PLA2 domain in parvoviruses

Sequence analysis of 21 *Pavovirinae* members revealed that they contain an 80 amino acid conserved region in their VP1up. 20 amino acids out of the 80 were fully conserved and an additional 40 showed some kind of physicochemical similarity. This domain also can be found on *Densovirinae*'s VP1ups (except *Brevidensovirus* members) even though they have no homology to the Parvovirinae members in their major capsid component. This remarkable stability suggested that this region could be responsible for some kind of biological function.

Searching protein databanks with stringent parameters did not result in homologous protein hits while searches with relaxed parameters produced thousand of proteins with weak homology to the region. Among them hundreds of secreted phospholipase A2 (sPLA2). PLA2s form a superfamily of key enzymes involved in physiological and pathological processes such as lipid membrane metabolism, signal transduction pathways, inflammation, acute hypersensitivity, and degenerative diseases. PLA2s catalyze the hydrolysis of phospholipid substrates at the 2-acyl ester (sn-2) position to release lysophospholipids and free fatty acids. In sPLA2, H48 functions as a general base and is assisted by D99 to deprotonate a catalytic water molecule that hydrolyzes the phospholipid ester. The adjacent D49, via its β-carboxyl group, and backbone carbonyl oxygens from Y28, G30, and G32 in the Ca2+ binding loop coordinate the catalytic Ca2+ cofactor involved in the stabilization of the transition state.

Interestingly, the catalytically important amino acids in sPLA2 matched some of the conserved amino acids of VP1up, which suggested that the parvoviral VP1 might have a PLA2 activity. Enzyme activity of purified VP1ups was measured with the mixed micelles assay. Four VP1up polypeptides of parvoviruses (PPV, *Gm*DNV, AAV and B19) from distant genera were expressed in order to demonstrate PLA2 activity. All four had PLA2 activity though with remarkable differences. PPV VP1up showed almost three order of manitude higher activity ((*k*cat/*k*M)app=(71.9∀9.4)x$10^5 M^{-1}s^{-1}$) than B19 VP1up (2,5∀0,2)x$10^4 M^{-1}s^{-1}$) and almost four order of magnitude higher than *GmDNV* VP1up (0,4∀0,03) x$10^4 M^{-1}s^{-1}$).

The enzymatic features of viral PLA2

The core 80 amino acid domain did not show optimal activity, and an additional 30 amino acids outside the conserved domain had a large impact on the PPV VP1up catalytic efficiency. sPLA2 inhibitors oleyloxyethylphosphorylcholine (OP) and manoalide (MA) also inhibited the VP1up PLA2 activity. MA concentration for 50% inhibition (IC50) in the radiolabeled E. coli PLA2 assay was 3.8∀0.4 μM and 2.0∀0.3 μM for B19 and PPV PLA2, respectively. The pH optimum of PPV PLA2 in the mixed micelles assay was 8.0 The Ca^{2+} concentration for optimum activity of the expressed PPV PLA2 was above 10 mM as for most sPLA2, whereas chelating $Ca^{2+}$by EDTA or EGTA abolished activity of both PPV and B19 pvPLA2. PPV and B19 pvPLA2 did not exhibit a substantial substrate preference among phosphatidylcholine, -ethanolamine, and -inositol phospholipids. Since PPV VP1up resigns inside intact capsid, only background PLA2 activity could be detected from PPV virion. However, when the PLA2 domain was exposed by dissociation of capsids, enzyme activity was measured close to levels of expressed VP1up, demonstrating that under specific conditions virion-associated PLA2 activity can be revealed.

The 3D structure of many group I/II and group III sPLA2s has been solved, and served to predict the viral PLA2

structure. Sequence alignments suggested that the structure of viral PLA2s is closer to that of the group III sPLA2s. Based on this model, potentially critical amino acids for the enzyme activity were targeted for site-directed mutagenesis. The same mutations on the infectious clone of PPV made it possible to study the relationship between pvPLA2 activity and virus infectivity.

Both enzyme activity and viral infectivity decreased dramatically when amino acids in the catalytic site (H41 and D42) were mutated. D63 was conserved in viral PLA2 suggesting that it corresponded to the catalytic D99 in sPLA2. Its replacement by A63 or N63 decreased strongly both enzyme activity and viral infectivity, supporting the predicted relative position of α helices in the 3D-structural model. P21 was conserved in the Ca^{2+} binding loop of all pvPLA2 but not among that of sPLA2. When P21pv was mutated to amino acids that occur often at that position in sPLA2 (R, W, L), both the enzyme activity and the infectivity were significantly reduced. The role of the K88, which is conserved in parvoviruses and bee venom PLA2, is unknown but essential since even conservative mutation (K88R) strongly lowered activity. Overall, the impact of mutations showed a direct correlation (R^2 = 0.898) between enzyme activity and viral infectivity.

The role of PLA2 in viral infection

As described above, both wt and PLA2 mutant infectious clones were effective in producing virions upon transfection. Hence, pvPLA2 is required before replication and packaging. Conversion of single-stranded parvoviral genomes into double-stranded DNA during normal infection is achieved in the nucleus by cellular DNA polymerases. Therefore viral PLA2 is implicated in the infection process prior to the presence of single-stranded genomes in the nucleus.

Experiments with S^{35} and immunhystochemically labeled PPV and AAV proved that the viral PLA2 is not necessary for the binding or the entry into the cell. In situ DNA hybridization was undertaken to investigate whether mutant viruses transferred their genome from the late endosomal/lysosomal compartment to the nucleus as efficiently as wt PPV. We were not able to detect replicating DNA of the HD and P21 mutants, which suggested that DNA from mutant viruses did not penetrate the nucleus. So, in conclusion viral PLA2 acts somewhere between the late endosome/lysosome and the nucleus.

To establish whether infectious DNA and active pvPLA2 needed to reside in the same particle complementation studies were undertaken. The addition during infection of 1 μM expressed wt PPV PLA2 or 0.4 μM snake or bee venom sPLA2 in *trans*, at a 10^4-fold excess to PLA2 virus mutants, could not rescue the virus. Wt virus with β-propiolactone-inactivated DNA but active pvPLA2 (70% activity of untreated) in 10-fold excess also failed to rescue mutant virus upon co-infection. Similar results was obtained with a VP1up/*LacZ* recombinant AAV-2 (rAAV-2) vector, which contains the *LacZ* reporter gene in a capsid carrying mutations in VP1up. A co-infection of VP1up/*LacZ* rAAV-2 vectors and wt AAV-2 did not restore the infectivity of the VP1up/*LacZ* rAAV-2 vectors. Active pvPLA2 and infectious DNA is therefore required in *cis* to obtain infection.

Alternative ORFs in parvoviruses

The PLA2 domain is more than less a common genetic characteristics of the *Parvoviridae* family, however exclusive genetical features can be found in every genus. For example,

the protein X ORF found in the genome of V9 erythrovirus (nucleotide 2586 to 2831), is found in all members of the *Erythrovirus* genus and is predicted in every case to contain two transmembrane helices. A large genus-specific ORF (nucleotide 2717 to 3340) can be found to overlap the VP ORF of all members of the *Dependovirus* genus. Genome analysis also revealed a small, conserved alternative ORF overlapping the amino-terminal portion of the VP2 ORF of all members of the *Parvovirus* genus. Among members of the *Parvovirus* genus, the ORF extends between 50 to 68 amino acids and starts with an ATG codon 4 nucleotides downstream of the VP2 protein start codon. Exceptionally in PPV, the ATG codon is positioned 7 nucleotides after the VP2 initiation codon. We designated this ORF as "small alternatively translated protein" (SAT)-ORF.

The characteristics of PPV SAT

The primary sequence of SAT proteins is not particularly conserved. However, all SAT proteins were predicted by different topology prediction programs to be membrane proteins containing a single membrane spanning helix in approximately the same position.

To demonstrate its existence and to identify the product of this ORF, GFP was inserted into the infectious clone of PPV in five different positions of the SAT ORF in all three open reading frames. All five positive control constructs in which the GFP was inserted into the VP frame resulted in GFP expression in PT cells. GFP expression could be detected by fluorescence in both the cytoplasm and the nucleus.

All five constructs in which GFP was inserted in frame of the SAT-ORF also expressed GFP, although the distribution of the protein in the cell showed marked differences, depending on the insertion points. The two constructs in which GFP was inserted before the predicted membrane-spanning helix showed cytoplasmic and nuclear distribution similar to that of the VP2-GFP constructs. However, fusion proteins in which the GFP was inserted at the 3' end or after the predicted membrane-spanning helix coding sequence were excluded from the nucleus and the cell membrane. To demonstrate that this localization of SAT was not the result of some unrecognized interaction of the fusion tag, another construct was made where the GFP fusion tag of the full-length SAT was replaced by a double FLAG tag. This fusion protein showed exactly the same localization as the GFP-labeled SAT.

The nuclear membrane staining and the characteristic pattern in the cytoplasm suggested that these proteins resided in the ER-nuclear membrane system. Co-transfection of the FLAG fusion protein construct with the pDsRed2-ER plasmid expressing the calreticulin-RFP fusion protein ER marker demonstrated that the two proteins co-localized and that SAT indeed localized in the ER-nuclear membrane compartment.

Mutations and deletions confirmed that a hydrophobic α-helix is the main determining factor of the confinement of SAT in the ER.

The first AUG in mRNAs usually initiates translation in eukaryotes. However, there are some exceptions to this rule. One of the exceptions is when two AUGs are just a few nucleotides from each other. The closeness of the potential start codons of the SAT proteins and VP2 indicated that they could be translated from the same mRNA. To investigate this hypothesis, we cloned the spliced versions of VP1-GFP and VP2-GFP fusion constructs into pcDNA3.1 vectors and expressed them in PT cells. The expressed proteins were immunoprecipitated and followed by Western blotting using an anti-GFP MAb. A single band could be detected on the blot when VP1 mRNA was expressed and GFP was fused to the VP

frame (corresponding to the appropriate size of the VP1-GFP fragment) No protein was detected from the VP1 mRNA when GFP was fused to the SAT frame. In contrast, proteins translated from both the VP and SAT frames of the VP2 mRNA, indicated that SAT is indeed translated from the VP2 mRNA. The ratio of the alternative translation of SAT was around 60% of the VP2 proteins.

Three methionines can be found in the PPV SAT-ORF and two different versions of SAT can be distinguished on the Western blot. These three methionines were mutated to investigate whether the SAT proteins are initiated at different AUGs. Mutation of the second and third AUG did not have any visible effect. However, mutation of the first AUG completely eliminated all protein translation from the SAT frame. This suggested that the multiple bands might be the result of posttranslational modifications rather then multiple initiations.

The function of SAT

To investigate the role of SAT in the PPV life cycle, mutant viruses were created by introducing mutations, which reduced or abolished SAT expression. These mutants lized the cells later and spread slower than the wild type virus. Complementation experiments were done to test whether the "slow-spreading" phenotype of the SAT- mutants was really due to the loss of the SAT. The SAT supplied in *trans* could complement the phenotype of the mutant virus demonstrating that the impaired spreading is the consequence of the loss of SAT.

The comparative sequence and genome analysis of the parvoviruses led to the discovery of a novel enzymatic domain and a novel protein among these viruses. The PLA2 is characteristic for almost the whole virus family while the SAT protein is an exclusive feature of the *Parvovirus* genus. The enzyme activity, the existence of SAT and their biological significance was proven experimentally. The viral PLA2 is essential for the productive virus infection, while the SAT facilitates the cell lysis and viral spreading. The discoveries and experiments discussed in this thesis contribute to the better understanding of the parvoviruses and have a high impact on the present and future research of the field.

PhD THESIS

Recognition of novel parvoviral functions by comparative sequence analysis

Zoltán Zádori

Montreal 2005