone d'ADNc infectieux donnant naissance au virus recombinant témoin. C'est avec celui-ci que les analyses comparatives des propriétés phénotypiques des virus recombinants mutants ont été effectuées. Les premières expériences de transfection des cellules BHK21 ont montré que les deux clones mutants provoquaient un effet cytopathique beaucoup plus retardé dans le temps que celui du clone sauvage témoin. Ceci n'était pas en corrélation avec les effets de ces mutations ponctuelles sur la protéine N *in vitro*. Pourtant leur authenticité a bien été vérifiée par les analyses de séquençage.

Il faut rappeler que la manipulation des clones infectieux comporte des risques potentiels de mutations spontanées aléatoires pouvant survenir lors des manipulations génétiques elles-mêmes, lors des passages des plasmides durant les transformations bactériennes et lors de la production de grandes quantités de plasmides à travers les différentes étapes de clonage et de sous clonage (Boyer et Haenni, 1994). Pour s'assurer que les différences phénotypiques pouvant être observées avec les analyses comparatives subséquentes soient dues à ces mutations autour du codon AUG IL, le contexte wild-type a été restauré génétiquement à partir de ces clones mutés eux-mêmes.

L'estimation de la stabilité des mutations apportées au virus a été également une étape indispensable dans la compréhension des résultats. Selon les analyses de séquençage, effectuées à différents passages viraux, il y avait restauration spontanée du génotype original (AUG IL wild-type) vers le 5^{ème} (pour le mutant abrogé) ou le 10^{ème} passage (pour le mutant AUG IL optimal) (Fig. 4 de l'article 3).

Les analyses des courbes de croissance virale et des profils de synthèse des protéines M et N aux premiers niveaux de passage ont montré que les deux clones mutants avec des AUG IL abrogé et optimal se comportaient de façon presque identique et avaient une croissance plus retardée et moins efficace que les virus recombinants ayant un génotype de codon IL wild-type, c'est-à-dire le virus recombinant témoin et les virus recombinants dont le codon AUG IL wild-type avait été restauré (Fig. 5 de l'article 3 et

données non-montrées). Par ailleurs, les résultats des plages de lyse montraient deux phénotypes distincts ; des plages minuscules et des plages larges. Cependant, tous les virus recombinants portant le génotype AUG IL wild-type montraient des plages larges homogènes, alors que les virus mutants montraient des populations hétérogènes dont une partie prédominante correspondait à des plages nettement plus petites (Fig. 6 de l'article 3). En considérant l'ensemble des résultats, nous avons avancé l'hypothèse que le phénotype des plages minuscules était dû aux mutations introduites autour du codon AUG IL. La meilleure façon de vérifier cela était de récupérer de l'ARN de cellules infectées à partir des deux types de plages de lyse pour ensuite en déterminer l'extrémité 5' par séquençage. Le résultat était autant spectaculaire qu'attendu ; le phénotype des plages minuscules correspondait au génotype de l'AUG IL abrogé ou au génotype de l'AUG dans un contexte optimal alors que le génotype de l'AUG IL wild-type (dans un contexte sous optimal) correspondait au phénotype des plages larges ou naturelles observées avec des cellules infectées par la souche Bucyrus du VAE.

Les différentes propriétés observées avec le virus recombinant avec un génotype AUG IL restauré expérimentalement étaient similaires à celles du virus avec un génotype restauré spontanément et naturellement ainsi qu' à celles du virus « wild-type ». Ceci est bien important car ainsi, l'on peut écarter des mutations spontanées autres que celles ciblées et qui auraient pu être responsables du phénotype des plages minuscules. La restauration du génotype AUG IL « wild-type » semblait être un processus progressif et rapide chez les virus recombinants mutants, puisque des plages moyennes et larges ont été observées aux premiers passages viraux alors qu'aux 15^{èmes} passages, aucune petite plage n'a été observée. Cependant, le fait que, même aux premiers passages, des plages de lyse larges étaient présentes, il s'avère ainsi difficile de préciser à quel cycle de réplication virale cette restauration est survenue.

Finalement, les plages de lyse minuscules signifient en principe un effet cytopathique beaucoup plus faible, produisant alors beaucoup moins de virus infectieux que les plages de lyse larges. Cela signifie que les mutations introduites rendant l'AUG IL abrogé ou dans un contexte optimal pour l'initiation de la traduction ont été responsables d'un effet d'atténuation de la réplication du VAE et de la réduction de l'infectivité virale en culture cellulaire.

Récemment Molenkamp et al. (2000) ont rapporté une étude similaire sur le rôle de la SL du VAE en construisant à partir d'un clone infectieux (van Dinten et al., 1997), un

clone mutant dans lequel l'ORF IL est détruit. Selon cette étude, le virus recombinant qui y est généré a présenté un phénotype similaire au virus type sauvage en culture cellulaire car : i) il a été capable de se répliquer et à synthétiser des ARNm sg; ii) le surnageant de culture cellulaire récupéré après transfection a produit des titres similaires que le virus recombinant sauvage; iii) la mutation originale introduite a été vérifiée par séquençage sur des produits de RT-PCR provenant de trois différentes plages de lyse. Les auteurs ont conclu que le produit de cet ORF-IL n'est pas requis pour la réplication virale et par conséquent, ne semble jouer aucun rôle dans le cycle vital viral au moins en culture cellulaire.

L'étude de Molencamp et al. (2000) ne semble pas être en contradiction avec la nôtre pour les raisons suivantes : 1) le peptide produit de la traduction de l'ORF IL n'a pas fait l'objet de cette recherche pour déterminer son éventuelle expression in vitro ou en culture cellulaire; 2) la suppression de l'AUG IL menant à la non expression du présumé peptide IL, ne fut pas suivie de l'inhibition de la réplication virale. Selon notre analyse, ceci est dû à la restauration rapide du génotype sauvage. Molencamp et al. (2000) ont probablement analysé des plages de lyse dont le phénotype était similaire à celui du virus sauvage. D'ailleurs, selon nos travaux, même le surnageant récupéré juste après transfection est composé d'un mélange de virus recombinants portant les deux génotypes (muté et restauré). Ainsi, selon nos résultats, si la restauration spontanée ne survient pas et que le virus muté reste inchangé même après plusieurs passages viraux, on peut présumer que la réplication virale serait remarquablement affectée ou même inhibée.

Le génotype de l'AUG IL abrogé correspond à une mutation G16A alors que celui de l'AUG IL dans un contexte optimal correspond à une mutation T11A. Ce sont des mutations ponctuelles ne ciblant pas la même position mais deux positions séparées de 5 nt. De plus, l'observation de l'alignement des SL des 11 souches et isolats de VAE utilisées dans notre étude comparative (Fig. 1 de l'article 2) renforce l'idée de la nécessité de maintenir stable cette région.

Ainsi, il s'avère, d'après les résultats obtenus, que :

(1) les bases autour du codon AUG IL du VAE correspondant à la séquence consensus 5' UGUAUGG 3' sont indispensables à la réplication et l'infectivité optimale du VAE;

(2) Les virus mutés sont soumis à des pressions sélectives pour restaurer et maintenir cette séquence consensus.

Le mécanisme par lequel le VAE a pu réparer les mutations introduites dans la SL du génome sont jusqu'à présent inconnues. Par contre, il est notoire que les virus à ARN sont soumis à des variations génétiques à travers des mutations des nucléotides ou des AA, des recombinaisons moléculaires homologues ou non homologues, et/ou des réassortiments génomiques, dépendamment des familles de virus et du type de leur génome (Domingo et Holland, 1997). Dans notre cas, l'on peut écarter le phénomène de la recombinaison pour au moins deux raisons. (i) Lors des étapes de transcription *in vitro*, les plasmides d'ADNc des différents clones infectieux provenaient de clones bactériens purs dont les séquences ont été vérifiées au niveau de la région leader du VAE juste avant les étapes de transfection. (ii) Les transcrits d'ARN ont été utilisés séparément de sorte qu'une seule espèce d'ARN a été électrotransfectée dans les cellules permissives. Ceci s'oppose donc au principe de la recombinaison qui requiert la présence d'au moins deux espèces d'ARN viral (Cavanagh, 1997).

Les virus à ARN ont les taux de mutations les plus élevés connus qui sont de l'ordre de 10^{-5} à 10^{-3} pour un génome de 10 kb. Ceci veut dire qu'après la réplication d'un ARN génomique, sa progéniture peut contenir de 0,1 à 10 mutations dispersées aléatoirement dans la molécule (Drake 1993; Domingo et Holland, 1997). Néanmoins, il est invraisemblable qu'après réversion d'une base dans la SL de deux virus mutés à deux positions différentes, un même génotype du virus wild-type soit restauré par l'effet tout simplement du hasard. Il doit donc exister une pression naturelle quelconque pour réparer sélectivement les mutations ponctuelles introduites aux positions 11 et 16 de la SL du VAE.

Le mécanisme de cette restauration reste à déterminer. Cependant, en excluant les phénomènes de recombinaison et de mutations spontanées aléatoires, les hyper mutations au sein des ARNs cellulaires et viraux ont également souvent été expliquées par des mécanismes tels les réactions d'édition d'insertions ou de substitutions (Scott, 1995; 1997). Cette édition est exercée par une enzyme cellulaire : la dsRAD pour « double-stranded RNA adenosine deaminase » (Bass, 1997) encore appelée l'«ADAR» (Liu et al., 1998). Celle-ci convertit un résidu adénosine en un résidu inosine au lieu d'un résidu uridine, causant ainsi une transition U à C dans le brin négatif complémentaire de l'ARN. Cela aboutit alors au remplacement dans le nouveau brin positif d'ARN du résidu original A par un G tel qu'il est dans le génotype « wild-type ».

Récemment, en utilisant le clone infectieux du VAE de van Dinten et al. (1997), Snijder et al. (1998) ont introduit des mutations ponctuelles dans la région 5' de l'ORF2a

codant pour la nouvelle protéine E du VAE. Ces mutations ont consisté à abroger les codons d'initiation AUG de l'ORF2a et l'ORFF2b. Cela résultait en une infectivité virale très affaiblie. Cependant, après transfection cellulaire et quelques passages viraux successifs, l'infectivité virale a été retrouvée et l'analyse des régions mutées a révélé que les codons AUG des deux ORF ont été restaurés. Les auteurs ont associé cette réversion génétique rapide à l'implication de la dsRAD. Chez le coronavirus murin VHM, l'édition d'ARN a déjà été évoquée dans une molécule déficiente interférente au lieu de la recombinaison pour expliquer la restauration au type sauvage d'un long ORF de 88 kDa dont la fonction était inconnue mais indispensable à sa réplication (Kim et al., 1993).

Plusieurs autres exemples de virus à ARN ont été rapportés où ce mécanisme d'édition d'ARN a été évoqué. Chez le VIH type 1, une structure d'une boucle d'épingle dans la région TAR a été montrée comme étant le substrat de l'activité de cette enzyme dans le noyau des cellules infectées (Sharmeen, et al., 1991). Plus récemment, Bourara et al. (2000) ont démontré que des ARNms du VIH-1 dans les cellules infectées chroniquement sont le siège de mutations de type G-à-A et C-à-U. Ceci est en relation avec le contrôle de la traduction des protéines impliquées dans la régulation de l'expression et la survie des cellules infectées chroniquement.

Chez le virus de l'hépatite delta, deux isoformes de protéines principales s'expriment à partir d'un seul ORF de l'ARN antigénomique par un mécanisme d'édition d'ARN dont la suppression compromet totalement la viabilité virale (Polson et al., 1996; 1998). L'ADAR exerce son action en changeant le codon stop d'une protéine de 24 kDa en un codon tryptophane pour qu'une autre protéine dont le codon stop est situé plus loin soit exprimée à partir de l'ORF étendu (Greeve et al., 1994).

L'édition d'ARN est connue depuis un peu plus d'une décennie (Bass 1997). Elle semble être une étape importante dans la réplication virale et a été rapportée chez plusieurs autres virus tels le virus de la rougeole (Horikami et Moyer, 1991), le parainfluenza type 3 (Galinski, et al., 1993), le virus de la maladie de New Castle (Steward, et al., 1993), le morbillivirus (Haas et al. 1995), les paramyxovirus (Hausmann et al. 1996) et le filovirus Ebola, (Sanchez et al., 1996).

La restauration du génotype « wild-type » 5' UGUAUGG 3' à partir des virus recombinants mutants du VAE utilisés dans notre travail peut ainsi être expliquée elle aussi par le phénomène d'édition soit lors de la duplication de l'ARN, de la transcription ou de la formation des structures secondaires ou dans toute autre étape durant laquelle l'ARN viral

est provisoirement en situation double brin. Bien que ceci requiert des travaux supplémentaires pour le prouver expérimentalement, cette restauration est la preuve tangible de l'importance de cette région dans la réplication virale du VAE en général.

CHAPITRE VII

CONCLUSION

L'extrémité 5' de la séquence leader du VAE a été déterminée. La SL complète de la souche de référence du VAE est donc de 206 nt sans le motif de jonction 5' UCAAC 3' situé à son extrémité 3'. Suite à cette détermination, un codon initiateur AUG a été retrouvé aux positions 14 à 16. Cet AUG ouvre un cadre de lecture avec un codon de terminaison situé aux position 125 à 127 à l'intérieur de la SL, d'où le terme d'ORF intraleader dont la taille est de 111 nt. Le peptide prédit avec cet ORF a été estimé à 37 AA avec un PM d'environ 4,6 kDa.

L'analyse de la SL complète à l'échelle primaire sans le motif de jonction de 11 souches et isolats a révélé que la plupart des souches avaient la même taille de 206 nt à l'exception de deux isolats dont la première base est délétée (205 nt) ou une base supplémentaire est insérée (207 nt). Les substitutions entre les différentes souches et isolats furent considérées comme étant des variations mineures puisqu'elles étaient éparpillées le long de la SL et ne représentaient pas des motifs particuliers pouvant susciter de l'intérêt. De plus, les pourcentages d'identité entre elles étaient élevés et se situaient entre 92,2 à 98,5 %. Cependant, la souche Vienne de VAE qui fut la seule souche européenne à être utilisée dans cette étude comparative montrait le plus faible pourcentage d'identité par rapport aux autres souches qui étaient toutes Nord-Américaines.

Les structures secondaires de ces souches et isolats ont aussi été prédites. Cinq patrons génétiques particuliers ont été obtenus pour l'ensemble des 11 souches et isolats. Tous ces patrons ont en commun la présence de trois boucles d'épingle (A, B et C) constituées par les mêmes positions de nt. Cependant ces patrons ont été séparés en deux groupes principaux I et II selon la présence ou l'absence d'une quatrième boucle d'épingle (D). Bien que certaines positions des bases étaient assujetties à des mutations, la forme de la structure n'était pas affectée pour ces 4 boucles d'épingle. Le fait que la souche Vienne ait formé un patron distinct confirme l'observation de sa divergence à l'échelle de la structure primaire mentionnée plus haut. Des études plus poussées sur les structures secondaires en effectuant des délétions des boucles d'épingle avec des combinaisons d'altérations de la boucle, de la tige, des deux ou de plusieurs ensembles, devraient être poursuivies pour déterminer le rôle de ces structures et des facteurs cellulaires et ou viraux pouvant s'y associer.

La comparaison des séquences SL de ces différentes souches et isolats ont aussi démontré une grande stabilité de l'ORF IL. Par ailleurs, le codon initiateur AUG IL était présent dans un contexte sous optimal pour l'initiation de la traduction.

Des plasmides recombinants contenant l'ADNc correspondant à l'ARN 7 sous le contrôle du promoteur de la polymérase du phage T7 ont été développés avec succès et ont été préparés de manière à contenir des mutations ponctuelles autour du codon d'initiation IL. À l'aide des techniques de transcription et de traduction *in vitro*, nous avons réussi à détecter une expression du peptide IL du VAE dans un système acellulaire. Cela représente la première démonstration et preuve expérimentale de la possibilité d'expression de peptide de la SL chez les nidovirus. L'expression *in vitro* du peptide IL du VAE a été confirmée en créant des mutations ponctuelles autour du codon d'initiation lui conférant des contextes différents pour l'initiation de la traduction. Lorsque le codon AUG IL était abrogé, l'expression du peptide IL n'a pas été détectée. Par contre, lorsque le contexte du codon AUG IL était gardé sous optimal "wild-type" ou modifié en un contexte optimal et non favorable, l'expression du peptide IL a été détectée *in vitro*. Finalement, les taux de synthèse de la protéine N ont doublé lorsque le codon IL était abrogé ou placé dans un contexte non favorable mais ont baissé lorsque le contexte a été optimisé. Par conséquent, le codon IL qui est dans un contexte sous optimal (le type sauvage) exerce un effet d'atténuation de la synthèse de la protéine capsidale *in vitro*. Cet effet pourrait s'exercer avec les autres protéines artévirales, vu la présence de la même SL dans chaque ARNm sg. Il faut cependant noter que les séquences intergéniques (entre la SL et le cadre de lecture de chaque protéine) sont différentes d'un ARN sg à un autre. De plus, malgré la présence d'une même région non codante 3' sur tous les ARNm sg, les séquences situées entre le cadre de lecture de chaque protéine et cette région non codante 3' varient au niveau de leur taille. Des études expérimentales au niveau de ces deux régions pourraient également être menées pour possiblement expliquer les différences de molarité des différentes protéines structurales synthétisées lors de l'infection virale ainsi que leur rôle dans d'autres fonctions virales. De plus, cette expression *in vitro* a démontré que la traduction des protéines artévirales se déroule selon le mécanisme « leaky scanning » dans lequel un faible pourcentage des ribosomes initie la traduction au niveau du premier codon AUG quand il est dans un contexte sous optimal par le ribosome et que l'autre plus grand pourcentage initie la traduction au niveau du prochain codon AUG situé en aval de l'ARN sg.

Un clone d'ADNc infectieux de VAE développé par des collaborateurs hollandais nous a été fourni pour y introduire le même type de mutations ponctuelles autour du codon initiateur AUG IL du VAE. Des expériences de transfection à l'aide des techniques

d'électroporation ont été réussies et ont généré des virus VAE recombinants portant trois types de codon d'initiation IL : le type sauvage (dans un contexte sous optimal pour le virus non muté), le type abrogé (avec une mutation G16U) et le type optimal (avec une mutation U11A). Les expériences de transfection des cellules BHK21 ont montré que les deux clones mutants provoquaient un effet cytopathique beaucoup plus retardé que celui du clone sauvage témoin. Selon les analyses de séquençage effectuées à différents passages viraux, il y eu une restauration spontanée et rapide du génotype original (5' UGUAUGG 3' correspondant aux positions 11 à 16 de la SL) vers le 5ième (pour le mutant abrogé) ou le 10ième passage (pour le mutant avec un AUG IL optimal).

Les analyses des courbes de croissance virale et des profils de synthèse des protéines M et N aux premiers niveaux de passages ont montré que les deux clones mutants se comportaient de façon presque identique et avaient une croissance plus retardée et moins efficace que les virus recombinants ayant un génotype de codon IL « wild-type ». Par ailleurs, les résultats des plages de lyse ont montré deux phénotypes distincts : des plages minuscules et des plages larges. Cependant, tous les virus recombinants portant le génotype AUG IL wild-type montraient des plages larges homogènes, alors que les virus mutants montraient des populations hétérogènes dont la partie prédominante correspondait à des plages nettement plus petites. Le phénotype des plages minuscules correspondait au génotype de l'AUG IL abrogé ou au génotype de l'AUG dans un contexte optimal alors que le génotype de l'AUG IL wild-type (dans un contexte sous optimal) correspondait au phénotype des plages larges ou naturelles.

Les mutations G16U et U11A ont donc été responsables d'un effet d'atténuation de la réplication du VAE et de la réduction de l'infectivité virale en culture cellulaire. Ces effets sont possiblement dus au changement de la structure nucléotidique primaire autour du codon AUG IL. Il n'est cependant pas impossible non plus que ces mutations soient à l'origine de la destruction de la structure secondaire au niveau de ce codon d'initiation IL. Cette structure pourrait aussi être le site de liaison de facteurs cellulaires et ou viraux et qui seraient nécessaires à l'initiation de la traduction ou la transcription des ARNm du VAE. Des expériences de mutagenèse dirigées, abrogeant des structures primaires d'une ou de plusieurs bases en même temps, mais sans affecter la structure secondaire, pourraient être préconisées pour répondre à cette question.

Par conséquent, les bases autour du codon AUG IL du VAE correspondant à la séquence consensus 5' UGUAUGG 3' semblent être indispensables à la réplication et

l'infectivité optimale du VAE. De plus, des mutations à son niveau semblent être soumises à des pressions sélectives pour restaurer et maintenir cette séquence consensus naturelle. Cette restauration suggère l'importance de cette région dans la réplication virale du VAE.

CHAPITRE VIII

BIBLIOGRAPHIE

Abastado, J. P., P. F. Miller, B. M. Jackson, and A. G. Hinnebusch. 1991. Suppression of ribosomal reinitiation at upstream open reading frames in amino acid-starved cells forms the basis for GCN4 translational control. *Mol. Cell Biol.* 11: 486-496.

Aiyar, A., D. Cobrinik, Z. Ge, H. J. Kung, and J. Leis. 1992. Interaction between retroviral U5 RNA and the T psi C loop of the tRNA(Trp) primer is required for efficient initiation of reverse transcription. *J. Virol.* 66: 2464-2472.

Aiyar, A., Z. Ge, and J. Leis. 1994. A specific orientation of RNA secondary structures is required for initiation of reverse transcription. *J. Virol.* 68: 611-618.

Anderson, D. J., and G. Blobel. 1983. Immunoprecipitation of proteins from cell-free translations. *Methods Enzymol.* 96:111-120.

Archambault, D., G. Laganiere, S. Carman, and G. St-Laurent. 1997. Comparison of nucleic and amino acid sequences and phylogenetic analysis of open reading frames 3 and 4 of various equine arteritis virus isolates. *Vet. Res.* 28:505-16.

Archambault, D., G. Laganière, and G. St-Laurent. 1998. Genetic variation and phylogenetic analysis of open reading frames 3 and 4 of various equine arteritis virus isolates. *Adv. Exp. Med. Biol.* 440:813-819.

Baker, S. C., and M. M. Lai. 1990. An *in vitro* system for the leader-primed transcription of coronavirus mRNAs. *EMBO J.* 9: 4173-4179.

Balasuriya, U. B., P. V. Rossitto, C. D. DeMaula, and N. J. MacLachlan. 1993. A 29K envelope glycoprotein of equine arteritis virus expresses neutralization determinants recognized by murine monoclonal antibodies. *J. Gen. Virol.* 74: 2525-2529.

Balasuriya, U. B., P. J. Timoney, W. H. McCollum, and N. J. MacLachlan. 1995. Phylogenetic analysis of open reading frame 5 of field isolates of equine arteritis virus and identification of conserved and nonconserved regions in the GL envelope glycoprotein. *Virology* 214: 690-697.

Balasuriya, U. B., J. F. Evermann, J. F. Hedges, A. J. McKeirnan, J. Q. Mitten, J. C. Beyer, W. H. McCollum, P. J. Timoney, and N. J. MacLachlan. 1998. Serologic and molecular characterization of an abortigenic strain of equine arteritis virus isolated from infective frozen semen and an aborted equine fetus. *J. Am. Vet. Med. Assoc.* 213: 1586-1589.

Balasuriya, U. B., H. W. Heidner, J. F. Hedges, J. C. Williams, N. L. Davis, R. E. Johnston, and N. J. MacLachlan. 2000. Expression of the Two Major Envelope Proteins of Equine Arteritis Virus as a Heterodimer Is Necessary for Induction of Neutralizing Antibodies in Mice Immunized with Recombinant Venezuelan Equine Encephalitis Virus Replicon Particles. *J. Virol.* 74: 10623-10630.

Bandyopadhyay, P. K., A. Pritchard, K. Jensen, and H. L. Lipton. 1993. A three-nucleotide insertion in the H stem-loop of the 5' untranslated region of Theiler's virus attenuates neurovirulence. *J. Virol.* 67: 3691-3695.

Baric, R. S., S. A. Stohlman, and M. M. Lai. 1983. Characterization of replicative intermediate RNA of mouse hepatitis virus: presence of leader RNA sequences on nascent chains. *J. Virol.* 48: 633-640.

Baric, R. S., S. A. Stohlman, M. K. Razavi, and M. M. Lai. 1985. Characterization of leader-related small RNAs in coronavirus-infected cells: further evidence for leader-primed mechanism of transcription. *Virus Res.* 3: 19-33.

Baric, R. S., C. K. Shieh, S. A. Stohlman, and M. M. Lai. 1987. Analysis of intracellular small RNAs of mouse hepatitis virus: evidence for discontinuous transcription. *Virology* 156: 342-354.

Bergman, A. M. 1913. Beiträge zur kenntnis der virusträger bei Rotlaufseuche, Influenza erysipelatosi, des Pferdes. *Z. Infektionskrankh.* 13 : 161-174.

Blair, E. D., C. C. Blair, and E. K. Wagner. 1987. Herpes simplex virus virion stimulatory protein mRNA leader contains sequence elements which increase both virus-induced transcription and mRNA stability. *J. Virol.* 61: 2499-2508.

Bourara, K., S. Litvak, and A. Araya. 2000. Generation of G-to-A and C-to-U Changes in HIV-1 transcripts by RNA Editing. *Science* 289: 1564-1566.

Breese, Jr, S. S., and W. H. McCollum. 1970. Electron-microscope characterization of equine arteritis virus. Pp 133-139. Dans J. T. Bryans, et H. Gerber (ed.) *Proceedings of the Second International Conference of Equine Infectious Diseases, Paris 1969*, S. Karger, Basel.

Brian, D. A., and W. J. M. Spaan. 1997. Recombination and coronavirus defective interfering RNAs. *Semin. Virol.* 8 : 101-111.

Brillon, A. Fontaine M., et Morallion R. 1967. Fièvre typhoïde du cheval. Artérite à virus. *Recueil Méd. Vét. Ecole Alfort.* 143: 17-27.

Brinton, M. A., E. I. Gavin, and A. V. Fernandez. 1986. Genetic variation among strains of lactate dehydrogenase-elevating virus (LDV). *J. Gen. Virol.* 67 : 2673-2684.

Brinton-Darnell M., and Plagemann P. G. 1975. Structure and chemical-physical characteristics of lactate dehydrogenase-elevating virus and its RNA. *J. Virol.* 16: 420-433.

Bryans, J. T., E. R. Doll, M. E. W. Crowe, and W. H. McCollum. 1957. The blood picture and thermal reaction in experimental viral arteritis of horses. *Cornell Vet.* 47:42-52.

Bürki, F. 1966. Further properties of equine arteritis virus. *Arch. Gesamte Virusforsch.* 19: 123-129.

Bürki, F. 1970. The virology of equine arteritis virus. Pp. 125-129. Dans J. T. Bryans, et H. Gerber (ed.) *Proceedings of the Second International Conference of Equine Infectious Diseases, Paris 1969*, S. Karger, Basel.

Bürki, F., H. Aspöck , and C. Kunz. 1972. Failure to propagate equine arteritis virus in an Aedine and Anopheline mosquito species. *Zentr. Bacteriol. Mikrobiol. Hyg. Abt. Orig.* 219: 109-111.

Bürki, F., A. Hofer, and N. Nowotny. 1992. Objective data plead to suspend import-bans for seroreactors against equine arteritis virus except for breeder stallions. *J. Appl. Anim. Res.* 1: 31-42.

Cao, J., and A. P. Geballe. 1995. Translational inhibition by a human cytomegalovirus upstream open reading frame despite inefficient utilization of its AUG cod. *J. Virol.* 69:1030-1036.

Carman, S., C. Rae, and E. Dubovi. 1988. Equine arteritis virus isolated from a standardbred foal with pneumonia. *Can. Vet. J.* 29: 937.

Cavanagh, D. 1997. Nidovirales: a new order comprising Coronaviridae and Arteriviridae. *Arch. Virol.* 142: 629-633

Cavanagh, D., and M. C. Horzinek. 1993. Genus Torovirus assigned to the Coronaviridae. *Arch. Virol.* 128: 395-396.

Chang, R. Y. R. Krishnan, and D. A. Brian. 1996. The UCUAAAC promoter motif is not required for high-frequency leader recombination in bovine coronavirus defective interfering RNA. *J. Virol.* 70: 2720-2729.

Chen, W., and R. S. Baric. 1995. Function of a 5'-end genomic RNA mutation that evolves during persistent mouse hepatitis virus infection in vitro. *J. Virol.* 69:7529-40.

Chen, Z, K. S. Faaberg, P. G. Plagemann. 1994. Detection of negative-stranded subgenomic RNAs but not of free leader in LDV-infected macrophages. *Virus Res.* 34: 167-177.

Chen, Z., K. S. Faaberg, and P. G. Plagemann. 1994. Determination of the 5' end of the lactate dehydrogenase-elevating virus genome by two independent approaches. *J. Gen. Virol.* 75: 925-930.

Chiou, H. C., C. Dabrowski, and J. C. Alwine. 1991. Simian virus 40 late mRNA leader sequences involved in augmenting mRNA accumulation via multiple mechanisms, including increased polyadenylation efficiency. *J. Virol.* 65: 6677-6685.

Chirnside, E. D., and W. J Spaan. 1990. Reverse transcription and cDNA amplification by the polymerase chain reaction of equine arteritis virus (EAV). *J. Virol. Methods* 30:133-40.

Chirnside E. D., C. M. Wearing, M. M. Binns, and J. A. Mumford. 1994. Comparison of M and N gene sequences distinguishes variation amongst equine arteritis virus isolates. *J. Gen. Virol.* 75: 1491-1497.

Chirnside, E. D., A. A. de Vries, J. A. Mumford, and P. J. Rottier. 1995a. Equine arteritis virus-neutralizing antibody in the horse is induced by a determinant on the large envelope glycoprotein GL. *J. Gen. Virol.* 76: 1989-1998.

Chirnside, E.D., P. M. Francis, A. A. de Vries, R. Sinclair, and J. A. Mumford. 1995b. Development and evaluation of an ELISA using recombinant fusion protein to detect the presence of host antibody to equine arteritis virus. *J. Virol. Methods.* 54: 1-13.

Chirnside E. D., P. M. Francis , and J. A. Mumford. 1995c. Expression cloning and antigenic analysis of the nucleocapsid protein of equine arteritis virus. *Virus Res.* 39: 277-288.

Clayton, H. 1987. 1986 outbreak of EAV in Alberta, Canada. *J. Equine Vet. Sci.* 7: 101.

Clark, I. 1892. Transmission of pink-eye from apparently healthy stallions to mares. *J. Comp. Pathol.* 5: 261-264.

Cobrinik, D., A. Aiyar, Z. Ge, M. Katzman, H. Huang, and J. Leis. 1991. Overlapping retrovirus U5 sequence elements are required for efficient integration and initiation of reverse transcription. *J. Virol.* 65: 3864-3872.

Coignoul, F. L., and N. F. Cheville. 1984. Pathology of maternal genital tract, placenta, and fetus in equine viral arteritis. *Vet. Pathol.* 21: 333-340

Cole, J. R., R. F. Hall, H. S. Gosser, J. B. Hendricks, A. R. Pursell, D. A. Senne, J. A. Pearson, and C. A. Gipson. 1986. Transmissibility and abortogenic effect of equine viral arteritis in mares. *J. Am. Vet. Med. Assoc.* 189: 769-771.

Collins, J. K., S. Kari, S. L. Ralston, D. G. Bennet, J. L. Traub-Dargatz, and A. O. McKinnon. 1987. Equine viral arteritis in a veterinary teaching hospital. *Prev. Vet. Med.* 4: 389-397.

Contag, C. H., E. F. Retzel, and P. G. W. Plagemann. 1986. Genomic differences between strains of lactate dehydrogenase-elevating virus. *Intervirology* 26: 228-233.

Conzelmann, K. K., N. Visser, P. Van Woensel, and H. J. Thiel. 1993. Molecular characterization of porcine reproductive and respiratory syndrome virus, a member of the arterivirus group. *Virology* 193: 329-339.

Cook, R. F., S. J. Gann, and J. A. Mumford. 1989. The effect of vaccination with tissue culture-derived vaccines on detection of antibodies to equine arteritis virus by enzyme-linked immunosorbent assay (ELISA). *Vet. Microbiol.* 20: 181-189.

Crawford, T. B., and W. C. Davis. 1970. Fluorescent and electron microscope studies of equine arteritis virus. *Fed. Proc.* 29: 286, abstr. 225.

Crawford, T. B., and J. B. Henson. 1973. Immunofluorescent, light microscopic and immunologic studies of equine arteritis virus, p. 282-302. Dans J. T. Bryans, et H. Gerber (ed.) *Proceedings of the Third International Conference of Equine Infectious Diseases, Paris 1972*, S. Karger, Basel.

Darnell, M. B., and P. G. Plagemann. 1972. Physical properties of lactic dehydrogenase-elevating virus and its ribonucleic acid. *J. Virol.* 10: 1082-1085.

den Boon, J. A., E. J. Snijder, E. D. Chirnside, A. A. F. de Vries, M. C. Horzine, and W. J. Spaan. 1991. Equine arteritis virus is not a togavirus but belongs to the coronaviruslike superfamily. *J Virol.* 65: 2910-2920.

den Boon, J. A., W. J. Spaan, and E. J. Snijder. 1995. Equine arteritis virus subgenomic RNA transcription: UV inactivation and translation inhibition studies. *Virology* 213: 364-372.

den Boon, J. A., M. F. Kleijnen, W. J. Spaan, and E. J. Snijder. 1996. Equine arteritis virus subgenomic mRNA synthesis: analysis of leader-body junctions and replicative-form RNAs. *J. Virol.* 70: 4291-4298.

Deregt, D., A. A. de Vries, M. J. Raamsman, L. D. Elmgren, and P. J. Rottier. 1994. Monoclonal antibodies to equine arteritis virus proteins identify the G_L protein as a target for virus neutralization. *J. Gen. Virol.* 75: 2439-2444.

de Vries, A. A. F., E. D. Chirnside, P. J. Bredenbeek, L. A. Gravestien, M. C. Horzinek, and W. J. M. Spaan. 1990. All subgenomic mRNAs of equine arteritis virus contain a common leader sequence. *Nucleic Acids Res.* 18:3241-3247.

de Vries, A. A. F., E. D. Chirnside, M. C. Horzinek, and P. J. M. Rottier. 1992. Structural proteins of equine arteritis virus. *J. Virol.* 66:6294-6303.

de Vries, A. A. F., E. D. Chirnside, M. C. Horzinek, and P. M. Rottier. 1993. Equine arteritis virus contains a unique set of four structural proteins. Pp. 245-253. Dans H. Laudeand J.-F. Vautherot (ed.). *Coronaviruses: molecular biology and virus-host interactions.* Adv. Exp. Med. Biol. , vol. 342. Plenum Press, New York.

de Vries A. F. A, S. M. Post, Raamsman M. J., M. C. Horzinek, and P. J. Rottier. 1995a. The two major envelope proteins of equine arteritis virus associate into disulfide-linked heterodimers. *J. Virol.* 69: 4668-4674.

de Vries, A. F. A., M. J. Raamsman, H. A. van Dijk, M. C. Horzinek, and P. J. M. Rottier. 1995b. The small envelope glycoprotein (Gs) of equine arteritis virus folds into three distinct monomers and a disulfide-linked dimer. *J. Virol.* 69: 3441-3448.

de Vries, A. F. A., M. C. Horzinek, P. J. M. Rottier, and R. J. de Groot. 1997. The genome organization of the nidovirales: Similarities and differences between Arteri- Toro-, and Coronaviruses. *Sem. Virol.* 8: 33-47.

Doll, E. R., J. T. Bryans, W. H. McCollum, and M. Crowe. 1957a. Isolation of a filterable agent causing arteritis of horses and abortion by mares. Its differentiation from the equine abortion (influenza) virus. *Cornell Vet.* 47 :3-41.

Doll, E. R., S. Knappenberger, and J. T. Bryans. 1957b. An outbreak of abortion caused by equine arteritis virus. *Cornell Vet.* 47 :69-75.

Doll, E. R., J. T. Bryans, and W. H. McCollum. 1968. Immunization against equine viral arteritis using modified live virus propagated in cell cultures of rabbit kidney. *Cornell Vet.* 58: 497-524.

Domingo, E., and J. J. Holland. 1997. RNA virus mutations and fitness for survival. *Annu. Rev. Microbiol.* 51: 151-178.

Donzé, O., and P. F. Spahr. 1992. Role of the open reading frames of Rous sarcoma virus leader RNA in translation and genome packaging. *EMBO J.* 11: 3747-3757.

Donzé, O., P. Damay, and P. F. Spahr. 1995. The first and third uORFs in RSV leader RNA are efficiently translated: implications for translational regulation and viral RNA packaging. *Nucleic Acids Res.* 23: 861-868.

Drake, J., W. 1993. Rates of spontaneous mutation among RNA viruses. *Proc. Natl. Acad. Sci. U.S.A.* 90: 4171-4175.

Elazhary, Y., C. Goulard, A. Vrins, et O. Fassi Fihri. 1990. Artérite virale équine au Québec: Étude sérologique. *Méd. Vét. Qué.* 20: 23-27.

Estes, P. C., and N. F. Cheville. 1970. The ultrastructure of vascular lesions in equine viral arteritis. *Am. J. Pathol.* 58: 235-253.

Faaberg, K. S., and P. G. Plagemann. 1997. ORF3 of lactate dehydrogenase-elevating virus encodes a soluble, nonstructural, highly glycosylated, and antigenic protein. *Virology* 227: 245-251.

Faaberg, K. S., M. R. Elam, C. J. Nelsen, and M. P. Murtaugh. 1998. Subgenomic RNA7 is transcribed with different leader-body junction sites in PRRSV (strain VR2332) infection of CL2621 cells. *Adv. Exp. Med. Biol.* 440: 275-279.

Fukunaga, Y., and W. H. McCollum. 1977. Complement fixation reactions in equine viral arteritis. *Am. J. Vet. Res.* 38: 2043-2046.

Fukunaga, Y., H. Imagawa, E. Tabuchi, and Y. Akiyama. 1981. Clinical and virological findings on experimental equine viral arteritis in horses. *Bull. Equine Res. Inst.* 18: 110-118.

Fukunaga, Y., R. Wada; K. Hirasawa; M. Kamada, T. Kumanomido; and Y. Akiyama. 1982. Effect of the modified Bucyrus strain of equine arteritis virus experimentally inoculated into horses. *Bull. Equine Res. Inst.* 19: 97-101.

Fukunaga, Y., R. Wada, T. Matsumura, T. Anzai, H. Imagawa, T. Sugiura, T., Kumanomido, T. Kanemaru, and M. Kamada. 1992. An attempt to protect persistent infection of equine viral arteritis in the reproductive tract of stallions using formalin inactivated virus vaccine, p. 239-244. Dans W. Plowright, P. D. Rossdale, and J. F. Wade (ed.). *Proceedings of the Sixth International Conference on Equine Infectious Diseases*, Cambridge 1991. R & W Publications, Newmarket, England.

Fukunaga, Y., H. Imagawa, T. Kanemaru, and M. Kamada. 1993. Complement-dependent serum neutralization with virulent and avirulent Bucyrus strains of equine arteritis virus. *Vet. Microbiol.* 36: 379-383.

Fukunaga, Y., T. Matsumura, T. Sugiura, R. Wada, H. Imagawa, T. Kanemaru, and M. Kamada. 1994. Use of the serum neutralisation test for equine viral arteritis with different virus strains. *Vet Rec.* 134 : 574-576.

Fütterer, J., K. Gordon, H. Sanfacon, J. M. Bonneville, and T. Hohn. 1990. Positive and negative control of translation by the leader sequence of cauliflower mosaic virus pregenomic 35S RNA. *EMBO J.* 6: 1697-1707.

Fütterer, J., Z. Kiss-Laszlo, and T. Hohn. 1993. Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* 73:789-802.

Galinski, M. S., R. M. Troy, and A. K. Banerjee. 1992. RNA editing in the phosphoprotein gene of the human parainfluenza virus type 3. *Virology* 186: 543-550.

Gerber, H., F. Steck, B. Hofer, L. Walther, and U. Friedly. 1978. Serological investigations on equine viral arteritis, Pp. 461-465. Dans J. Bryans and H. Gerber (ed.), *Proceedings of the Forth International Conference on Equine Infectious Diseases*, Lyon 1976. Veterinary Publications Inc., Princeton, N. J.

Gilbert S. A. , P. J. Timoney, W. H. McCollum, and D. Deregt. 1997. Detection of equine arteritis virus in the semen of carrier stallions by using a sensitive nested PCR assay. *J. Clin. Microbiol.* 35 :2181-3.

Glaser, A. L., A. A. F. de Vries, M. J. B. Raamsman, M. C. Horzinek, and P. J. M. Rottier. 1999. An infectious cDNA clone of equine arteritis virus: a tool of future fundamental studies and vaccine development. *In* U. Wernery, J. F. Wade, J. A. Mumford, and O.- R. Kaaden (ed), *Proceedings of the Eighth International Conference on Equine Infectious Diseases*, Dubai 1998, p.166-176. R & W Publications, Ltd., Newmarket, England.

Godeny, E. K., L. Chen L, S. N. Kumar, S. L. Methven, E. V. Koonin, and M. A. Brinton. 1993. Complete genomic sequence and phylogenetic analysis of the lactate dehydrogenase-elevating virus (LDV). *Virology* 194: 585-596.

Godeny, E. K., L. Zeng, S. L. Smith, and M. A. Brinton. 1995. Molecular characterization of the 3' terminus of the simian hemorrhagic fever virus genome. *J. Virol.* 69: 2679-2683.

Godeny, E. K., A. A. de Vries, X. C. Wang, S. L. Smith, and R. J. de Groot. 1998. Identification of the leader-body junctions for the viral subgenomic mRNAs and organization of the simian hemorrhagic fever virus genome: evidence for gene duplication during arterivirus evolution. *J. Virol.* 72: 862-867.

Golnik, W., A. Moraillon, and J. Golnik. 1986. Identification and antigenic comparison of equine arteritis virus isolated from an outbreak of epidemic abortion of mares. *J. Vet. Med.* 33: 413-417.

Greeve, J., D. Hartwig, E. Windler, and H. Greten. 1994. Requirements for editing in the genomic RNA of hepatitis delta virus. *Biochimie.* 76: 1209-1216.

Haas, L., M. D. Baron, B. Liess, and T Barrett. 1995. Editing of morbillivirus P gene transcripts in infected animals. *Vet. Microbiol.* 44: 299-306.

Hackett, P. B., R. B. Petersen, C. H. Hensel, F. Albericio, S. I. Gunderson, A. C. Palmenberg, and G. Barany. 1986. Synthesis *in vitro* of a seven amino acid peptide encoded in the leader RNA of Rous sarcoma virus. *J. Mol. Biol.* 190: 45-57.

Harlow, E., and D. Lane. 1988. *Antibodies: a laboratory manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y.

Harry, T. O., and W. H. McCollum. 1981. Stability of viability and immunizing potency of lyophilized, modified equine arteritis live virus vaccine. *Am. J. Vet. Res.* 42: 1501-1505.

Hausmann, S, J. P. Jacques, and D. Kolakofsky. 1996. Paramyxovirus RNA editing and the requirement for hexamer genome length. *RNA* 2: 1033-1045.

Hay, N., and Y. Aloni. 1985. Attenuation of late simian virus 40 mRNA synthesis is enhanced by the agnoprotein and is temporally regulated in isolated nuclear systems. *Mol. Cell Biol.* 5:327-1334.

Hedges, J. F., U. B. Balasuriya, P. J. Timoney, W. H. McCollum, N. J. MacLachlan. 1996. Genetic variation in open reading frame 2 of field isolates and laboratory strains of equine arteritis virus. *Virus Res.* 42: 41-52.

Hedges, J. F., U. B. Balasuriya, S. Ahmad, P. J. Timoney, W.H. McCollum, T. Yilma, and N. J. MacLachlan. 1998. Detection of antibodies to equine arteritis virus by enzyme linked immunosorbant assays utilizing G(L), M and N proteins expressed from recombinant baculoviruses. *J. Virol. Methods.* 76:127-137.

Hedges, J. F., U. B. Balasuriya, and N. J. MacLachlan. 1999a. The open reading frame 3 of equine arteritis virus encodes an immunogenic glycosylated, integral membrane protein. *Virology* 264. 92-98.

Hedges, J. F., U.B. Balasuriya, P. J. Timoney, W. H. McCollum, and N. J. MacLachlan. 1999b. Genetic divergence with emergence of novel phenotypic variants of equine arteritis virus during persistent infection of stallions. *J. Virol.* 73: 3672-3681.

Hierholzer, J. and R. Killington. 1996. Quantitation of virus; plaque assay. *In* B. W. Mahy and H. O. Kangro (*ed*). *Virology Methods Manual*. pp. 38-39. Academic Press, London:

Hirzmann, J., D. Luo, J. Hahnen, and G. Hobom. 1993. Determination of messenger RNA 5'-ends by reverse transcription of the cap structure. *Nucleic Acids Res.* 21: 3597-3598.

Hiscox, J. D. H. Pocock, and P. Britton. 1993. Control of TGEV mRNA transcription. *Adv. Exp. Med. Biol.* 342: 99-104.

Hiscox, J. A., K. L. Mawditt, D. Cavanagh, and P. Britton. 1995. Investigation of the control of coronavirus subgenomic mRNA transcription by using T7-generated negative-sense RNA transcripts. *J. Virol.* 69: 6219-6227.

Hofmann, M. A., R. Y. Chang, S. Ku, and D. A. Brian. 1993a. Leader-mRNA junction sequences are unique for each subgenomic mRNA species in the bovine coronavirus and remain so throughout persistent infection. *Virology* 196: 163-171.

Hofmann, M. A., S. D. Senanayake, and D. A. Brian. 1993b. A translation-attenuating intraleader open reading frame is selected on coronavirus mRNAs during persistent infection. *Proc. Natl. Acad. Sci. U.S.A.* 90: 11733-11737.

Holyoak, G. R., T. V Little, W. H. McCollum, and P. J. Timoney. 1993. Relationship between onset of puberty and establishment of persistent infection with equine arteritis virus in the experimentally infected colts. *J. Comp. Pathol.* 109:29-46.

Horikami, S. M., and S. A. Moyer. 1991. Synthesis of leader RNA and editing of the P mRNA during transcription by purified measles virus. *J. Virol.* 65:5342-5347.

Horzinek, M., J. Maess, and R. Laufs. 1971. Studies on the substructure of togaviruses. II Analysis of equine arteritis, rubella, bovine viral diarrhea, and hog cholera viruses. *Arch. Gesamte Virusforsch.* 33: 306-318.

Hou-Jong, M.-H. J., S. H. Larsen, and A. Roman. 1987. Role of the agnoprotein in regulation of simian virus 40 replication and maturation pathways. *J. Virol.* 61:937-939.

Huntington, P. J., P. M. Ellis, A. J. Forman, and P. J. Timoney. 1990. Equine viral arteritis. *Aust. Vet. J.* 67: 429-431.

Hwang, Y. K., and M. A. Brinton. 1998. A 68-nucleotide sequence within the 3' noncoding region of simian hemorrhagic fever virus negative-strand RNA binds to four MA104 cell proteins. *J. Virol.* 72: 4341-51.

- Hyllseth, B. 1969. A plaque assay of equine arteritis virus in BHK-21 cells. Arch. Gesamte Virusforsch. 28: 26-33.
- Hyllseth, B. 1970. Buoyant density studies on equine arteritis virus. Arch. Gesamte Virusforsch. 30: 97-104.
- Hyllseth, B. 1973. Structural proteins of equine arteritis virus. Arch. Gesamte Virusforsch. 40: 177-188.
- Hyllseth, B., U Pettersson. 1970. Neutralization of equine arteritis virus: enhancing effect of guinea pig serum. Arch. Gesamte Virusforsch. 32: 337-347.
- Inoue, T., R. Yanagawa, and M. Shinagawa. 1975. Immunofluorescent studies on the multiplication of equine arteritis virus in Vero and equine dermis (NBL-6) cells. Jpn. J. Vet. Sci. 37: 569-575.
- Jaksch, W., M. Sibalin, E. Taussig, L. Pichler, and F. Bürki. 1973. Natural cases and experimental transmissions of equine-virus-arteritis in Austria. Dtsch. Tieraertzl. Wochenschr. 80: 374-380.
- Jameson, B. A., and H. Wolf. 1988. The antigenic index: a novel algorithm for predicting antigenic determinants. Comput. Appl. Biosci. 4:181-186.
- Jang, S. K., M. V. Davies, R. J. Kaufman, and E. Wimmer. 1989. Initiation of protein synthesis by internal entry of ribosomes into the 5' nontranslated region of encephalomyocarditis virus RNA *in vivo*. J. Virol. 63: 1651-1660.
- Jang, S. K., T. V. Pestova, C. U. Hellen, G. W. Witherell, and E. Wimmer. 1990. Cap-independent translation of picornavirus RNAs: structure and function of the internal ribosomal entry site. Enzyme 44: 292-309.
- Jay, G., S. Nomura, C. W. Anderson, and G. Khoury. 1981. Identification of the SV40 agnogene product: a DNA binding protein. Nature 291: 346-349.

Jeong, Y. S., and S. Makino. 1994. Evidence for coronavirus discontinuous transcription. *J Virol.* 68:2615-2623.

Jones, T. C. 1969. Clinical and pathologic features of equine viral arteritis. *J. Am. Vet. Med. Ass.* 155: 315-317.

Jones, T. C., Doll, E. R., and J. T. Bryans. 1957. The lesions of equine arteritis virus. *Cornell Vet.* 47 :52-68.

Kheyar, A., G. St-Laurent, and D. Archambault. 1996. Sequence determination of the extreme 5' end of the equine arteritis virus leader region. *Virus Genes* 12:291-295.

Kheyar, A., Martin S., G. St -Laurent, P. J. Timoney, and W. H. McCollum, and D. Archambault. 1997. Expression cloning and humoral immune response to the nucleocapsid and membrane proteins of equine arteritis virus. *Clin. Diagn. Lab. Immunol.* 6 :648-652

Kheyar, A., G. St-Laurent, M. Diouri, and D. Archambault. 1998. Nucleotide sequence and genetic analysis of the leader region of Canadian, American and European equine arteritis virus isolates. *Can. J. Vet. Res.* 62:224-230

Kim, Y. N., M. M. Lai, and S. Makino. 1993. Generation and selection of coronavirus defective interfering RNA with large open reading frame by RNA recombination and possible editing. *Virology* 194: 244-253.

Kinney, R. M., G.-J. Chang, K. R. Tsuchiya, J. M. Sneider, J. T. Roehrig, T. M. Woodward, and D. W. Trent. 1993. Attenuation of Venezuelan equine encephalitis virus strain TC-83 is encoded by the 5'-noncoding region and the E2 envelope glycoprotein. *J. Virol.* 67:1269-1277

Klavano, G. G. 1987. Equine viral arteritis. *Can. Vet. J.* 28: 281.

Kobiler, D., C. M. Rice, C. Brodie, A. Shahar, J. Dubuisson, M. Halevy, and S. Lustig. 1999. A single nucleotide change in the 5' noncoding region of Sindbis virus confers neurovirulence in rats. *J. Virol.* 73:10440-10446.

Kondo, T., Y. Fukunaga, K. Sekiguchi, T. Sugiura, and H. Imagawa. 1998. Enzyme-linked immunosorbent assay for serological survey of equine arteritis virus in racehorses. *J. Vet. Med. Sci.* 60 :1043-1045.

Konings, D. A., and R. R. Gutell. 1995. A comparison of thermodynamic foldings with comparatively derived structures of 16S and 16S-like rRNAs. *RNA* 1: 559-574.

Konishi, S., H. Akashi, A. Sentsui, and M. Ogata. 1975. Characterization of the virus and trial survey on antibody with Vero cell culture. *Jpn. J. Vet. Sci.* 37: 259-267.

Kozak, M. 1986. Regulation of protein synthesis in virus-infected animal cells. *Adv. Virus Res.* 31: 229-292.

Kozak, M. 1987a. An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. *Nucleic Acids Res.* 15: 8125-8148.

Kozak, M. 1987b. Effects of intercistronic length on the efficiency of reinitiation by eukaryotic ribosomes. *Mol. Cell Biol.* 7: 3438-3445.

Kozak, M. 1988. Leader length and secondary structure modulate mRNA function under conditions of stress. *Mol. Cell Biol.* 8: 2737-2744.

Kozak, M. 1989a. Circumstances and mechanisms of inhibition of translation by secondary structure in eukaryotic mRNAs. *Mol. Cell Biol.* 9: 5134-5142.

Kozak, M. 1989b. Context effects and inefficient initiation at non-AUG codons in eukaryotic cell-free translation systems. *Mol. Cell Biol.* 9: 5073-5080.

Kozak, M. 1989c: The scanning model for translation: an update. *J. Cell Biol.* 108. 229-241.

Kozak, M. 1990. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* 87: 8301-8305.

- Kozak, M. 1991a. A short leader sequence impairs the fidelity of initiation by eukaryotic ribosomes. *Gene Expr.* 1: 111-115.
- Kozak, M. 1991b. Effects of long 5' leader sequences on initiation by eukaryotic ribosomes *in vitro*. *Gene Expr.* 1: 117-125.
- Kozak, M. 1991c. An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* 115:887-903.
- Kozak, M. 1994. Features in the 5' non-coding sequences of rabbit alpha and beta-globin mRNAs that affect translational efficiency. *J. Mol. Biol.* 235: 95-110.
- Kozak, M. 1999. Initiation of translation in prokaryotes and eukaryotes. *Gene* 234:187-208.
- Kuo, L. L., J. T. Harty, L. Erickson, G. A. Palmer, and P. G. Plagemann. 1991. A nested set of eight RNAs is formed in macrophages infected with lactate dehydrogenase-elevating virus. *J. Virol.* 65: 5118-5123.
- Kuo, L., Z. Chen, R. R. Rowland, K. S. Faaberg, and P. G. Plagemann. 1992. Lactate dehydrogenase-elevating virus (LDV): subgenomic mRNAs, mRNA leader and comparison of 3'-terminal sequences of two LDV isolates. *Virus Res.* 23: 55-72.
- Lai, M. M. 1986. Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing. *Bioessays.* 5: 257-260.
- Lai, M. M. 1990. Coronavirus: organization, replication and expression of genome. *Annu. Rev. Microbiol.* 44: 303-333
- Lai, M. M., R. S. Baric, P. R. Brayton, and S. A. Stohlman. 1984. Characterization of leader RNA sequences on the virion and mRNAs of mouse hepatitis virus, a cytoplasmic RNA virus. *Proc. Natl Acad. Sci. USA.* 81: 3626-3630.

Lai, M. M., and D. Cavanagh. 1997. The molecular biology of coronaviruses. *Adv. Virus Res.* 48: 1-100.

Lai, M. M., C. D. Patton, and S. A. Stohlman. 1982. Further characterization of mRNA's of mouse hepatitis virus: presence of common 5'-end nucleotides. *J. Virol.* 41: 557-565.

Lai, M. M., C. D. Patton, R. S. Baric, and S. A. Stohlman. 1983. Presence of leader sequences in the mRNA of mouse hepatitis virus. *J. Virol.* 46: 1027-1033.

Lepage, N., G. St-Laurent, S. Carman, and D. Archambault. 1996. Comparison of nucleic and amino acid sequences and phylogenetic analysis of the Gs protein of various equine arteritis virus isolates. *Virus Genes* 13: 87-91.

Lewin, B. 1994. Messenger RNA is the template. Pp 253-276. Dans Lewin B. (ed) *Genes*. 5th ed. Oxford University Press. New York. USA.

Li, K., Z. Chen, and P. G. Plagemann 1999. High-frequency homologous genetic recombination of an arterivirus, lactatedehydrogenase-elevating virus, in mice and evolution of neuropathogenic variants. *Virology*.258: 73-83.

Little, T. V., G. R. Holyoak, W. H. McCollum, and P. J. Timoney. 1991. Output of equine arteritis virus from persistently infected stallions is testosterone-dependent, Pp. 225. Dans W. Plowright, P. D. Rossdale, and J. F. Wade (ed.). *Proceedings of the Sixth International Conference on Equine Infectious Diseases*, Cambridge 1991. R & W Publications, Newmarket, England.

MacLachlan N. J., U. B. Balasuriya, J. F. Hedges, T. M. Schweidler, W. H. McCollum, P. J. Timoney, P. J. Hullinger, and J. F. Patton. 1998. Serologic response of horses to the structural proteins of equine arteritis virus. *J. Vet. Diagn. Invest.* 10: 229-236.

Maess, J., E. Reckzo, and H. O. Böhm. 1970. An investigation of the morphology of equine arteritis virus. Pp. 130-132. Dans J. T. Bryans, and H. Gerber (ed.) *Proceedings of the*

Second International Conference of Equine Infectious Diseases, Paris 1969, S. Karder, Basel.

Magnusson, P., B. Hyllseth, and H. Marusik. 1970. Morphological studies on equine arteritis virus. *Arch. Gesamte Virusforsch.* 30: 105-112.

Makino, S., J. G. Keck, S. A. Stohlman, and M. M. Lai. 1986a. High-frequency RNA recombination of murine coronaviruses. *J. Virol.* 57: 729-737.

Makino, S., S. A. Stohlman, and M. M. Lai. 1986b. Leader sequences of murine coronavirus mRNAs can be freely reassorted: evidence for the role of free leader RNA in transcription. *Proc. Natl. Acad. Sci. USA.* 83: 4204-4208.

Makino S., L. H. Soe, C. K. Shieh, and Lai M. M. 1988. Discontinuous transcription generates heterogeneity at the leader fusion sites of coronavirus mRNAs. *J. Virol.* 62: 3870-3873.

McCollum, W. H. 1969. Development of a modified virus strain and vaccine for equine viral arteritis. *Am. J. Vet. Res.* 155: 318-322.

McCollum, W. H. 1970. Vaccination for equine arteritis virus. Pp. 143-151. Dans J. T. Bryans, and H. Gerber (ed.) *Proceedings of the Second International Conference of Equine Infectious Diseases, Paris 1972, S. Karder, Basel.*

McCollum, W. H. 1976. Studies of passive immunity in foals to equine viral arteritis. *Vet. Microbiol.* 1 : 45-54.

McCollum, W. H. 1981. Pathological features of horses given avirulent virus intramuscularly. *Am. J. Vet. Res.* 42: 1218-1220.

McCollum, W. H. 1986. Responses of horses vaccinated with modified avirulent-live equine arteritis virus propagated in E. Dermis (NBL-6) cell line to nasal inoculations with virulent virus. *Am. J. Vet. Res.* 47: 1931-1934.

McCollum, W. H., E. R. Doll, J. C. Wilson, and C. B. Johnson. 1961. Propagation of equine arteritis virus in monolayer cultures of equine kidney. *Am. J. Vet. Res.* 22: 731-735.

McCollum, W. H., E. R. Doll, and J. C. Wilson. 1962a. The recovery of virus from horses with experimental cases of equine arteritis using monolayer cell cultures of equine kidney. *Am. J. Vet. Res.* 23: 465-469.

McCollum, W. H., E. R. Doll, J. C. Wilson, and J. Cheatham. 1962b. Isolation and propagation of equine arteritis virus in monolayer cell cultures of rabbit kidney. *Cornell Vet.* 52: 452-458.

McCollum, W. H., and T. W. Swerczek. 1978. Studies of an epizootic of equine viral arteritis in racehorses. *Equine Vet. J.* 2:293-299.

McCollum, W. H., and P. J. Timoney. 1984. The pathogenic qualities of the 1984 strain of equine arteritis virus, Pp. 34-84. Dans: Proceedings of the Grayson Foundation International Conference of Thoroughbred Breeders Organizations, Dromoland Castle, Ireland 1984.

McKenzie, J. 1990. Compulsory control program for equine viral arteritis. *Surveillance* 17: 19-20.

McKnight, K. L., D. A. Simpson, S.-C. Lin, T. A. Knott, G. M. Polo, D. F. Pence, D. B. Johannsen, H. W. Heidner, N. L. Davis, and R. E. Johnston. 1996. Deduced consensus sequence of Sindbis virus strain AR339: mutations contained in laboratory strains which affect cell culture and in vivo phenotypes. *J. Virol.* 70:1981-1989.

Mendez, A., C. Smerdou, A. Izeta, F. Gebauer, and L. Enjuanes. 1996. Molecular characterization of transmissible gastroenteritis coronavirus defective interfering genomes: packaging and heterogeneity. *Virology* 217: 495-507.

Mertz, J. E., A. Murphy, and A. Barkan. 1983. Mutants deleted in the agnogene of simian virus 40 define a new complementation group. *J. Virol.* 45: 36-46.

Meulenberg, J. J., E. J. de Meijer, and R. J. Moormann. 1993a. Subgenomic RNAs of Lelystad virus contain a conserved leader-body junction sequence. *J. Gen. Virol.* 74 : 1697-1701. *Virology* 192: 62-72.

Meulenberg, J. J., M. M. Hulst, E. J. de Meijer, P. L. Moonen, A. den Besten, E. P. de Kluyver, G. Wensvoort, R. J. Moormann. 1993b. Lelystad virus, the causative agent of porcine epidemic abortion and respiratory syndrome (PEARS), is related to LDV and EAV. *J. Gen. Virol.* 74: 1697-1701

Meulenberg, J. J., J. N. Bos-de Ruijter, R. van de Graaf, G. Wensvoort, and R. J. Moormann. 1998. Infectious transcripts from cloned genome-length cDNA of porcine reproductive and respiratory syndrome virus. *J. Virol.* 72: 380-387.

Milligan, J. F., and O. C. Uhlenbeck. 1989. Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol.* 180: 51-62.

Molenkamp, R., B. C. Rozier, S. Greve, W. J. Spaan, and E. J. Snijder. 2000. Isolation and characterization of an arterivirus defective interfering RNA genome. *J. Virol.* 74: 3156-3165.

Molenkamp, R., H. van Tol, B. C. Rozier, Y. van Der Meer, W. J. Spaan, and E. J. Snijder. 2000. The arterivirus replicase is the only viral protein required for genome replication and subgenomic mRNA transcription. *J Gen Virol.* 81:2491-2496.

Morris, D. R., and A. P. Geballe. 2000. Upstream open reading frames as regulators of mRNA translation. *Mol. Cell. Biol.* 20:8635-8642.

Moustakas, A., T. S. Sonstegard, and P. B. Hackett. 1993a. Alterations of the three short open reading frames in the Rous sarcoma virus leader RNA modulate viral replication and gene expression. *J. Virol.* 67: 4337-4349.

Moustakas, A., T. S. Sonstegard, and P. B. Hackett. 1993b. Effects of the open reading frames in the Rous sarcoma virus leader RNA on translation. *J. Virol.* 67: 4350-4357.

Mumford, J. A. 1985. Preparing for equine arteritis. *Equine Vet. J.* 17: 6-11.

Murphy, T. , P. J. Timoney, and W. H. McCollum. 1988. Analysis of genetic variation among strains of equine arteritis virus. Pp.3-12. Dans: D. G. Powell (ed.) *Proceedings of the Fifth International Conference on Equine Infectious Diseases*, Lexington 1987. The University Press of Kentucky, Lexington, Kentucky.

Murphy, T., W. H. McCollum, P. J. Timoney, B. W. Klingeborn, B. Hyllseth, W. Golnik, and B. Erasmus. 1992. Genomic variability among globally distributed isolates of equine arteritis virus. *Vet. Microbiol.* 32: 101-115.

Nagy, P. D., and A. E. Simon. 1997. New insights into the mechanisms of RNA recombination. *Virology* 235: 1-9

Nakagawa, H., T. Inomoto, and A. K. Rustgi. 1997. A CACCC box-like cis-regulatory element of the Epstein-Barr virus ED-L2 promoter interacts with a novel transcriptional factor in tissue-specific squamous epithelia. *J. Biol. Chem.* 272: 16688-16699

Neu, S. M., P. J. Timoney, and W. H. McCollum. 1988. Persistent infection of the reproductive tract in the stallions persistently infected, with equine arteritis virus, Pp. 149-154. D. G. Powell, (ed.). *Proceedings of the Fifth International Conference on Equine Infectious Diseases*, Lexington 1987. The University Press of Kentucky, Lexington, Kentucky.

Ng, S. C., J. E. Mertz, S. Sanden Will, and M. Bina. 1985. Simian virus 40 maturation in cells harboring mutants deleted in the agnogene. *J. Biol. Chem.* 260:1127-1132.

Okumura K., G. Sakaguchi, S. Takagi, K. Naito, T. Mimori, and H. Igarashi. 1996. Sp1 family proteins recognize the U5 repressive element of the long terminal repeat of human T cell leukemia virus type I through binding to the CACCC core motif. *J. Biol. Chem.* 271: 12944-12950.

Oleksiewicz, M., B., A., Botner, J. Nielsen, and T. Storgaard. 1999. Determination of 5'-leader sequences from radically disparate strains of porcine reproductive and respiratory syndrome virus reveals the presence of highly conserved sequence motifs. *Arch. Virol.* 144: 981-987.

Paton, D. J., I. H. Brown, S. Edwards, and G. Wensvoort. 1991. Blue ear disease of pigs. *Vet. Rec.* 128: 617.

Patton, J. F., U. B. R. Balasuriya, J. F. Hedges, T. M. Schweidler, P. J. Hullinger, and N. J. MacLachlan. 1999. *Arch. Virol.* 144: 817-827.

Pedersen, K. W., Y. van der Meer, N. Roos, and E. J. Snijder. 1999. Open reading frame 1a-encoded subunits of the arterivirus replicase induce endoplasmic reticulum-derived double-membrane vesicles which carry the viral replication complex. *J. Virol.* 73: 2016-2026.

Pelletier, J., M. E. Flynn, G. Kaplan, V. Racaniello, and Sonenberg. 1988. Mutational analysis of upstream AUG codons of poliovirus RNA. *J. Virol.* 62: 4486-4492.

Plagemann, P. G., and V. Moennig. Lactate dehydrogenase-elevating virus, equine arteritis virus, and simian hemorrhagic fever virus: a new group of positive-strand RNA viruses. *Adv. Virus Res.* 1992. 41:99-192.

Polson, A. G., B. L. Bass, and J. L. Casey. 1996. RNA editing of hepatitis delta virus antigenome by dsRNA-adenosine deaminase. *Nature* 380: 454-456.

Polson, A. G., H. L. Ley, B. L. Bass, and J. L. Casey. 1998. Hepatitis delta virus RNA editing is highly specific for the amber/W site and is suppressed by hepatitis delta antigen. *Mol. Cell Biol.* 18: 1919-1926.

Pooggin, M. M., T. Hohn, and J. Futterer. 1998. Forced evolution reveals the importance of short open reading frame A and secondary structure in the cauliflower mosaic virus 35S RNA leader. *J. Virol.* 72: 4157-4169.

Portfield, J. S., J. Casals, M. P. Chumakov, S. Y. Gaidamovich, S. Hannoun, I. H. Holmes, M. C. Horzinek, M. Mussgay, N. Oker-Blom, and D. W. Trent. 1978. Togaviridae. *Intervirology* 9: 129-148.

Pottie, A. 1888. The propagation of influenza from stallions to mares. *J. Comp. Pathol.* 1:37-38.

Pritchard, A., M. A. Calenoff, S. Simpson, K. Jensen, and H. Lipton. 1992. A single base deletion in the 5' noncoding region of Theiler's virus attenuates neurovirulence. *J. Virol.* 66: 1951-1958.

Prickett, M. E., W. H. McCollum, and J. T. Bryans. 1973. The gross and microscopic pathology observed in horses experimentally infected with equine arteritis virus, Pp. 265-272. Dans : J. T. Bryans, and H. Gerber (ed.) *Proceedings of the Third International Conference of Equine Infectious Diseases, Paris 1972*, S. Karger, Basel.

Radwan, A. I., and D. Burger, 1973a. The complement-requiring neutralization of equine arteritis virus by late antisera. *Virology* 51:71-77.

Radwan, A. I., and D. Burger, 1973b. The role of sensitizing antibody in the neutralization of equine arteritis virus by complement or anti-IgG serum. *Virology* 53: 366-371.

Radwan, A. I., and B. Crawford. 1974. The mechanisms of neutralization of sensitized equine arteritis virus by complement components. *J. Gen. Virol.* 25: 229-237.

Ramina A., L. Dalla Valle, S. De Mas, E. Tisato, A. Zuin, M. Renier, V. Cuteri, C. Valente, and F. M. Cancellotti. 1999. Detection of equine arteritis virus in semen by reverse transcriptase polymerase chain reaction-ELISA. *Comp. Immunol. Microbiol. Infect. Dis.* 22: 187-97

Resnick, J., and T. Shenk. 1986. Simian virus 40 agnoprotein facilitates normal nuclear location of the major capsid polypeptide and cell-to-cell spread of virus. *J. Virol.* 60:1098-1106.

Rinaldo, C. H., T. Traavik, and A. Hey. 1998. The agnogene of the Human Polyomavirus BK Is Expressed. *J. Virol.* 72:6233-6236.

Rohll, J. B., N. Percy, R. Ley, D. J. Evans, J. W. Almond, and W. S. Barclay. 1994. The 5'-untranslated regions of picornavirus RNAs contain independent functional domains essential for RNA replication and translation. *J. Virol.* 68: 4384-4391.

Romero, J. R., and H. A. Rotbart. 1995. Sequence analysis of the downstream 5' nontranslated region of seven echoviruses with different neurovirulence phenotypes. *J. Virol.* 69: 1370-1375.

Sagliocco, F. A., M. R. Vega Laso, D. Zhu, M. F. Tuite, J. E. McCarthy, and A. J. Brown. 1993. The influence of 5'-secondary structures upon ribosome binding to mRNA during translation in yeast. *J. Biol. Chem.* 268: 26522-26530.

Sagripanti, J. L., R. O. Zandomeni, and R. Weinmann. 1986. The cap structure of simian hemorrhagic fever virion RNA. *Virology* 151: 146-150.

Saito, A., T. Kanno, Y. Murakami, M. Muramatsu, and S. Yamaguchi. 1996. Characteristics of major structural protein coding gene and leader-body sequence in subgenomic mRNA of porcine reproductive and respiratory syndrome virus isolated in Japan. *J. Vet. Med. Sci.* 58: 377-380.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sanchez, A., S. G. Trappier, B. W. Mahy, C. J. Peters, and S. T. Nichol. 1996. The virion glycoproteins of Ebola viruses are encoded in two reading frames and are expressed through transcriptional editing. *Proc. Natl. Acad. Sci. U.S.A.* 93: 3602-3607.

Sawicki, S. G., and D. L. Sawicki. 1990. Coronavirus transcription: subgenomic mouse hepatitis virus replicative intermediates function in RNA synthesis. *J. Virol.* 64: 1050-1056.

Sawicki, S. G., and D. L. Sawicki. 1995. Coronaviruses use discontinuous extension for synthesis of subgenome-length negative strands. *Adv. Exp. Med. Biol.* 380: 499-506.

Sawicki, S. G., and D. L. Sawicki. 1998. A new model for coronavirus transcription. *Adv. Exp. Med. Biol.* 440:215-219.

Schägger, H., and G. von Jagow. 1987. Tris-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* 166:368-379.

Schreiber, S. S., T. Kamahora, and M. M. Lai. 1989. Sequence analysis of the nucleocapsid protein gene of human coronavirus 229E. *Virology* 169: 142-151.

Scott, J. 1995. A place in the world for RNA editing. *Cell* 81:833-836.

Scott, J. 1997. RNA editing. Message change for a fat controller. *Nature* 387: 242-243.

Sedman, S.A., and J. E. Mertz. 1988. Mechanisms of synthesis of virion proteins from the functionally bigenic late mRNAs of simian virus 40. *J. Virol.* 62: 954-961.

Sedman, S. A., P. J. Good, and J. E. Mertz. 1989. Leader-encoded open reading frames modulate both the absolute and relative rates of synthesis of the virion proteins of simian virus 40. *J. Virol.* 63: 3884-3893.

Sedman, S. A., G. W. Gelembiuk, and J. E. Mertz. 1990. Translation initiation at a downstream AUG occurs with increased efficiency when the upstream AUG is located very close to the 5' cap. *J. Virol.* 64: 453-457.

Sekiguchi K., S. Sugita, Y. Fukunaga, T. Kondo, R. Wada, M. Kamada, and S. Yamaguchi. 1995. Detection of equine arteritis virus (EAV) by polymerase chain reaction (PCR) and differentiation of EAV strains by restriction enzyme analysis of PCR products. *Arch. Virol.* 140 :1483-91.

Senanayake, S. D., and D. A. Brian. 1997. Bovine coronavirus I protein synthesis follows ribosomal scanning on the bicistronic N mRNA. *Virus Res.* 48: 101-105.

Senanayake, S. D., M. A. Hofmann, J. L. Maki, and D. A. Brian. 1992. The nucleocapsid protein gene of bovine coronavirus is bicistronic. *J. Virol.* 66: 5277-5283.

SenGupta, D. N., B. Berkhout, A. Gatignol, A. M. Zhou, and R. H. Silverman. 1990. Direct evidence for translational regulation by leader RNA and Tat protein of human immunodeficiency virus type 1. *Proc. Natl. Acad. Sci. U.S.A.* 87:7492-7496.

Senne, D. A., J. E. Perraon, and E. A. Carbrey. 1985. Equine viral arteritis: a standard procedure for the virus neutralization test and comparison of results of a proficiency test performed at five laboratories, Pp. 29-34. Dans: *Proceedings of the 89th Annual Meeting of the United States Animal Health Association, Milwaukee, WI.*

Sethna, P. B., S. L. Hung, and D. A. Brian. 1989. Coronavirus subgenomic minus-strand RNAs and the potential for mRNA replicons. *Proc. Natl. Acad. Sci. USA.* 86: 5626-5630.

Sharmeen, L., B. Bass, N. Sonenberg, H. Weintraub, and M. Groudine. 1991. Tat-dependent adenosine-to-inosine modification of wild-type transactivation response RNA. *Proc. Natl. Acad. Sci. U.S.A.* 88: 8096-8100.

Shen, S., J. Kwang, W. Liu, and D. X. Liu. 2000. Determination of the complete nucleotide sequence of a vaccine strain of porcine reproductive and respiratory syndrome virus and identification of the Nsp2 gene with a unique insertion. *Arch. Virol.* 145: 871-883.

Shinagawa, M., R. Yanagawa, and T. Inoue. 1976. Growth of equine arteritis virus in cells derived from infectious canine hepatitis virus-induced hamster tumor and transformed cells. *Jpn. J. Vet. Sci.* 38: 25-32.

Shieh, C. K., L. H. Soe, S. Makino, M. F. Chang, S. A. Stohlman, and M. M. Lai. 1987. The 5'-end sequence of the murine coronavirus genome: implications for multiple fusion sites in leader-primed transcription. *Virology* 156: 321-330.

Sleat, D. E., R. Hull, P. C. Turner, and T. M. Wilson. 1988. Studies on the mechanism of translational enhancement by the 5'-leader sequence of tobacco mosaic virus RNA. *Eur. J. Biochem.* 175: 75-86.

Slobodskaya, O.R., A. P. Gmyl, S. V. Maslova, E. A. Tolskaya, E. G. Viktorova, and V. I. Agol. 1996. Poliovirus neurovirulence correlates with the presence of a cryptic AUG upstream of the initiator codon. *Virology* 221: 141-150.

Smith, S. L., X. Wang, and E. K. Godeny. 1997. Sequence of the 3' end of the simian hemorrhagic fever virus genome. *Gene* 191: 205-210.

Snijder, E. J., J. Ederveen, W. J. Spaan, M. Weiss, and M. C. Horzinek. 1988. Characterization of Berne virus genomic and messenger RNAs. *J. Gen. Virol.* 69: 2135-2144.

Snijder, E. J., and Horzinek M. C. 1993. Toroviruses: replication, evolution and comparison with other members of the coronavirus-like superfamily. *J Gen. Virol.* 74: 2305-2316.

Snijder, E. J., and W. J. Spaan. 1995. The Coronaviridae. Pp. 239-255S. Dans: G. Siddell (ed.). Plenum, New York.

Snijder, E. J., J. A. den Boon, P. J. Bredenbeek, M. C. Horzinek, R. Rijnbrand, and W. J. Spaan. 1990a. The carboxyl-terminal part of the putative Berne virus polymerase is expressed by ribosomal frameshifting and contains sequence motifs which indicate that toro- and coronaviruses are evolutionarily related. *Nucleic Acids Res.* 18: 4535-4452.

Snijder, E. J., M. C. Horzinek, and W. J. Spaan. 1990b. A 3'-coterminal nested set of independently transcribed mRNAs is generated during Berne virus replication. *J. Virol.* 64: 331-338.

Snijder, E. J., J. A. den Boon, M. C. Horzinek, and W. J. Spaan. 1991. Characterization of defective interfering RNAs of Berne virus. *J. Gen. Virol.* 72: 1635-1643.

Snijder, E. J., A. L. Wassenaar, and W. J. Spaan. 1992. The 5' end of the equine arteritis virus replicase gene encodes a papainlike cysteine protease. *J. Virol.* 66: 7040-7048.

Snijder, E. J., A. L. Wassenaar, and W. J. Spaan. 1994. Proteolytic processing of the replicase ORF1a protein of equine arteritis virus. *J. Virol.* 68: 5755-5764.

Snijder, E. J., A. L. Wassenaar, W. J. Spaan, and A. E. Gorbalenya 1995. The arterivirus Nsp2 protease. An unusual cysteine protease with primary structure similarities to both papain-like and chymotrypsin-like proteases. *J. Biol. Chem.* 270 :16671-16676.

Snijder, E. J., A. L. Wassenaar, L. C. van Dinten, W. J. Spaan, and A. E. Gorbalenya. 1996. The arterivirus nsp4 protease is the prototype of a novel group of chymotrypsin-like enzymes, the 3C-like serine proteases. *J. Biol. Chem.* 271 :4864-4871.

Snijder, E. J., H. van Tol, K. W. Pedersen, M. J. B. Raamsman, and A. A. F. de Vries. 1999. Identification of a novel structural protein of arteriviruses. *J. Virol.* 73:6335-6345.

Spaan, W., H. Delius, M. Skinner, J. Armstrong, P Rottier, S. Smeekens, B. A. van der Zeijst, and S. G. Siddell. 1983. Coronavirus mRNA synthesis involves fusion of non-contiguous sequences. *EMBO J.* 2: 1839-1844.

Spaan, W. J. M., J. A. den Boon, P. J. Bredenbeck, E. D. Chirnside, A. F. H. Noten, E. J. Snijder, A. A. F. de Vries, and M. C. Horzinek. 1990. New aspects of positive strand RNA viruses Pp. 12-19. Dans: M. A. Brinton and F. X. Heinz, (ed.) American Society for Microbiology, Washington, D.C.

Stadejek, T., H. Björklund, C. R. Bascunana, I. M. Ciabatti, M. T. Scicluna, D. Amaddeo, W. H. McCollum, G. L. Autorino, P. J. Timoney, D. J. Paton, B. Klingeborn and S. Belàk. 1999. Genetic diversity of equine arteritis virus. *J. Gen. Virol.* 80: 691-699.

Starick E. 1998. Rapid and sensitive detection of equine arteritis virus in semen and tissue samples by reverse transcription-polymerase chain reaction, dot blot hybridization and nested polymerase chain reaction. *Acta Virol.* 42 :333-339.

Starick E. 1999. [Detection of equine arteritis virus (EAV) in stallions--a contribution to the improvement of EAV diagnosis.]. *Berl. Munch. Tierarztl. Wochenschr.* 112 :10-3

Steward, M., I. B. Vipond, N. S. Millar, P. T. Emmerson. 1993. RNA editing in Newcastle disease virus. *J. Gen. Virol.* 74: 2539-2547.

Stewart, S. R., and B. L. Semler. 1998. RNA structure adjacent to the attenuation determinant in the 5'-non-coding region influences poliovirus viability. *Nucleic Acids Res.* 26:5318-5326.

St-Laurent G., G. Morin, and D. Archambault. 1994. Detection of equine arteritis virus following amplification of structural and nonstructural viral genes by reverse transcription-PCR. *J. Clin. Microbiol.* 32 :658-665.

St-Laurent, G., N. Lepage, S. Carman, and D. Archambault. 1997. Genetic and amino acid analysis of the GL protein of Canadian, American and European equine arteritis virus isolates. *Can. J. Vet. Res.* 61: 72-76.

Stohlman, S. A., R. S. Baric, G. N. Nelson, L. H. Soe, L. M. Welter, and R. J. Deans. 1988. Specific interaction between coronavirus leader RNA and nucleocapsid protein. *J. Virol.* 62: 4288-4295.

Tahara, S. M., T. A. Dietlin, C. C. Bergmann, G. W. Nelson, S. Kyuwa, R. P. Anthony, and S. A. Stohlman. 1994. Coronavirus translational regulation: leader affects mRNA efficiency. *Virology* 202: 621-630.

Terpstra, C., G. Wensvoort, and J. M. Pol. 1991. Experimental reproduction of porcine epidemic abortion and respiratory syndrome (mystery swine disease) by infection with Lelystad virus: Koch's postulates fulfilled. *Vet. Q.* 13: 131-136.

Tesar, M., S. A. Harmon, D. F. Summers, and E. Ehrenfeld. 1992. Hepatitis A virus polyprotein synthesis initiates from two alternative AUG codons. *Virology* 186: 609-618.

Timoney, P. J. 1984. Clinical, virological and epidemiological features of the 1984 outbreak of the equine arteritis virus in the thoroughbred population in Kentucky, USA, Pp. 24-33. Dans: Proceedings of the Grayson Foundation International Conference of Thoroughbred Breeders Organizations, Dromoland Castle, Ireland.

Timoney, P. J., and W. H. McCollum. 1988. Equine viral arteritis: epidemiology and control. *J. Equine Vet. Sci.* 8: 54-59.

Timoney, P. J., and W. H. McCollum. 1990. Equine viral arteritis: current clinical and economic significance. *Proc. Am. Assoc. Equine Pract.* 36: 403-409.

Timoney, P. J., and W. H. McCollum. 1993. Equine viral arteritis. *Vet. Clin. North Am. Equine Pract.* 9: 295-309.

Timoney, P. J., W. H. McCollum, T. W. Murphy, A. W. Roberts, A. W. Willard, and J. G. Carswell. 1987a. The carrier state in equine arteritis virus infection in the stallion with specific emphasis on the venereal mode of transmission. *J. Reprod. Fertil. Suppl.* 35:95-102.

Timoney, P. J., W. H. McCollum, and A. W. Roberts. 1987b. Detection of the carrier state in stallions persistently infected with equine arteritis virus. *Proc. Am. Assoc. Equine Pract.* 32: 57-65.

Timoney, P. J., W. H. McCollum, A. W. Roberts, and T. W. Murphy. 1986. Demonstration of the carrier state in natural acquired equine arteritis infection in the stallion. *Res. Vet. Sci.* 41 : 279-280

Timoney, P. J., N. W. Umphenour, and W. H. McCollum. 1988. Safety evaluation of a commercial modified live equine arteritis virus vaccine for use in stallions, Pp. 17-29. Dans: D. G. Powell (ed). Proceedings of the 5th International Conference on Equine Infectious Diseases, Lexington 1987. The University Press of Kentucky, Lexington, Kentucky.

Tsukamoto, K., S. Konishi, and M. Ogata. 1982. Effects of *Mycoplasma orale* on growth of equine arteritis virus in vero cells. *Jpn. J. Vet. Sci.* 44: 989-991.

Tsukiyama-Kohara, K., N. Iizuka, M. Kohara, and A. Nomoto. 1992. Internal ribosome entry site within hepatitis C virus RNA. *J. Virol.* 66: 1476-1483.

Vaala, W. E., A. N. Hamir, E. Dubovi, P. J. Timoney, and B. Ruiz. 1992. Fatal congenitally acquired equine arteritis virus infection in a neonatal thoroughbred. *Equine Vet. J.* 24:155-158.

van Berlo, M. F., M. C. Horzinek, and B. A. M. van der Zeijst. 1980. Temperature-sensitive mutants of equine arteritis virus. *J. Gen. Virol.* 49: 97-104.

van Berlo, M. F., M. C. Horzinek, and B. A. M. van der Zeijst. 1982. Equine arteritis virus-infected cells contain six polyadenylated virus-specific RNAs. *Virology* 118 :345-352.

van Berlo, M. F., P. J. Rottier, M. C. Horzinek and B. A. M. van der Zeijst. 1986a. Intracellular equine arteritis virus (EAV)-specific RNAs contain common sequences. *Virology* 152: 492-496.

van Berlo, M. F., P. J. Rottier, W. J. M. Spaan, and M. C. Horzinek. 1986b. Equine arteritis virus-induced polypeptide synthesis. *J. Gen. Virol.* 67: 1543-1549.

van Marle, G., J. C. Dobbe, A. P. Gulyaev, W. Luytjes, W. J. M. Spaan, and E. J. Snijder. 1999. Arterivirus discontinuous mRNA transcription is guided by base pairing between sense and antisense transcription-regulating sequences. *Proc. Natl. Acad. Sci. USA* 96:12056-12061.

van der Meer, Y., H. van Tol, J. K. Locker, and E. J. Snijder. 1998. ORF1a-encoded replicase subunits are involved in the membrane association of the arterivirus replication complex. *J. Virol.* 72 :6689-6698.

van der Most, R. G., R. J. de Groot, and W. J. Spaan. 1994. Subgenomic RNA synthesis directed by a synthetic defective interfering RNA of mouse hepatitis virus: a study of coronavirus transcription initiation. *J. Virol.* 68: 3656-3666.

van der Vossen, E. A., L. Neeleman, J. F. Bol. 1993. Role of the 5' leader sequence of alfalfa mosaic virus RNA 3 in replication and translation of the viral RNA. *Nucleic Acids Res.* 21: 1361-1367.

van der Zeijst, B. A. M., and M. C. Horzinek. 1975. The genome of the equine arteritis virus. *Virology* 68: 418-425.

van Dinten, L. C., A. L. Wassenaar, A. E. Gorbalenya, W. J. Spaan, and E. J. Snijder. 1996. Processing of the equine arteritis virus replicase ORF1b protein: identification of cleavage products containing the putative viral polymerase and helicase domains. *J. Virol.* 70: 6625-6633.

van Dinten, L. C., S. Rensen, A. E. Gorbalenya, and E.J. Snijder. 1999. Proteolytic processing of the open reading frame 1b-encoded part of arterivirus replicase is mediated by nsp4 serine protease and is essential for virus replication. *J. Virol.* 73: 2027-2037.

van Dinten, L. C., H. van Tol, A. E. Gorbalenya, and E.J. Snijder. 2000. The predicted metal-binding region of the arterivirus helicase protein is involved in subgenomic mRNA synthesis, genome replication, and virion biogenesis. *J. Virol.* 74: 5213-5223.

van Nieuwstadt A. P., J. J. Meulenberg, A. van Essen-Zanbergen, A. Petersen-den Besten, R. J. Bende, R. J. Moormann, and G. Wensvoort. 1996. Proteins encoded by open reading frames 3 and 4 of the genome of Lelystad virus (Arteriviridae) are structural proteins of the virion *J. Virol.* 70: 4767-4772.

Vrins, A., P. Guay et J. Longtin. 1987. L'artérite virale équine: une entité présente au Québec. *Méd. Vét. Qué.* 17: 6-11.

Wada, R., Y. Fukunaga, T. Kondo, and T. Kanemaru. 1995. Ultrastructure and immunocytochemistry of BHK-21 cells infected with a modified Bucyrus strain of equine arteritis virus. *Arch. virol.* 140: 1173-1180.

Wakiyama, M., I. Hirao, I. Kumagai, K. Watanabe, and K. Miura. 1992. Relationship between the 5'-leader sequence of mRNAs and the translation efficiency in a eukaryotic *in vitro* system. *Nucleic Acids Symp. Ser.* 27: 159-160.

Wang, X. C., S. L. Smith, and E. K. Godeny. 1998. Organization of the simian hemorrhagic fever virus genome and identification of the sgRNA junction sequences. *Adv. Exp. Med. Biol.* 440: 281-287.

Wassenaar, A. L., W. J. Spaan, A. E. Gorbalenya, and E. J. Snijder. 1997. Alternative proteolytic processing of the arterivirus replicase ORF1a polyprotein: evidence that NSP2 acts as a cofactor for the NSP4 serine protease. *J Virol.* 71: 9313-22.

Wensvoort G., E. P. de Kluyver, J. M. Pol, F. Wagenaar, R. J. Moormann, M. M. Hulst, R. Bloemraad, A. den Besten, T. Zetstra, and C. Terpstra. 1992. Lelystad virus, the cause of porcine epidemic abortion and respiratory syndrome: a review of mystery swine disease research at Lelystad. *Vet. Microbiol.* 33: 185-193.

Westaway, F. G., M. A. Brinton, S. Y. Gaidamovich, M. C. Horzinek, A. Igarashi, L. Kääriäinen, D. K. Lvov, J. S. Portfield, P. K. Russell, and D. W. Trent. 1985. *Togaviridae*. *Intervirology* 24: 125-139.

Willson, J. C., E. R. Doll, W. H. McCollum; and J. Cheatham. 1962. Propagation of equine arteritis virus previously adapted to cell culture of horse kidney in monolayer cultures of hamster kidney. *Cornell Vet.* 52: 205-208.

Zeegers, J. J., W., B. A. M. van der Zeijst, and M. C. Horzinek. 1976. The structural proteins of equine arteritis virus. *Virology* 73: 200-205.

Zeng, L., E. K. Godeny, S. L. Methven, and M. A. Brinton. 1995. Analysis of simian hemorrhagic fever virus (SHFV) subgenomic RNAs, junction sequences, and 5' leader. *Virology* 207: 543-548.

Zhang, X, and M. M. Lai. 1994. Unusual heterogeneity of leader-mRNA fusion in a murine coronavirus: implications for the mechanism of RNA transcription and recombination. *J Virol.* 68 :6626-6633.

Zuker, M., D.H. Mathews, and D.H. Turner. 1999. Algorithms and Thermodynamics for RNA Secondary Structure Prediction: A Practical Guide In J. Barciszewski and B. F. C. Clark, eds. *RNA Biochemistry and Biotechnology*. Pp. 11-43. Kluwer Academic Publishers, NATO ASI Series.

CHAPITRE IX

APPENDICE

AUTRES CONTRIBUTIONS

- **KHEYAR A**, S Martin, PJ Timoney, HW McCollum, G St-Laurent, D Archambault. 1997. Expression cloning and humoral immune response to the nucleocapsid (N) and membrane (M) proteins of equine arteritis virus. *Clinical Diagnostic and Laboratory Immunology* 4:648-652.

- **KHEYAR A**, G St-Laurent, M Diouri, J Dufresne and D Archambault. 1997. Sequence determination and genetic analysis of the leader region of various equine arteritis virus isolates. *Advances in Experimental Medicine and Biology* 440: 805-812.