Scuola Normale Superiore

Pisa

Ana Backovic

Capsid Assembly and Single Stranded DNA Genome Formation of Adeno-Associated Virus Type2 in Yeast Cells

Ph.D Thesis in Molecular Biology

Supervisor: D.Phil. Alvaro Galli

Pisa, 2011

"I keep six honest serving-men:

(They taught me all I knew)

Their names are What and Where and When

And How and Why and Who"

Rudyard Kipling

Here comes the gratitude before all gratitudes; to somebody who was a great person, professor and leader, whose deep love for science gave the birth to this project and whose never stumbling belief, pushed it forward until today. For this, and for showing me that science is the land of wonders where even a mistake seems an improvement, I dedicate this dissertation to my sorely missed professor Arturo Falaschi.

Abstract

Saccharomyces cerevisiae has provided an array of genetic tools to study unknown aspects of viral life cycles, supporting replication of many different RNA or DNA viruses (e.g. Tombusviruses or Papillomaviruses). It also provides means for up-scalable, cost- and time-effective production of various virus-like particles (e.g. Human Parvovirus B19 or Rotavirus) and as such represents a useful tool for vaccine development. To extend the utility of the *S. cerevisiae* expression system, we expressed AAV2 structural and nonstructural proteins in yeast cells, using both authentic AAV2 and heterologous yeast promoters. For the first time, we described the assembly of AAV2 virus-like particles from yeast-expressed AAV2 structural proteins. To do this we used AAV p40 promoter, whose activity in yeast cells resembled the one of yeast glycolytic promoters, resulting in the synthesis of the most abundant capsid protein VP3 when transformed yeast cells were grown on glucose as a carbon source. The expression of other two VPs was induced from yeast, galactose inducible pGal1 promoter. Simultaneous production of all three VPs was achieved by growing the yeast cells in the medium containing both glucose and galactose, while their relative production levels were further optimized by varying amounts of each carbon source in the induction medium, followed by the fine tuning of the induction time.

Moreover, we investigated the ability of the yeast *Saccharomyces cerevisiae* to carry out the replication of a recombinant rAAV2. When a plasmid harboring the rAAV2 genome in which the cap gene was replaced with the *S. cerevisiae* URA3 gene, was co-transformed in yeast with a plasmid expressing Rep68 from constitutive yeast promoter pADH, a significant number of URA3+ clones were scored (more than 30-fold over controls). Molecular analysis of low molecular weight DNA revealed that the single stranded DNA is formed, in Rep68 and ITR dependent manner, and that the plasmid is entirely replicated. The ss DNA contained the ITRs, URA3 gene and also vector sequences suggesting that ss rAAV genomes were not obtained by the canonical AAV replication mechanism.

These results could open new prospects for using yeast cell in two ways: (i) as a model system for studying viral and cellular factors involved in AAV2 capsid assembly and packaging of rAAV ss genomes; and (ii) as a novel cell factory for developing superior recombinant rAAV production technologies.

Table of Contents

List of figures vi

List of tables vii

1. Introduction 1

1.1. Adeno-Associated Virus Biology and Life Cycle 1

1.1.1. Virus classification and structure 1

1.1.2. AAV2 genomic organization 2

1.1.3. Productive (replicative) and latent AAV virus infection 4

1.2. AAV-based vectors in gene therapy 9

1.2.1. Brief introduction to gene therapy 9

1.2.2. Properties of AAV-based vectors 10

1.2.3. Strategies for improving properties of AAV2 vectors 12

1.2.4. Historical overview of current methods for AAV-vector production 14

1.3. Budding yeast, *Saccharomyces cerevisiae*, as a cell factory for viral proteins and model organism for studying virus biology 17

1.3.1. Cell factories for recombinant pharmaceuticals. Why yeast? 18

1.3.2. *Saccharomyces cerevisiae* as a cell factory for VLPs 20

1.3.3. *Saccharomyces cerevisiae* supports replication of different RNA and DNA viruses 22

2. PROJECT I: Expression of AAV structural and non-structural proteins and assembly of virus-like particles in yeast *S. cerevisiae* 25

2.1. The aim of the project: 25

2.2. The background of the project 25

2.2.1. Critical components of AAV2 capsid assembly in the recombinant background 25

2.2.2. Regulation of Cap protein expression in the natural background 26

2.3. Results 27

2.3.1. Expression of AAV structural and nonstructural proteins from natural viral promoters and influence of Rep proteins on capsid protein expression in yeast *S.cerevisiae* 27

2.3.2. Inducible Yeast Promoter for Regulation of AAV2 Structural Protein Expression 35

2.3.3. Simultaneous expression of AAV2 capsid proteins and modulation of their stoichiometry toward successful capsid assembly. 38

2.3.4. Assembly of yeast-cell derived recombinant capsid proteins 44

2.4. Conclusions and future prospective 50

3. PROJECT II: Formation of AAV single stranded DNA genome from a circular plasmid in *Saccharomyces cerevisiae* 56
3.1. The aim of the project 56
3.2. The background of the project 56
3.2.1. AAV replication in the recombinant background 56
3.3. Results 56
3.3.1. Rep68 increases the frequency of yeast clones containing the AAV genome. 57
3.3.2. Expression of Rep proteins 58
3.3.3. AAV ssDNA formation in yeast 60
3.3.4. Expression of Adenovirus helper factors E1b55K and E4orf6 in yeast cells and their influence upon rAAV ssDNA formation in the presence of Rep 66
3.4. Conclusion and future prospective 68
4. To sum up… 73
5. Materials & methods 74
5.1. Strains, media, yeast transformation and cultivation 74
5.2. Plasmid construction 75
5.3. RNA isolation and Retro-Transcription 77
5.4. Analysis of protein expression. 78
5.5. AAV VLP extraction and purification 79
5.6. Transmission electron microscopy 80
5.7. DNA isolation and Southern blotting 80
References 82
Acknowledgements 102

List of figures

Figure 1.1: AAV genome organization (left) and Inverted Terminal Repeat (right). 2

Figure 1.2: AAV life cycle 5

Figure 1.3: Mechanism of AAV DNA replication. 6

Figure 1.4: AAV vector production and rAAV genome processing in the targeting cell nuclei 10

Figure 1.5: Different AAV serotypes target different groups of tissues.....12

Figure 1.6: Transient-transfection based protocols for rAAV production from HEK-293 cells.....15

Figure 2.0: Cap gene organisation. 26

Figure 2.1: Expression of AAV2 Cap and Rep proteins from natural promoters in yeast. .28

Figure. 2.2 A: Scheme of the plasmid Yep-Rep78p40Cap 30

Figure 2.2 B: Cap expression from p40 promoter in presence of Rep 78, expressed from the constitutive ADH1 promoter. 30

Figure 2.3: Distribution of AAV2 Rep proteins to soluble and insoluble protein fraction obtained by glass-bead extraction method.....31

Figure 2.4: Recovery of AAV2 proteins by ”Optimized post-alkaline” extraction.....33

Figure 2.5: Reverse Transcription (RT)-PCR-based analysis of p40mRNAs. 34

Figure 2.6: Expression of AAV2 structural proteins from the galactose-inducible promoter pGal1 36

Figure 2.7 A, B: VP1 & VP3 expression in double-transformed yeast cells, listed in Table 2.1. 39

Figure 2.8: The “short induction” strategy for preventing decrease of VP3 protein. 41

Figure 2.9 A: “High glu-high gal” - strategy for VP1:VP3 ratio optimization: 42

Figure 2.9 B: “Low glu-high gal” - strategy for optimization of VP1:VP3 ratio:.....43

Figure. 2.10: Concentrating VP assembly products by high-speed ultracentrifugation through 40% sucrose-cushion.....45

Figure.2.11. Purification of yeast cell-derived AAV2 capsid like-structures by CsCl-gradient ultracentrifugation. 49

Figure 3.1. Schematic representation of plasmids. 57

Figure 3.2: The frequency of colonies carrying rAAV genome increase when Rep68 is expressed. 58

Figure 3.3: Western blot analysis of the Rep protein expression in yeast. 59

Figure 3.4: AAV replication in yeast.....62

Figure 3.5: S1 and Mung Bean nuclease sensitivity. 63

Figure 3.6: Characterization of newly replicated DNA. 64

Figure 3.7: ssDNA formation is dependent on ITRs and Rep68. 66

Figure 3.8: Effect of the expression of E4orf -E1b55k on ss DNA formation. 67

Figure 3.9: Model of ssDNA formation from a plasmid containing rAAV genome in yeast. 70

List of tables

Table 1.1: S. cerevisiae cell factory for production of “virus-like” pharmaceuticals.....20

Table 1.2: Viruses that replicate in yeast.....23

Table 2.1: Constructs used to cotransfrorm RSY12 cells for simultaneous expression of VP1 &VP3.....38

Table 2.4. Comparison of the G/C content in the genomes: Total GC composition and the GC content at the 1st, 2nd, and 3rd codon position in each genome are listed as indicated55

1. Introduction

This work communicates two separate research lines that merge to open new prospective for adopting yeast *Saccharomyces cerevisiae* as a novel model-host organism for studying still elusive aspects of Adeno-associated virus (AAV) biology. The first part relates to using yeast cell expression system for production of AAV structural proteins from the cognate and heterologous yeast promoters and analyses the ability of these proteins to assemble in authentic virus-like particles (VLPs) inside of yeast cell. The second part investigates the aptitude of the *Saccharomyces cerevisiae* replication machinery to carry out the replication of a recombinant AAV (rAAV) genome in the presence of yeast-cell expressed AAV Rep 68 protein. Together, the results obtained from these two research lines evaluate the properties of yeast *S. cerevisiae* as a novel, putative production host for rAAV-based vectors.

1.1. Adeno-Associated Virus Biology and Life Cycle

1.1.1. Virus classification and structure

Adeno-Associated Virus (AAV) belongs to *Parvoviridae* family, a family of the smallest and the simplest viruses, whose single stranded DNA (ss-DNA) is enclosed in a non-enveloped icosahedral capsid. These viruses infect numerous species, starting from insects (subfamily *Densovirinae*) to highly evolved vertebrates, including man (subfamily *Parvovirinae*). The vertebrate subfamily is divided in 3 genera, Parvo-, Erythro- and Dependo-virus, whose members share similar genomic structure, but are distantly related in sequence (~15% identity) [1]. The first two genera, Parvovirus (Canine parvovirus and murine one, the minute virus of mice-MVM), Erythrovirus (human prototype B19) are pathogens that can replicate on their own in the host cell and hence, are called “autonomous”. By contrast, non-pathogenic Dependoviruses, with AAV as a type-species, are dependent upon “helper” virus co-infection for completion of their replicative cycle. In particular, the name “AAV” derives from first identification of this virus, nearly fifty years ago, in a form of a super-infection (contaminant or a satellite virus) of the cells infected with Adenovirus (Ad) [2-4]. Although Ad helper functions have been most extensively studied (reviewed in [5]), several DNA viruses are later shown to provide helper functions, such as herpesvirus, human papillomavirus and vaccinia virus [6-9].

To date, over 100 AAV variants have been isolated from adenovirus preparations (stocks) of human and non-human primate tissues while new ones are continuously emerging; around 14 serotypes (variants of a known serology) have been well characterized [10, 11] and 11 members recognized by the International Committee for Taxonomy of viruses, including AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV-6, avian AAV, bovine AAV, canine AAV, equine AAV, and ovine AAV (available at:

http://www.ictvdb.org/Ictv/index.htm). All these variants share similar virion structures and genomic organization. The virion shells of icosahedral symmetry (T=1) enclose single-stranded, approximately 4.7kb DNA genomes, comprised of two functional regions-terminal repeats and two viral genes (ORFs), rep and cap gene. Small, 60 subunit capsids, 20-25 nm in diameter, are made of three types of protein subunits, VP1, VP2 and VP3. For attachment and internalization, all AAV serotypes use ubiquitously expressed cell receptors , such as heparin sulfate proteoglycans-HSPGs (AAV2/3) [12-14] or sialic acid (AAV4/5) [15] (reviewed in [11, 16, 17]). However, important differences between serotypes consisted in the variations in amino acid compositions of their capsid proteins, result in the usage of different cell-surface receptors and coreceptors, causing different cell entry pathways and distinct tissue tropism among serotypes. For instance, both AAV1, AAV5 and AAV6, utilize N-linked sialic acid [15, 18] while O-linked 2,3-sialic acid serves as a binding receptor for AAV4 [15]. Ever since the first infectious clone of AAV serotype 2 (AAV2) was established in 1982 [19], its genome has been studied most thoroughly among all serotypes and AAV2-based vectors quickly gained great popularity in gene therapy applications. Its crystal structure has been recently published [20].

1.1.2. AAV2 genomic organization

For more than 25 years (since AAV2 clones were first established) all knowledge about AAV2 biology accumulated around the 4.7kb, single-stranded DNA AAV2 genome, comprising two large genes, *rep* and *cap*. All Parvoviruses share the same genome organization where the 5' *rep* gene, located on the left half of the genome, codes for nonstructural proteins, essential for viral replication, while 3' *cap* gene, the right half located, codes for structural proteins of the capsid (fig. 1.1 left). All AAV2 mRNAs terminates at a single poly-adenylation site at map unit 96 [21].

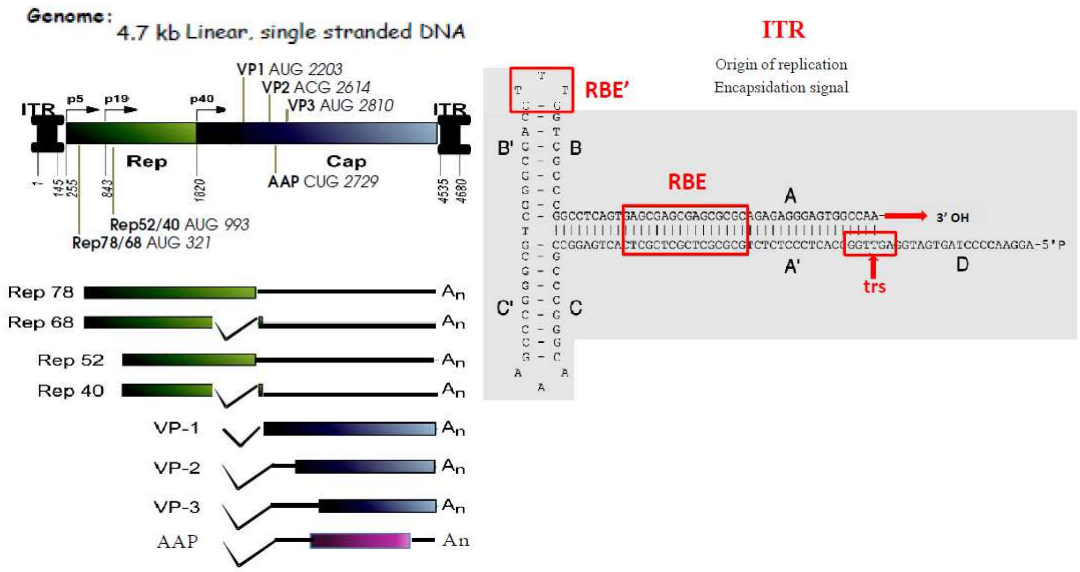


Figure 1.1: AAV genome organization (left) and Inverted Terminal Repeat (right).

Top-left part: The AAV-2 genome divided into 3 segments: rep, cap and ITRs and the nucl. positions indicated are: viral promoters, p5, p19 and p40, and translational start sites of Cap and Rep proteins. *Middle-left*: Rep proteins (boxes) and corresponding mRNAs (with polyA tails). *The bottom-left*: p40 encoded capsid proteins VP1, VP2 and VP3 and regulatory protein of capsid assembly, AAP, and corresponding mRNAs. *Right*: Secondary structure of an AAV-2 ITR showing the RBEs (RBE, GAGCGAGCGAGCGCGC and RBE', CTTTG) and the TRS (AGTTGG).

ITRs. AAV genes are flanked by 145 nucleotide-long inverted terminal repeats (ITRs). The first 125 nucleotides of the ITR have multipalindromic structure whose complementary regions folds upon themselves to maximize base pairing and forms aT-shaped hairpin structure, while the other 20 bases, called the D sequence, remain unpaired (fig.1.1- right). ITRs contain all *cis*-acting sequences necessary for packaging, replication, integration of the viral genome in the host cell DNA and for the subsequent rescue from the integrated state (reviewed elsewhere [22, 23]). In addition, it's been proposed that ITRs possess some transcriptional promoter activity, although relatively weak [24]. A key role of the ITR is that T-hairpin secondary structure provides 3'-hydroxyl group for the initiation of viral replication, thus being the origin of replication and serving as a primer for second-strand synthesis by a host DNA polymerase. This activity and the *cis*-acting ITRs' function in integration necessitate, *in trans*, activity of the two large Rep proteins, Rep78 and 68, which specifically bind at the Rep binding elements, RBE and RBE' of ITRs and nick in a strand- and site-specific manner (endonuclease activity) at the terminal resolution site, TRS located downstream of RBEs [22].

The rep gene = rep ORF, by the use of two promoters, p5and p19 (map units 5 and 19), encodes four nonstructural proteins, longer Rep78/-68 and shorter Rep52/-40, respectively, designated by apparent molecular weights. Rep68 and 40 differ from Rep78 and 52 due to an alternative splicing which replaces 92 residue amino-acid element in the later ones with 9 element in the former ones, while amino-acid sequence in the central portion of the gene is identical for all Rep proteins, translated from the same ORF (fig.1.1-left). Rep proteins are involved in a number of processes of viral life cycles , starting from the control of viral gene expression, over regulation of AAV-2 DNA replication (reviewed in [22, 25]). Large Rep proteins, Rep 78 and 68, perform the activities required for AAV DNA metabolism and site-specific integration [26-28] by means of their specific site-specific DNA binding and endonuclease activities as well as nonspecific ATP-ase and helicase activities, common to all four Rep proteins [29-33]. Large Rep proteins coordinate the viral gene expression mainly by transactivation of p19 and p40 promoter and repression of p5 promoter, both through the interaction with the RBEs in the TR and within the p5 promoter [34]. Repression of p5 by Rep68 and 78 is well described during productive infection, when the presence of helper proteins enable its derepression and consequential activation of the other two AAV promoters [34-37]. Conversely, during latent infection, AAV transcripts are not detected [38], suggesting that p5 repression may be attributed to the negative effect of some cellular factors such as YY1 [39]. Besides, large Rep proteins have been shown to affect the expression of various cellular and viral genes although the exact

level of action is often unknown [40-43]. Moreover, inhibition of the cell proliferation or S-phase arrest by Rep 78 is the activity mainly associated with the protein's ability to induce cellular DNA damage in the cell [44, 45]. Probably related to its ability to arrest the cell cycle, Rep has been shown to promote p53-independent apoptosis [46]. Finaly p19 promoter products, Rep 52 /40 play an essential role in the generation and accumulation of ss progeny genomes used for packaging in preformed capsids [47, 48].

Cap gene. For more than 25 years it was believed that AAV2 *cap* gene = *cap* ORF, where three capsid proteins VP1 (87 kDa), VP2 (73 kDa) and VP3 (62 kDa), with common C-terminus, are generated from the same p40 promoter and overlapping sequences in a single ORF (ORF1). Only recently, a new AAV2 gene product has been identified to derive from the alternative ORF2 of the cap gene, coding for a nonstructural viral protein involved in promotion of capsid assembly [49] (Fig.1.1-left, bottom part). The relative ratio of single protein subunits in mature virions is approximately 1:1:10 for VP1, 2 and 3 respectively, reflecting the intracellular stoichiometry of the three VPs inside the producing cell. This, in turn, is regulated on transcriptional and translational level. The two splicing events use the same splice donor site, but two different acceptor splice sites, among which the one closer to p40 promoter is poorly used, resulting in lower abundance of VP1 coding mRNA and thus, lower intracellular VP1 protein level. VP2 and VP3 are translated from the shorter mRNA whose greater abundance is due to the use of conventional splice acceptor site. Lower intracellular level of VP2 is the result of the poor translation initiation from uncommon ACG initiator, while high expression of VP3 protein owes to the usage of common AUG initiator in the favorable Kozak context [50, 51].

Although little is known about AAV capsid assembly *in vivo*, generally accepted concept of AAV virion assembly suggests the two step mechanism where empty capsids are firstly formed into which singles stranded DNA genomes are then introduced [52]. This concept is based on early studies which demonstrated that Cap protein localized in the punctuate nucleolar structures in the absence of Rep or AAV DNA [36, 53]. The nucleolus may provide factors/chaperones that help assembly process, the notion further supported by the discovery of two nucleolar proteins, nucleolin and nucleophosmin, found to associate with AAV capsids *in vitro* [54, 55]. The recent finding of Sontag et al. complemented this concept, showing that a newly identified AAV protein, AAP, not only stimulates transport of VP proteins to the nucleolus, but also, in some way, promotes the assembly process itself [49].

1.1.3. Productive (replicative) and latent AAV virus infection

Viral entry and trafficking to the nucleus. On their way to the nucleus where AAV2 has to deliver its genome in order to start replication or latent integration, virions have to pass multiple barriers: receptor binding, cell entry, intracellular trafficking, endosomal release, viral uncoating or nuclear entry. AAV2 infection on the cellular level starts with multiple contact with glycoprotein receptor (HSPG), the primary receptors for AAV2 binding, and

is further assisted by various coreceptors that are thought to be essential for its internalization and first post-internalization events (reviewed in [56]). These include: fibroblast growth factor receptor-1 [57], hepatocyte growth factor receptor [58], αvβ5 and αvβ1 integrin [12, 59] and the 37/67 kd laminin [60], that act as co-receptors. It's currently believed that virions enter the cell by endocytosis in a clathrin & dynamin-dependent process, facilitated by integrins and other coreceptors, and are rapidly transported toward perinuclear area via endosomal compartment: through early, late, recycling endosomes and lysosomes [56, 61]. When and how exactly, endosomal release occurs is still under investigation and is thought to be cell type specific [11]. However, recent studies have identified certain capsid components necessary for subcellular trafficking, thought to be prone to conformational changes and critical for passage from one cellular compartment to another. In particular, higher acidity of endosomes is thought to trigger a capsid conformational change which brings out to the capsid surface two functional regions located near the N termini of VP1 and VP2 (stretches of hydrophilic, basic regions-BRs), hidden in the capsid interior during assembly process. These elements, phospholipase A2 (PLA2) domain and putative nuclear localization signals (NLS) in BRs are thought to be crucial for vesicular disruption followed by endosomal escape and nuclear entry, respectively [62-65]. Although it's known that mutations in these regions have severe negative impact on AAV2 infectivity, their direct role in trafficking has not yet been demonstrated for AAV2. After endosomal escape, capsids can be either degraded by proteasomes or imported into nucleus. Although several groups have separately demonstrated that nucleus may be the place of AAV2 uncoating, there is still lot of contradiction regarding the mechanism and efficacy of virion/genome import into the nucleus, being the focus of current investigations (reviewed in [11]). Once inside the nucleus, AAV will undergo lytic or latent life cycle depending on the physiological state of the host cell, permissive or non permissive cellular environment for AAV productive infection as reviewed in the Figure 1.2.

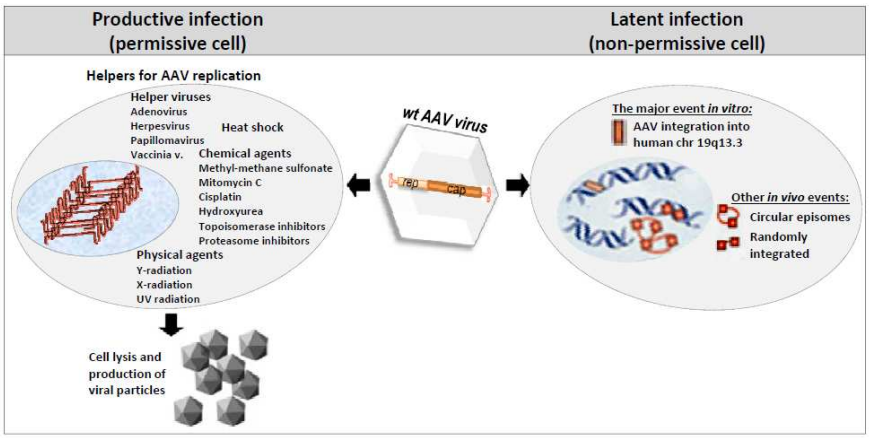


Figure 1.2: AAV life cycle

Replicative cycle. Infection with another virus (one of above listed AAV helpers) or genotoxic stress caused by various physical and chemical agents are the factors that determine the permissive cell-state that support conversion of single-stranded AAV2 genome to a duplex DNA. This is done by the host DNA polymerase acting on the 3' end of the viral ITR. The resulting incomplete double-stranded ds-monomer (the form with the priming ITR closed in a hairpin and the other one replicated) is transcriptionally competent thus yielding Rep78/68 proteins, which are necessary to bind RBE and RBE' of priming ITR and make a nick at TRS site, allowing ITR to unwind and the synthesis of ds-monomer can be completed (schematically presented in the fig. 1.3). Ds-monomer is essential molecular form both for viral protein expression and for second round of self-priming replication which will generate double-stranded DNA dimer molecules by strand-displacement mechanism.

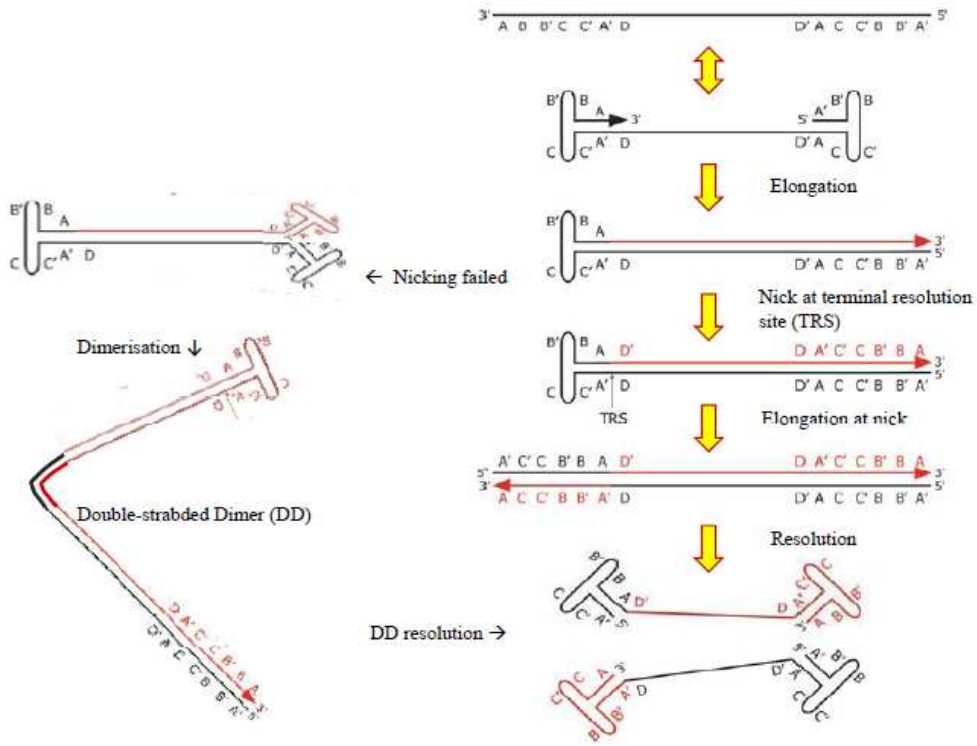


Figure 1.3: Mechanism of AAV DNA replication.

Letters indicate specific sequences in the inverted terminal repeats (AA', B-B', C-C', D-D' denote complementary sequences). Input viral DNA is in black and newly synthesized DNA in red. The arrow indicates the 3'OH available as a primer for DNA synthesis. TRS: terminal resolution site.

The two most abundant replicative intermediates, replicating-form RF- monomer and RF-dimer serve as reservoir of ss viral DNA which, at some point, is displaced from these templates, probably along with ss DNA encapsidation in preformed capsids. AAV capsids package both positive (+ polarity) and negative DNA strand (- polarity) with equal efficiency (reviewed elsewhere: [17, 22, 48]).

All the steps involved in AAV DNA replication have been reproduced in cell/free systems using over-expressed recombinant Rep proteins [66, 67]. These studies further strengthened the concept that in the presence of Rep, AAV replication can rely mostly, if not uniquely, on direct help from cellular factors, since purified cellular proteins, such as replication protein A (RPA), replication factor C (RFC), proliferating cell nuclear antigen (PCNA), minichromosome maintenance (MCM) proteins, and DNA polymerase δ (Pol δ), were sufficient to replicate the AAV genome in vitro. Studies conducted on the contribution of helper viruses to AAV replication are important not only to identify helper activities that can be used to produce recombinant AAV vectors but also to understand how AAV adapts its replication strategy to the helper virus and to the nuclear environment in general.

Adenovirus (Ad) helper functions have been most extensively studied among AAV helpers and four of its proteins: E1A, E1B, E2A, E4orf6, as well as one virus-associated RNA (VAI RNA), have been identified as indispensable for AAV replication. These Ad functions help AAV productive life cycle by stimulating viral gene expression and by enhancing AAV genome replication, mainly indirectly (reviewed in [5]). The E1A gene product relieves repression from AAV p5 promoter, probably through direct binding of the YY1 element [39]. Similarly, the Ad DNA-binding protein (DBP), the E2a gene product, was reported to be able to activate transcription from the p5 promoter [68]. Moreover, early studies showed that Adenovirus DBP protein is able to modestly enhance the processivity of AAV genome replication in vitro [69]. The E4orf6 gene has been shown to be essential for replication on several levels. First, it was demonstrated that its product can overcome the rate-limiting step of second-strand synthesis of the single-strand virus genome [70]. More recently, it was shown that the adenovirus proteins E1b55k and E4orf6 can stimulate AAV genome replication by degrading the cellular Mre11/Rad50/Nbs1 (MRN) complex that restricts AAV genome replication during adenovirus coinfection [71,72]. The function of the virus-associated RNA in AAV DNA replication remains somewhat elusive and has been suggested as a translational enhancer of AAV RNAs [73].

Herpes Simplex Virus type 1or 2 co-infection with AAV results in AAV titers similar to those obtained in the presence of Ad and the very first studies in this field also indicated that this virus enabled more rapid AAV life cycle than the one with Ad [5, 6], which was later confirmed in experiments with recombinant AAV plasmids [74]. Early studies showed that the HSV-1 helicase-primase (HP) complex (UL5/8/52) and DBP (ICP8) were sufficient to replicate rAAV-2 plasmids [75] and it was later shown that the helicase activity, but not primase activity, of HP complex was required for this effect [76]. More recently, a comprehensive study of HSV-1 helper activities demonstrated that the HSV-1 immediate-early proteins ICP0, ICP4, and ICP22 could stimulate rep gene expression and that HSV-1 DNA polymerase encoded by UL30, along with its associated processivity factor (UL42), although not strictly required, significantly increased AAV replication

levels induced in the presence of the HP complex and ICP8 [9, 77]. Only recently, a novel HSV1 protein has been shown to be directly associated with Rep proteins within AAV replication centers (RC), further demonstrating that this viral exonuclease plays a critical role during AAV replication by enhancing the formation of discrete AAV replicative forms, suitable for packaging, thus increasing the titer of infective AAV particles [78]. Having identified five HSV-1 factors in direct association with Rep proteins opposing to a single adenoviral helper protein [79], this study also points out a more direct role of HSV-1 in AAV replication than that of adenovirus, indicating that AAV may be able to differentially adapt its replication strategy to a nuclear environment induced by a helper.

Latent cycle. Under non-permissive conditions (in the absence of helper functions or stress) AAV genomes that reached the nucleus of the infected cell have the capacity to establish long-term latency by integrating into a specific locus, accordingly named AAVS1, on chromosome 19 (19q13.3-qter) (see fig. 1.2, right half). Upon nuclear entry, initial conversion of ss AAV genomes to ds forms enables basal expression of Rep 78 and 68, the only viral trans acting factors essential for integration, who form the bridge between ITR and p5 promoter- located Rep Binding Sites (RBS) on one side and the homologous sequences located in the AAVS1 locus on the other [22, 48, 80]. Various RBS-like elements have been identified in human genome, but the vicinity of RBE-like and TRS-like region in 33 bp region of AAVS1 make it almost identical to the analogous regions in ITR which probably enable viral integration via semi-homologous recombination between the two regions (reviewed in [17]). Another proposed model for imprecise integrational mechanism is based on simultaneous replication of AAV genome and AAVS1 region, both mediated by AAV large Rep proteins with intrinsic tendency of a Rep-mediated replication complex to switch template strands. In any case, if latently infected cells are subsequently infected with a helper virus, the AAV genome is rescued and re-enters the replicative phase of the life cycle [81]. However, recent studies with rAAV plasmids have shown that in contrast to AAVS1, replication at the viral terminal resolution site does not appear necessary for the integration event to occur [82] and the viral sequences required for site-specific integration has been identified. This element, designated p5IEE for ‘p5 promoter sequence residing integration efficiency element’, is the sole cis requirement for high-efficiency integration of AAV genome into AAVS1 [83]. Although approximately 0.1% of infecting wt AAV genomes integrate at AAVS1, this mechanism could nonetheless play a role in the natural history of the virus, and AAV proviruses have been identified in AAVS1 after natural infection [84]. The wild-type AAV genome can also integrate at sites in the genome other than AAVS1, as demonstrated by the identification of a provirus on chromosome 1q31.1 in human tonsillar tissue and in multiple different chromosomes *in vitro* [85, 86]. These events were presumably Rep-independent and thus analogous to the integration of rep-deficient vectors described in the next section. Remarkably, latent AAV genomes have been detected in various tissues derived from humans and non-human primates, including bone marrow, brain, spleen,

colon, heart, liver, lymph node and kidney (reviewed in [87]). While the presence of these latent genomes is consistent with integrated proviruses, in some cases episomal genomes may also persist for prolonged periods [88]. A simplified scheme of the latent AAV infection is shown in the figure 1.2, right part.

1.2. AAV-based vectors in gene therapy

1.2.1. Brief introduction to gene therapy

Gene therapy uses variety of nucleic acids as therapeutic agents ("genes as drugs"). Depending on the desired function and the disease profile, the therapeutic can be in a form of protein coding DNA /non-coding DNA / RNA or analogous molecules, such as: mRNA targeting antisense oligos, siRNA, decoys, ribozymes etc. Their usage is wide, starting from prevention (e.g when used as adjuvants for vaccination), over gene augmentation in acute treatments, to compensation for "genetic loss-off function" in chronic diseases. Moreover, therapeutic genes can provide cytotoxicity in cancer cells or be used for correction of endogenous genes, which presents extremely important branch of gene therapy yet, still limited to experimental studies in cell-culture. In each gene therapy approach the choice of a therapeutic gene (transgene), a target cell/organ and a delivery route have to be matched in the best way to provide maximal **efficiency, specificity and persistence** of gene transfer, while **minimizing toxicity** associated with: gene administration method and/or the transgene itself. Gene therapy uses three major routes of delivery: 1. *ex-vivo*, gene transfer is performed in laboratory, in isolated hematopoietic or stem cells, followed by re-administration of "transgenic" cells back into the patient; 2. *In-vivo* systemic delivery (intra-venous/arterial/peritoneal); and 3. *In-vivo* topical delivery, directly in tissues hardly reachable by systemic delivery (e.g brain, eye, joints) or where systemic delivery could cause unwanted effects (e.g tumors). Naked nucleic acid therapeutics are generally difficult to deliver and are destined to rapid elimination from the circulation for couple of reasons: rapid clearance by serum nucleases, lack of organ-specific distribution and low efficiency of cellular uptake. The use of virus-based vectors is currently the best way for improving delivery efficiency and long term transgene expression; while most gene therapy strategies are coalescing around two types of viral vector: lentivirus vectors for *ex vivo* gene transfer and AAV vectors for *in vivo* transfer into postmitotic tissues [89]. Nevertheless, these viral vectors have various limitations: i) inherent packaging constraints limit the size of the therapeutic gene; ii) safety concerns related to: immunotoxicity (harmful immune response) of capsids, envelope proteins, residual viral genes and risks associated with high probability of mutagenic integration (retroviral & lentiviral vectors); iii) transcriptional silencing of integrated vectors leading to reduced transgene expression over time. In addition, even weak immune responses generated against vectors can dramatically decrease transgene expression upon re-administration. Similarly, pre-existing (memory) immunity against common viruses in

human population, such as Ad and AAV is one of the main limitations of these vectors. Thus, it is imperative to develop efficient producing and engineering technologies to render the vectors capable of escaping immune recognition and avoid inflammation.

(Literature used for this section: [17, 90-93])

1.2.2. Properties of AAV-based vectors

Several marked features of AAV life cycle such as the ability of these viruses to infect non-replicating cells, capacity to integrate its proviral DNA into a specific site of the human genome and almost non-immunogenic nature of AAV infection, indicate how well these viruses adapted to long-term persistence and coexistence with the host cell without causing any deleterious effect. These aspects continue to inspire still growing use of AAV viruses in gene transfer technology. The first successful cloning of AAV2 in early 1980s led to more than two decades of constant effort in improving recombinant technologies for AAV-based vector production followed by extensive studies of unknown aspects of AAV biology (particularly serotype 2). So far, AAV vectors have shown therapeutic efficacy in a range of animal models (mice, pigs, dogs, ship etc.), owing to a number of properties which make them preferential with respect to other viral vectors. First of all, these vectors are "gutless", constructed by complete deletion of all viral genes, rep and cap, and insertion of "therapeutic" gene (gene of interest) between ITRs, which are the only viral *cis*-acting element indispensible for replication and packaging or recombinant AAV genomes, as long as rep, cap and helper functions are provided *in trans* by the packaging cell [94] (see fig. 1. 4, left part). Since only 300 nucleotides of viral DNA are retained in the vectors, the risk of their recombination with wt virus is minimal. Further, such "gutless" AAV vectors, deprived of the previously mentioned *cis* integrational element and *trans* acting integrational factors (p5IEE element and rep68/78 proteins, successively) are not capable of integrating their genome site-specifically. Once they enter the nucleus of a targeting cell, these vectors may persist in couple of molecular states (Fig. 1.4 - right part).

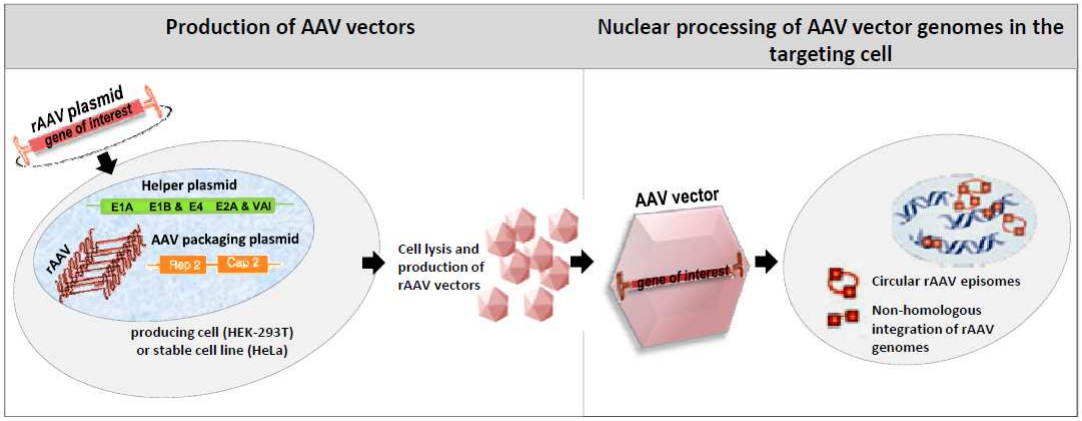


Figure 1.4: AAV vector production and rAAV genome processing in the targeting cell nuclei

One common outcome is the conversion of the AAV genome to a double-stranded circular episome either by second-strand synthesis or complementary strand pairing [95, 96]. These episomes can be further converted to high-molecular-weight multimeric concatamers [97], formed by the recombination of monomer genomes [98], and probably are the main source of long-term transgene expression, particularly in non-dividing cells [85, 97]. Such a transductional mechanism of AAV vectors almost eliminates the risk of insertional mutagenesis or silencing, seen for integrating retroviral vectors. Accordingly, in the absence of the immune response to a therapeutic gene, or the pre-existing immunity to common AAV serotypes, transgene expression from AAV vectors in vivo can last for months or years. Such a success was specifically documented in AAV-based gene transfer approaches in animal models [99], which lack memory immunity to most of the human AAV serotypes used in these approaches. However, a large body of data demonstrates that AAV vectors can also integrate at non-homologous sites in the host genome, both *in vitro* and *in vivo*, either as a single-copy proviruses or concatamers [84, 87, 100]. More recent studies focused on analysis of large number of vector : chromosome junctions (e.g. Inagaki et al. characterized approx.1000 integration sites from liver, heart and skeletal muscle [101]) and provided insights into AAV vector integration process. The integration of AAV vectors occurred preferentially at specific sites, 'hot spots', in the genome such as: ribosomal DNA repeats that encode ribosomal RNAs [102, 103], CpG islands and within 1kb of transcription start sites [101-103]; segmental duplications, satellite DNA and palindromes [101, 102]. Moreover, in mouse hepatocytes, around 60 % of AAV vector integrations occurred in active genes [103]. These integration hot spots may represent regions of genomic instability and together with the evidence that AAV vectors integrates at chromosomal double-strand breaks [104] (generated by the endonuclease I-Sce I, or by treatment with etoposide and γ –irradiation), suggest the model of AAV vector integration. According to Deyle and Russell, these data suggest that the nonhomologous integration of AAV vectors is not a random event and that the preferential integration at certain sites presumably reflects availability of free chromosomal ends that can ligate to AAV, such as regions prone to double-strand breaks or other forms of DNA damage [87]. Even though a rare event, the nonhomologous AAV vector integration can be associated with chromosomal deletions and rearrangements [105], where the most concerning, possible genotoxic effect is the malignant transformation of a transduced cell. The majority of experimental data in animal models have demonstrated that AAV vectors are safe in this regard. Still, some convincing evidence for AAV-induced tumorigenesis derives from a follow up study with mice with mucopolysacharidosis VII (resulting from mutations in β-glucuronidase) injected with the vector carrying human β-glucuronidase gene [106]. Both mutant and non-mutant mice injected with the same vector had an increased incidence of hepatocellular carcinoma (~50% compared with ~8% in controls) and showed the occurrence of four independent tumors containing vector proviruses in the same locus (but not in the surrounding normal liver tissue) suggesting that insertional mutagenesis by AAV

vectors was the cancer trigger, although these loci were not previously associated with hepatocellular carcinoma. Besides, AAV vectors also integrate into tissues other than the liver, heart, skeletal muscle [101], thus pointing to the need for further research to fully assess the risk associated with AAV vector genome integration into the host genome. (Other literature used for this section: [16, 17, 80, 107].

1.2.3. Strategies for improving properties of AAV2 vectors

Albeit transduced with different efficiencies, skeletal muscle, hart, liver, brain, retina and lungs are the main targets for AAV2-based gene transfer approaches, while many cell types and relevant diseased tissues still remain non-permissive to transduction by AAV2 based-vectors (and vectors based on other serotypes). The reasons for these differences in tissue transduction efficiencies go beyond the receptor choice and internalization and reflect cell-type related differences (limitations) at various points essential for efficient transgene expression, such as : viral subcellular trafficking to the nucleus, nuclear entry, uncoating and second-strand synthesis (reviewed in [17, 80]). The last limitation was partially overcome with development of **self-complementary (sc) AAV vectors**, which once in the nucleus can fold upon themselves and elicit rapid expression of a therapeutic gene without the need for inefficient second-strand synthesis [108]. Moreover, packaging of AAV2 based-recombinant genomes into capsids of other serotypes, which show different tissue tropism from the one of AAV2 serotype, **"pseudotyping technology"** (fig.1.5), can meliorate efficiency of transduction for some tissues, but only slightly widen the range of AAV2 permissive tissues [17, 109, 110].

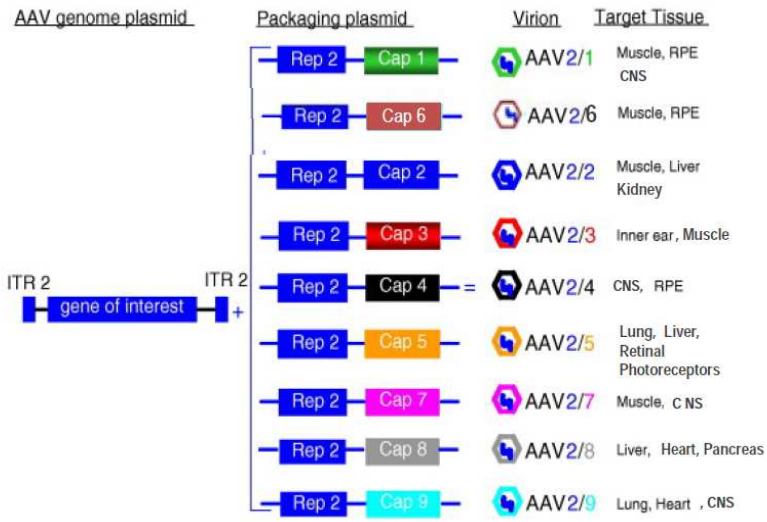


Figure 1.5: Different AAV serotypes target different groups of tissues.
Information taken from www.aaveye.eu and Ref. [17]

Beside pseudotyping, other similar production technologies were developed in order to overcome transduction limitations of AAV2 vectors and further expanding tissue tropism.

These technologies tend to synergistically combine properties of naturally occurring serotypes and further improve them by introducing certain modifications (e.g. targeting peptides), leading to generation of **mosaic, chimeric and targeting vectors (direct and indirect)**. The mosaic vectors are composed of mixed capsid subunits of various AAV serotypes [111], while chimeric vectors are produced by domain swapping approach among multiple parental serotypes, involving either entire capsid loops or parts thereof or individual residues [112, 113]. For example, domain swapping between serotypes 1 and 2 yielded a chimeric vector who could be purified by heparin affinity chromatography (purification method designed for AAV2 serotype), had increased transduction efficiency of muscle cells (AAV1 property) and was less efficiently recognized by anti-AAV2 neutralizing antiserum [113]. Furthermore, **genetically engineered capsids** can lead to more restrictive targeting to a diseased tissue. Numerous groups have developed novel strategies to engineer "designer" AAVs tailored for better transduction of clinically relevant organs (reviewed in detail in ref [11, 114, 115]) These strategies can be grouped in "direct", physical capsid modifications approaches or "indirect" approaches that base on chemical binding of targeting molecules (ligands) to the viral surface, forming a conjugate ideally able to retarget the capsid to a refractory cell type. Examples of such ligands are: bispecific antibodies [116], avidin-coupled ligands [117] or cellular receptors. In the direct targeting approach, targeting ligands are genetically inserted into the viral capsid proteins [118, 119]. The insertion site has to be chosen in a way to avoid interference with assembling, packaging and infectivity properties of modified virions [120]. In this view, N-terminus of VP1 and VP2 was shown to be a good choice for such modifications [121]. In addition, conferring selective tissue targeting often demands elimination of broad tissue tropism, since a considerable safety issues are raised if non-target tissues are transduced. It's been observed that systemically routed AAV vectors accumulate in the liver and spleen (mainly associated with HSPG receptor recognition of AAV vectors), thus reducing the transduction efficiency in non-liver target tissues [122, 123]. Some success in combining vector de-targeting from liver and spleen was achieved by insertion of targeting ligands in positions of or close to HSPG binding domains [124, 125].

The aforementioned example of domain swapping between AAV1 and AAV2 serotypes demonstrate that the use of different AAV serotypes present one way for avoiding vector clearance by pre-existing anti-AAV2 immunity to vector capsids. However, results in humans indicate that AAV capsid-primed CD8+T cells are likely to cross-react with alternate serotypes thus urging further development of new strategies for circumventing immunotoxicity [126]. One of possible strategies is introduction of mutations in capsid epitopes recognized by anti AAV2-neutralizing antibodies (A20). Mutations are either results of clever guesses [127] or the use of **"evolutionary strategies",** where combinatorial libraries of mutated cap regions are subjected to high-throughput screenings for selection of desired characteristics [125, 126, 128].

(For detailed reviews of all aforementioned strategies see Ref: [10, 11, 80, 126]).

Two inherent limitations of AAV vectors are the small, aprox.5kb, packaging capacity of the capsids and infrequent integration in quiescent target cell-genomes. The first limitation can be overcome, in a part, by skilful engineering of transgenes and their regulatory regions (minimizing the size) and by use of trans-splicing approach for expression of transgenes greater than 9kb [129]. Although rare and mostly related to chromosomal rearrangements, integration events point to the need for fully assessing the associated risk in the future. Finally, long-term expression of the therapeutic level of transgene is the major challenge associated with clinical usage of AAV vectors. As indicated from large number of AAV -based pre-clinical and phase I (less phase II and III) clinical trials, these challenges are tissue-specific and embrace cell-type dependant variables related to administration rout and predominant immune responses directed to the vector or to the transgene product [130].

1.2.4. Historical overview of current methods for AAV-vector production

Extensive studies on AAV-vectors in the last decade resulted in a variety of currently existing methods for their production. However, the rising number of AAV vector-based gene therapy trials that require high vector doses, over 10^{13} genome copies (g.c.)/kg of body weight [131], impose the need for creating new production technologies with better "yield versus cost" parameters.

When the first foreign gene was expressed from AAV2 vectors in mammalian cells in the early 1980's, the stocks of virions carrying recombinant AAV2 genome were obtained by **transfection of Ad2 infected HEK-293 cells** (a human Adenovirus E1A and E1B genes - transformed cell line) with two bacterial plasmids, one carrying rAAV genome and the other one providing missing cap gene in trans [132]. Soon later, concerns surrounding the use of replication-competent vectors lead to elimination of rep gene form recombinant genomes and providing it *in trans* together with cap gene, while infection with Adenovirus was still used for providing helper functions. Frequent contamination of AAV2 stocks with Ad urged cloning of Ad genes into bacterial plasmid, thus eliminating infection [94, 133]. Concerns regarding generation of replication competent vectors were reduced by placing *rep* and *cap* genes in opposite orientations. The **triple transfection-based** AAV production method was the subject of further improvements associated with extensive studies on minimal helper function requirements, leading to currently used protocols for HEK-293T production system which involves following components: (1) vector containing recombinant rAAV genome - the transgene expression cassette flanked by two ITRs (AAV2 derived), (2) plasmid providing Rep and Cap proteins *in trans*, and (3) plasmid containing adenovirus genes encoding E1, E2A, E4, and virus-associated RNA (VA-I RNA) comprising minimal requirements for inducing cell permissivity to productive AAV replication [5]. The next upgrading introduced to this system was placing AAV2 rep and

cap genes together with Ad-helper genes in a single "helper" plasmid, thus diminishing number of transfecting constructs from three to two (fig.1.6). The same procedure is used for production of "pseudotyped" AAV2 vector genomes; instead of AAV2 cap gene sequence, the helper construct comprises of serotype-specific *cap* and AAV2 *rep* expression cassettes [134].

Vector generation in a producing cell involves following steps: Rep and VP proteins are expressed from helper plasmid and the type2-ITR flanked transgene cassette is rescued from the vector genome and replicated in Rep78/68 dependent manner followed by encapsidation of ssDNA into preformed capsids of desired serotype (reviewed elsewhere [17, 135]). The vector yield from transfected HEK293 cells was further improved with fine tuning of Rep78-Rep52 expression level with better results obtained when Rep78 was inferior to Rep52. Although advanced, the productivity of transient transfection-based method (titers of viral stocks obtained) was still limited by transfection efficiency of the cell line utilized. So far, the best results are obtained with HEK293 cells when calcium phosphate precipitation is utilized for transient transfection (reviewed in [135, 136]).

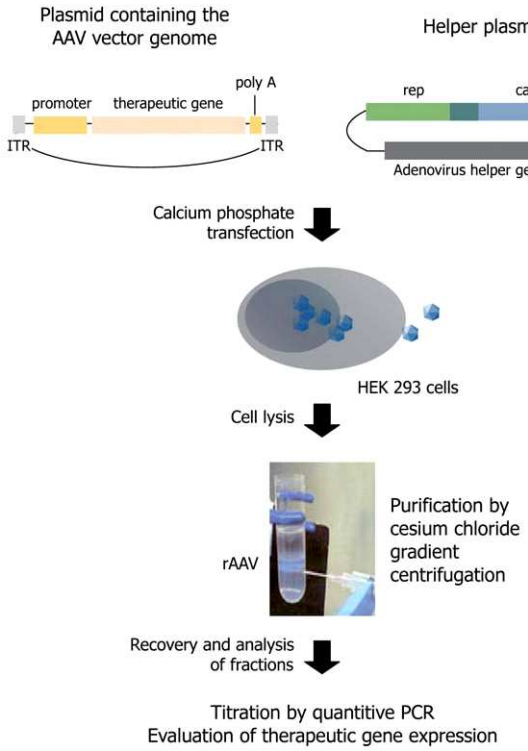


Figure 1.6: Transient-transfection based protocols for rAAV production from HEK-293 cells. The system presented is using two plasmid components: 1. AAV vector itself, harboring transgene expression cassette, flanked by (ITRs); plasmid encoding for: Rep, Cap and adenoviral proteins providing helper functions. Twenty-four hours after transfection, cells are lysed and the vectors are purified by standard cesium chloride ultracentrifugation. Figure reproduced from ref.[17].

To overcome transfection limitations, **infection-based protocols using recombinant viral vectors** are developed. Packaging cells (employing different cell lines) are infected with

Adenovirus, Herpesvirus or Baculovirus-based vectors (reviewed in [135]), carryng both rAAV genome and helper functions. Although high rAAV titers were obtained with Herpesvirus - Baby-Hamster-Kidney cell line (BHK) double-infection system [137], this method is still not widely accepted. On the contrary, recombinant Baculovirus infection of evolutionary distant insect cells has gained lot of popularity since originally established [138]. The original system consisted of three different baculoviruses: the 1st one carrying rep genes, the 2nd one-cap genes and the 3rd one carrying ITR flanked transgene.

Finally, great effort has been made for creating **"packaging" and "proviral", stable cell lines** (mammalian), considered as the best systems for scaling up manufacturing procedure. Packaging cell lines contain rep and cap genes integrated in the host genome, while the second one contain integrated recombinant AAV genome. The best currently existed stable-cell line systems are based on HeLa cell-lines, but BHK and 293T cells have been also successful so far. Opposing the HeLa packaging cell line, the major difficulty in generating a 293-based AAV producer cell line is the E1A mediated activation of AAV p5 promoter, which control AAV Rep78/68 proteins, which are known to be cytotoxic if constantly expressed [46]. Thus, this difficulty is overcome by creating cell lines with highly inducible Rep systems [139-141]. However, a main limitation which renders stable-cell line approach insufficiently efficient is the multi-step and time-consuming procedure. Due to multiple steps of transfection and/or infection and selections, several weeks to months are needed to produce a high-yield cell line [135, 136, 140, 141].

Since AAV particles are very resistant to harsh manipulation and solvents, **conventional methods for purification of AAV vectors** are based on cesium chloride density gradient ultracentrifugation, which provides efficient separation of empty capsids from genome containing virions [17] (Fig.1.5-bottom). Still, it's been observed that longer exposure of particles to CsCl can cause a certain decrease in vector infectivity, leading to development of alternative purification strategies. Zolotukhin et al. described the use of nonionic iodixanol gradients followed by ion-exchange or heparin-affinity column chromatography for the purification of AAV2 and other serotypes able to bind heparan sulfate [142].

So far the most commonly used rAAV production method, based on transient transfection of mammalian cells, holds one crucial limitation: **the use of adherent cell cultures** which make this methods difficult to scale-up. Growing vector requirements for large animal-preclinical and human-clinical trials practically cannot be met with the use of adherent cell-solid supports, such as tissue culture plates (around 50-15 cm plates needed for 293-cell production of 10^{14} rAAV particles with current transfection methods). This issue is being addressed with new **technologies using shake flasks and bioreactor-based suspension cultures** of HEK-293 cells, suitable for scaling-up, although expensive and manpower demanding [143, 144]. Aside of great improvements achieved in traditional production in mammalian cells, **Baculovirus-Sf9 suspension culture system** has been the subject of active optimization in the last years and hence, developed into a reliable and

scalable process. Although producing as much rAAV particles per Sf9 cell as per mammalian cells, the original system was mainly limited to small-scale production for two reasons: (i) it required high MOI of each of the three Baculoviruses (Bacs) for efficient co-infection and (ii) proliferation of infected cells had to be limited due to high instability of rep Bac helper and the consequent loss of Rep protein expression following passage 2/3 [145]. These deficiencies were alleviated with optimal expression of Rep78 and Rep52 from a single polycistronic transcript and placing the Cap ORF into the same baculovirus expression vector (BEV), in the opposite orientation with respect to the Rep cassette [146]. Stabile expression of the VP and Rep proteins from consolidated "RepCap"- Bac was observed through at least six amplification passages. This enabled couple-fold increase in the cell-density of suspension cultures, while keeping the same particle per cell yield. Further advances involved optimization of feeding and growth conditions as well as improvements in processing protocols for efficient recovering of rAAV vectors from large volumes and biomass (40-100l bioreactor cultures). Currently, Bac-Sf9 system can meet requirements of commercializing rAAV production, offering lower production costs, high productivity and relatively easy scalability [147, 148].

Efficient Rep78-dependent rAAV replication and production of biologically active vectors in invertebrate cells inspired us to explore the permissivity of the most simple, eukaryotic, single-cell organism, the yeast *Saccharomyces cerevisiae*, for rAAV genome replication on one side and AAV virus-like particles (VLPs) assembly on the other. In view of the greater complexity of the cell biology and genetics of metazoans, yeast has already demonstrated its usefulness for virus research, which will be discussed in the following section together with the advantages of yeast-based heterologous expression system over the other currently applicable systems for foreign protein production (viral proteins in particular).

1.3. Budding yeast, *Saccharomyces cerevisiae*, as a cell factory for viral proteins and model organism for studying virus biology

So far, structural and nonstructural viral proteins have been expressed from practically all recombinant cell-expression systems, from bacteria to human cells. The producer cell choice depends on the characteristics of foreign protein/protein complex (e.g. VLPs) on one hand and from downstream applicability of the produced protein/VLPs on the other, such as:

1) Basic research (e.g. Structure–function analysis [149]); 2) Biomedical applications: diagnostic research and seroepidemiological studies employing various pathogenic viruses; production of vaccines and vaccine candidates, either in a form of surface viral protein or peptide or multisubunit VLPs (in detail in the following sections, table 1.1); 3) production of VLP for nanotechnology applications (e.g. nanoparticles for DNA and drug delivery [150]).

While small scale production satisfies the needs of basic research, commercialized biomedical applications, designed for use in humans and animal models, demand usage of safe, easy to scale up and possibly, low-cost, recombinant expression systems.

The first part of this section addresses some of the major advantages of the high yield-yeast expression system over other heterologous systems, underlying the notable importance of this organism in the field of vaccine production, mostly based on efficient VLPs assembly in yeast cellular background. The second part provides a brief overview of the *S. cerevisiae* permissivity to replication of many lower and higher eukaryotic viruses and of a further use of this phenomena in elucidating still unknown aspects of viral life cycles. Together, they suggest the potential of extending current usage of yeast-cell factories to production of viral vectors for gene transfer approaches.

1.3.1. Cell factories for recombinant pharmaceuticals. Why yeast?

Since the beginnings of recombinant DNA (rDNA) technologies in the late 70's, large set of methodological platforms have been developed for production of foreign proteins of an important therapeutical, industrial and research values in both eukaryotic and prokaryotic systems. The first recombinant protein-based pharmaceutical to enter the market in early 80's (Food & Drug Administration of USA (FDA) approved) was human insulin derived from recombinant *Escherichia coli. E.coli* was the first microorganism used for production of recombinant therapeutics saving uncountable lives, albeit, this expression system was often facing obstacles related to frequent aggregation of overexpressed proteins, thought to be associated with aberrant protein folding and/or stress response of the host to the foreign product [151]. The attempt to produce recombinant Hepatitis Virus B small (s) surface antigen-sHBAg in *E.coli* expression system failed due to instability of the protein and its deleterious effect upon the host cell. In addition, this first recombinant subunit vaccine was abortive since it failed to induce immune response in "immunized" test animals [152]. Soon after, the platform of heterologous systems for expression of recombinant pharmaceuticals incorporated: bakery yeast *Saccharomyces cerevisiae,* other less common yeast species (*Pichia pastoris, Hansenula polymorpha*), mammalian and insect cells and more recently, transgenic plants and animals adapted to the same purpose. The preference in using bacterial and yeast cell-factories over the other expression platforms is hold, on one hand, in a simple microbial cultivation and instrumentation procedures, and, on the other hand, in a genetic flexibility of these microorganisms, apt to engineering and manipulation [151]. Stable, autonomously replicating plasmid-based protein expression from high-density microbial cell cultures enabled inexpensive, controllable and highly scalable production of polypeptides of interest. However, the aforementioned drawbacks of *E. coli* expression platform mainly consisted in its inability or inefficiency to authentically process foreign proteins and introduce certain post-translational modifications (PTMs) present in most proteins processed by mammalian cells (e.g. glycosylation or phosphorylation) [153, 154]. Among PTMs, glycosylation is the most common and

important one, playing a crucial role in protein folding, processing, stability, final biological activity, tissue targeting, serum half-life and immunogenicity of the protein. Therefore, failures of bacterial cell in e.g disulfide bond formation, correct proteolytic processing and/or introduction of correct PTMs, may result in a production of insoluble, unstable or inactive proteins. Advances in genetic engineering enabled some of these hurdles, such as protein misfolding and aggregation, to be overcome by stable introduction of various chaperons and foldases [154, 155]. Otherwise, yeast is generally considered advantageous microbial cell-based expression system since its eukaryotic intracellular environment is suitable for the most of authentic posttranslational processing events, typical of higher metazoans (and most importantly of mammalian cells); whereas the low-cost production from this organism is the advantage over higher eukaryote expression systems [156]. Furthermore, the long history of *S. cerevisiae* industrial applicability in brewing, winemaking and baking demonstrate the absolute absence of hazardous pyrogens, pathogens or viral inclusions in the final products obtained from these cells. Accordingly, this microorganism has been classified as a GRAS (generally regarded as safe) [157, 158]. The era of its successful utilization in pharmaceutical biotechnology started already in the early 80's, when the first commercialized recombinant vaccine against Hepatitis B (sHBAg) was made in *Saccharomyces cerevisiae*, soon after the production in *E.coli* failed [159]. Its wide utilization has been further improved with genetic modifications of selected *S. cerevisiae* strains, facilitated by the availability of its complete genome sequence, published in 1996 [160]. Subsequent sophisticated approaches of genetic engineering enabled manipulation of entire yeast metabolic pathways performing gene deletions or introductions of single/multiple genes, creating "improved" strains with novel characteristic [158]. For example, addition of humanized N-glycans of the intermediate mannose type or even the complex types provides the option to produce biopharmaceuticals with human protein modifications [161]. In the last decade, heterologous expression systems for biopharmaceutical production based on *S. cerevisiae* and other yeast species are given the highest commercial value [162]. In a big part this is because *S.cerevisiae* is an extensive tool-kit of plasmids for high-level expression from large number of wild type (wt) and artificial, regulative and constitutive yeast promoters, of different strengths and substrate dependences [163, 164]. Besides, new regulation systems continue to be added to the existing expression platform and together with a range of engineerized "gracious strains" [165] continue to offer more refined transcriptional- to -posttranslational control. Shortly, yeast cell allows refined control of "when" (e.g. triggering expression by means of inducible promoters), "where" (intracellular accumulation vs. secretion) and "how much" (easily scaled-up expression) of the recombinant product will be obtained.

1.3.2. *Saccharomyces cerevisiae* as a cell factory for VLPs

Since yeast derived HBV vaccine was certified for human use in early 80's, high-yield production from S. cerevisiae system has been reported for large number of viral structural proteins derived from both prokaryotic and eukaryotic viruses. The viral biopharmaceuticals include: 1. surface proteins, useful candidates for recombinant subunit vaccines but also used for structural analysis [156]; 2. viral core and nucleocapsid proteins, mostly used for diagnostic tests and seroepidemiological studies; 3. authentic and chimeric VLPs, which are of the greatest interest to this work (1-3 reviewed in the table 1.1).

Table 1.1: S. cerevisiae cell factory for production of "virus-like" pharmaceuticals

Fammily & Genus	Virus species (+1parasite)	Downstream application	The type of "virus-like" pharmaceuticals
Hepadnaviridae	Hepatitis B	-FDA approved vaccine; [159, 166, 167] last generation vac. "RecombivaxHB" [168] -Drug delivivery nanoparticles [150]	
Polyomaviridae	Human papilloma virus (PAP) type: 6,11,16,18	FDA approved vaccine- "Gardasil [169-171]	Surface protein assembled in Autentic Virus-Like Particles (VLPs):
	Bovine PAP virus[172] Cottontail Rabit PAP [173]	Vaccine candidates	
Polyomaviridae	Human polyomavirus[174]	Diagnostics Seroepidemiology	
Parvoviridae	Human Parvovirus B19 [175]		
Leviviridae	Bacteriophage Q3[176] Bacteriophage MS2 [177]	"Nano"-drug delivery of: Peptides (e.g -antigens) and mRNAs (vaccines)	
Retroviridae	HIV-1[178]	Vaccine candidates	Authentic &chimeric VLPs Triple-layered VLPs
Reoviridae	Rotavirus [179]		
Paramixoviridae	Human Parainfluenza Virus (HPIV) type1,3 [180] Henipaviruses [181] Menangle Virus [182] Tioman Virus [183] Measles Virus [184] Mumps Virus [185] Sendai Virus [186]	Vaccine candidates Diagnostics Seroepidemiology	Core and nucleocapsid proteins assembled in Nucleocapsid - like Particles (NLPs)
Bunyaviridae	Hantaviruses [187]		
Orthomyxoviridae	Influenza A virus [188]	Vaccine candidates	Surface protein: Influenza Hemaglutinin Surface glicoproteins
Rhabdoviridae	Rabies virus [189, 190]		
Herpesviridae	Herpessimplex virus [191]		

In the recombinant background, conformationally authentic VLPs are formed from the viral structural proteins expressed in the same cell, that spontaneously assemble into highly ordered structures, in absence of viral genetic material or non-structural viral proteins [184, 192]. As mentioned above, the architectural characteristics of the particles reflect the ones of the native virions enabling their application in various research and medical fields: vaccinology, diagnostics and seroepidemiology, material science and nanotechnology. As shown in the table 1.1, *S. cerevisiae* produced VLPs have been mostly used as diagnostic antigens and vaccines for human and animal diseases.

However, yeast is not the major source of existing VLPs that have been produced so far for more than 30 human and animal pathogenic species. Till recently, only simple VLPs have been efficiently produced in yeast cells. They are composed of single nucleocapsid protein layer or a chimeric protein assembled in one layer [193]. Among them, of the greatest importance are two FDA approved vaccines, currently used for human vaccination worldwide: against hepatitis B and against human papillomavirus type 16 and 18 (see table1.1).

General advantages and current advances in yeast-cell expression technologies (discussed above) hold promises to overcome such challenges. This was recently proved in practice with the first successful production of triple-layered Rotavirus-like particles in S.*cerevisiae* [179]. So far, Rotavirus like particles have been efficiently produced from IC-BEV system which still gives much higher VLP titers than yeast [179, 194].

However yeast cell offers low-cost, safe and easy to scale-up production while insect cell culture requires high cost media and highly skilled personnel for manipulation with pathogenic baculovirus, suggesting that this may be the start point for yeast overtaking the race with IC-BEV system. In this view, it is worth of mentioning that number of yeast species for production of VLPs and other pharmaceuticals has been increasing.

Currently, beside *S. cerevisiae*, methylotrophic strains *Pichia pastoris* and *Hansenula polymorpha* are well recognized producers while dimorphic *Arxula adeninivorans* represents a novel platform that still has to establish itself for biomedical applications [157, 184, 195]. Finally, couple of studies have demonstrated that *S. cerevisiae* not only assembles VLPs but also supports packaging of the recombinant DNA which comprises the cognate viral encapsidation elements. This was firstly shown for *S. cerevisiae* assembled Bovine papillomavirus-type 1 particles, able to replicate and package viral ds DNA, and thus, form infections virions in yeast cells [196] (see below). In addition, MS2-phage VLPs were able to package heterologous, HIV-1gag mRNA inside the yeast cell [177]. These two examples may resemble extended arms of *S. cerevisiae* expression system applicability toward production of DNA and RNA delivery vectors.

1.3.3. *Saccharomyces cerevisiae* supports replication of different RNA and DNA viruses

Beside its prominent medical importance in the view of vaccine production and public protection from two deadly pathogens, *S. cerevisiae* has been also used in the basic virus research for understanding complex biology of numerous pathogenic virusus and clarifying the function of the individual proteins from HIV1 or *Hepatitc C* [197, 198]. As the cell biology and genetics of higher eukaryotes are highly complex, a simpler eukaryotic organism, such as bakery yeast, was a good choice for virus propagation toward elucidation of still unknown aspects of viral life cycles. This owes to its highly conserved, fundamental biochemical pathways and more than 40% conserved sequences homology with known human genes (www.yeastgenome.org). Besides, commercially available yeast gene-deletion library, covering around 85% of all yeast genes (excluding essential ones) was a valuable tool for screening host factors that support or inhibit viral replication (reviewed in [199]). Furthermore, the natural *S. cerevisiae* viruses, two viruses with ssRNA genomes and two viruses with double stranded RNA genomes [200], were useful tool for better understanding of life cycles of RNA viruses in general, and for identifying host factors involved in their replication. For instance, one of the common negative replication regulators for these viruses and Brome mosaic virus (+RNA) was shown to be involved in a degradation of non-polyadenylated mRNAs, thus playing a role in natural yeast replicons [199].

Until now, many RNA or DNA viruses that infect plants *(Bromoviridae*, *Tombusviridae*, *Geminiviridae* and *Avsunviroidae* familiy), animals (such as the Flock House, Nodamura and *Bovine Papillomaviruses*) or humans (Human Papilloma Virus), are shown to be efficiently replicated in yeast (reviewed in [199, 201]) (see table 1.2).

As presented in the table, both + strand RNA genomes of the three virus groups and closed circular (single stranded) RNA of the viroid group (the smallest known pathogenic agents) can be replicated in a nonconventional host, *Saccharomyces cerevisiae*. The first one among them to be described was Brome Mosaic Virus (BMV), whose triple, coding mRNA genome was efficiently replicated and encapsidated in the yeast cell. Couple of subsequent studies of the underlying mechanism and host players in BMV replication revealed many important aspects the life cycle of this virus, such as the process of template selection and replication complex assembly in the vesicles of endoplasmic reticulum [202, 203]. Similarly, studies of Tombusvirus replication revealed that CIRV members of this family replicate in association with intracellular membranes of mitochondria [204], whereas CNV replicates at peroxisomal membranes. Moreover, cis acting elements of viral RNAs, fundamental for replication, were also identified in yeast studies [205]. Finally, the only higher eukaryotic RNA virus, able to both replicate and assemble infections virions inside of *S. cerevisiae* cell belongs to the *Nodaviridae* family of animal viruses.

Studies of RNA viruses' replication in yeast, using traditional yeast mutagenic analysis and genome-wide screening approaches upon yeast deletion libraries, resulted in the ground-breaking identification of multiple host factors required for viral RNA replication and recombination. Moreover, the features of RNA viruses' replication in yeast cells, such as existence of organelle-like RNA replication factories, were the recapitulation of the ones in their corresponding natural hosts, including higher eukaryotic viruses [199, 206, 207].

So far, two large families of DNA viruses, *Papillomaviridae* and *Geminiviridae* (Table 1.2), with circular dsDNA and ssDNA genomes, respectively, have been described to efficiently replicate in *S. cerevisiae.*

Table 1.2: Viruses that replicate in yeast

	Family	Virus	Genome	Natural host
RNA viruses	*Bromoviridae*	Brome Mosaic Virus (BMV) [202, 203, 208]	(+) RNA	Plants
	Tombusviridae	Carnation Italian RingspotVirus (CIRV) [204, 209] Cucumber Necrosis Virus (CNT) [205] Tomato Bushy Stunt Virus (TBSV) [210]	(+) RNA	Plants
	Nodaviridae	Flock House virus [206] Nodamura virus [211]	(+) RNA	Animals
	Avsunviroidae	Avocado Sunblotch Viroid [212]	ss RNA circular	Plants
DNA viruses	*Papillomaviridae*	Human Papillomavirus (HPV) [213]	dsDNA circular	Humans
		Bovine Papillomavirus (BPV) [172, 196]	dsDNA circular	Animals
	Geminiviridae	Mung bean yellow mosaic India virus [214]	dsDNA circular	Plants

The first one includes oncogenic Human papillomavirus (HPV) and Bovine papillomavirus (BPV). Recombinant genomes of HPV-types 11,16,18 and 31 and BPV-type 1 replicate stably as episomes in *S. cerevisiae*, in the absence of any viral proteins, completely depending on the host replication machinery (highly conserved among yeast and humans) [213]. The striking finding of another research group was the ability of higher eukaryotic Bovine papillomavirus to fully complete its lyfe cycle in the yeast cellular environment, starting from infection of the yeast protoplast, over viral replication, particle formation and genome packaging into viral shells, thus forming infectious BPV-1 particles [172].

Further development of new yeast-based replication platforms for large number of pathogenic or deadly viruses will help understanding complex virus–host interactions providing new targets for antiviral drug development. Moreover, these studies, together

with the possibility of creating “humanized yeast systems” by coexpressing human proteins essential for viral replications, encourage design of a novel, yeast-based replication systems for the viruses used as gene delivery vehicles in gene therapy approaches.

2. PROJECT I: Expression of AAV structural and non-structural proteins and assembly of virus-like particles in yeast *S. cerevisiae*

2.1. The aim of the project:

Stable expression of AAV structural proteins in *S. cerevisiae* cell and optimization of their relative molar ratios as a prerequisite for highly ordered interactions of capsid protein subunits toward generating icosahedral capsid like structures.

2.2. The background of the project

2.2.1. Critical components of AAV2 capsid assembly in the recombinant background

So far, AAV2 capsid assembly has been studied in two, evolutionary far, eukaryotic cellular systems – mammalian and insect cell. The common concept accepted in these systems is that assembly process does not require the expression of viral genes other than cap gene [53, 215], the general rule of VLP assembly [184]. In order to define minimal requirements for production of infective AAV virions, the three capsid proteins have been expressed separately and in combination and their assembling capacity investigated in different recombinant backgrounds, including *in vitro* studies of capsid protein oligomerization and formation of capsid-like structures [216] and *in vivo* studies in the two aforementioned systems, *Spodoptera frugiperda* Sf9 insect cells - baculovirus system [215, 217] and mammalian cells (mainly HEK-293 and HeLa cell-lines) (reviewed elsewhere, [218]). In these studies the requirement for each capsid protein in the formation of intact particles was assessed in the absence of other AAV proteins, adenovirus (Ad) helper function and packageable AAV genomes. Although with some contradictions, these studies uniformly proposed either VP1 or VP2 to be essential for capsid assembly, probably due to their ability to localize major capsid protein, VP3, to the nucleus, the site of capsid assembly [53, 217]. The main contradiction in these studies related to the sufficiency of VP3 alone in forming capsid shells ([215-217, 219], reviewed in: [48, 49]) has been resolved recently with a discovery of a novel AVV2 regulative protein (AAP), indispensible for VP3 assembly in virus like particles in the absence of other two VPs [49].

VP3 constitutes the surface topology of AAV2 capsids, while the largest capsid protein VP1 is shown to be essential for viral infectivity. Later, this was explained by the presence of N terminally-located phospholipase domain (PLA2) and nuclear localization sequences (NLS), thought to be necessary for endosomal escape [56, 62, 220] and nuclear entry [65,

221, 222], respectively. On the contrary, the other scarce capsid protein, VP2, is shown to be non-essential and can be omitted when infective virions are to be produced in the recombinant background [219, 221].

In order to study the permissiveness of yeast *S.cerevisiae-in vivo* system to AAV2 viral particle assembly we analyzed the expression of AAV capsid proteins both from their natural p40 promoter and from yeast cell-derived promoters, with the purpose of finding the best expression system that would promote VP protein self-assembly into AAV VLPs. The rational for testing AAV promoters in yeast comes from studies of other viruses showing that a number of wt viral gene control elements retain their functionality in *S. cerevisiae* [202]. Even more complex elements, such as Hepatitis C virus IRES (internal ribosomal entry site), do not require natural (human) *trans* regulators for its activity in S. cerevisiae. All proteins necessary for internal initiation of translation from this element are encoded by S. cerevisiae genome [197]

2.2.2. Regulation of Cap protein expression in the natural background

The three AAV2 capsid proteins, VP1, VP2, and VP3, are expressed from two mRNAs, differentially spliced from 2.6kb p40 mRNA precursor. Translated from overlapping ORFs, these proteins differ only in their N-terminal region as shown in the Figure 2.0. Different estimations of VP1:VP2:VP3 ratios in AAV2 virus shells have been published so far, ranging between1:1:8 and 1:1:20, respectfully (reviewed in [147]). The capsid composition is thought to reflect the stoichiometric ratio of the three VPs inside the producing cell [51], which in turn is regulated on transcriptional and translational level. In mammalian cells, the two largest AAV non-structural proteins, Rep 78 and Rep 68, have been shown to regulate expression from the p40 promoter and this regulation is even more pronounced in the presence of the helper virus [34, 36, 223, 224].

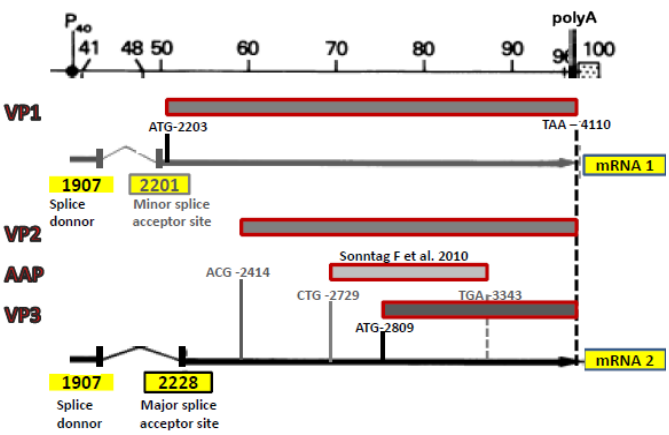


Figure 2.0: Cap gene organisation. The top part of the picture shows the portion of the AAV2 genome on a scale of 100 map units (1map unit is approx. 47 nucleotide) comprising *cap* gene from p40 promoter (map unit 40) till common polyA site (map unit 96). Underneath, 2 *cap* mRNAs are presented as arrows, with indicated intron–exon organization and translational start and stop codones for proteins encoded by these mRNAs. Three structural VP proteins and a non-structural protein, AAP, are shown above corresponding mRNAs, in boxes.

So far identified *cis*-acting elements required for p40 transactivation by Rep proteins are: p5 promoter and ITR located RBEs, p19 CArG-like element and p40 proximal Sp1 site

[34]. However, differential effects of large Rep proteins on all three AAV promoter activities in 293, HeLa and other cell lines imply the crucial role of cellular factors in Rep mediated p40 transcriptional regulation [36]. These observations have to be taken into consideration when AAV-2 capsids are produced in the recombinant background where capsid proteins are expressed from Rep-Cap expression plasmids lacking terminal repeats. Accordingly, in the novel yeast-cell expression system, we were curious to test the potential influence of Rep proteins (Rep78 or Rep68 in particular) on the Cap protein expression from the wt p40 promoter, both when Rep proteins were expressed from viral p5 and p19 promoter in their natural genomic configuration, and when expressed from a heterologous yeast promoter.

2.3. Results

2.3.1. Expression of AAV structural and nonstructural proteins from natural viral promoters and influence of Rep proteins on capsid protein expression in yeast *S.cerevisiae*

Protein Expression from wt AAV promoters

AAV viruses lack the cognate RNA polymerase, depending on the host transcriptional machinery for expression of its proteins. Remarkable homology between *S. cerevisiae* and human transcriptional apparatus [225] increases the possibility that AAV promoters could be functional in this simple eukaryote. To test this hypothesis we made a yeast-cell compatible multi-copy plasmid containing unmodified AAV2 cap gene with all regulative elements: p40 promoter, intron element and polyadenylation signal. To do this the entire VP expression cassette located between 1428 and 4495 nucleotide position of AAV2 genome (numbering is as for the sequence under GenBank accession no. AF043303.1) was cloned into the yeast episomal YEplac181 (Leu2 gene as a selection marker) The resulting plasmid, nominated Yplac p40Cap, has yeast *2-micron* origin of replication which constantly provide the cell with 20-50 copies of the recombinant gene per haploid yeast genome [226] and thus, might compensate for the weak activity of foreign p40 promoter and other heterologous control elements. The schematic representation of the Yplac p40Cap construct is shown in the Figure 2.1-A, left part.

Regulation of VP protein expression by large Rep proteins in mammalian cells is mainly achieved through transcriptional regulation of p40 promoter activity and post-transcriptional maturation of its precursor mRNA [227] (although not as prominent as in the presence of adenovirus and ITRs). As mentioned before, the first process is dependent, *in cis*, upon sequences associated with p5 and p19 promoter elements [223], while the second one is mostly pronounced when the AAV2 intron element is paired to its natural promoter and extended polyadenylation signal [227]. To keep the same regulation network in yeast cell, we introduced rep and cap gene (in their wt genomic configuration) into the

yeast vector, thus keeping p5-RBRs, p19, p40, intron and AAV2 polyadenylation element in the authentic relative positions. The resulting construct, named YplacRepCap (fig.2.1B-left) allowed simultaneous analysis of protein expression from all AAV2 promoters.

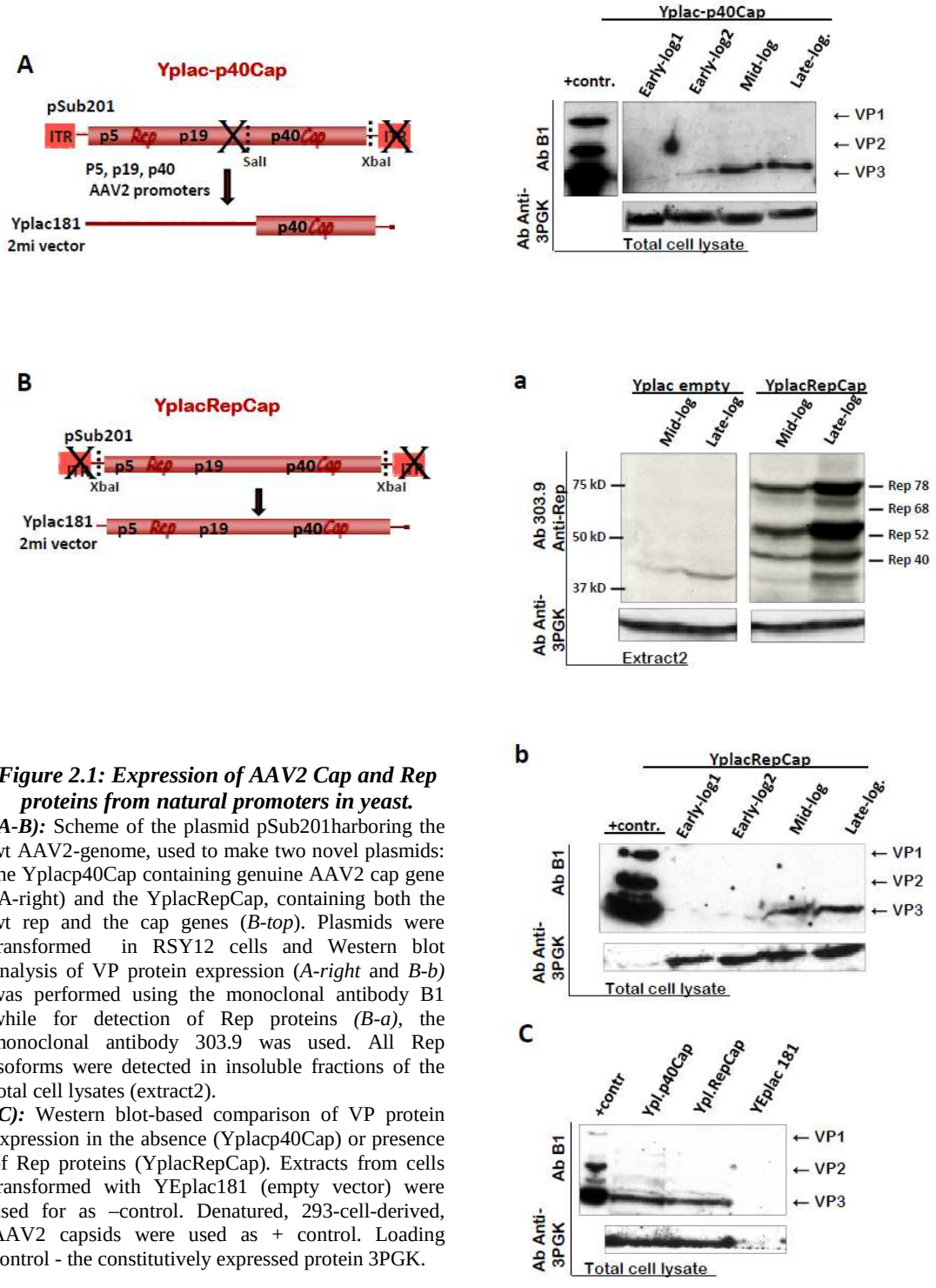


Figure 2.1: Expression of AAV2 Cap and Rep proteins from natural promoters in yeast.

(A-B): Scheme of the plasmid pSub201harboring the wt AAV2-genome, used to make two novel plasmids: the Yplacp40Cap containing genuine AAV2 cap gene (A-right) and the YplacRepCap, containing both the wt rep and the cap genes (*B-top*). Plasmids were transformed in RSY12 cells and Western blot analysis of VP protein expression (*A-right* and *B-b)* was performed using the monoclonal antibody B1 while for detection of Rep proteins *(B-a)*, the monoclonal antibody 303.9 was used. All Rep isoforms were detected in insoluble fractions of the total cell lysates (extract2).

(C): Western blot-based comparison of VP protein expression in the absence (Yplacp40Cap) or presence of Rep proteins (YplacRepCap). Extracts from cells transformed with YEplac181 (empty vector) were used for as –control. Denatured, 293-cell-derived, AAV2 capsids were used as + control. Loading control - the constitutively expressed protein 3PGK.

In yeast, all Rep isoforms are detected when their expression is regulated by viral, p5 and p19 promoters (YplacRepCap construct) (fig. 2.1.B), while VP3 that is the only capsid protein detected from p40 promoter, both in the absence (Yplacp40Cap) or presence of Rep proteins (YplacRepCap and Yep-Rep78p40Cap vector) (fig. 2.1-C and fig. 2.2-B-b).

Plasmids Yplacp40Cap and YplacRepCap, and the original vector YEplac181 were transformed in the haploid strain RSY12 of *S. cerevisiae*. This strain has a double auxotrophy (Leucine and Uracile), so both URA3 and LEU2 gene markers can be used to select for transformant yeast cells. Growth curves of shake-flask cultures of Yplacp40Cap, YplacRepCap and YEplac181control plasmids-transformed cells were similar (data not shown), with doubling time of about 2.7 h for exponentially growing cells, when glucose was used as the carbon source. During the culture time span, $2x10^8$ cells were collected at 4 different time points and subjected to protein extraction under denaturizing conditions to yield total cell protein extracts. The first two time points correspond to early exponential (log) growth phase, the third one to mid-log phase and the forth one to the late-log/early-stationary phase. Cap protein expression was analyzed by Western blot analysis of total cell lysates at each of these time points (fig. 2.1, parts A and B-b) ("optimized post-alkaline" extraction method was used and will be described in the next section). Part A defines capsid proteins found in Yplacp40Cap-cell lysates, while part B-b shows total cell extracts from YplacRepCap-transformants**.** The only capsid protein expressed from p40 promoter in both Yplacp40Cap and YplacRepCap transformed cells was VP3. VP3 accumulated with time, following exponential cell growth and biomass accumulation (the maximal amount of protein was extracted from mid and late-exponential growth phases, fig.2.1, parts A and B). The monoclonal antibody B1, we used, is widely used for detection of native VP proteins [53, 228]. The absence of VP1 and 2 was confirmed in both immunoblotting and in immunofluorescence (data non shown), by means of another monoclonal antibody, A69 which specifically binds N-terminal epitope common to VP1 and VP2 [53, 228]. Although the figure 2.1 shows overall VP3 concentration in total cell extract of transformed cells, majority of this protein was found in insoluble cell aggregates (see later).

In parallel, Rep protein expression was analyzed in extracts derived from YplacRepCap-transformed clones, collected in two different time points, mid-log and late-log phase. Western blot analysis showed the presence of all 4 Rep polypeptides (detected by the monoclonal antibody cl.303.9 which recognizes all four Rep species [53, 228]), accumulated to the max level by the late-log phase (fig.2.1 part B-a). Although to a smaller extent than VP3, majority of Rep proteins, particularly p5 products, were recovered from

insoluble protein compartment (extract2) (see later). Since insoluble fraction is richer in all Rep proteins, only Western blot result for this fraction is presented in the fig. 2.1-B-a.

Fig.2.1 part C shows capsid proteins recovered from total cell lysates derived from the same number of +Rep(YplacRepCap) and –Rep clones (Yplacp40Cap), collected in the same mid-late log phase of both cultures (calculated from $O.D_{600}$). Apparently, there was no significant difference in cap protein expression in these two cell clones, neither qualitative (VP3 was the only cap protein detected) nor quantitative (relative VP3 amount estimated by the band intensity in the Western blot). This fact demonstrated that there was no influence of Rep proteins on VP protein expression from p40 promoter in yeast, when both rep and cap genes were present in their natural genomic configuration.

Cap expression in the presence of constitutively expressed Rep78

To confirm whether Rep expression does not alter the VP protein expression in yeast cells, we also analyzed VP expression pattern when the Rep78 expression is driven by the constitutive yeast promoter pADH1. Rep78 expression cassette was cloned upstream of the cap gene in Yplacp40Cap resulting in Yep-Rep78p40Cap plasmid (Fig.2.2.A).

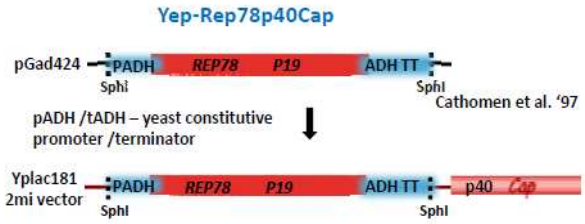


Figure. 2.2 A: Scheme of the plasmid Yep-Rep78p40Cap

pGRep78 construct, harboring the Rep78 cDNA under yeast promoter pADH1,was used to make plasmid Yep-Rep78p40Cap that contains ADH-Rep78 expression cassette and p40-regulated cap gene.

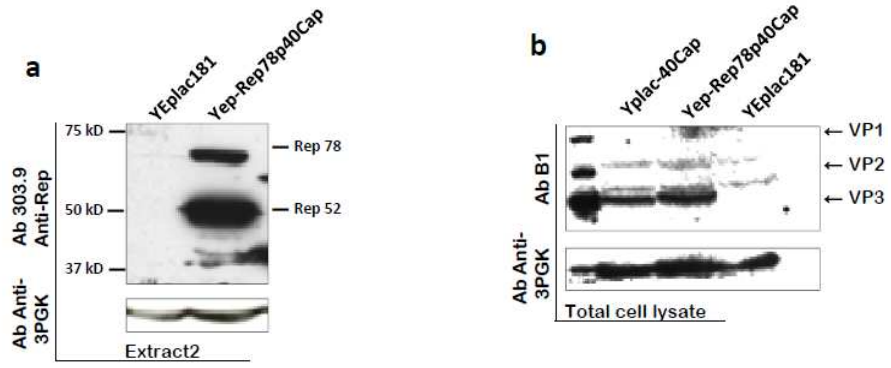


Figure 2.2 B: Cap expression from p40 promoter in presence of Rep 78, expressed from the constitutive ADH1 promoter.

(a) Yep-Rep78p40Cap plasmid was transformed in RSY12 cells. Western blot analysis of Rep protein expression identified Rep isoforms in the extract2 (insoluble aggregates) of total cell lysate. *(b)* Western blot analysis of VP proteins in extract derived from Yep-Rep78p40Cap transformed cells (in the presence of Rep proteins) and from Yplac40Cap clones (in the absence of Rep). Extracts from cells transformed with empty vector, YEplac181, were used as negative control. Loading control: Constitutive yeast protein 3PGK. Denatured, 293-cells derived AAV2 capsid were used as + control for defining VP molecular weights.

As shown in the scheme, Rep78 cassette derives from p.G.Rep78 expression vector (published by Cathomen et al. [229]; the courtesy of Weitzman), and consists of Rep78 cDNA, flanked by constitutive yeast promoter ADH1 (truncated version) and ADH terminator. Western blot analysis of rep proteins in extracts derived from yeast cells transformed with Yep-Rep78p40Cap construct showed that not only Rep78 protein is expressed, but also p19 promoter works and, therefore, Rep 52 is produced from this expression cassette (Fig 2.2, B-a). Since p19 promoter is the inherent part of Rep78 cDNA, the presence of Rep52 is consistent with previous demonstration of p19 promoter activity from YplacRepCap construct. Western blot results for cap protein expression shown in the Figure 2.2, part B-b, demonstrate again that VP3 is the only capsid protein expressed in the cells that were constitutively producing Rep78 from the yeast promoter. While expression of Rep78 from p5 could not been detected in the end-points of the culture growth, the early exponential and the late stationary growth phase, Rep 78 expression from constitutive yeast promoter was detected in all phases of the culture growth (data not shown). However, as shown in the Figure 2.2, part B–b, Rep proteins in Yep-Rep78p40Cap-transformed cells did not significantly change cap protein expression form p40 promoter with respect to no-Rep VP3 producing cells (Yplacp40Cap transformed).

Solubility of VP3 capsid protein and of Rep proteins in *S. cerevisiae* protein extract

The very first analysis of the Rep protein expression from YplacRepCap construct showed that Rep proteins localized to both soluble and insoluble compartment of the yeast cell protein extracts prepared by mechanical-glass bead lysis. In this method, soluble fraction was obtained by lysis in the buffer with physiological salt concentrations and a low concentration of a non-ionic detergent (0.5% Triton X-100), mostly yielding yeast cytoplasmic proteins. Insoluble fraction, including majority of nuclear proteins, intrinsic membrane proteins and receptors, were recovered from the sediment obtained by high speed centrifugation of the crude lysate, after incubation with high strength buffer (1% Sodium Dodecyl Sulfate (SDS) and 1% Deoxycholic Acid (DOC)-sodium salt). As shown in the Figure 2.3, Rep52 protein was almost uniformly present in both fractions, while small amount of Rep 78 was detected only in the insoluble protein fraction. Such a distribution might be the reminiscent of the function-related cellular distribution of Rep proteins, typical for mammalian cells [43].

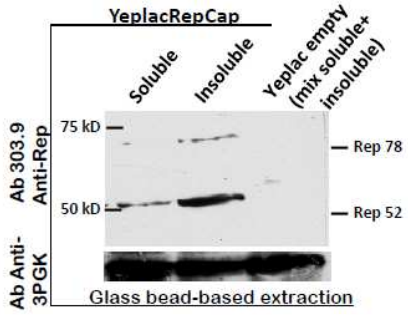
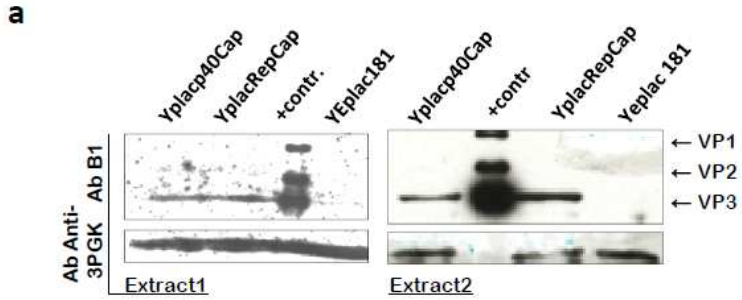


Figure 2.3: Distribution of AAV2 Rep proteins to soluble and insoluble protein fraction obtained by glass-bead extraction method.

Western blot analysis showing differentially distributed Rep52 and Rep78 to soluble and insoluble protein fraction of extracts derived from YplacRepCap transformed yeast cells. Extracts from cells transformed with YEplac 181 (empty vector) were used as -control. Loading control: Constitutive yeast protein 3PGK.

.

On the contrary, no VP3 protein was found in the soluble fraction of extracts derived from either Yplacp40Cap or YplacRepCap transformed cells, while only small amounts of VP3 was recovered from the pellet upon SDS-DOC solubilisation (data not shown). However, this result was poorly reproducible, which led to adoption of another protein extraction method, more efficient in extracting maximum amount of proteins from relatively small number of cells. The published "post-alkaline" lysis procedure guarantees such an advantage over the standard use of glass-bead extraction [230]. Although improving Cap expression analysis, ''post-alkaline'' lysis did not provide absolute reproducibility of the data until further optimized with the second extraction from the sediment obtained after purification of the total lysate obtained by this method. This sediment contained protein aggregates and proteins associated with the cellular membranes, which required harsher conditions to be liberated from the cellular structures and efficiently transferred to solubilised state. We achieved this by combining solubilising properties of various ionic and non-ionic detergents in elevated concentrations, further assisted by sonication. Only with this second round of extraction VP3 protein was recovered with 100% reproducibility. The obtained fraction of cellular proteins was designated as extract 2 in the presented figures. This was deduced from Western blot analysis in Figure 2.4-a, which showed that majority of VP3 proteins localized to extract 2, while only small amounts of this protein were recovered in the first round of extraction (extract1), which follows the procedure of Kushnirov [230]. This result strongly demonstrates that VP3 accumulated in insoluble aggregates within *S. cerevisiae* cells, in concordance with results from Sf9 and HeLa cells [215, 216]. Together these three evolutionary distant, eukaryotic expression systems strongly support an intrinsic propensity of AAV capsid proteins to associate with cellular structures. Moreover, with optimization of "post-alkaline" procedure (addition of the second extraction cycle) we were able to extract all AAV2 proteins expressed in RSY12 yeast cells with greater efficiency, including non-structural ones.



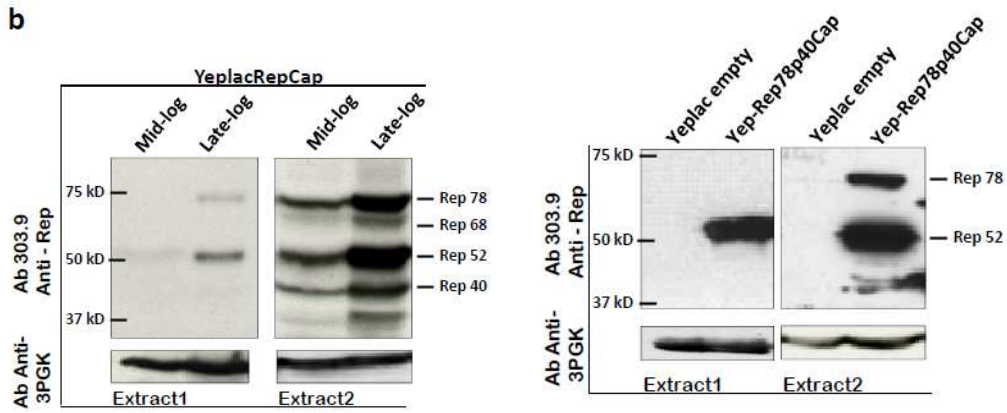


Figure 2.4: Recovery of AAV2 proteins by "Optimized post-alkaline" extraction.

a) Western blot analysis of cap proteins in total cell extracts obtained by optimized post-alkaline extraction method from Yplacp40Cap and YplacRepCap-transformed cells, revealed majority of VP3 proteins in the extract2 (insoluble aggregates). Extracts from YEplac181 transformed cells were used as -control. Denatured, 293 cell-derived, AAV2 capsids used as + control. Loading control: Constitutive yeast protein 3PGK;
b) YplacRepCap cells differentially grown to mid or late-log phase *(b-right)* and YepRep78p40Cap cells grown to late-log phase *(b-left)* were subjected to "Optimized post-alkaline" protein extraction for Western blot analysis of the Rep protein distribution to two proteic fractions obtained by this method.

Figure 2.4-b shows Western blot result for distribution of Rep proteins to extracts 1&2 of total cell lysates derived from YplacRepCap (fig 2.4-b, left) and YepRep78p40Cap (fig 2.4-b, right) transformed clones. Uniformly for both clones, majority of large Rep proteins localized to the extract 2, while small Rep isoforms were almost equally distributed to both extracts. Low amounts of Rep68 and Rep 40 were detected only after optimization of the extraction method, while they could not be detected with glass-bead-based protein extraction. Although poorly expressed comparing to their non-spliced relatives, the presence of Rep68 and Rep40 suggests that, albeit weakly, the major AAV intron (bounded by eukaryotic-cell conventional donor and acceptor splice sites) is recognized by yeast splicing machinery. Poor splicing of foreign intron elements (e.g. in *Drosophila* gene) has already been documented for the *S. cerevisiae* expression system [231].

In some cases (e.g.ADE8 promoter of *Drosophila*) foreign promoters are found to give aberrant transcriptional initiation in *S. cerevisiae* [232]. To prove the existence of correctly initiated p40mRNAs in YplacP40Cap–transformed yeast cells, we performed PCR amplification of retro-transcribed mRNA pool derived from this clone, using specifically designed primer pairs. The upstream primer starts with p40 transcriptional start site (nucleotide position 1853) and is identical to the first 20 nucleotide sequence of 5' untranslated region (5'UTR-Exon1), common to all three putative, correctly initiated forms of p40 RNAs: the non-spliced precursor and two differentially spliced p40mRNA species. The downstream primer is complementary to 18bp region common to all three VP ORFs with the nucleotide position 2758. PCR conditions were optimized in a manner to provide the highest specificity of the assay, excluding nonspecific binding to other regions of Cap ORF, while non-specific pairing with yeast mRNAs (cDNAs), if any, could be disregarded

by comparison with the control PCR (mRNA pool derived from cells transformed with YEplac181-empty vector). In principal, such primer design could distinguish between all three putative mRNAs, corresponding to three PCR products of different lengths, 923, 575 and 549 bp long. PCR products are separated and analyzed upon acrylamide gel electrophoresis. However, the only p40 product detected under most stringent PCR conditions was 2.6 kb precursor mRNA (988bp band) (fig.2.5), while other shorter species were hardly detectable and difficult to interpret by this assay.

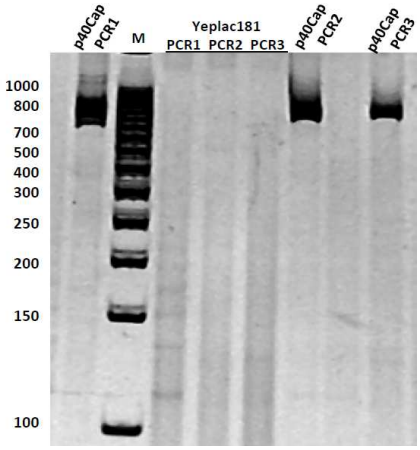


Figure 2.5: Reverse Transcription (RT)-PCR-based analysis of p40mRNAs.

Total cellular RNA was extracted from Yplacp40Cap and YEplac181 (-contr.) transformed clones, and corresponding mRNA pools were retro-transcribed and subjected to primer specific PCR under 3 different conditions: PCR1 of the lowest stringency and PCR3 of the highest one. The PCR products are separated by acrylamide gel electrophoresis and stained with EtBr.- DNA Marker used: 50bp-1000 bp ladder.

This result, although not obtained with a very sophisticated analysis, suggests that VP3 protein derives from the 2.6 kb p40-precursor in yeast cells, rather than from its spliced product. If so, its expression could be interpreted in the context of the "scanning translation initiation" model in eukaryotes. According to this model 40S ribosomal subunits scan from the 5' end of an mRNA until reaching an AUG codon suitable for initiation. In these terms, the environment of VP1 initiator (GNNAUGG) (as well as of other AUGs present in 5'UTR) might be suboptimal in yeast cell, alike uncommon VP2 ACG, causing ribosomes to bypass till reaching the VP3 AUG resided in the favorable "Kozak-like" context. However, 2.6kb-p40 precursor is shown to be very poor template of VP protein expression, both in *in vitro* and *in vivo* experiments in HeLa cells, probably due to multiple short ORFs present in 5'UTR [233]. Taking this into account, we cannot exclude the possibility that even a low amount of spliced p40 mRNAs (too low to be detected by RT-PCR-based analysis), may still be the main source of VP3 protein in yeast.

Nevertheless, we did not investigate deeper into the RNA profile of AAV2 proteins expressed in yeast cells, but rather focused on finding other strategies for expression of the missing AAV2 structural proteins (VP1 and VP2, or at least VP1) required for assembly of AAV2-VLPs.

2.3.2. Inducible Yeast Promoter for Regulation of AAV2 Structural Protein Expression

Since VP2 was shown to be non essential for infectivity of the AAV2 virions [219, 221], simultaneous expression of VP1 and VP3 in yeast cells would be the primary condition for developing yeast-cell based system for production of the AAV2-like particles with biological properties of the wt capsids.

Potential reasons for missing VP1 protein expression from p40 AAV promoter are numerous; low activity of the foreign viral promoter, absence of splicing, inefficient transport of non-spliced mRNA to the cytoplasm and/or poor stability of this transcript, inefficient translation from non-spliced mRNAs, increased proteolytic degradation of this protein in yeast cellular environment etc. The use of yeast promoters and terminators should circumvent at least some of these challenges. On the way to creating efficient VP-protein expression system, we created various constructs based on strong galactose inducible promoter pGal1. To this purpose, we used commercially available multi-copy vector pYes2, whose multiple cloning site (MCS) is flanked by efficient yeast gene expression regulators, Gal1 promoter and cyc terminator. This vector harbors URA3 gene for constant selection in auxotrophic RSY12 strain. Figure 2.6-A shows a panel of pYes2-based constructs in the top-to-bottom order in which each one will be described further in the text.

pYes2-Gal1 promoter-based constructs

Initially, we tested weather all AAV2 capsid proteins from overlapping ORFs could be produced simultaneously when their transcription was regulated by yeast Gal1 promoter instead of viral p40 promoter. Cap gene sequence downstream of p40, including intron element (the sequence between map pos.1882 and 4495 of AAV2 genome in GenBank, accession no. AF043303.1) was introduced in MCS of pYes2 vector and the novel vector was designated **pYesIntronCap**. With this strategy, we wanted to avoid the separation of VP ORFs in three different cDNAs, since the VPs expressed in this manner did not assemble well in tissue culture cells [216]. Besides, elements present in the AAV2 intron region (possibly RNAs) contribute to VP protein expression when induced from the CMV promoter (Kleinschmidt J., personal communication). The possible analogy with human cells was the rationale for keeping AAV2 intron sequence between Gal1 promoter and VP1 start site. RSY12 cells transformed with pYesIntronCap and pYes2 empty vector (as negative control) were actively grown on glucose for one or two days before the induction of Cap gene expression was triggered upon carbon source exchange in the growth medium from glucose to galactose. At the end of glucose growth, a small fraction of actively growing cells (10^8) was collected and lysed for protein extraction, while the rest of the culture was subjected to 4h and 8h induction and then processed for Western blot analysis to assess the expression of cap proteins (fig.2.6-B).

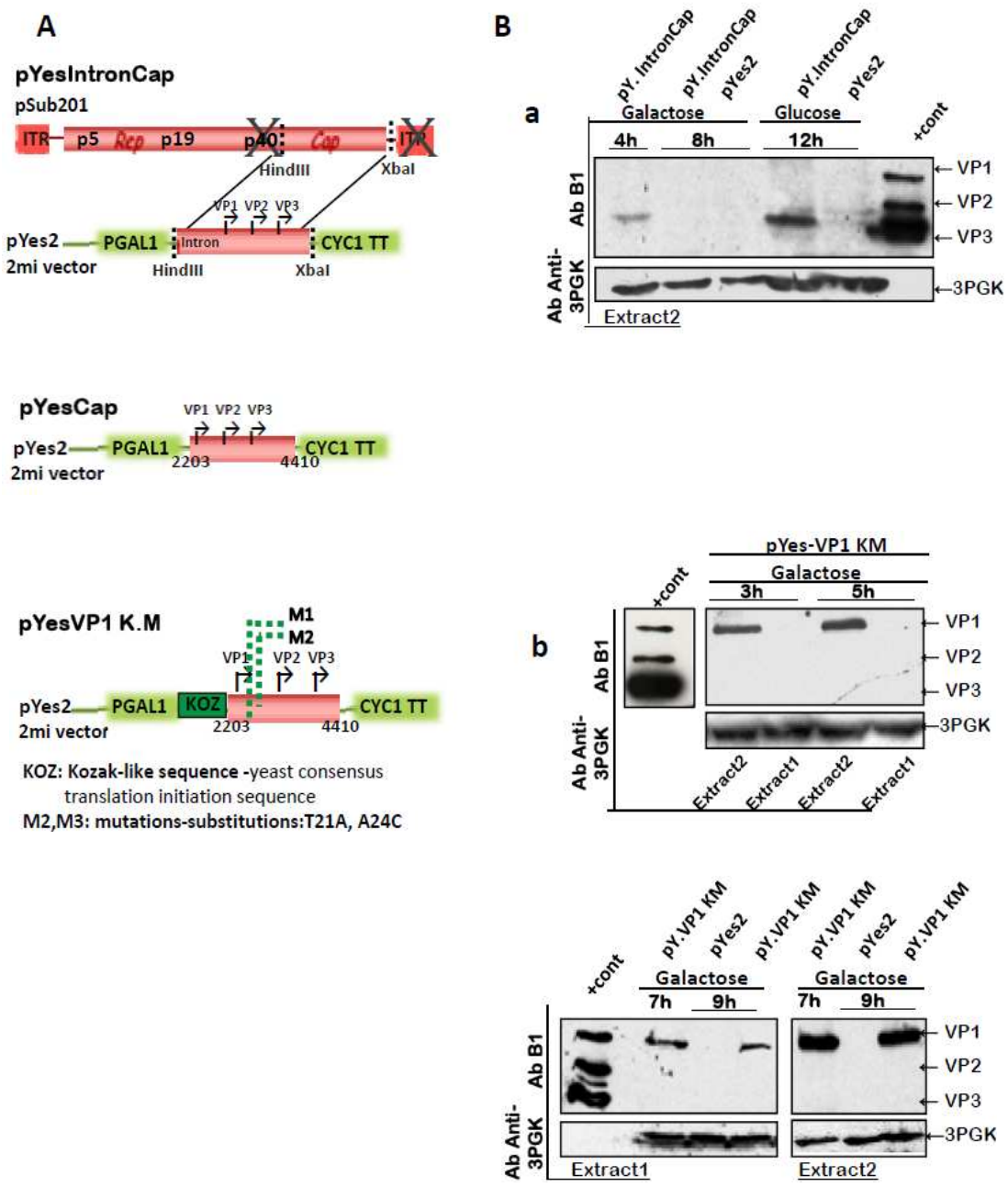


Figure 2.6: Expression of AAV2 structural proteins from the galactose-inducible promoter pGal1
(A): Panel of pYest2 vector-based constructs, carrying differences in cap transcription units were used to transform RSY12 cells. *(B)*: Western blot analysis of VP protein expression. After pre-induction growth in glucose, transformed clones were exposed to galactose for different induction times. pYesIntronCap expressed VP3 was detected in the extract 2 of glucose grown cells and to a smaller extent in the extract 2 of 4h Gal-grown cells, while no Vp3 was detected after 8h of Gal induction *(a)*. pYesVP1 K.M-expressed VP1 was found in extract 2 after 3h and 5h of galactose-induction (b-top), and in both, extracts 1 and 2, after 7h and 9h induction (b-bottom). Extracts from cells transformed with empty vector, pYes2, were used as -control.

During the glucose growth, Gal1 promoter activity is on the basal level which can differ from one yeast strain to another ("leaky" or non repressed promoter phenotype is strain-dependent) [234]. Surprisingly, the only protein expressed from this construct was again VP3 and its max amount was detected in the extracts of the cells grown on glucose. After the carbon source switch to galactose, some background VP3 level was still observed after 4h of galactose growth and completely vanished after 8h of induction (fig.2.6B-a). VP3 was detected only in the extract2. Decrease instead of expected increase in protein level after glucose-galactose nutrient switch could not be explained in terms of toxicity of overexpressed capsid proteins since no toxicity-related phenotype (e.g. decreased growth rate or abolished growth) was observed for gal-grown cells. Instead, galactose metabolism seemed to be detrimental for VP protein expression. We proposed one explanation for the observed "galactose effect", based on the hypothesis that VP3 protein is expressed from the spliced Gal1 mRNA form. In yeast cells, big environmental changes, sensed as stressful, can cause smaller or bigger alterations in splice factors availability and once cells adapt to new conditions different splicing pattern is fixed [235]. Carbon source exchange with water-wash step in between (temporary starvation) is stressful for the yeast cell and even a small alterations in a pool of splice factors could abolish already poor splicing of the foreign AVV2 intron.

To test this hypothesis, we created another vector, named **pYesCap**, in which the *CAP* coding region without the upstream intron element (the sequence between VP1 ATG, the nucleotide pos.- 2203, and common VP Stop codon, pos.- 4410 in the wt AAV2 genome sequence), was cloned downstream to the Gal1 promoter. In the absence of splicing, the main protein product expected to be synthesized upon induction, is VP1. However, if VP3 were produced from non-spliced mRNA, then also the pYesCap plasmid could give rise to some VP3 protein level. Instead, neither VP1, nor VP3 was detected from this vector regardless the carbon source used (data not shown). The absence of VP3 expression in glucose grown cells supports the hypothesis that intron element (the spliced pGal1 mRNA) was essential for VP3 protein expression from pYesIntronCap construct. Yet, the absence of VP1 synthesis after Galactose induction could have been explained only after we introduced following <u>modifications in the VP1 cDNA</u>:

1. Introduction of yeast Kozak-like element right upstream of VP1 start-site (yeast translation initiation consensus sequence designed according to pYes2-vector manual, Invitrogen),
2. Elimination of an out-of-frame ATG, located between VP1 and VP2 initiation codons,
3. Removal of the major AAV2 splice acceptor (position 2227), downstream of VP1 ATG

Although not as efficient as consensus Kozak sequence in mammalian cells, the yeast-cell homologous element is thought to provide up to 3 fold increase of translation initiation

[226, 236, 237]. The second two modifications are already shown to contribute to improved VP expression in Baculovirus-Insect cell system [138]. Elimination of the neighboring out-of -frame ATG is supposed to prevent interference with the choice of the right ORF, while removal of the splice acceptor should prevent aberrant splicing of VP1 mRNA. The resulting vector harboring these modifications was named **pYesVP1K.M.** Western blot analysis of the total protein extracts derived from RSY12 cells transformed with this vector (obtained by "optimized post-alkaline" extraction method) showed that no cap protein was expressed after 12-15 h growth on glucose (data not shown), while galactose induction caused gradual increase in VP1 protein expression with max level achieved around 7-8h after induction (fig. 2.6B-b bottom). VP1 was the only capsid protein expressed in these cells. In the first 5 hours of induction (fig.2.6B-b top) no VP1 protein was detected in the extract 1 (not shown), whereas a minor portion of VP1 could be found n this fraction only after protein accumulation with longer induction times (7h and 9h in fig.6B-bottom). Just like VP3, majority of VP1 protein was recovered in the extract 2 (insoluble fraction) of the total cell-lysate.

2.3.3. Simultaneous expression of AAV2 capsid proteins and modulation of their stoichiometry toward successful capsid assembly.

In order to determine assembling abilities of VP1 and VP3 proteins in RSY12 strain background, yeast cells were subjected to co-transformation with two plasmid combinations (table 2.1):

1. In the first experimental setting yeast cells were cotransformed with pYesVP1.KM and YplacRepCap and selected double-transformants were named "+Rep - double"
2. In the second setting cells were cotransformed with pYesVP1.KM and Yplacp40Cap; these transformants had the same genetic material for capsid protein expression as "+Rep-double", but were lacking in rep gene, thus named "-Rep - double"

Table 2.1: Constructs used to cotransform RSY12 cells for simultaneous expression of VP1 &VP3 protein

Pairs of plasmids	Labelling of double-transformed clones
pYesVP1 K.M + Yeplac-p40Cap	*" - Rep" double*
pYesVP1 K.M + Yeplac-RepCap	*" + Rep" double*
pYes2 + Yeplac181 - empty vector	– control

Based on the findings with single transformed clones, the expected expression profile for the double transformants would be as follows:

Continuous expression of VP3 protein and Rep proteins during the whole culture (glucose + galactose), and boost of the VP1 expression after the switch from glucose to galactose-based growth. Yet, Western blot analysis of total-cell lysates derived from "+Rep" and "-Rep- double" clones, grown for 12h in glucose medium (non-inductive growth phase) + 8h in galactose medium (induction phase), showed one discrepancy from the expected VP1-VP3 expression pattern: As expected, only VP3 protein was produced by actively dividing glucose-cultures (both, "+Rep" and "-Rep"- double) (fig.2.7-A); however, after 8h of induction in galactose, induced VP1 synthesis was followed by an unexpected decrease in the relative amount of VP3 protein – it was hardly detectable in the cell lysate of the "+Rep-double" clone, while no VP3 was detected in the "-Rep-double" cell lysate, fig.2.7-A (the complete VP3 absence observed for "-Rep-double" clone was probably due to inferior amount of proteins loaded, as seen from the signal of the loading control 3PGK).

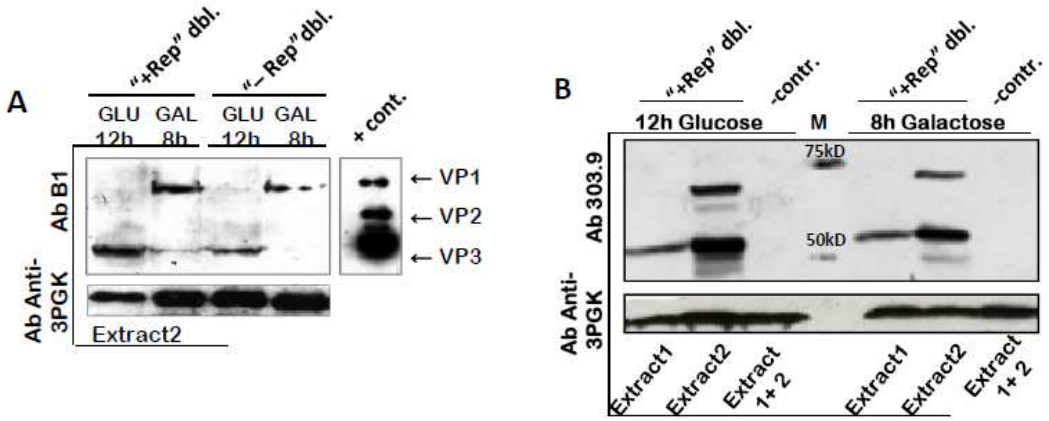


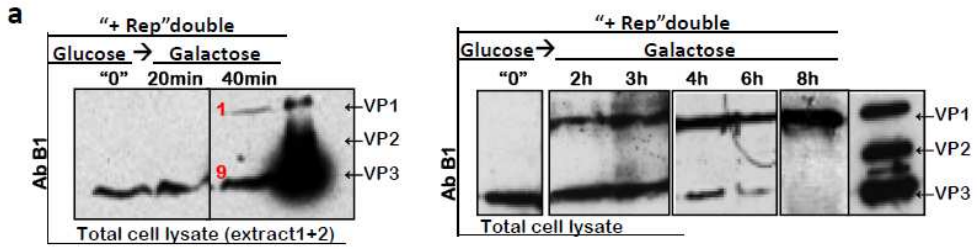
Figure 2.7 A, B: VP1 & VP3 expression in double-transformed yeast cells, listed in Table 2.1.
After 12h growth on glucose, "+Rep" and "-Rep" double cells were transferred to induction growth on galactose. (A) Cells were collected for protein extraction at the end of the glucose culture and after 8h of galactose induction and VP protein expression was analyzed by Western blot (B) Rep protein expression before and after galactose growth in "+Rep"double clones was analyzed by Western blot. Extracts from cells transformed with empty vectors, yEplac181 and pYes, were used as negative control.

However, Western blot analysis of Rep protein in "+Rep-double" clones showed that both p5 and p19 Rep isoforms were "constitutively" expressed throughout the whole culture regardless the carbon source used for the growth (fig.2.7-B) and their distribution to protein fractions was the same as seen before for the cells transformed with YplacRepCap; Rep78 was recovered in the extract 2 (insoluble fraction), while Rep52 was detected in both extracts; small amount of Rep68 and 40 were found only in the second extract. Probably the most interesting observation from this analysis was the loss of Rep68 during galactose growth, which may support the general hypothesis that splicing was negatively affected by the carbon source switch. A minor decrease in the Rep78 and 52 level after

galactose growth was probably the result of decreased growth rate in this carbon source (doubling time in galactose is almost twice as much the one in glucose, 4.5 h vs 2.7 h).

Decrease in VP3 protein level during culture growth on galactose suggests that galactose catabolism either prevented *de novo* synthesis of this protein or decreased its expression rate, resulting in VP3 "dilution" in the growing cell population.

Taken this into account, induction phase on galactose was gradually decreased and the situation improved. Figure 2.8 shows relative amounts of VP1 and VP3 proteins at different time points of galactose-grown "+Rep-double" yeast cell clones. Time "0h" corresponds to the end of the pre-induction growth phase, in 2% glucose medium, when only VP3 protein was accumulated. After glucose growth, cells were pelleted, water washed and transferred to 5% galactose medium. At different time points of galactose-induction (indicated in the figure) cells were collected for "optimized" cell protein extraction procedure, yielding two protein fractions, united prior to gel electrophoresis and Western blot analysis, for more correct analysis of VP1:VP3 ratios at each time point. VP1 expression was observed already 40 min after galactose induction (fig. 2.8, part-a, right) and its fast increase in the first 3h was followed with little or no reduction of VP3 protein in this time frame, observed by comparison of VP3 band intensities (fig. 2.8-a, left). With the induction time extended over 3 hours, a gradual increase in VP1 expression was followed by a decrease in the VP3 level, which was no longer observed at the end of galactose culture (8h), when VP1 expression reached its maximum (fig. 2.8-a, left). pYesVP1.KM + Yplacp40Cap double-transformants ("- Rep - double" clone) showed similar VP1:VP3 expression pattern over the time span of glucose + galactose growth (data not shown). The gradual increase of the relative VP1:VP3 ratio was calculated from corresponding band intensities at each of the indicated times and the values are presented in tabular format (fig. 2.8-b). This ratio most closely resembled the wt VP stoichiometry when galactose induction was reduced to 40 minutes from the moment when VP1 protein was first detected. The importance of the "correct" stoichiometry of capsid protein subunits has been already discussed and well reported to influence external capsid topology and thus, infectivity of produced particles, rather than efficiency of assembly *per se* [219]. Nevertheless, the weak point of the **"short induction" strategy,** in terms of VP1:VP3 optimization, is the difficulty in determining the exact time point at which VP1 expression homogenously set in for all cells in the population. This was observed by incomplete experimental reproducibility when the same growth conditions were repeatedly applied. In some cases cells responded to the same, "optimal" induction time (40min.) with either superior level of VP1 expression (comparing to the one shown in fig. 2.8-a) or with no detectable VP1 level (data not shown).



b Table: The gradual increase of the relative VP1:VP3 ratio with extended induction time

Induction Time	20 min	**40 min**	2h	3h	4h	6h	8h
VP1/VP3	0	**0.11**	0.9	1.4	2.2	5.2	++

Figure 2.8: The "short induction" strategy for preventing decrease of VP3 protein.
After 17-24h growth on glucose, "+Rep"-double cells were transferred to induction growth on galactose. Cells were collected for protein extraction and Western blot analysis of VP1 and VP3 expression at the end of the glucose culture ("0" time point) and at various time points of galactose induction *(a)*. Relative VP1/Vp3 ratios at each time point were determined by band densitometry and obtained values are shown in the Table. The best ratio, 1:9=0.11, was obtained after 40 min of galactose induction *(b)*.

Induction time up to 2-3h falls within the latency-phase of the culture (lack of cell divisions), typical of glucose-galactose nutrient switch, since, at the beginning of galactose consumption yeast *gal*-genes are repressed by intracellular glucose concentrations [238]. Likewise, repression of recombinant Gal promoter-regulated genes is still not completely relieved shortly after glucose depletion and small alterations in culture conditions such as altered oxygenation / pH, residual glucose in the culture medium, etc. (poorly controllable for shake-flask cultures), could cause either "premature or postponed" derepression of gal-regulated promoters. Moreover, water wash-step at the end of the glucose growth, during which, some cells in the population are temporarily starved and some are deoxygenized, can further increase variations in cell responses to substrate transition [238]. Hence, working with cultures at "transition state" becomes difficult and unsustainable. In addition, water wash-step between medium exchanges is inconvenient for scaling up the production system.

To avoid these challenges, we tried **the strategy of "mixed cultures"** (glucose + galactose), upon which, fine tuning of VP1:VP3 ratio was achieved by varying starting concentrations of glucose and galactose in "inductive" medium and by regulating induction time. We tested this approach under two different nutrient conditions: 1. high glucose (1.5%) and high galactose (2.5%) and 2. low glucose (0.5%) and high galactose (5%). In the absence of glucose, galactose is immediately metabolized for the cell culture maintenance resulting in its reduced availability for recombinant protein expression which consequently achieves its maximal level after only couple of hours of galactose induction

[165]. This is why in our system, the maximal VP1 expression is already achieved after 7-8 h of induction (see above) and is not further induced unless new galactose is added to the medium. When both nutrients are present in high concentrations (**high glucose, high galactose**), glucose, as a preferred carbon source, is used by the cell until its external concentrations are completely exhausted, after which yeast starts utilizing galactose for the maintenance [239]. Concordantly, during the first 9h of growth under high glucose (1.5%) and high galactose (2.5%) conditions, double-transformed cultures ("- Rep -double" and "+Rep-double") passed through 3 cell divisions (assessed by optical density measurement), with the growth-rate typical of glucose metabolism (with doubling time- 2.7h). Concordantly with glycolitic-like phenotype of p40 promoter (active in glucose), this growth was followed by *de novo* VP3 protein synthesis. This was observed by Western blot analysis (fig. 2.9, A), which showed that relative VP3 level increased after 9h of induction in high glu-high gal medium, accompanied with 3 cell divisions, comparing to its level before induction - "0" time point (the number of cells/wet weight used for protein extraction and amount of protein loaded on the gel was constant for all time points).

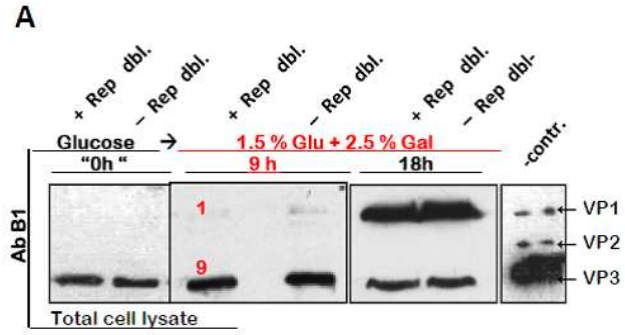


Figure 2.9 A: "High glu-high gal" - strategy for VP1:VP3 ratio optimization:
Western blot analysis of VP1:VP3 expression. *(A):* Glucose-grown "+ Rep" and "-Rep" double cultures were "induced" in presence of high glucose (1.5%) and high galactose (2.5%). "0" time point refers to the end of the preinductive, glucose growth phase, and the beginning of the mixed-culture. Cells were lysed before and after 9h or 18h of induction. The best VP1:VP3 ratio was detected for 9h induction time.

This increase in Vp3 level is consistent with increased metabolic rate following elimination of old medium and addition of fresh glucose (1.5% in the mixed culture), together with cell dilution upon medium exchange, known to boost culture growth. VP1 was detected 9h after the beginning of the culture (fig. 2.9, A) and VP1:VP3 ratio at this point was quite optimal (around 1:9). Although insignificant for the "correct" VP1:VP3 optimization, induction extended over 9h (this point can be considered as the beginning of VP1 expression under high glu-high gal growth conditions), resulted in more prominent VP1 increase and higher expression pick (18h lanes in fig.2.9A) with respect to its max expression achieved after 8h of induction in 5% galactose medium (see fig.2.8, a-8h). Beside, VP1 accumulation was followed by more controlled decrease of VP3 protein which could still be detected at the moment of VP1 expression pick (18h). This result confirmed that glucose addition to galactose medium enabled yeast cell a more rational

usage of galactose in the mixed medium, the strategy already seen in other yeast expression systems [165, 239]. Furthermore, in this system, VP1 expression was tightly controlled both by glucose repression and galactose upregulation subsequent to glucose exhaustion. In mixed cultures, glucose repression is alleviated much before the actual galactose utilization [238], suggesting more uniform "transition state" and thus, more uniform onset of VP1 accumulation for cells in the population.

Finally, based on these results we hypothesized that by limiting glucose to low levels in induction medium (**low glu-high gal**) the optimal VP1:VP3 ratio could be obtained at earlier time points comparing to high glu-high gal conditions. To test this hypothesis, "+Rep-double" 12h-grown glucose cultures were transferred for induction to the medium with the starting composition of 0.5% glu + 5% gal. The VP1:VP3 ratio was analyzed in the same way as described above. Western blot analysis was done for five different induction times where the longest one was 9h, equal to the minimal time required for VP1 detection under high glu-high gal conditions (fig. 2.9, B-top).

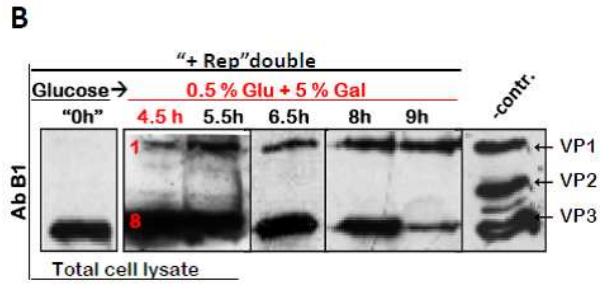


Table: The gradual increase of the relative VP1:VP3 ratio with extended induction time

Induction time	**4.5h**	5.5h	6.5h	7.5h	9h
VP1/VP3	**0.13**	0.23	0.3	0.5	1.4

Figure 2.9 B: "Low glu-high gal" - strategy for optimization of VP1:VP3 ratio:
Western blot analysis of VP1:VP3 expression *(B-top):* "+Rep"double culture was "induced" in the presence of low glucose (0.5%) and high galactose (5%) and cells were lysed before induction("0h") and after 5 different periods of induction. *(B-bottom):* the VP1:VP3 ratios are presented in the Table. The best ratio, 1:8 = 0.13, was detected after 4.5h induction.

The best VP1:VP3 ratio was obtained for 4.5h induction period, which according to band intensities was approx. 1: 8. In this time frame, initial growth on limited glucose concentration enabled the cell culture to divide after approximately 3h from the "mixed culture" onset while preserving *de novo* VP3 synthesis. The 1st division was followed by delayed 2nd division – aprox. 8h after the culture onset, consistent with the switch from glucose to galactose metabolism (doubling time on galactose is 5h). Thus, the gradual increase of relative VP1:VP3 ratio, calculated from band intensities and schematically presented in the table (Fig.2.9, B-bottom), probably resulted from VP1 accumulation prior

to 2nd cell division. After the 2nd division, VP3 dilution in the cell population in the absence of its *de novo* synthesis contributed to more prominent VP1:VP3 increase.

On the whole, the benefit of "low glu-high gal" over "high glu-high gal" strategy was obtaining preferential VP1-VP3 expression pattern in two times shorter period (4.5h instead of 9h). 4.5h induction on 0.5% glucose + 5% galactose was considered as "optimal induction". Together, 12h preinductive-logarithmic glucose growth plus 4.5h of the inductive growth were considered as "optimal growth conditions" under which the favoured VP1:VP3 ratio was achieved by: continuous VP3 accumulation, from the beginning of the glucose culture till the end of the mixed culture, while VP1 protein was produced during the inductive phase of the mixed culture.

2.3.4. Assembly of yeast-cell derived recombinant capsid proteins

High-speed ultracentrifugation through 40% sucrose-cushion.

To test the permittivity of yeast-cellular environment for VP1 and VP3 assembling in icosahedral virus-like particles, "+Rep-double" and "-Rep-double" clones were grown under "optimal" conditions. 3g of yeast cell biomass (~400 x 10^8) of each clone was lysed and processed by the method compatible with VLP extraction (under non-denaturing conditions), while a small aliquot of induced cells (1 x 10^8) was processed by optimized post-alcaline extraction method for Western blot analysis of relative intracellular levels of VPs in the total-cell lysates.

As seen before, individually expressed capsid proteins accumulated mainly as part of insoluble aggregates in total-cell lysates and were recovered after couple of sonication cycles in buffer with high concentrations of ionic and anionic detergents which consequently denatured the proteins. To enhance protein solubilisation, without damaging VLP structure, we used mild non-ionic detergents (DOC and NP-40) and elevated salt concentrations (0.5M NaCl) in the extraction buffer (DNS buffer). Large yeast biomass was disrupted by glass beads homogenization (see chapter 4). Besides, all so far used, AAV2 production systems, in vitro, Sf9 insect and HEK293 cells, showed that VP assembly intermediates and virus particles achieve better solubilisation in buffers of high-ionic strength [240]. Yeast cells were lysed by extended glass beads homogenization + short sonication in DNS buffer. In order to enrich putative virus-like structures in the soluble fraction, 15.000 g-centrifugation sediment of the crude lysate was subjected to the second extraction with elevated number of sonication cycles. United extracts were cleared from cell debris and subjected to ultracentrifugation through 40% sucrose cushion. Resulting pellet and supernatant fractions were analyzed by Western blot analysis for VP protein composition (fig. 2.10, A). VP proteins were found in the ultracentrifuge pellet and not in the supernatant fraction, consistent with their large size and multimeric organization.

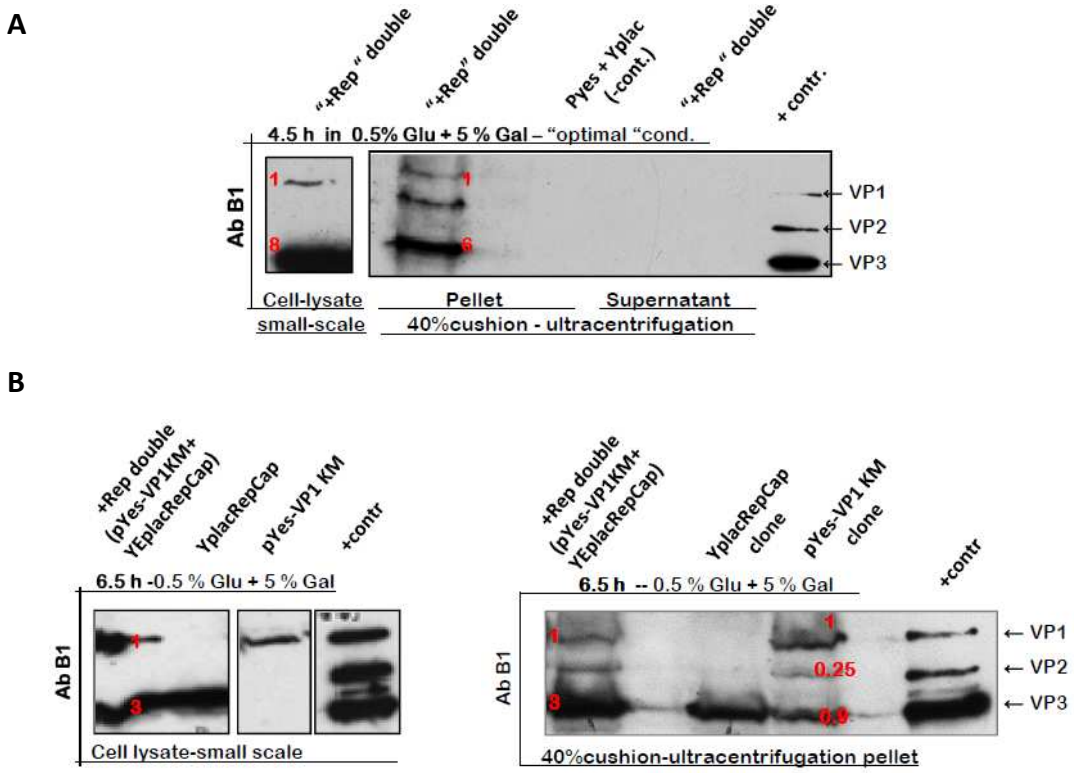


Figure. 2.10: Concentrating VP assembly products by high-speed ultracentrifugation through 40% sucrose-cushion.

(A): "+Rep " double yeast cells induced for 4.5h in 0.5 % Glu + 5 % Gal medium were subjected to a small scale-protein extraction (under denaturing conditions) and to a large scale, non-denaturing extraction. After 3h ultracentrifugation in 40% sucrose cushion of the native extract, supernatant and pellet fraction of the native protein extract were analyzed by Western blot for the presence of VP proteins. VP1, 2 and 3 were detected only in the pellet fraction, in the relative ratio ~1:2:6. VP1 and VP3 recovered in the total cell lysates were in the relative ratio 1:8. *(B):* YplacRepCap-transformed, pYes-VP1 KM- transformed and cells cotransformed with both plasmids ("+Rep double") were induced for 6.5 h in 0.5 % Glu + 5 % Gal medium. Small and large scale-protein extracts were processed as described in *(A).* Relative frequencies (numbers in red) of VP1, VP2 and VP3 in the ultracentrifugation pellets were analyzed and compared with relative amounts of VP proteins in the corresponding total cell lysates.

Surprisingly, beside VP1 and VP3, Western blot analysis of the pelleted material for the first time revealed the presence ofVP2 protein. This result was interpreted in terms of low overall VP2 expression, whose concentration in the bulk of yeast proteins (in the total cell-lysate) was under detection limit of Western blot analysis. Nevertheless, putative localized distribution of capsid proteins inside the yeast cell could increase "local" VP concentrations, allowing inter-VP homo and/or hetero oligomerization. Moreover, after 4.5h induction, VP1: VP3 ratio, detected in the ultrac. pellet (aprox.1:6, according to band intensities) was similar to the one in the total cell lysate (aprox.1:8). Though, VP2 was almost as frequent as VP1.

To confirm that the ratio of VPs in the pellet resembled their relative intracellular levels, cells were grown under "suboptimal" conditions that were shown to result in VP1: VP3

ratio higher from the "optimal" one. For instance, cells were induced for 6.5 h in 0.5% glu + 5% gal medium, previously shown to result in VP1:VP3 ratio 1 : 3.3 (see fig. 2.9, B-bottom). In addition, in the same experiment, we investigated the origin of VP2 protein expression since both constructs used for VP1 and VP3 expression comprised unmodified VP2 ORF. To do this, yeast cells transformed with single plasmids, VP1 expressing construct (pYesVP1K.M) and Vp3 expressing construct (YplacRepCap), were grown under the same conditions as double-transformed yeast cells (+Rep double), with 6.5h induction period. After induction, all three cultures were processed for "capsid purification" and subjected to 40% sucrose cushion ultracentrifugation. In parallel with the ultracentrifugation pellet analysis (fig.2.10, B-right), the total cell-extracts of single-transformed cells (processed by "optimized post-alkaline" extraction) were analyzed by Western blot to monitor the relative intracellular levels of VPs (fig.2.10, B-left).

This analysis yielded the following results:

1. In the ultracentrifugation pellet derived from pYes2VP1K.M – transformed cells, all three VP proteins were detected (fig. 2.10, B-right). As expected, the amount of VP2 in this pellet was inferior to VP1, but the level of VP3 protein was higher than expected from read-through translation mechanism (the relative VP1:VP2:VP3 ratio in this pellet was 1: 0.25: 0.9). Besides, as seen before, the absence of VP2 and VP3 from the total cell protein extracts (fig.2.10-B-left) confirmed their low overall intracellular concentration in the bulk of yeast proteins. However, oligomerizing /assembling properties of capsid proteins enabled the visualization of the two scarce VPs after assembly products were concentrated by ultracentrifugation;

2. The only capsid protein uniformly found in the ultracentrifugation pellet derived from YplacRepCap transformed cells (fig.2.10-B-right) **and in the corresponding total protein extract** (fig.2.10-B-left) **was VP3.** The presence of VP3 in the pellet suggests that this protein alone is able to form super-molecular-weight oligomers or higher order-structures, concentrated by high-speed ultracentrifugation through 40% sucrose cushion;

3. After 6.5 h induction, the relative VP1:VP3 ratio in the ultracentrifugation pellet derived from double-transformed yeast cells ("+Rep"-double) was almost identical to their intracellular stoichiometry. In detail, under suboptimal growth conditions (6.5h in induction) the relative VP1:VP3 level in the total cell lysate (deduced from band densitometry) was 1:3 (fig.2.10-B-left), while the VP1:VP2:VP3 ratio in the corresponding ultracentrifugation pellet was 1: 0.5: 3.3, respectively (fig.2.10-B-right).

Putting together the results obtained for +Rep double clones after 4.5h and 6.5h induction (characterized by great excess of VP3), we can bring out the conclusion that capsid assembly intermediates consisted of VP1 and VP3 in the relative ratio similar to their relative intracellular levels. However, upon high-speed ultracentrifugation VP2 was almost as efficiently recovered as VP1 protein, although its intracellular concentration was lower

and thus, undetectable by Western blot. Moreover, we detected almost the same amount of VP1 and VP3 in the pellet derived from pYesVP1KM-transformed cells, which is in the contradiction with the overall concentration of the three capsid proteins in this cell clone. In addition, we can say that VP3 alone was able to form super-molecular weight structures, suggesting that the main AAV2 capsid constituent, VP3, could be the limiting factor in highly-ordered interactions of cap proteins in yeast cellular background. Another factor that may be important for shaping VP protein interactions and assembly properties is the local concentration of the three VPs in cellular compartments permissive to VP assembling.

Fractionation of assembled capsid proteins by ultracentrifugation in CsCl gradient

To further characterize types of assembly-products concentrated by 40% sucrose cushion ultracentrifugation, a bigger culture volume of "+Rep-double" cells was grown under "optimal" conditions and 10g of yeast biomass were processed for "capsid purification" by glass-bead homogenization and subjected to 40% sucrose-cushion ultracentrifugation. Assembly products concentrated in the pellet were resuspended in DN buffer and DNAse digested to improve solubilisation. After clearance of aggregated, "junk" material, 6 ml of the supernatant ("CsCl input") was subjected to fractionation by CsCl density gradient centrifugation. 0.5% of the total input volume was analyzed in Western to determine relative amounts of capsid proteins in the "input" material (fig. 2.11-A). In the same figure, capsid proteins detected in the "junk fraction" represent the fraction of assembly products which failed to solubilise and were lost as aggregated material. These proteins were disaggregated only after resuspension in standard protein loading buffer followed by high-temperature incubation. Two different volumes of denatured "junk" fraction were analyzed for better estimation of the relative amount of lost material, suggesting that approximately 5-10% of the total VP protein yield recovered by the first ultracentrifugation was lost due to poor solubilisation.

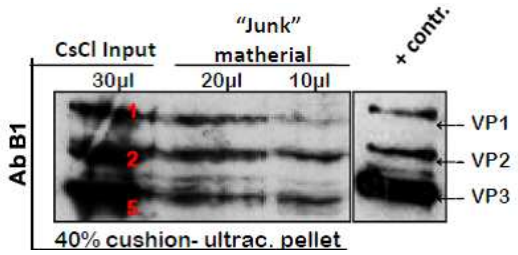
After 48h of CsCl fractionation, 13 different fractions of gradually increasing densities (f1 is the fraction of the minimal density (1.370) and f13 of the maximal one (1.475)) were collected, dialyzed against PBS, and each fraction was analyzed for VP protein composition in the Western blot (fig.2.11-B). Small aliquots of two or three adjacent fractions of similar density (1+2, 3+4, 5+6+7, 8+9, and 10+11) were united for transmission-electron microscopy (TEM)-based detection of putative capsid-like structures (fig.2.11-C). Fraction 13 (named "pellet") was not analyzed.

Western blot identification of VP-containing fractions

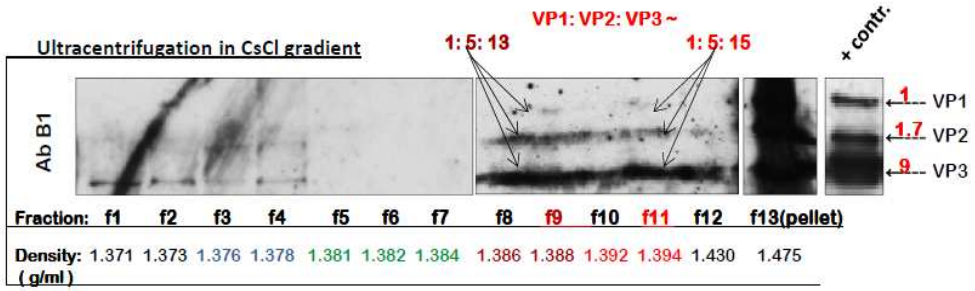
For the highest assay sensitivity, maximal volume of each fraction was loaded on SDS-page gel and blots were incubated with both B1 and A69 Ab (with partial stripping step in between-see chapter 4) (fig.2.11-B). The analysis revealed all three capsid proteins over the range of fractions, in different combination. In the 1st fraction (one of the lowest

density), only VP3 protein was detected, probably resembling VP3 homo-oligomers of the smallest molecular weight; fractions 2 and 3 contained almost equal amounts of VP2 and VP3; the 4th fraction contained all three VP proteins, whereas no VPs were detected in 5-7 fractions; finally, 8-11 fractions contained all three VPs. Relative VP frequencies (1:2:7) in f8-f11 most closely mirrored and the ones found in 293-cell derived AAV2 vectors (1:1.5:8) used as positive control (a kind gift from Mauro Giacca's lab). VP3 was the only capsid protein detected in f12. Very weak band intensity in each fraction, comparing to the last fraction- f13 ("pellet"), suggested that lot of input material was lost in the form of heavy aggregates.

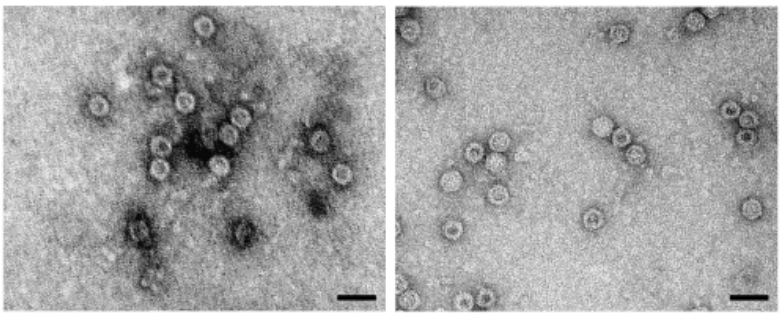
A



B



C



D

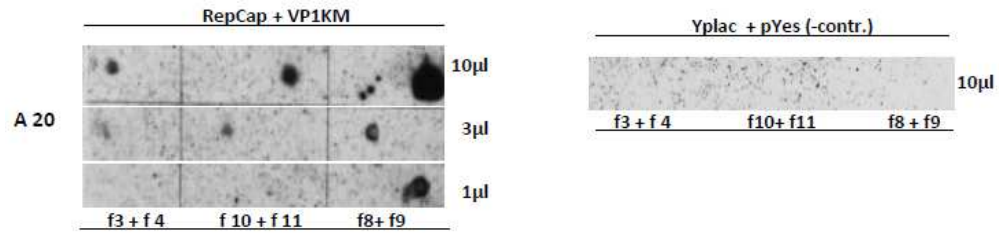


Figure.2.11. Purification of yeast cell-derived AAV2 capsid like-structures by CsCl-gradient ultracentrifugation.

Native protein extracts derived from "+Rep" double yeast cells induced for 4.5h in 0.5 % Glu + 5 % Gal were subjected to 40% sucrose cushion-ultracentrifugation. Pelleted material was resuspended in capsid-resuspension buffer (see chapter 4) and cleared from insoluble aggregates = "junk" by high-speed centrifugation.

(A): Western blot analysis, using mAb B1, shows relative frequencies (numbers in red) of VP proteins in 0.5% (30µl) of the native extract, "Input", and relative amounts of VP proteins lost as insoluble aggregates in "Junk" fraction.

(B): The Input was subjected to CsCl –gradient ultracentrifugation for 48 h. 13 fractions of different densities were recovered and analyzed for the presence of VP proteins by Western blot. Structures recovered in fractions 8-11 had VP compositions that most closely resembled the one of wt capsids.

(C): Fractions of similar densities (indicated in the same colour) were united and subjected to TEM analysis. Capsid-like structures of ~20nm size were identified in fraction f8+f9 (*i*) and compared with 293T –derived AAV2 empty capsids (*ii*). Scale bar is 40 nm.

(D): 3 fraction pairs that gave positive results in TEM were spotted on the nitrocellulose membrane in three quantities indicated on the right side bar and analyzed for the presence of AAV capsids with the capsid-specific mAb A20 antibody. The strongest signal (which indicates the greatest number of capsids) was detected in the fraction f8+f9. As negative control of the assay, the same number of cells cotransformed with empty vectors, Yplac181 and pYes, were processed as described in (A) and the obtained CsCl fractions of the corresponding densities were incubated with A20.

Electron microscopy analysis of CsCl fractions

After fractions of similar densities were united, four resulting "fractions" were analyzed by negative staining and electron microscopy. Groups of capsid-like structures were observed in 2 out of four fractions, f8+f9 (fig. 2.11-C) and f10+f11 fraction, while in the fraction f3+f4 only singular globular structures, of the correct size (≈20nm) were detected in a very low frequency (data not shown). As shown in the Western blot results from these fractions, f8-f11 contained three capsid proteins in the relative frequencies similar to the ones of wt AAV-2 capsid. Micrographs in the figure 2.11-C show that yeast-cell derived capsid-like structures (left) were morphologically similar to the ones obtained from 293 cells (right) (HEK293T cell-produced empty AAV2 capsids are a kind gift from Jurgen Kleindchmidt).

The density of capsid positive fractions is far over the expected ones for empty AAV capsids produced in mammalian cells (approximately 1.32 g/ml) [241]. This could be explained in terms of a "heavier yeast extract" since all obtained fractions had higher densities than the corresponding fractions derived from AAV producing 293T cells (e.g. the lightest fractions were 1.37 for yeast and approximately 1.28 g/ml for 293T cells) [241]. This "heavy extract" might be the consequence of cellular impurities (e.g. remaining

chromatin; although the extract was treated with DNase the extract was still viscous when loaded on the sucrose cushion), ineffectively removed in the upstream processing steps. These steps are extremely laborious since a large cell biomass is used for capsid production (around one billion of cells). Beside, the sucrose remaining from the first ultracentrifugation and high concentration of salts in the resuspention buffer increase the density of assembly intermediate-containing solution. In such an ambient AAV capsids may tend to aggregate and thus separate as the fraction of higher density. Finally, an additional CsCl gradient centrifugation of capsid positive fractions should overcome the impurity and aggregation problem and unable better characterization of capsids. Yet, this step would significantly decrease the overall capsid concentration and requires improvements in the upstream steps and in total yield of capsids proteins. Another possible explanation for the unusual densities of capsid positive fractions could be the presence of the variegated capsid population with empty and full/partially full capsids. Several groups have already demonstrated that AAV capsid can encapsidate different DNA impurities. These includes both viral *cap* and/or *rep-cap* sequences [242, 243] as well as pieces of a host cell genome or prokaryotic DNA of plasmids used for viral production. Besides, Kronenberg et al. have shown that only full AAV2 capsids, but not the empty ones, externalize VP1-VP2 N-terminus, using the energy of the DNA encapsidation into preformed capsids [63]. According to these studies, an indirect way to see weather yeast produced AAV capsids contain also full particles, might be to test their reactivity with N-terminus specific antibody (e.g. Ab A1, [53]). However this approach is hardly feasible in the case of low capsid yield since the percentage of DNA containing particles may be too low to be detected with such an approach.

Immunoreactivity of yeast-derived particles with anti-capsid A20 antibody

To support electron microscopy results, 3 possibly “VLP-positive” fractions were analyzed in Ab-A20-based dot-blot assay for testing immunoreactivity of yeast-derived particles with anti-capsid A20 antibody (Fig. 2.11-D). A20 is a widely used monoclonal antibody with the affinity for conformational capsid epitope and does not bind native capsid proteins [53, 216]. As shown in the Figure, all three fractions showed reactivity to A20 antibody, with the highest concentration of A20-reactive virus-like particle in f8+f9 fraction.

2.4. Conclusions and future prospective

A considerable body of data about yeast-expressed viral structural proteins, able to assemble into authentic and chimeric virus-like particles (VLPs), has been reported up to date, including the recent success in *S. cerevisiae*-based production of VLPs of human Parvovirus, B19, the “autonomous” relative of AAV virus. This encouraged us to consider the possibility of using yeast *S. cerevisiae* as an *in vivo* cell system for production of AAV2 VLPs, with physical and biochemical properties similar to those of AAV vectors produced in mammalian cells. To do this, two requirements had to be satisfied: the

simultaneous expression of at least two out three AAV structural proteins, VP1 and VP3 that form the structure of infective virions, and, secondly, the intracellular stoichiometry of these proteins should be similar to the one found in AAV vectors produced by mammalian cells. The co-expression of VP1 and VP3 in RSY12 strain was achieved by stable cell co-transformation with two high copy plasmids, one carrying "VP1 expression cassette" under inducible yeast promoter Gal1" and the other one harbouring "VP3 expressing", genuine AAV cap gene under its cognate p40 promoter. VP3 protein expression from AAV p40 promoter in yeast cells resembled the expression profile of yeast glycolitic-promoters characterized by positive correlation with the cell growth rate during glucose-based growth and expression cessation upon carbon source exchange from glucose to galactose [226]. On the other hand, the pGal1 galactose inducible promoter, guiding VP1 protein expression, is a strong, inducible yeast promoter which yielded the first detectable amounts of VP1 protein, already 40 minutes after removal of repressor (glucose) and addition of the inducer (galactose) to the growth medium. Upon this minimal induction time, the co-transformed yeast cells synthesized small amount of VP1 protein while sustaining high VP3 intracellular level accumulated during the pre-induction growth on glucose. Although not exactly simultaneous, but rather consecutive expression of VP1 and VP3 proteins yielded the "correct" stoichiometry of their final intracellular concentrations. However, the "short induction" strategy was poorly reproducible since it was based on marginal VP1 expression in the first moments of induction which, in turn, was prone to fluctuations even if the same inducing and growth conditions were used. Certain determinants of the induction start point, such as residual glucose concentration in the growth medium and actual state of the cells in the moment of induction onset (e.g stress caused by instantaneous starvation and deoxygenation during the water-wash step between nutrient exchanges) are hardly controllable with the flask-shake cultures, causing poor reliability of this strategy. These lead to adoption of another strategy "glucose + galactose, mixed culture approach", the strategy commonly used for high-yield production of heterologous proteins in yeast expression system [165]. Some general characteristics of this strategy introduced significant improvements into our expression system such as: elimination of stressful and inconvenient water-wash step, sustained p40 promoter activity and better control of Gal1 promoter both by galactose-based upregulation and glucose-based repression. This resulted in simultaneous expression of VP1 and VP3 as long as glucose was present in the medium, thus enabling fine tuning of the relative VP protein levels by varying both induction time and the amount of glucose in inductive medium. The best results in terms of: *i*) "optimal, wt-like" VP1:VP3 ratio and *ii*) experimental reproducibility and convenience (the shortest induction time), was obtained with keeping residual glucose (0.5%) concentration in reach galactose (5%) inductive medium. The "low glu-high gal" strategy seemed to satisfy all requirements for efficient assembly of capsid proteins in multisubunit VLPs. To check this, we adopted a standard, two step-procedure for AAv capsid purification, composed of: and encouraged us to proceed to the next step, capsid

purification. This was done by implementation of a standard 2 step procedure that included: 1) precipitation of assembly intermediates from native protein extracts by high-speed ultracentrifugation through 40% sucrose-cushion, followed by 2) capsid purification by fractionation in CsCl gradient. Already the first step revealed the unexpected presence of VP2 protein in VP1-VP3 containing assembly products. 40% sucrose cushion ultracentrifugation of the extract derived from yeast clone transformed with the "VP1 expressing construct" showed that small amounts of VP2 and VP3 protein (not-detectable in the total cell lysates) were expressed from this construct by means of "slipping ribosome" mechanism of alternative translation initiation from VP2-ACG and VP3-ATG start sites comprised in the "polycistronic" VP1 cDNA. Although expressed in low concentrations by this clone, VP2 and 3 were able to perform homo-oligomerization and/or hetero-oligomerization with the more abundant VP1 protein generating high molecular weight structures, precipitated after high-speed ultracentrifugation.

Going back to assembly intermediates derived from co-transformed yeast cells, their characterization in the 2nd step CsCl ultracentrifugation, revealed only small percentage of AAV capsids with properties similar to the wt ones, obtained in couple of fractions of CsCl gradient. These properties included VP composition, capsid architecture (assessed by electron microscopy), physical, biochemical (high CsCl stability and stability during various processing steps) and immunological properties (reactivity to A20 antibody). However, great percentage of assembly intermediates concentrated in the 1st step ultracentrifugation were actually recovered in a form of insoluble aggregated material. The similar situation was seen in the first Baculovirus-Insect cell VP-expression system made by Ruffing et al. [215] where 3 VPs were expressed from separated ORFs (cDNAs), each one carrying mutations in the other two VP-start codons. Later on, the formation of insoluble aggregates in this system was circumvented by producing VPs from overlapping ORFs (a single transcriptional unit), suggesting that mutations introduced were probably the reason of poor capsid assembly or of their aggregation [138]. However, no such mutations were present in VP expressing constructs in our system.

A common propensity of unassembled VP monomers to associate with cellular structures has been documented in various recombinant backgrounds including: aforementioned Baculovirus-Sf9 system, *in vitro* system and HeLa cells, and is very prominent in our *S. cerevisiae* expression system. Besides, L1 monomers of HPV16 expressed and assembled in VLPs in *yeast P.pastoris*, showed the same tendency [195]. L1 association with cellular structures was probably exhibited through hydrophobic regions of unassembled monomers. Cellular membranes could serve as sites of "molecular crowding" which could assist monomer assembly in VLPs. Majority of cellular VP1 and VP3 protein expressed by yeast, were recovered from insoluble cellular aggregates. This suggests that yeast-expressed capsid proteins have similar membrane-binding affinities. However, even small differences in these affinities (due to different N terminal extensions of VPs) could result in such

relative concentrations of the three VP proteins, in these “crowding zones”, different from their overall intracellular stoichiometry that we document upon their denaturizing extraction. The impact of such alterations on the final concentration of assembled capsid probably gains significance in the case of low overall production levels of VPs in our expression system. Besides, it is also possible that AAV particles of non-wt stoichiometry could have different surface properties, with certain hydrophobic regions exposed on the capsid surfaces that may cause VLPs to attach to cellular membranes. Normally, wt AAV virions are soluble while their hydrophobic residues are buried in the capsid interior [244].

Furthermore, foreign protein localization to insoluble cellular aggregates is often associated with its ubiquitination and further targeting to proteasome for degradation. We tend to exclude this possibility for VP proteins expressed in RSY12 yeast strain, since there was no difference in the total amounts and the degradation profile of VP1 and VP, when expressed in HT393 mutant strain that lacks couple of main proteasomal subunits. These comparative experiments were performed in the starting phase of this project and due to mentioned lack of differences the mutant strain was abandoned.

Finally, certain steps in the downstream processing (particle purification) could be additional limiting factors for the final capsid yield. Critical in this view is the glass bead-based mechanical cell rupture with unknown affects upon AAV capsid adherence to these artificial surfaces. However, since this is one of the standard procedures for VLP purification form yeast cells, it is rather the efficiency of the cell rupture which is the critical factor of the capsid yield.

To improve production properties of our yeast cell-based AAV VLP expression system various techniques can be employed, some of them being already under way.

Overcoming methodological obstacles:

Our *S. cerevisiae* VP expression system is dependent on the fine setting of growth conditions during two production phases: pre-inductive 2% glucose growth (more than 12h) and 4.5h inductive growth in low glucose (0.5%)- high galactose (5%) medium. With the use of flask cultures and a shaking incubator, the indicated amounts of each nutrient present actually the starting concentrations and cannot be maintained as such during the whole culture, significantly affecting the final protein yield. Beside, due to a glucose exhaustion and accumulation of the metabolic by-products at the end of the pre-inductive phase, the cells have to be pelleted and culture medium exchanged, instead of simple addition of galactose. Albeit less stressful than water-wash starvation, the harvesting step can still cause significant environmental fluctuations which can be differentially sensed by the cells in the population creating a sort of “variegated” system [245]. In such a system, following galactose induction could have different kinetics among cells in the population. If the residual glucose and high galactose levels could be kept constant in all cells of the population during extended culture growth, a prolonged ”induction” would result in a

multifold increase of both VP1 and VP3 levels. This can be achieved by implementation of novel technologies such as the chemostat continuous culture system to provide steady-state conditions and constant growth rates [246]. This can be achieved by adjusting the flow rate of fresh medium (with precise amounts of each nutrient) into the growth chamber, providing a defined environment amenable to high-yield production [247]. The other system which offers the same benefits of constant feeding, acidity and culture oxygenation, is so called fed-batch cultivation in bioreactors (fermentors) [179].

Expression of AAV assembly activating protein (AAP) in yeast:

Another strategy to test in order to meliorate the efficiency of AAV capsid assembly in yeast cellular environment is to test weather AAP protein, co-expressed with VPs and/or overexpressed from the separate plasmid, could exhibit the same assembly promoting effect seen in mammalian cells [64]. Since expression from CTG initiator is very rare and inefficient in yeast cells, AAP was not expected to be expressed in its natural genomic configuration (e.g. from YplacRepCap construct). Indeed, when we cloned the authentic cDNA of AAP gene in the separate plasmid, under strong Ga1 promoter, no protein was detected. Instead, we have recently changed the AAP's CTG initiator to the conventional ATG one, in the same plasmid, in frame with the FLAG tag sequence for the visualisation of the fusion protein. Protein expression analysis is currently under investigation.

"Yeast-izing" AAV *cap* gene:

When a recombinant protein is not expressed to a sufficient level, the first "solution" that comes to a head of an average researcher is: "let's change the promoter"! This is not surprising since "expression = promoter" used to be a concept of recombinant DNA technologies for many years. However, translational and post-translation regulations of protein expression significantly differ between evolutionary distant species, such as yeast and other metazoan. Differences in the abundance of iso-accepting tRNAs that resemble the relative frequencies of synonymous codons, are observed between organisms, and also between the tissues of a given organism, and are postulated as one explanation for "why the same gene is more or less efficiently translated in different organisms and organs" [248-250]. The codon usage bias between the viral and the host cell genome is shown to be an important determinant of viral protein expression within the host cell [251].

Virus genomes frequently have a G+C content significantly different from that of their host species, reflecting a different codon usage pattern [251, 252]. Similarly, the GC composition of AAV genome and *S. cerevisiae* genome is tabulated. Table shows that GC content in AAV2 and *S. cerevisiae* genomes are quite different ranging from 40.65–67.53% and from 36.64–44.58%; frequencies in the 3rd letter of the codons are strikingly different suggesting great codon usage bias between these two species.

Table2.4. Comparison of the G/C content in the genomes: Total GC composition and the GC content at the 1st, 2nd, and 3rd codon position in each genome are listed as indicated

	Total content (%)	1st (%)	2nd (%)	3rd (%)
S. cerevisiae	39.77	44.58	36.64	38.10
AAV2	54.73	56.01	40.65	67.53

Analysis of codon frequencies between yeast and Cap coding region revealed that 14 most frequently used codons in AAV cap gene are more than twice as frequent as the ones in yeast genes. Low abundance of iso-accepting tRNA for these codons, at least in part, underlies low expression of the VP proteins in this system. Codon optimization of cap gene sequences, or "yeast-izing" the AAV genes, so that they can reflect more closely the codon usage of the host (yeast), should maximize the expression of the "optimized" gene and is currently under investigation in our lab.

3. PROJECT II: Formation of AAV single stranded DNA genome from a circular plasmid in *Saccharomyces cerevisiae*

3.1. The aim of the project

In the first part of this work, we demonstrated for the first time, that yeast is able to form AAV2 capsid like particles. As the final aim of the entire lab project is to use yeast as a host to produce rAAV vectors able to infect human cells, we have also studied the ssDNA formation and replication in *Saccharomyces cerevisiae* and the mechanism underlying this process.

3.2. The background of the project

3.2.1. AAV replication in the recombinant background

The most important function (and the origin of the name) of the two largest Rep isoforms, Rep78 and Rep68, is the regulation of AAV-2 replication during the proliferative phase of AAV2 life cycle. They perform this function via their interaction with Rep-binding element (RBE) and terminal resolution site (trs) sequences located within the ITRs. Besides, these two Rep isoforms guide site-specific integration of AAV genome during the latent phase of the virus life cycle [28-31]. The small Rep protein species, Rep52 and Rep40 are involved in the generation and accumulation of ssDNA viral genomes from double-stranded replicative intermediates [32, 33, 47, 48]. Since functions of two large and two small Rep proteins are overlapping, one of each (Rep78 and 52 or Rep68 and 40) has to be provided *in trans* to enable replication and packaging of 'gutless' AAV vectors (ITRs flanking the gene of interest) in the recombinant background. Although helper virus functions are compulsory for rAAV2 replication in mammalian cells, this is not the case with Insect cells. Either these cells, or Baculovirus used to introduce rAAV genomes, seem to provide all necessary factors for efficient rAAV replication in the presence of Rep proteins [138, 145, 146]. Previous studies using yeast two/one hybrid system have shown that yeast expressed Rep 78/68 proteins retain their RBE site specific, binding activity, which is the first step in establishing their role in promoting replication of rAAV genomes [229]. Weather this role can be completed upon rAAV introduction in yeast cells is the leading question of this project and will be addressed in this chapter.

3.3. Results

(The most of results in this section have been recently published by our group; Ref [253])

3.3.1. Rep68 increases the frequency of yeast clones containing the AAV genome.

To explore whether the yeast *Saccharomyces cerevisiae* is able to sustain the rescue of the AAV genome from a plasmid containing the ITRs, we constructed the pAAVRepURA3 vector containing the wild type rep gene and the URA3 marker flanked by the ITRs (Fig.3.1A) and the pAAVpokURA vector which does not contain any AAV sequence, but only URA3 gene and the stuffer sequence, pok (Fig. 3.1B).

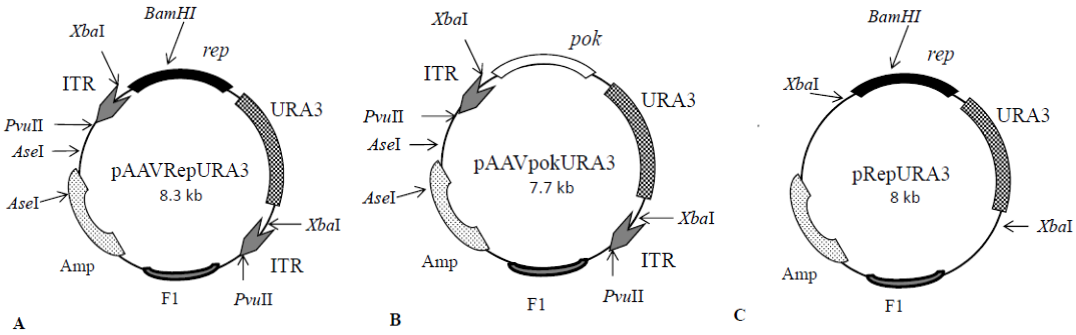


Figure 3.1. Schematic representation of plasmids.

(A) The vector pAAVRepURA contains wild type *Rep* gene and the yeast *URA3* marker replacing the *Cap* gene flanked by the two ITRs. B) The vector pAAVpokURA contains a stuffer sequence, pok, instead of *Rep* gene. The ITRs are the only AAV sequence present in the vector. C) The vector pRepURA has the same sequences as pAAVRepURA without ITRs.

With the vector pAAVpokURA, we wanted to determine whether the internal Rep sequence could affect ssDNA formation. These vectors do not share any homology to the genome of the RSY12 strain which harbors the complete deletion of the chromosomal URA3 gene [254]. Moreover, as pAAVRepURA3 and pAAVpokURA do not carry any yeast replication origin such as ARS or 2 micron, thus, they cannot give rise to URA3 colonies unless the vector integrates in the yeast genome or replicates extra chromosomally. The pAAVRepURA3 and pAAVpokURA were co-transformed into the strain RSY12 with the plasmid containing the Rep68 expression cassette, or with the control plasmid pGAD424. The plasmids pGAD424 and pGRep68 carry the LEU2 gene as yeast selection marker. As shown by counting the number of colonies grown onto the plates with medium lacking both uracil and leucine (Fig. 3.2A), the frequency of URA3+LEU2+ colonies increased from $3.05\pm1.1x10^{-3}$ (pAAVRepURA3 and pGAD424) to $91.7\pm43.6x10^{-3}$ (pAAVRepURA3 with pG.Rep68).

Similarly, the frequency of URA3+LEU2+ colonies rose from $18.1\pm5.2x10^{-3}$ (pAAVpokURA with pGAD424) to 63.3±19.4x10-3 (pAAVpokURA with pG.Rep68) (Fig. 3.2B).

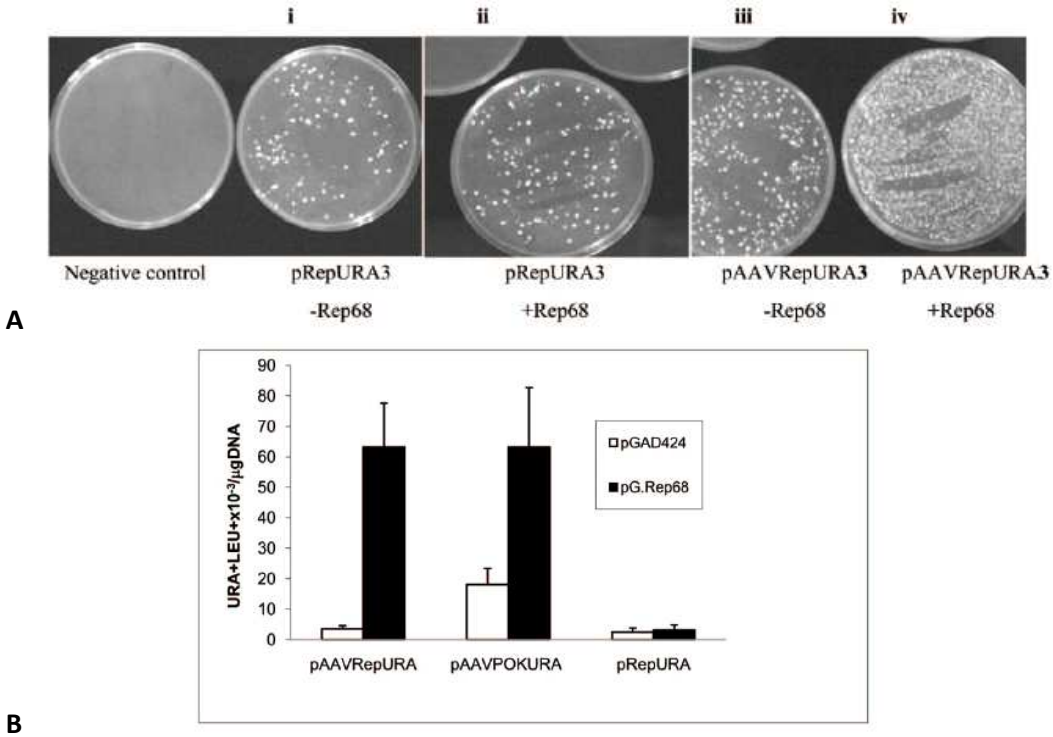


Figure 3.2: The frequency of colonies carrying rAAV genome increase when Rep68 is expressed. The plasmid pAAVRepURA3, pAAVpokURA linearized with *Pvu*II and the control plasmid, pRepURA3, digested with *Xba*I, were co-transformed with the plasmid pGAD424 or pG.Rep68 containing the *LEU2* marker gene (A, B). (A) Representative plates comparing colonies obtained from the transformed yeast RSY12 strain with plasmid pRepURA and pGAD424 (i), pRepURA and pG.Rep68 (ii), pAAVRepURA and pGAD424 (iii); pAAVRepURA and pG.Rep68 (iv). (B) Transformed yeast cells were scored for their ability to form colonies on selective medium lacking leucine and uracil. The frequency was calculated as described in material and methods. Results are the mean of 4 independent experiments±standard deviation.

The expression of Rep68 did not change the frequency of URA3+LEU2+ colonies when yeast was transformed with the plasmid pRepURA3 which does not contains the ITRs (fig 3.2B). These results suggest that the effect of the Rep68 is related to the presence of ITRs and that Rep68 may be important for AAV replication and/or rescue in yeast as it is in human cells.

3.3.2. Expression of Rep proteins

Rep protein expression from natural AAV2 promoters, p5 and p19, preserved in rep gene of pAAVRepURA constructs.

Similarly to YplacRepCap plasmid, where Rep proteins are expressed from p5 and p19 promoter in their wt genomic configuration, the same result was obtained for pAAVRepURA vector where complete Rep coding region was preserved. Western blot analyses (Figure 3.3A) showed that the distribution of Rep proteins to protein fractions obtained by “optimized post-alkaline” extraction was the same. As seen before, Rep52 and

Rep40 are present both in the extract 1 (fig.3.3A, lane1) and extract 2 (lane2), while Rep78 is observed only in the second protein extract (Fig 3.3A, lane2). Rep68 is not observed.

Rep68 protein expression from Constitutive Yeast ADH1 Promoter

The properties and schematic of pG.Rep78 was already represented in the chapter 2.3 (see fig.2.2A). pG.Rep68 was made in the same way as p.G Rep78, by placing Rep68 cDNA under truncated version of yeast ADH1 promoter. Originally, these plasmids were made for the use in yeast one-hybrid system where expression of large Rep proteins was demonstrated indirectly by Rep78/68 mediated transcriptional activation of reporter gene expression, while, according to the author, these proteins couldn't be detected by the Western Blot analysis. In our hands, in RSY12 strain background, Western Blot analysis showed significant expression of large Rep proteins from both constructs.

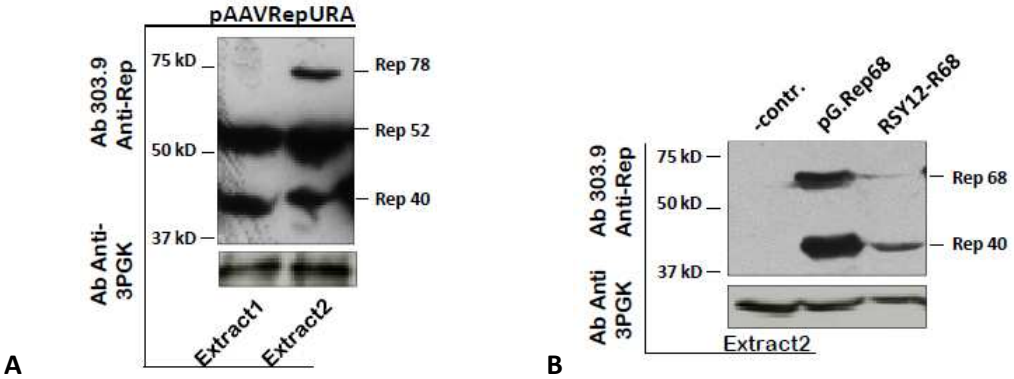


Figure 3.3: Western blot analysis of the Rep protein expression in yeast.

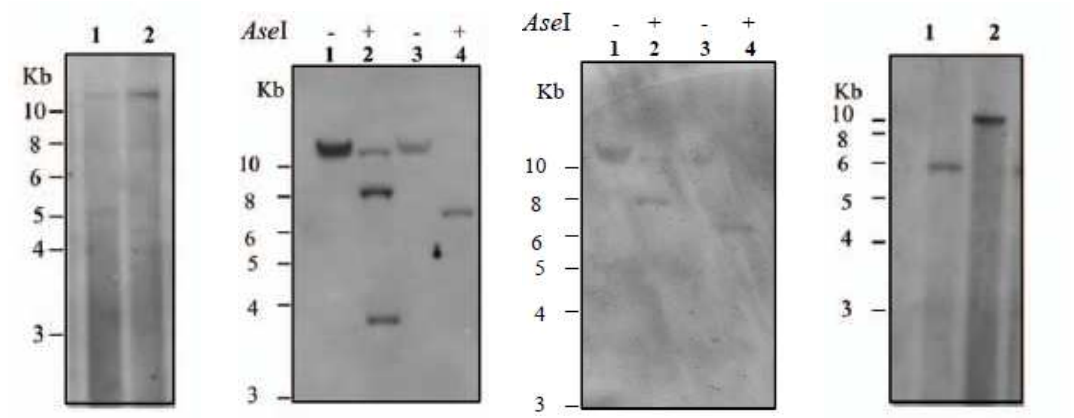
(A) Distribution of Rep proteins in two protein fractions (extract1 and extract2/insoluble fraction) derived from cells transformed with pAAVRepURA. (B) Analysis of Rep proteins present in the extract2 of lysates derived from cells transformed with pG.Rep68 plasmid (lane2) and from RSY12-R68 clone carrying 2 integrated copies of ADH-Rep68 expression cassette (lane3). Extracts from pGAD (empty vector) transformed cells were used as -control. Loading control: Constitutive yeast protein 3PGK.

Results for pG.Rep68 are shown in the figure 3.3B. Besides, just like for pG.Rep78, p19 regulated coding region comprised in Rep68 cDNA gave rise to Rep 40. As expected, Rep68 was mainly found in insoluble protein fraction, extract2 (fig.3.3B, lanes2), while Rep40 was almost equally distributed between the first and the second protein extract (fig.3.3B, lanes1, 2). In the lane 3 of the fig.3.3B we present the relative amount of the rep proteins expressed form the yeast clone carrying the pADH-Rep68-tADH expression cassette randomly integrated in the yeast genome. Comparing to RSY12 cells transformed with multicopy (more than 20 copies) plasmid harboring the Rep gene, apparently smaller amount of Rep proteins corresponds to 2 copies of Rep68 cDNA detected by the Southern blot analysis of the genomic DNA derived from this cell clone (Southern blot data not shown). However, this cell clone stably expressed rep proteins after many repeated growth cycles in the complete growth media (without selection).

3.3.3. AAV ssDNA formation in yeast

The AAV genome inserted into a plasmid vector can initiate a productive AAV replication when it is transfected in human cells that are simultaneously or subsequently infected with a helper virus. The AAV genome is released from a circular plasmid in a way that is similar to the rescue of the integrated AAV provirus in latent phase [19]. It has also been observed that the rescue of the AAV genome in HeLa cells extracts is more efficient when the Rep68 protein is expressed [26]. We, therefore, checked if Rep proteins expressed from pAAVRepURA were sufficient to rescue AAV genome from the circular plasmid in yeast. To do so, low molecular weight (Mr) DNA from URA3+ yeast clones transformed with the pAAVRepURA was analyzed by Southern blot and probed with URA3 gene to check for the presence of rescued ssDNA, which is expected to be about 3kb (Fig. 3.4A). As we observed a band with a molecular weight higher than 10kb, we concluded that rescue did not occur and that the URA+ phenotype is determined by integration of the vector in the genomic DNA. We therefore, analyzed the genomic DNA extracted from these clones, after digestion with *Ase*I. This enzyme cuts the pAAVRepURA backbone in two sites but not the AAV sequence (Fig. 3.4G). The analysis revealed the presence of multiple (fig 3.4B and C, lane 2, or single integrated genomes (fig 3.4B and C, lane 4). This result suggests that a certain level of Rep proteins might be necessary to keep the pAAVRepURA plasmid replicating autonomously inside the yeast cell. The Rep protein expression can be directed by the AAV natural promoters p5 and p19, as already shown, but if AAV replication in yeast occur in Rep dependent manner, we hypothesized that *rep* gene should be located in another vector in order to provide a certain amount of Rep to initiate the AAV genome replication. It is likely that the pAAVRepURA vector is maintained as an episome only when Rep proteins reached quite high level of expression. To verify this hypothesis, we transformed yeast cells that have the pAAVRepURA integrated in the genome with the plasmid pG.Rep68. Two of the transformed URA3+LEU2+ clones were subjected to the extraction of low Mr DNA from and analyzed by Southern blot. This analysis revealed the presence of only a band of ~6kb in one clone (Fig. 3.4D, lane 1) and a band with a molecular weight higher than 10kb in the another clone (Fig. 3.4D, lane 2). The band of 6kb could be due to a plasmid excision event occurring by intrachromosomal recombination that has been reported to occur at high frequency in haploid yeast strain [255, 256]. We, then, investigated the possibility that AAV ssDNA rescue took place after the co-transformation of pG.Rep68 and pAAVRepURA. Under these experimental conditions, yeast may have a sufficient level of Rep proteins, necessary to induce AAV ssDNA rescue from the transformed plasmid. Low Mr DNA was extracted from several independent clones, digested with *Dpn*I and analyzed by Southern blot using the URA3 as a probe. *DpnI* digestion allows discriminating between DNA replicated in yeast from that replicated in bacteria because it cleaves only double-stranded sequences methylated or hemimethylated by *Dam* methylase. As shown in the figure 3.4E in two clones transformed with the pAAVRepURA and the empty pGAD,

we mainly observed that the highest band is digested by *Dpn*I (Fig. 3.4E, lane 2 and 4). This band may correspond to the transformed plasmid; the other band, higher than 10kb, resistant to *Dpn*I digestion, corresponds to genomic DNA; in fact, probing low Mr DNA from the same clones with the genomic marker ADE2, the same band was observed (data not shown). In the two clones, no band corresponding to ssDNA was observed. On the contrary, when we analyzed the low Mr DNA extracted from yeast cells co-transformed with pAAVRepURA and pG.Rep68, four main bands were observed: a band higher than 10kb that is genomic DNA, a band of about 10kb, and two smaller bands of about 5.5kb and 2.5-3kb (Fig. 3.4F, lanes 1, 3, 5, 7). The *Dpn*I restriction did not change the band pattern, but determined a decrease in intensity of the 5.5kb band and an increase in intensity of the 10kb band consistent with a nicking activity of DpnI (Fig. 3.4F, lanes 2, 4, 6, 8). Since the plasmid pAAVRepURA contains 27 DpnI sites (Fig. 3.4G) and some of them are really close to each other (less than 10 nucleotides apart), nicking of one or more sites may result in a modification of the tertiary DNA structure leading from a supercoiled circle (5.5kb) to a nicked one (10kb). Moreover, pAAVRepURA has a region of approximately 2kb including the URA3 gene, which is free of DpnI sites (Fig 3.4G). Thus, if DpnI restriction had occurred, we would have observed the 2kb band corresponding to URA3 gene, as seen in the case of the plasmid pAAVRepURA extracted from bacteria (Fig.3.4E, lane 5). These results strongly suggest the presence of newly replicated episomal DNA in yeast clones. The 2.5-3kb band was not restricted by DpnI suggesting that it is a product of the new replication and could be the ssDNA rAAV2 genome (Fig 3.4F, lanes 2, 4, 6, 8). The molecular weight corresponds to the ssDNA progeny of AAV observed by Wang et al. [257]. Further demonstration that the three bands are the product of new replication of the plasmid DNA is presented in figure 3.5A and B. Low Mr DNA extracted from one clone shown in Fig.3.4G is digested with the MboI restriction enzyme which recognizes the same palindromic sequence as DpnI but cleaves only non-methylated DNA, such as the DNA replicated in yeast (Fig. 3.5A).



A B C D

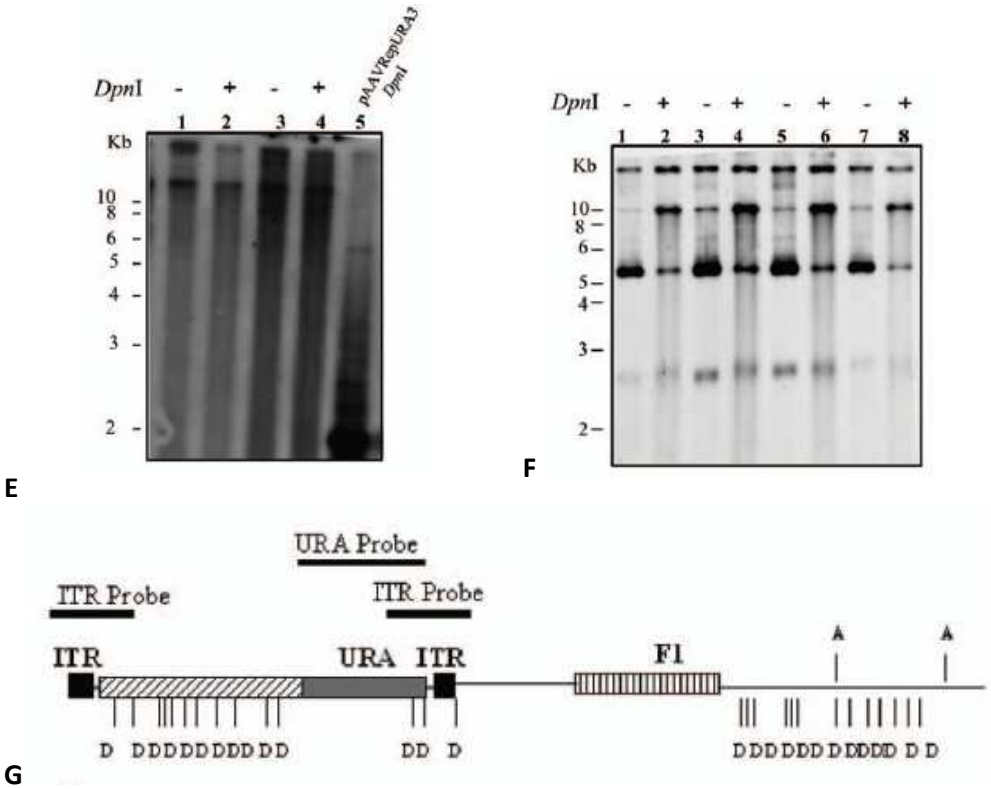


Figure 3.4: AAV replication in yeast.

(A) Southern blot analysis of low Mr DNA of two different yeast clones URA3+ (lane 1 and 2) derived from transformation with pAAVRepURA3 using the URA3 marker gene as probe. (B, C) Southern blot analysis of genomic DNA of the same two clones as in B undigested (lane 1 and 3), and digested with *Ase*I (lane 2 and 4) probed with URA3 (B) or ITR probe (C). (D) Low Mr DNA of the two yeast clones URA3+LEU2+ derived from transformation with pAAVRepURA3 and successively transformed with pG.Rep68 (lane 1 and 2) analyzed on Southern Blot using the URA3 probe. (E) Low Mr DNA of yeast clones URA3+LEU2+ derived from co-transformation of pAAVRepURA with control plasmid pGAD424. Lanes 1 and 3 show the undigested DNA and lanes 2 and 4, DNA digested with *Dpn*I and subjected to Southern blot analysis using URA3 marker gene as probe. Lane 5 shows the result of *Dpn*I digestion of the pAAVRepURA plasmid. *Dpn*I was performed in order to demonstrate that AAV DNA replicated in yeast. (F) Low Mr DNA of yeast clones URA3+LEU2+ derived from co-transformation of pAAVRepURA with plasmid pG.Rep68. Lanes 1, 3, 5, 7 show the undigested DNA and lanes 2, 4, 6, 8 DNA digested with *Dpn*I and subjected to Southern blot analysis using URA3 marker gene as probe. (G) Schematic representation of *Dpn*I/*Mbo*I (indicated with D) and *Ase*I (indicated with A) restriction map of pAAVRepURA plasmid. The *Dpn*I/*Mbo*I restriction endonucleases do not cut in the URA3 gene.

The not digested DNA shows the same bands when probed with URA3 (Fig 3.5A, lane 1) or ITR (Fig 3.5B, lane 1) suggesting that it contains the complete sequence of recombinant AAV. Digestion with *Mbo*I enzyme degraded completely the 10kb and the 5.5kb bands and produced the 2kb band detected with URA3 probe (Fig.3.5A, lane 3) and not with ITR probe (Fig.3.5B, lane 3), suggesting that these bands represent mainly newly replicated DNA. The 2.5-3kb molecule is only partially digested by *Mbo*I and not completely degraded as the other bands. This observation does not rule out the possibility that the band represents the ssDNA because it has been demonstrated that some restriction enzymes, including *Mbo*I, are able to cut ssDNA [258, 259].

Characterization of ssDNA

In order to ascertain that the lowest band corresponds to ssDNA AAV genome, we exposed the extracted DNA to the single-stranded DNA-specific S1 or *Mung Bean* nuclease. Low Mr DNA digested with these two enzymes was analyzed by Southern blot and probed with URA3 or ITRs. As shown in figure 3.5, in the lanes where the DNA samples treated with the S1 nuclease were loaded (Fig. 3.5A, lane 4; fig 3.5B, lane 4) the 2.5-3kb band is not present. The same result was observed when the DNA was treated with *Mung Bean* nuclease (Fig. 3.5C, lanes 2 and 4). Importantly, when the probe "ITRs" was used in the Southern blot (Fig. 3.5B), the three bands were again obtained, demonstrating that these replication-essential DNA elements are present.

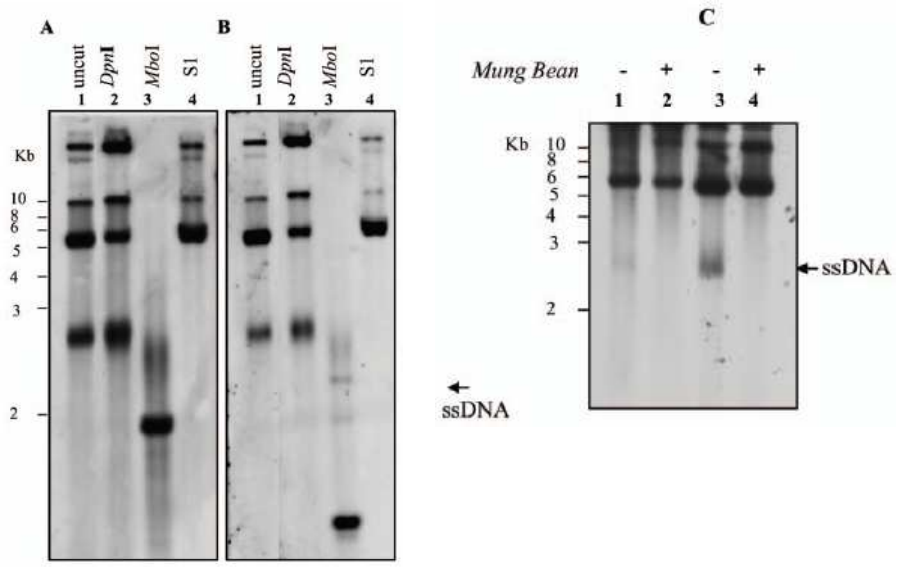


Figure 3.5: S1 and Mung Bean nuclease sensitivity.
(A, B) Low Mr DNA from a URA3+LEU2+ clone derived from co-transformation of pAAVRepURA with plasmid pG.Rep68, was digested with *Dpn*I (loaded in the lane 2), *Mbo*I (loaded in the lane 3) , S1 nuclease (in the lane 4) and was analyzed on Southern Blot using URA3 probe (A) or ITR probe (B). (C) Southern blot analysis of low Mr of twoURA3+LEU2+ clone derived from co-transformation of pAAVRepURA with plasmid pG.Rep68 was digested with *Mung Bean* nuclease (in the lane lane 2 and 4) and compared with the not digested DNA (lane 1, 3). DNA was detected using the URA3 probe. The arrows indicate the ssDNA which disappeared after digestion with S1 and *Mung Bean* nuclease.

Characterization of the replicated rAAV

During the rescue of AAV from plasmid in 293T cells, it is possible to isolate the linear double-length double-stranded DNA molecules and the DNA double-stranded single copy [257]. According to the AAV DNA replication model, the secondary structure formed by the ITRs provides a free 3' hydroxyl group for the initiation of viral DNA replication via a self-priming strand-displacement mechanism involving leading-strand synthesis and double-stranded replicative intermediates monomer formation (Fig.3.6F,i).

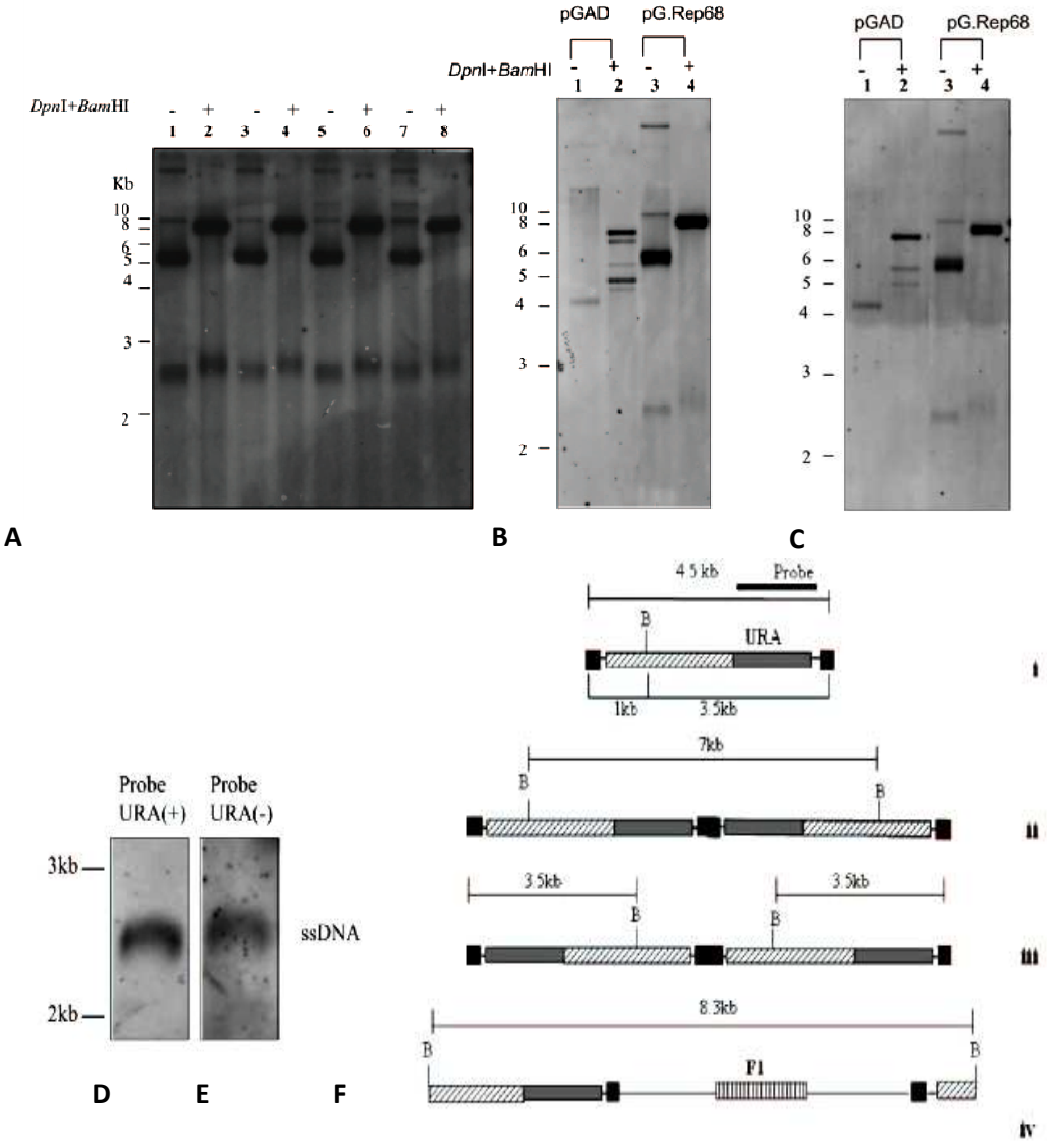


Figure 3.6: Characterization of newly replicated DNA.

(A) Low Mr DNA of four clones derived from yeast co-transformed with pAAVRepURA and pG.Rep68 digested with *Dpn*I and *Bam*HI (lane 2, 4, 6, 8) and not digested (lane 1, 3, 5, 7) was analyzed on Southern Blot and probe with the URA3 gene. *Bam*HI cuts once in AAV construct. (B, C) Low Mr DNA of one clone obtained from the cotransformation of pAAVRepURA with pGAD not digested (lane 1) and digested with *DpnI* and *BamHI* (lane 2); the low Mr DNA of one out four clones analyzed in panel A undigested (lane 3) and digested with *DpnI* and *BamHI* (lane 4) was analyzed on Southern Blot and detected with URA3 probe (B) and F1 probe (C). (D, E) Equal amount of low Mr DNA of one 1 of 4 clones analyzed in panel A was loaded in two gel slots and transferred on nylon membrane. The membrane was cut in two pieces corresponding to gel slots and further cut to leave only ssDNA. The ssDNA was detected using as probe a 100-mer oligonucleotide complementary to the filament (+) of the URA gene, URA(+) (D), or an oligonuc. probe complementary to filament (-) of the URA gene, URA(-) (E), to determine the polarity of ssDNA. (F) Restriction maps of predicted replicative intermediates: i) linear monomer, ii) tail-to-tail dimer, iii) head-to-head dimer, iv) linear vector. B indicates position of *Bam*HI. The sizes of the fragments liberated following *Bam*HI cleavage and recognized by the URA probe are shown next to corresponding structures. ITRs are indicated as black boxes.

These structure is resolved at the trs by site-specific nicking of the parental strand opposite of the original 3' end position (i.e., at nucleotide 125) leading to the formation of ssDNA. When nicking does not occur, elongation proceeds through the covalently closed hairpin structure generating linear double-length double-stranded molecules with either a tail-to-tail configuration (Fig.6F,ii) or a head-to-head configuration (Fig.3.6F,iii) [22]. To determine the structure of molecules replicated in yeast, we digested low Mr DNA with BamHI that cuts pAAVRepURA only once in rep sequence at 1kb (Fig. 3.6F, i). If the newly replicated molecules would be generated according to AAV replication model observed in human cells, the BamHI digestion of low Mr DNA detected with URA3 probe, would produce a band of 3.5kb in case of linear monomer (Fig.3.6F, i); a band of 7kb in case of tail to tail dimer (Fig.3.6F,ii); a band of 3.5kb in case of head to head (Fig.3.6F,iii). As shown in Fig.3.6A, no one of these molecules was produced, but rather we found only one band of ~ 8kb (lanes 2, 4, 6, 8) which is the size of the linearized plasmid containing the backbone and AAV genome (Fig. 3.6F, iv).To confirm this hypothesis, we digested with BamHI the low Mr DNA isolated from a URA3+LEU2+ clone carrying pAAVRepURA plus pGAD (Fig. 3.6B, lane 2; fig. 3.6C, lane 2) or pAAVRepURA plus pG.Rep68 (Fig.3.6B, lane 4; fig.3.6C, lane 4) and hybridized with the F1 probe (Fig.3.6B) that recognizes F1 origin of replication, located in the backbone of the plasmid, or with the probe URA3 (Fig.3.6C) that recognized the URA3 sequence in the AAV vector. As shown in the figure 3.6B and 6C, the bands obtained with F1 probe are coincident with those obtained with URA3 probe. The results of the BamHI digestion produce a band of ~8kb recognized by URA3 and F1 (Fig 3.6B lane 4; Fig 3.6C, lane 4). This observation indicates that the band of 5.5kb corresponds to the supercoiled pAAVRepURA and the band of 10kb corresponds to the nicked circular form. Moreover, the ssDNA molecules band is recognized by both probes indicating the presence of two different molecules, one containing AAV sequences and one with vector sequences. To determine the polarity of ssDNA, we detected the ssDNA with two oligonucleotides probes complementary to filament +, URA(+) (Fig. 3.6D) or to filament - URA(-) (Fig. 3.6E). As shown in figure 3.6D and E both the plus and minus strands of ssDNA are generated.

ssDNA formation is independent of AAV Rep sequence and dependent on ITRs.

Recombinant AAV vectors do not contain AAV genome sequences but only the transgene that can be used for gene therapy in human cells. It has been shown that a region of the AAV2 rep gene acts as a *cis*-acting Rep-dependent element able to promote the replication of transiently transfected plasmids [260, 261]. To investigate whether the formation of ssDNA was dependent on the presence of rep sequence in the vector, we constructed the pAAVpokURA that has URA3 gene and a stuffer sequence (denominated “pok”) flanked by the two ITRs (Fig. 3.1B). In total, the sequence that corresponds to the recombinant AAV genome is 4.1kb. This vector was transformed in yeast clone expressing Rep68 from an integrated copy of pG.Rep68 plasmid. Low Mr DNA was isolated, treated with S1

nuclease to ascertain the presence of ssDNA and Southern Blot analyzed. As shown in the fig. 3.7A lane1, the ssDNA is produced.

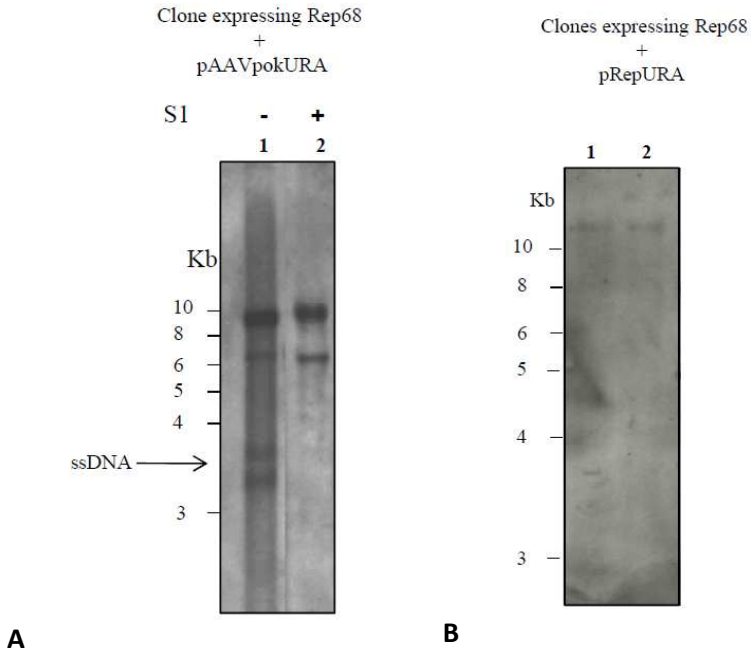


Figure 3.7: ssDNA formation is dependent on ITRs and Rep68.

Southern blot of low Mr DNA obtained from a clone expressing Rep68 transformed with pAAVpokURA (A) or pRepURA (B) and probed with URA3. (A) Lane 1 shows low Mr DNA undigested and lane 2, low Mr DNA subjected to S1 nuclease. (B) Low Mr DNA of two different clones (lane 1 and 2) expressing Rep68 and transformed with pRepURA.

The two bands corresponding to ssDNA could represent two different secondary structures of the molecule. This suggests that the formation of ssDNA is not dependent on RBE in the rep gene, but only on ITRs. To verify if the ITRs are involved in the formation of the newly replicated molecules and the ssDNA, we constructed the vector pRepURA. This vector contains the same sequences as pAAVRepURA but not the ITRs (Fig. 3.1C). pRepURA was transformed in the yeast clone expressing Rep68. Two resulting URA3+LEU2+ clones were subjected to low Mr DNA extraction and subsequent Southern blot analysis (Fig.3.7B, lanes 1, 2). As shown in figure 3.7B, neither circular DNA molecule nor ssDNA was generated in these clones, in the presence of Rep68.

3.3.4. Expression of Adenovirus helper factors E1b55K and E4orf6 in yeast cells and their influence upon rAAV ssDNA formation in the presence of Rep

In human cells productive AAV-2 replication occur only in the permissive cellular environment characterized either by the presence of one of the helper viruses, such as Adenovirus (Ad) , Herpes Simplex Virus (HSV) or Human papilloma virus (HPV) or by the presence of various physical and chemical agents causing genotoxic stress (see chapter 1.1.3). The helper activities provided by Ad were the most extensively studied and

previous studies have delineated a minimal set of Ad proteins required for replication that includes: E1A, E1b55K, E4orf6, and DNA-binding protein (DBP). E1b55K and E4orf6 act as a complex to enhance rAAV transduction and wt AAV replication through overcoming barriers to second-strand synthesis of the genome by still elusive mechanism [70]. E1b55K/E4orf6 has also been shown to form a cullin-containing ubiquitin ligase to promote degradation of the Mre11 complex, DNA ligase IV, p53 and Bloom Helicase during Ad infection [262, 263]. Degradation of these host cell targets prevents apoptosis, DNA damage signaling, and viral genome processing during Ad infection [263, 264]. In addition, E1b55k is the only Adenovirus protein found to associate directly with Rep68/78 [79].

In order to create the best cellular conditions for the AAV ss DNA formation in yeast, we expressed E4orf6 and E1b55k Ad proteins under two galactose inducible promoters from Gal1-10 family and analyzed their putative influence on AAV replication intermediates, in the presence of Rep 68 protein in RSY12 yeast cells. Both E4orf6 and E1b55k genes were inserted in a single plasmid pEsc-G418 under control of Gal1 and Gal10 promoters, respectively, and the resulting construct pEscE4E1b was used to transform yeast clone that carries Rep68 gene integrated in its genome. As shown in the figure 3.8 two proteins were expressed not only in presence of galactose but also under non-inductive conditions (when glucose is used as a carbon source).

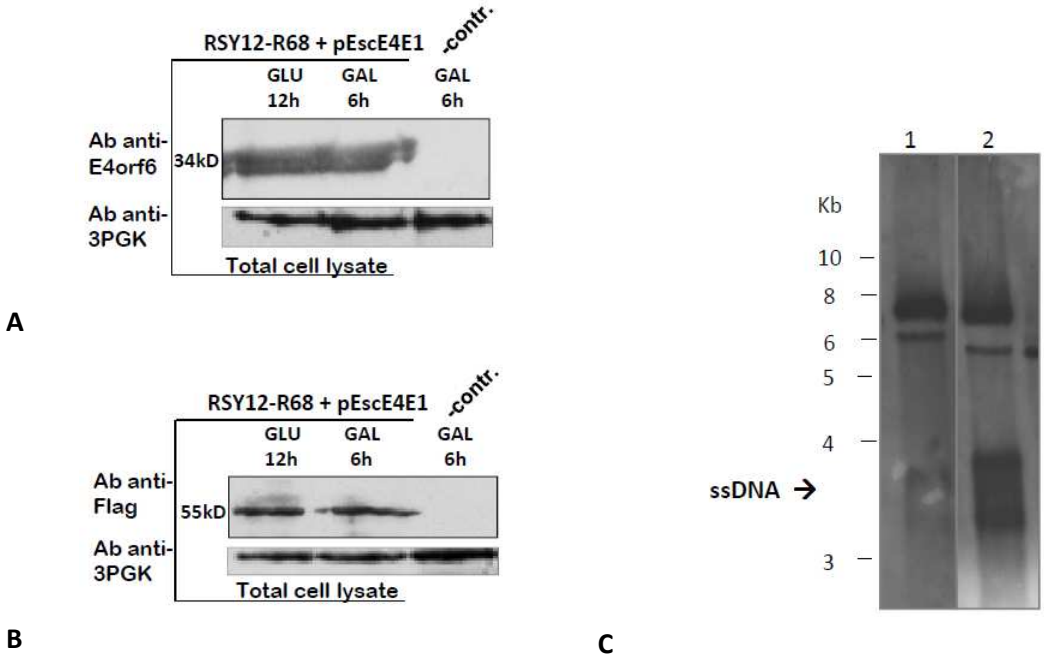


Figure 3.8: Effect of the expression of E4orf -E1b55k on ss DNA formation.

(A, B) Western blot analysis of E4orf6 and E1b55k protein expression in the RSY12-R68 clone carrying endogenous ADH-Rep68 gene and transformed with pEscE4E1 contrast, using the mAb against E4orf6 (A) or the mAb anti Flag antibody which recognizes Flag tag fused to E1b55k (B). Cells were grown over-night in glucose (glu) and then switched for 6h induction in galactose (gal) medium. E4orf6 and E1b55k expression

was analyzed at the end of both glu and gal growth phases. Protein extract from RSY12 clone transformed with the empty vector pEscG418, collected after growth in galactose were used as a negative control (-contr.). Loading control: Constitutive yeast protein 3PGK. *(C)* Southern blot analysis of low molecular weight DNA isolated from the RSY12-R68 clone transformed with the pAAVpokURA vector (lane 1), and from RSY12-R68 clone transformed with both pESCE4E1 plasmid and pAAVpokURA vector (lane 2). Both clones were grown in galactose. URA3 gene was used as probe.

A yeast clone expressing Rep68 and E4orf6-E1b55k was subsequently transformed with the pAAVpokURA vector and the transformed cells were grown in presence of galactose before the isolation of Hirt DNA. The production of ssDNA was analyzed by Southern blot (Fig. 3.8C, lane 1). In presence of E4orf6-E1b55k the amount of ssDNA increased with respect to the clone that lacks Ad helper functions, while other replication intermediates remained unchanged (Fig. 3.8C, lane 2).Weather this effect in yeast is a reminiscent of the MRN complex degradation promoted by these proteins in human cells, is the subject of current investigations.

3.4. Conclusion and future prospective

The results reported here establish the novel finding that the AAV ssDNA is formed in yeast from a double stranded circular plasmid containing the ITRs. Moreover, the plasmid is maintained throughout the generations as a double stranded circle in the absence of the 2 micron replication origin or ARS sequence. Initially, we simply constructed an AAV vector containing the wild type rep sequence and the URA3 marker between the two ITRs, in order to select yeast clones carrying the rAAV. Interestingly, we demonstrated that Rep protein expression occurs in yeast starting from the AAV promoter p5 and p19 (Fig. 3.3A). The failure to obtain the ssDNA genome when yeast was transformed only with pAAVRepURA vector is likely due to the absence of Rep protein at the moment of transformation; in fact, when Rep protein was expressed from pG.Rep68 plasmid, transformed together with high amount of AAV vector, the frequency of URA3+LEU2+ colonies was increased. The transforming DNA carrying the rAAV could follow different ways: it could be recognized by the DNA nucleases and degraded, or be channeled in the recombination process and consequently integrated in the host cell genome. The transforming rAAV plasmid could also be replicated and the AAV genome be rescued from the plasmid. The latter process can occur only in the presence of Rep68 [26]. To test the hypothesis that ssDNA may be generated by plasmid replication, we cotransformed the rAAV vector with pG.Rep68, allowing the expression of Rep68 and Rep40 (Fig. 3.3B) to a sufficient level in the moment of cell transformation with high concentration of rAAV. In these conditions, ssDNA was generated, as shown in figure 3.4F. Moreover, we also demonstrated that the ssDNA is formed only when the ITRs are present (Fig. 3.7B, lane 1 and 2). By using both in vivo and in vitro experimental systems, the principal features of the conventional AAV rescue model have been shown to include the synthesis of duplex linear replicative forms (monomers and dimers) that are self priming by virtue of terminal hairpin palindromes [257, 265, 266]. Two sort of mechanisms have been proposed to

explain rescue of AAV in human cells: rescue may be carried out by repair cellular nucleases [21, 267, 268] or it may be coupled to DNA replication [21], [26, 269]. It has been observed that rescue of the AAV genome from a plasmid may be carried out by a Holliday structure-resolving activity in vitro [270] and *in vivo* [271]. In any case, the episomal DNA is not produced by the "AAV rolling hairpin" type of DNA replication since the analysis of the structures of low Mr DNA molecules (Fig. 3.6) didn't reveal the canonical AAV replicative intermediates, but rather the replicated, supercoiled and nicked, circular plasmid and the ssDNA. Moreover, these molecules contain not only ITRs and URA3 sequences, but also vector sequences. The fact that not only AAV sequences but also plasmid backbone sequences are displaced as ss DNA, prone to be packaged in AAV capsids, could limit the potential of the yeast system for studying rAAV genome packaging and/or its application for AAV vector production. However, a similar problem has been already encountered in mammalian cells where a significant portion of the encapsidated DNA is derived from the ITR-containing rAAV vector plasmid, arising from 'reverse' packaging from the ITRs [272]. It was demonstrated that this problem could be at least diminished, if not resolved, by designing 'oversized' plasmid backbone, that is, exceeding the 4.7 kb natural packaging limit of AAV. We believe that the use of the same strategy can prevent potential packaging of yeast plasmid backbones in AAV capsids. The persistance of the whole ds rAAV plasmid and of its ss backbone sequences suggests that the entire plasmid is replicated as circle from which ssDNA is released. To explain the production of ssDNA genomes without generating linear duplex intermediate, we propose a model outlined in the figure 3.9. According to our model, Rep68 binds the ITR at the RBE element and nicks one strand at the level of trs (Fig. 3.9A). The nick creates a free 3'-OH end, from which DNA replication is initiated (Fig. 3.9B) [32]. As the replication passes through the body of ITR (Fig. 3.9C), the partially single stranded template strand will fold into a hairpin configuration (Fig. 3.9D). This folding of the ITR template may be supported by Rep binding to the tip of the hairpin loop of the ITR [268]. This event produces a strand switch and the template is likely to be the nicked strand and not the complementary one. The replication fork proceeds through the vector sequence and terminates when reaches the nicked ITR. A second nick by Rep protein occurs at the level of the newly synthesized ssDNA that is, subsequently, displaced by the helicase activity of Rep proteins (Fig. 3.9D). The ssDNA molecule containing the entire plasmid sequence is nicked by Rep protein (Fig. 3.9E). The final products are two ssDNA molecules containing only one ITR; one with vector sequences and one with rep and URA3 genes (Fig. 3.9F). Our model predicts that the resulting ssDNA genome with AAV sequences has a complete ITR only at the 5' end while the 3' end carries only a D element. The missing ITR, however, can be repaired via a gene correction mechanism described by Samulski et al. [21].

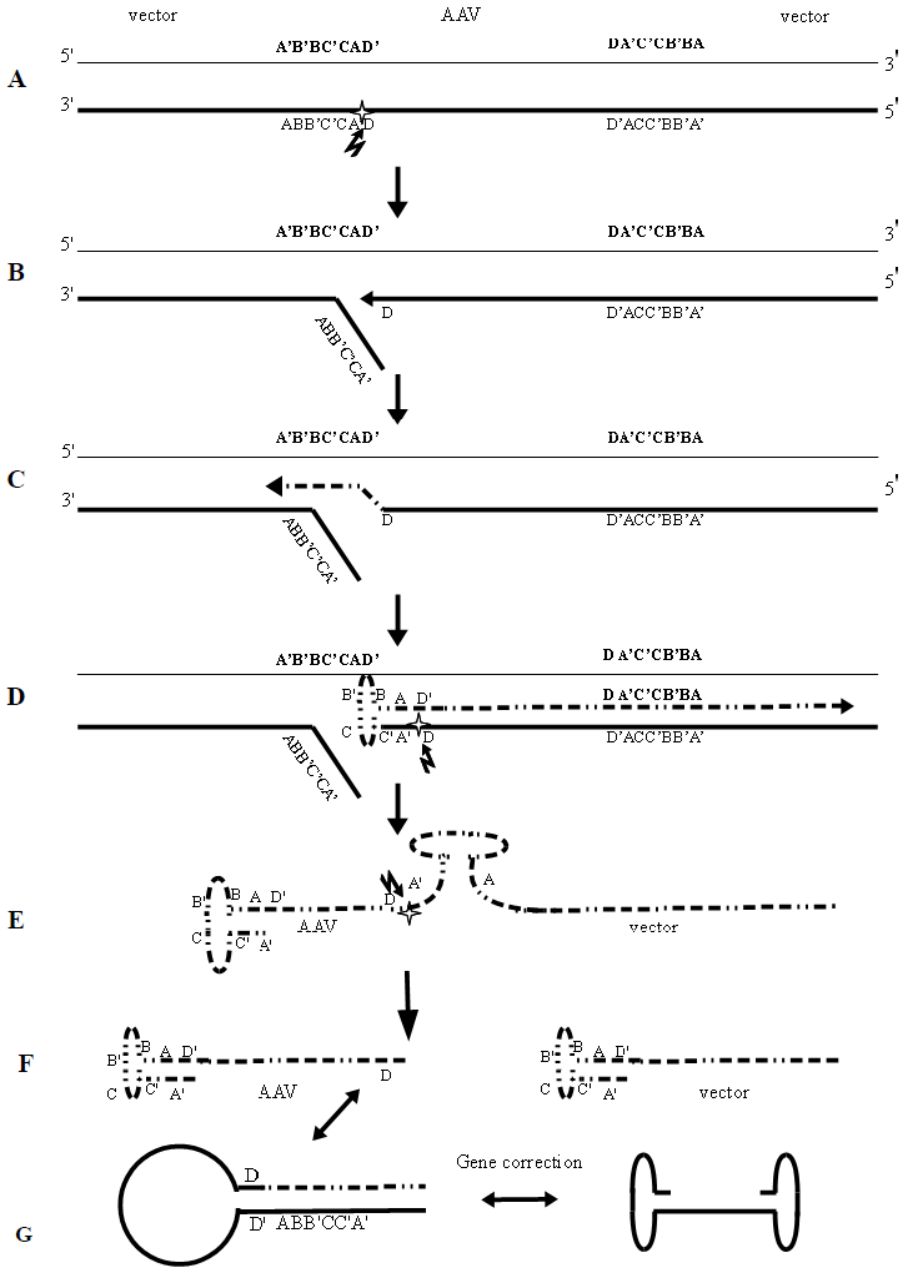


Figure 3.9: Model of ssDNA formation from a plasmid containing rAAV genome in yeast.

(A) For the sake of clarity, the plasmid caying ITRs is depicted as linear molecule. Rep nicks at the *trs* (white star) (B). A replication complex is assembled and replication commences through the TR towards the vector (C). The new synthesized ITR fold into a hairpin conformation displacing the replication strand. Such displacement determines a template switch so that the originally nicked strand is copied during replication. The replication passes through the other ITR and proceeds into the vector sequence (D). After replication fork has completed a full circle, Rep produces a second nick and the newly synthesized DNA is displaced as ssDNA (E). The new ssDNA is nicked again and two ssDNA containing only one complete ITR are formed (F). The missing ITR is repaired by gene correction mechanism (G).

This event involves the formation of a single stranded panhandle DNA intermediate in which both D elements are inverted and therefore can anneal to initiate the ITR repair synthesis (Fig. 3.9G) [273]. One ssDNA molecule has a complete ITR only at the 5' end, while the other one carries ITR sequences without D element. If the ITRs do not fold into a hairpin conformation, the replication continues on the backbone vector sequences and the plasmid is completely replicated as observed in vitro [26];[274]. In this case the ITR is used as a cellular origin of DNA replication [274]. We also validated our model by determining the polarity of the ssDNA. By using specific ssDNA probes, we demonstrated the both the plus and the minus AAV strands are formed in yeast (Fig. 3.6D, E). Moreover, this study demonstrated that, in yeast, the ITRs represent the only sequences required for ssDNA AAV formation from an episomal plasmid (Fig 3.7A, B). Why does the yeast produce such high level of ssDNA? Is that due to an accumulation process of the newly formed ssDNA or is the mechanism leading to ssDNA formation more efficient in Saccharomyces cerevisiae than in human cells? The answer to these questions is not very easy but understanding which cellular factors are involved in the ssDNA formation may be crucial for AAV encapsidation. It is known that yeast is very efficient in DNA repair mediated by homologous recombination. Yalkinoglu et al. have shown that a plasmid containing the extreme left half of AAV with a single ITR could replicate in cells treated with genotoxic reagents suggesting a possible role of DNA damage repair in this process [275]. Moreover, in cultured human cells, a marked increase in rAAV transduction efficiency is obtained by treating cells with agents that affect genomic DNA integrity or metabolism, such as UV irradiation, hydroxyurea (HU), topoisomerase inhibitors and several chemical carcinogens [276-279]. It has been demonstrated recently that MRE11 complex, which is primarily involved in DNA double strand break repair, binds the incoming ssAAV genome and poses a barrier to AAV and that the helper function provided by the adenoviral proteins E1b55K/E4orf6 involves the degradation of MRE11 complex [72, 280]. These observations suggest that in human cells MRN complex can have a role in the rescue of AAV from plasmid. The higher efficiency of the non-homologous end-joining apparatus in mammalian cells, related to the strong pressure to repair double-strand breaks in these large genomes, probably explains the ability of AAV to grow in these cells only when that apparatus is partially saturated by double-strand break events. This requirement is apparently not present in yeast that yet seemingly contains all the cellular factors necessary for AAV genome replication in the right balance. In yeast, double strand breaks are mainly repaired by homologous recombination so the MRN complex proteins are available for other processes and could be immediately engaged in ssDNA formation. The formation of ssDNA does not occur immediately after pAAV plasmid transformation and required the replication of the plasmid. In fact, we did not detect any ssDNA in yeast cells 3 or 24 hours after transformation (data not shown). Therefore, we can speculate that the formation of ssDNA initiates after the plasmid has been replicated. It could be interesting to study AAV ssDNA genome in yeast cells with

different genetic background. Thus, the utilization of this organism could allow not only to study the cellular requirements for viral production, but also, not less importantly, to overcome the present limitations in currently existing systems for commercial production of gene therapy used rAAV vectors; yeast, as a novel host, is safe, cheap and amenable to easy manipulation.

We can thus conclude that the AAV genome can replicate in *Saccharomyces cerevisiae*; this is not surprising because yeast can replicate DNA genomes of numerous viruses, such as Human papillomavirus [213] and the ssDNA genome, Indian Mung Bean Yellow Mosaic virus [214].

4. To sum up…

Taken together, the results of this work raise the importance of the budding yeast *Saccharomyces cerevisiae* both as a genetic system for studying various aspects of AAV vector biology and as a biotechnological tool, a cell factory, for a potential production of AAV gene therapy vectors. Using the simplicity and plasticity of yeast molecular biology, in the first part of the work, we explored different strategies to induce the synthesis of AAV2 capsid proteins to a sufficient level and in the correct stoichiometry to promote capsid assembly. The presented results clearly demonstrate that the assembly process occurs, albeit not as efficiently as seen for many other viruses, stimulating further investigations toward finding better conditions and/or discovering novel assembly promoting factors that could meliorate the current state of art. Accordingly, we are currently investigating whether over-expression of AAV assembly activating protein (AAP) in yeast cells could give desired results. The second part of the work demonstrates that *S. cerevisiae* can form single stranded DNA AAV2 genomes starting from the circular plasmid. The model we propose for underlying mechanism differs greatly from what is seen in other AAV host cells. This, however, is not as surprising as it appears to be since what we currently know about AAV replication in the presence of different helper viruses is that this virus is able to adopt to the novel "replication-supportive environment" in a way to optimally exploit the available replication factors. In this view, the genome-wide screening for the yeast cellular factors, regulating/being directly involved in the AAV replication, are the subject of future investigations. In addition, we are currently focused on studying the way in which Adenoviral proteins, E4 orf6 and E1b55, increase the amount of ss AAV DNA formation, without significantly changing the pattern of other rAAV replicative intermediates. This research line should also reveal some yeast DNA repair proteins, potentially involved in the aforementioned process. Finally, a significant effort is to be put in studying the ability of yeast cell to support packaging of ss rAAV genomes in preformed capsids.

5. Materials & methods

5.1. Strains, media, yeast transformation and cultivation

Bacterial recombinants were screened in *Escherichia coli* XL10Gold cells.

Two haploid, auxotrophic yeast strains were used, RSY12 and HT393. The “wild type” strain RSY12 (*MAT**a** leu2-3,112 his3-11,15URA3::HIS3*) has a complete deletion of URA3 gene which was replaced with the HIS3 gene [281]. The “mutant” strain HT393 *(MAT**a** leu2, ura3, pra1, prb1, prc1, cps1, pre1*) is deficient in several proteases including parts of the proteasome, vacuolar proteases and metalloproteases [282].

Comparative protein expression analysis for all AAV2 proteins was done by transforming each of the following constructs independently in two yeast strains, RSY12 and HT393:

1. 6 VP expressing: YEplacRepCap, YEplacp40Cap, Yep-Rep78p40Cap, pYesCap, pYesIntronCap, pYesCap
2. 3 Rep expressing: YEplacRepCap, Yep-Rep78p40Cap, pG.Rep68

The same amount of plasmid DNAs were transformed into yeasts using the standard high-efficiency, lithium chloride transformation method with single-stranded DNA as carrier [283]. Single transformants were selected onto SC-uracil (SC-URA) or SC-leucine (SC-LEU) plates and the selection continued with the following growth in the corresponding liquid medium.

Since no significant difference in relative protein VP/Rep expression levels was assessed by Western blot analysis for each strain couple of transformed constructs, for e.g between RSY12- YEplacRepCap and HT393- YEplacRepCap etc., only Western blot results obtained for RSY12 strain are presented in this work. Accordingly, all following cotransformation experiments for AAV particle production and purification, as well as for recombinant AAV vector genome replication studies used only RSY12 strain. Again, the same amount of plasmid DNAs were co-transformed into yeast using lithium chloride method. Cotransformed RSY12 cells were selected on SC-uracil-leucin (SC-URA-LEU) plates (and selection continued in the corresponding liquid medium). The frequency of URA3+LEU2+ colonies was calculated dividing the number of URA3+LEU2+ colonies by the number of LEU2+ total transformants.

Complete (YPAD) and synthetic complete (SC) mediums were prepared according to the standards. Induction media were prepared with either 5% galactose or different glucose + galactose concentration as stated in the Results section of the chapter 2.

Yeast was cultivated aerobically, in flasks, at 30° C, under constant orbital shaking (180 rpm). Different growth cycle experiments were started with a small volume inoculum (less than 100 µl per 50 ml of culture medium) of stationary phase yeast cells (grown for 24h from a single colony in 10 ml of SC) and cultivated overnight to obtain an early exponential (log) growth phase cell suspension of absorbance at 600 nm ($O.D_{600}$) 0-7-1.0, equivalent to ~0.3 mg/ ml-' cell dry weight. For mid-log phase cells were further grown to reach $O.D_{600}$ 3.5- 4.0 and for late-log/early stationary phase cell were grown till optical density was ~8. If not processed immediately after the grown cycle, cells were pelleted and stored at -70°C or liquid nitrogen frozen and then -70°C.

The RSY12 clone with integrated copies of ADHRep68- expression cassette, was obtained by cell transformation with the SnaBI -ADHRep68-Leu- PvuII DNA fragment excised from pG.Rep68 plasmid [229] with SnaBI/PvuII enzymes. In this way, 2 micron sequence was eliminated in order to avoid autonomous replication upon putative recircularisation, while Leu was kept for selection of integrants.

5.2. Plasmid construction

All AAV2 sequences in the following constructs derive from pSub201 plasmid which contains complete wild-type AAV-2 genome cloned into a plasmid backbone [19].

pYEplac181(Leu2), pYes2(Ura3) and pGAD 424(Leu2) are three commercially available, yeast-cell compatible vectors with *ura3/leu2* genes for selection in yeast .

Constructs with pYEplac181vector backbone:

For construction of the pYplacRepCap plasmid entire AAV2 genome without ITRs was cut out from pSub201, using Xba I sites for digestion, and placed in the same restriction site of the pYEplac181 vector. For construction of the pYplacp40Cap plasmid, unmodified AAV2 *Cap* gene was cut out from pSub201 using Sal I and Xba I sites and cloned in the same sites of YEplac181 vector. To make the Yep-Rep78p40Cap plasmid, ADHRep78 expression cassette was excised from pG.Rep78 plasmid [229], using flanking SphI sites and cloned upstream of Cap ORF in *Sph* I site of pYplacp40Cap.

The plasmids pGAD424, pG.Rep78 and pG.Rep68 were kindly provided by M. D. Weitzman [229]. The pG.Rep78 and pG.Rep68 constructs contain Rep68cDNA and Rep78cDNA, respectfully, downstream of the constitutive yeast promoter ADH1 in pGAD424 plasmid backbone.

Cloning in the pYes2 expression vector, downstream of inducible yeast promoter, Gal1:

For construction of the pYesIntronCap plasmid, AAV2 Intron + Cap ORF sequence (without p40 promoter) was cut out from pSub201 using HindIII and XbaI restriction sites and cloned in the same sites comprised in the multiple cloning site of pYes2 vector.

For construction of the pYesCap plasmid, VP1 cDNA(comprising all three overlapping VP cDNAs) (map positions 2203-4410 in wild-type AAV2 genome in the database, GenBank accession no. AF043303.1) was PCR amplified from pSub201 using following primers:

5' - ctt ATGGCTGCCGATGGTTATCTT-3' and

5' - ctggTTACAGATTACGAGTCAGGTA -3'

The sequence corresponding to the wild-type AAV2 genome is capitalized. PCR product was polyA tailed to enable compatible end ligation with polyT tailed intermediate vector pGEM®-T Easy (Promega Corporation). After ligation, PCR clones were sequenced and the correct VP sequence was NotI excised from the pGEM®-T Easy and cloned into NotI site of pYes2 MCS.

The pYesVP1K.M plasmid was constructed by cloning modified VP1 cDNA under Gal1 promoter. The VP1cDNA was PCR amplified from pSub201vecotr using following primers:

5'-ca**ggatcc***aaaca*ATGGCTGCCGACGGTTATCTACCCGATTGGCTC-3' and

5'-cc**ggatcc**TTACAGATTACGAGTCAGGTATCTGGTGC-3'.

Bam HI sites (at each primer end, underlined and bolded letters) were introduced outside the VP1 expression cassette for cloning in pYes expression plasmid. Beside *Bam* HI restriction site, upper primer also included 6 nucleotide long yeast Kozak sequence (italicized letters) upstream of the VP1 ATG and three modifications (underlined letters) relative to the genuine VP sequence in the database. The sequence corresponding to the wt VP ORF is capitalized. After direct cloning in pYes2 *Bam* HI site desired clone contained: Kozak sequence upstream of ATG start site and three mutations with respect to genuine AAV2 sequence. The first T to C mutation at position 11 eliminates an out-of frame ATG codon by creating an ACG triplet instead. The second modification is inactivation of the major AAV splice acceptor site, by substituting T to A at position 21 and A to C at position 24 (numbering is as for the seq. in the database)

Construction of vectors containing recombinant AAV genome:

The plasmids pAAVRepURA3 and pAAVpokURA are derivates of pSub201 plasmid.

The pAAVRepURA3 plasmid maintains the entire rep gene and has an internal deletion of cap gene. The URA3 marker gene, excised from pMA150 vector by using BamHI sites for restriction digestion [284], was cloned in the blunt ApaI/KpnI sites of pSub201 after filling in with Klenow polymerase. The pAAVpokURA contains only the URA3 gene and a stuffer sequence in order to increase the distance between ITRs. The stuffer sequence, named pok, consists of 2.5kb-long mouse pokemon gene. pAAVpokURA was made by cloning the pok sequence and URA3 gene (pMA150 derived) in pMCSsub vector. The pMCSsub was constructed by substituting the multiple cloning site of the pCMV-MCS

vector (Stratagene, La Jolla, CA,USA) with a multiple cloning site containing NotI XbaI XhoI EcoRI BamHI HindIII SalI XbaI NotI. The URA3 gene was cloned in the HindIII site of pMCSsub, obtaining pMCSsubURA. The pok sequence was cloned in XhoI and EcoRI sites of pMCSsubURA obtaining pMCSubpokURA. Pok and URA3 were then cut out with XbaI and cloned in XbaI site of pSub201.

The pRepURA plasmid was constructed by XbaI digestion of the pAAVRepURA plasmid to eliminate ITRs, and cloning the XbaI-RepURA-XbaI fragment in XbaI site of pMCSsub.

To obtain the plasmid pEsc-G418, LEU2 selection marker of the commercial vector pEsc-Leu (Stratagene) was substituted with the gene Kan-MX, whose product provide the transformed cells resistant to antibiotic G418. Kan-MX (~ 1500 bp) was cut out from another comercial plasmid pYMN1 (Euroscarf) using EcoRV and PstI restriction sites and cloned in EcoRV and BsteII (isoschizomero di PstI) sites of pEs-Leu. The mentioned digestion of the pEsc-Leu resulted in elimination of approx 700bp of LEU2 gene fragment causing its inactivation.

The gene E4orf6 was amplified by PCR from pDG plasmid [94], using following primer pairs:

5‘CTA**GGATCC**ATGACTACGTCCGGCGTT 3’ and

5‘CTA**GTCGAC**CATGGGGGTAGAGTCATA 3’

Restriction sites Bam HI and Sal I (at each 5’ primer end, respectively; shown in underlined and bolded letters) were introduced outside the E4orf6 expression cassette for cloning in respective sites in pESC-G418 expression plasmid, downstream of Gal1 promoter.

The gene E1b55k was amplified by PCR from genomic DNA derived from 293T cells using following primer pairs:

5‘ACT**GCGGCCGC**ATGGAGCGAAGAAACCCATC 3’ and

5‘ TTA**ACTAGT**CCATCTGTATCTTCATCGCTAGA 3’

Restriction sites NotI and SpeI (at each 5’ primer end, respectively; shown in underlined and bolded letters) were introduced outside the E1b55k expression cassette for cloning in respective sites in pESC-G418 expression plasmid, downstream of Gal10 promoter and in frame with tag FLAG ORF.

5.3. RNA isolation and Retro-Transcription

Total yeast RNA was extracted from 5 x 10^7 of actively growing cells (grown to early log phase) following the method published by Schmitt *et al.* [285]. 3 µg of total RNA was

retrotranscribed using QuantiTect Reverse Trascription Kit of Qiagen (using oligo-dT primers) according to the manufacture's recommendations. Resulting cDNA was subjected to PCR amplification using following pair of primers:

5'-AGTTGCGCAGCCATCGACGT-3' and 5'-TGGCTGTCCGAGAGGCTG- 3'

PCR products were analyzed upon 5% acrilamide gel-electrophoresis, for separation of nucleic acids less than 1000 bp in length, and visualized by Etidium-Bromide (EtBr) staining.

5.4. Analysis of protein expression.

For analysis of protein expression by gel electrophoresis, total cellular polypeptides were obtained by means of two different extraction methods:

1. **Glass-bead based cell rupture** followed by separation of soluble and insoluble protein fraction: 10 millions of cells grown till mid-exponential growth phase were harvested, washed and resuspended in the disruption buffer DB (30 mM Tris-HCl, pH 7.8, 2 mM EDTA, 0.5% Triton X-100, 1 mM PMSF; proteinase inhibitors cocktail added) and disrupted with equal volume of glass beads (400 nm diameter). Disruption at 4°C was performed by 10 cicles of 30' vortex + 30' ice incubation. Crude lysates were cleared by centrifugation (2000xg for 5 min, 4°C). The insoluble protein fraction was separated from soluble one by another centrifugation (10,000×g, 40 min, 4 ◦C). Insoluble proteins, from the pellet, were extracted twice in the extraction buffer (DB exempted of EDTA and supplemented with 1%SDS and 1%DOC) for 1 h at 4 ◦C agitation.

2. **"Optimized post-alkaline extraction"** method consists of two cycles of protein extraction. 100-200 x 10^6 of actively growing yeast cells were pelleted and subjected to the first round of extraction by the method published by Kushnirov [230]. Briefly, cells were pelleted and shortly exposed to mild alkali conditions (0.1M NaOH) and lysed by heat at 95°C for 5 min in 50µl of 1x SDS-PAGE loading buffer (4x buffer: 150mM TrisHCl pH6.8, 50% glycerol, 10% SDS, 10% β-mercaptoethanol, 0.06% bromophenol blue). After eliminating cellular debris by centrifugation, the first protein extract named "first extract" was ready for analysis by gel electrophoresis. For the "second extraction", cellular debris was treated with a variant of a standard harsh RIPA buffer (500mM NaCl, 50mM Tris-HCl, 1mM EDTA, 1% Triton, 1% DOC, 1% SDS) and put forward to a quick sonication (Soniprep 150). Precisely, the remaining pellet was resuspended in 140 µl of RIPA buffer (proteinase inhibitors added), incubated on ice for 20 min and then for another 20 min on the rotation at 4°C. After sonication (5 cycles of 10 seconds on and 20 seconds off; 15 microns amplitude), the second protein fraction, named "second extract" (insoluble aggregates), was separated from remaining cellular impurities by quick centrifugation and analyzed by gel electrophoresis.

Proteins were resolved on 10% polyacrylamide gel in SDS–Tris–glycine buffer. Following SDS-PAGE proteins were electro-transferred to Hybond-C Extra nitrocellulose membrane (Amersham Biosciences, UK) in methanol-Tris-glycine buffer. The blots were blocked with: 1) 10% milk or 2) 5% in TBS-T (Tween 0.1%) for 1 h and then, 3) 6 3% milk 1%BSA in PBS 1x for 1 h, 4) 5% milk in TBST overnight. After blocking blots were incubated overnight at 4°C with the following primary antibodies: 1) anti-Rep monoclonal mAb 303.9, anti.VP1,2 mAb A69 and anti-capsid mAb A20. All named antibodies were diluted 1:50 in TBS-T with 5% milk; 2) anti-VP1,2,3 diluted 1:200 in 2.5% milk TBS-T. All listed Abs are from Progen Biotechnik GmbH, Heidelberg, Germany. For detection of the protein E4orf6 we used 3) mAb-Anti E4orf6 diluted 1:20 in PBS solution supplemented with 3% milk and 1%BSA (a kind gift from D. Ornelles). Incubation was done overnight. For detection of E1b55k we used 4) mAb anti-Flag (Sigma) diluted 1:1000 in TBST-5% milk and incubated for 1h, room temperature (RT). For loading control the same blots were incubated for 1h with primary mouse antibody anti-3 phosphoglycerate kinase (3PGK, Molecular Probes) diluted 1:5000 in 1% milk in TBS-T. Secondary antibodies conjugated to horseradish peroxidase (IgG-HRP) (Santa Cruz Biotechnology, Inc.) were used for detection of specific antibody binding (diluted 1:2500 in the same solution of the corresponding primary antibody). The blots were stained with the Super Signal West Femto chemiluminescent substrate (Thermo Fisher Scientific, Inc.).

5.5. AAV VLP extraction and purification

For this purpose actively growing 25ml 2% glucose culture was diluted in 250 ml and re-grown on glucose for another 12-16h until reached OD_{600} 4 (mid-log phase). At this point the cells were collected and the media replaced with 1l of induction media (0.5% Glu+5% Gal). After 4.5 h cells were harvested and subjected to disruption:

Mechanical lysis under non-denaturizing conditions: 10g of yeast dry biomass was resuspended in 20ml of DNS disruption buffer of following composition: 500mM NaCl, 10mMTris(pH8), 1mM EDTA, 0.3% NP-40, 0.3% deoxycholate (DOC), supplemented with protease-inhibitor cocktail and homogenized in the presence of 400nm glass beads 10 cycles of 40sec vortex-homogenization. The crude lysate was cleared from cell-debris and undisrupted cells by quick-low speed centrifugation. After clearance, the crude lysate was subjected to 1h high speed centrifugation (20 000 x g), which was followed by the second extraction from the obtained pellet. The second extraction was performed in the same way as the first one with additional sonication step to enhance solubilization of yeast produced AAV VLPs. This was done in 10 sonication cycles (Soniprep 150; 15sec on and 30sec off; 15microns amplitude). After sonication the cell lysate was centrifuged again for 1h at 20000g at 4°C. This extended clearance was supposed to eliminate most efficiently aggregated, insoluble material. The resulting, second extract (supernatant), was united with the first extract and subjected to 200,000 x g, 3h ultracentrifugation through 40% sucrose

cushion (40% sucrose in TE-BSA buffer: 10mM Tris pH8, 1mM EDTA and 0.01%BSA). After the centrifugation, AAV particle-containing sediment was resuspended in 5ml Resuspension Buffer of physiological salts concentrations (150 mM NaCl, 10 mM Tris-HCl pH 8, 1 mM EDTA, 5 mM $MgCl_2$), quickly sonicated and left in agitation for 15 h at 4 °C. To improve solubilisation of assembly products, it was further subjected to 30min 37C DNase digestion (purity grade I, Benzonase, Novagen), stopped by 10mM EDTA. For digestion, $MnCl_2$ and $MgCl_2$ were added the extract to 1mM final concentrations. At the end, extract was cleared from aggregated material by 10 min, 15000 x g centrifugation at 4°C and supernatant was loaded on CsCl gradient for AAV VLPs purification. CsCl was added to the sample at final density of 1.4 g/cm^3, and ultracentrifugation was performed in a SW40-Ti rotor (48 h, 39,000 rpm, 18°C; ultracentrifuge). 13 fractions of 1.371 to 1.430 g/cm3 densities (refractometer readings) were collected, dialyzed against PBS buffer, and each fraction was analyzed for VP composition by Western blot. Fractions with similar densities were united and analyzed by negative staining electron microscopy.

5.6. Transmission electron microscopy

After purification by CsCl centrifugation suspensions of yeast-produced AAV VLPs were placed on 400-mesh carbon coated grids. The sample was stained with five drops of 2.5% (w/v) of uranyl acetate solution and incubated for 2 minutes at room temperature. The excess solution was removed by blotting the edge of each grid onto filter paper and the grid was air dried for 30 minutes. Samples were examined with a Jeol 100 SX transmission electron microscope with 80000 magnification.

5.7. DNA isolation and Southern blotting

Genomic DNA was isolated using the Master pure yeast DNA purification Kit (Epicentre Biotechologies). **Low molecular weight (Mr) DNA was isolated** from 5 to 10 independent URA3+LEU2+ clones harboring the AAV genome by using the yeast plasmid miniprep kit (Zymo Research Orange, CA) as previously reported [286]. The *Dpn*I resistance assay was performed to check that the plasmid had replicated in yeast. When the plasmid DNA is replicated and methylated by bacteria, it becomes sensitive to *Dpn*I digestion and it is thus distinguishable from DNA synthesized in eukaryotic cells which is not methylated and, therefore, resistant to digestion [287]. By contrast, *Mbo*I recognizes the same sequence as *Dpn*I, but cuts only DNA not methylated. The DNA samples were digested with *Dpn*I or *Mbo*I for 24 h and loaded on 1% agarose gel. The DNA was electrophoresed, transferred to nylon membrane (Roche) and probed with DIG-labelled URA3 obtained by PCR from the plasmid YEplac211 with following primers:

pURA-R 5’GATACCCGGGACCTGACGTCTAAG3’ and

pURA-F 5’ACATCCCGGGTGACGGGCTTGTCTGCTCCC3’.

The ITR probe was obtained by DIG random primed labelling kit (Roche) using as template gel purified ITRs obtained digesting pSub201 with XbaI and PvuII. To determine the presence of the plasmid backbone, we also constructed the F1 probe that was obtained by PCR performed with nucleotide mix containing DIG-dUTP from pSub201 plasmid with the following primers:

F1 forward: 5'CAA CCA TAG TAC GCG CCC 3';

F1 reverse: 5' CAA TAA ATC ATA CAG GCA AGG 3'.

To determine the polarity of ssDNA we used two synthetic oligonucleotides of 100 mer as probe (IDT integrated DNA technologies). The probe for the plus (+) strand, named URA(+) has the following sequence:

5'CTTCCAACAATAATAATGTCAGATCCTGTAGAGACCACATCATCCACGGTTC
TATACTGTTGACCCAATGCGTCTCCCTTGTCATCTAAACCCACACCGG 3'.

To determine the presence of the minus (-) strand, we used the synthetic URA(-) probe with the following sequence:

5'CCGGTGTGGGTTTAGATGACAAGGGAGACGCATTGGGTCAACAGTATAGAA
CCGTGGATGATGTGGTCTCTACAGGATCTGACATTATTATTGTTGGAAG 3'.

The probes were labeled using the DIG oligonucleotide 3'-End Labeling kit, 2nd generation (Roche), according to manufacturer's instructions. Equal amount of DNA was loaded on two gel slots, electrophoresed and transferred on nylon membrane positively charged (Roche). Since the 3'end labeled oligonucleotide probes are less sensitive than the PCR labeled probes, we hybridized only the ssDNA to increase the signal.

S1 and Mung Bean reaction

To identify the presence of ssAAV genome, low Mr DNA samples were digested with S1 nuclease (Promega) or *Mung Bean* Nuclease (New England Biolabs) which cut only ssDNA. Briefly, 10µl of low Mr DNA were digested with 0.1U of S1 nuclease for 30 min at 37°C in the provided buffer and inactivated at 70°C for 10 min. For *Mung Bean* Nuclease, 10µl of low Mr DNA were digested with 0.1 U of *Mung Bean* for 30 min at 37°C in the provided buffer. The reaction was stopped with STOP solution (5mM EDTA, 30mM TrisHCl pH 7.5) followed by heating at 65°C for 10 min.

References

1. Abebe, A., et al., *Timing of the HIV-1 subtype C epidemic in Ethiopia based on early virus strains and subsequent virus diversification.* AIDS, 2001. **15**(12): p. 1555-61.
2. Berns, K.I. and R.M. Linden, *The cryptic life style of adeno-associated virus.* Bioessays, 1995. **17**(3): p. 237-45.
3. Atchison, R.W., B.C. Casto, and W.M. Hammon, *Adenovirus-Associated Defective Virus Particles.* Science, 1965. **149**: p. 754-6.
4. Hoggan, M.D., N.R. Blacklow, and W.P. Rowe, *Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics.* Proc Natl Acad Sci U S A, 1966. **55**(6): p. 1467-74.
5. Geoffroy, M.C. and A. Salvetti, *Helper functions required for wild type and recombinant adeno-associated virus growth.* Curr Gene Ther, 2005. **5**(3): p. 265-71.
6. Buller, R.M., et al., *Herpes simplex virus types 1 and 2 completely help adenovirus-associated virus replication.* J Virol, 1981. **40**(1): p. 241-7.
7. Walz, C., et al., *Interaction of human papillomavirus type 16 and adeno-associated virus type 2 co-infecting human cervical epithelium.* J Gen Virol, 1997. **78 (Pt 6)**: p. 1441-52.
8. Schlehofer, J.R., M. Ehrbar, and H. zur Hausen, *Vaccinia virus, herpes simplex virus, and carcinogens induce DNA amplification in a human cell line and support replication of a helpervirus dependent parvovirus.* Virology, 1986. **152**(1): p. 110-7.
9. Alazard-Dany, N., et al., *Definition of herpes simplex virus type 1 helper activities for adeno-associated virus early replication events.* PLoS Pathog, 2009. **5**(3): p. e1000340.
10. Wu, Z., A. Asokan, and R.J. Samulski, *Adeno-associated virus serotypes: vector toolkit for human gene therapy.* Mol Ther, 2006. **14**(3): p. 316-27.
11. Buning, H., et al., *Recent developments in adeno-associated virus vector technology.* J Gene Med, 2008. **10**(7): p. 717-33.
12. Summerford, C. and R.J. Samulski, *Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions.* J Virol, 1998. **72**(2): p. 1438-45.
13. Zolotukhin, S., et al., *Recombinant adeno-associated virus purification using novel methods improves infectious titer and yield.* Gene Ther, 1999. **6**(6): p. 973-85.

14. Handa, A., et al., *Adeno-associated virus (AAV)-3-based vectors transduce haematopoietic cells not susceptible to transduction with AAV-2-based vectors.* J Gen Virol, 2000. **81**(Pt 8): p. 2077-84.

15. Kaludov, N., et al., *Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity.* J Virol, 2001. **75**(15): p. 6884-93.

16. Flotte, T.R., *Gene therapy progress and prospects: recombinant adeno-associated virus (rAAV) vectors.* Gene Ther, 2004. **11**(10): p. 805-10.

17. Giacca, M., *Vectors based on the Adeno-associated virus (AAV),* in *Gene Ther*. 2010, Springer-Verlag Italia: Milan. p. 112-123.

18. Wu, Z., et al., *Single amino acid changes can influence titer, heparin binding, and tissue tropism in different adeno-associated virus serotypes.* J Virol, 2006. **80**(22): p. 11393-7.

19. Samulski, R.J., et al., *Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells.* Proc Natl Acad Sci U S A, 1982. **79**(6): p. 2077-81.

20. Xie, Q., et al., *The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10405-10.

21. Samulski, R.J., et al., *Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV.* Cell, 1983. **33**(1): p. 135-43.

22. Goncalves, M.A., *Adeno-associated virus: from defective virus to effective vector.* Virol J, 2005. **2**: p. 43.

23. Nyberg, K., et al., *Workshop on long-term follow-up of participants in human gene transfer research.* Mol Ther, 2004. **10**(6): p. 976-80.

24. Haberman, R.P., T.J. McCown, and R.J. Samulski, *Novel transcriptional regulatory signals in the adeno-associated virus terminal repeat A/D junction element.* J Virol, 2000. **74**(18): p. 8732-9.

25. Berns, K.I. and C. Giraud, *Biology of adeno-associated virus.* Curr Top Microbiol Immunol, 1996. **218**: p. 1-23.

26. Ward, P., et al., *Adeno-associated virus DNA replication in vitro: activation by a maltose binding protein/Rep 68 fusion protein.* J Virol, 1994. **68**(9): p. 6029-37.

27. Surosky, R.T., et al., *Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome.* J Virol, 1997. **71**(10): p. 7951-9.

28. Weitzman, M.D., et al., *Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA.* Proc Natl Acad Sci U S A, 1994. **91**(13): p. 5808-12.

29. McCarty, D.M., et al., *Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein.* J Virol, 1994. **68**(8): p. 4988-97.

30. Wu, J., M.D. Davis, and R.A. Owens, *Factors affecting the terminal resolution site endonuclease, helicase, and ATPase activities of adeno-associated virus type 2 Rep proteins.* J Virol, 1999. **73**(10): p. 8235-44.

31. Zhou, X., et al., *Biochemical characterization of adeno-associated virus rep68 DNA helicase and ATPase activities.* J Virol, 1999. **73**(2): p. 1580-90.

32. Smith, R.H. and R.M. Kotin, *The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity.* J Virol, 1998. **72**(6): p. 4874-81.

33. Collaco, R.F., et al., *A biochemical characterization of the adeno-associated virus Rep40 helicase.* J Biol Chem, 2003. **278**(36): p. 34011-7.

34. Pereira, D.J. and N. Muzyczka, *The adeno-associated virus type 2 p40 promoter requires a proximal Sp1 interaction and a p19 CArG-like element to facilitate Rep transactivation.* J Virol, 1997. **71**(6): p. 4300-9.

35. Pereira, D.J. and N. Muzyczka, *The cellular transcription factor SP1 and an unknown cellular protein are required to mediate Rep protein activation of the adeno-associated virus p19 promoter.* J Virol, 1997. **71**(3): p. 1747-56.

36. Weger, S., et al., *Control of adeno-associated virus type 2 cap gene expression: relative influence of helper virus, terminal repeats, and Rep proteins.* J Virol, 1997. **71**(11): p. 8437-47.

37. Lackner, D.F. and N. Muzyczka, *Studies of the mechanism of transactivation of the adeno-associated virus p19 promoter by Rep protein.* J Virol, 2002. **76**(16): p. 8225-35.

38. Clark, K.R., et al., *Cell lines for the production of recombinant adeno-associated virus.* Hum Gene Ther, 1995. **6**(10): p. 1329-41.

39. Shi, Y., et al., *Transcriptional repression by YY1, a human GLI-Kruppel-related protein, and relief of repression by adenovirus E1A protein.* Cell, 1991. **67**(2): p. 377-88.

40. Horer, M., et al., *Mutational analysis of adeno-associated virus Rep protein-mediated inhibition of heterologous and homologous promoters.* J Virol, 1995. **69**(9): p. 5485-96.

41. Chiorini, J.A., et al., *Inhibition of PrKX, a novel protein kinase, and the cyclic AMP-dependent protein kinase PKA by the regulatory proteins of adeno-associated virus type 2.* Mol Cell Biol, 1998. **18**(10): p. 5921-9.

42. Labow, M.A., L.H. Graf, Jr., and K.I. Berns, *Adeno-associated virus gene expression inhibits cellular transformation by heterologous genes.* Mol Cell Biol, 1987. **7**(4): p. 1320-5.

43. Antoni, B.A., et al., *Adeno-associated virus Rep protein inhibits human immunodeficiency virus type 1 production in human cells.* J Virol, 1991. **65**(1): p. 396-404.

44. Saudan, P., J. Vlach, and P. Beard, *Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb.* EMBO J, 2000. **19**(16): p. 4351-61.

45. Berthet, C., et al., *How adeno-associated virus Rep78 protein arrests cells completely in S phase.* Proc Natl Acad Sci U S A, 2005. **102**(38): p. 13634-9.

46. Schmidt, M., S. Afione, and R.M. Kotin, *Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53.* J Virol, 2000. **74**(20): p. 9441-50.

47. King, J.A., et al., *DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids.* EMBO J, 2001. **20**(12): p. 3282-91.

48. Timpe, J., et al., *Mechanisms of adeno-associated virus genome encapsidation.* Curr Gene Ther, 2005. **5**(3): p. 273-84.

49. Sonntag, F., K. Schmidt, and J.A. Kleinschmidt, *A viral assembly factor promotes AAV2 capsid formation in the nucleolus.* Proc Natl Acad Sci U S A, 2010. **107**(22): p. 10220-5.

50. Becerra, S.P., et al., *Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript.* J Virol, 1988. **62**(8): p. 2745-54.

51. Van Vliet, K.M., et al., *The role of the adeno-associated virus capsid in gene transfer.* Methods Mol Biol, 2008. **437**: p. 51-91.

52. Myers, M.W. and B.J. Carter, *Assembly of adeno-associated virus.* Virology, 1980. **102**(1): p. 71-82.

53. Wistuba, A., et al., *Subcellular compartmentalization of adeno-associated virus type 2 assembly.* J Virol, 1997. **71**(2): p. 1341-52.

54. Qiu, J. and K.E. Brown, *A 110-kDa nuclear shuttle protein, nucleolin, specifically binds to adeno-associated virus type 2 (AAV-2) capsid.* Virology, 1999. **257**(2): p. 373-82.

55. Bevington, J.M., et al., *Adeno-associated virus interactions with B23/Nucleophosmin: identification of sub-nucleolar virion regions.* Virology, 2007. **357**(1): p. 102-13.

56. Harbison, C.E., J.A. Chiorini, and C.R. Parrish, *The parvovirus capsid odyssey: from the cell surface to the nucleus.* Trends Microbiol, 2008. **16**(5): p. 208-14.

57. Qing, K., et al., *Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2.* Nat Med, 1999. **5**(1): p. 71-7.

58. Kashiwakura, Y., et al., *Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection.* J Virol, 2005. **79**(1): p. 609-14.

59. Asokan, A., et al., *Adeno-associated virus type 2 contains an integrin alpha5beta1 binding domain essential for viral cell entry.* J Virol, 2006. **80**(18): p. 8961-9.

60. Akache, B., et al., *The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9.* J Virol, 2006. **80**(19): p. 9831-6.

61. Ding, W., et al., *rAAV2 traffics through both the late and the recycling endosomes in a dose-dependent fashion.* Mol Ther, 2006. **13**(4): p. 671-82.

62. Girod, A., et al., *The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity.* J Gen Virol, 2002. **83**(Pt 5): p. 973-8.

63. Kronenberg, S., et al., *A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini.* J Virol, 2005. **79**(9): p. 5296-303.

64. Sonntag, F., et al., *Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus.* J Virol, 2006. **80**(22): p. 11040-54.

65. Grieger, J.C., S. Snowdy, and R.J. Samulski, *Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly.* J Virol, 2006. **80**(11): p. 5199-210.

66. Ni, T.H., et al., *Cellular proteins required for adeno-associated virus DNA replication in the absence of adenovirus coinfection.* J Virol, 1998. **72**(4): p. 2777-87.

67. Nash, K., W. Chen, and N. Muzyczka, *Complete in vitro reconstitution of adeno-associated virus DNA replication requires the minichromosome maintenance complex proteins.* J Virol, 2008. **82**(3): p. 1458-64.

68. Chang, L.S. and T. Shenk, *The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters.* J Virol, 1990. **64**(5): p. 2103-9.

69. Ward, P., et al., *Role of the adenovirus DNA-binding protein in in vitro adeno-associated virus DNA replication.* J Virol, 1998. **72**(1): p. 420-7.

70. Ferrari, F.K., et al., *Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors.* J Virol, 1996. **70**(5): p. 3227-34.

71. Lilley, C.E., R.A. Schwartz, and M.D. Weitzman, *Using or abusing: viruses and the cellular DNA damage response.* Trends Microbiol, 2007. **15**(3): p. 119-26.

72. Schwartz, R.A., et al., *The Mre11/Rad50/Nbs1 complex limits adeno-associated virus transduction and replication.* J Virol, 2007. **81**(23): p. 12936-45.

73. Nayak, R. and D.J. Pintel, *Adeno-associated viruses can induce phosphorylation of eIF2alpha via PKR activation, which can be overcome*

by helper adenovirus type 5 virus-associated RNA. J Virol, 2007. **81**(21): p. 11908-16.

74. Toublanc, E., et al., *Identification of a replication-defective herpes simplex virus for recombinant adeno-associated virus type 2 (rAAV2) particle assembly using stable producer cell lines.* J Gene Med, 2004. **6**(5): p. 555-64.

75. Weindler, F.W. and R. Heilbronn, *A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication.* J Virol, 1991. **65**(5): p. 2476-83.

76. Slanina, H., et al., *Role of the herpes simplex virus helicase-primase complex during adeno-associated virus DNA replication.* J Virol, 2006. **80**(11): p. 5241-50.

77. Geoffroy, M.C., et al., *Herpes simplex virus type 1 ICP0 protein mediates activation of adeno-associated virus type 2 rep gene expression from a latent integrated form.* J Virol, 2004. **78**(20): p. 10977-86.

78. Nicolas, A., et al., *Identification of rep-associated factors in herpes simplex virus type 1-induced adeno-associated virus type 2 replication compartments.* J Virol, 2010. **84**(17): p. 8871-87.

79. Nash, K., et al., *Identification of cellular proteins that interact with the adeno-associated virus rep protein.* J Virol, 2009. **83**(1): p. 454-69.

80. Daya, S. and K.I. Berns, *Gene therapy using adeno-associated virus vectors.* Clin Microbiol Rev, 2008. **21**(4): p. 583-93.

81. Linden, R.M., et al., *Site-specific integration by adeno-associated virus.* Proc Natl Acad Sci U S A, 1996. **93**(21): p. 11288-94.

82. Young, S.M., Jr. and R.J. Samulski, *Adeno-associated virus (AAV) site-specific recombination does not require a Rep-dependent origin of replication within the AAV terminal repeat.* Proc Natl Acad Sci U S A, 2001. **98**(24): p. 13525-30.

83. Philpott, N.J., et al., *A p5 integration efficiency element mediates Rep-dependent integration into AAVS1 at chromosome 19.* Proc Natl Acad Sci U S A, 2002. **99**(19): p. 12381-5.

84. McCarty, D.M., S.M. Young, Jr., and R.J. Samulski, *Integration of adeno-associated virus (AAV) and recombinant AAV vectors.* Annu Rev Genet, 2004. **38**: p. 819-45.

85. Schnepp, B.C., et al., *Characterization of adeno-associated virus genomes isolated from human tissues.* J Virol, 2005. **79**(23): p. 14793-803.

86. Drew, H.R., L.J. Lockett, and G.W. Both, *Increased complexity of wild-type adeno-associated virus-chromosomal junctions as determined by analysis of unselected cellular genomes.* J Gen Virol, 2007. **88**(Pt 6): p. 1722-32.

87. Deyle, D.R. and D.W. Russell, *Adeno-associated virus vector integration.* Curr Opin Mol Ther, 2009. **11**(4): p. 442-7.

88. Gao, G., et al., *Adeno-associated viruses undergo substantial evolution in primates during natural infections.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 6081-6.

89. Mingozzi, F. and K.A. High, *Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges.* Nat Rev Genet, 2011. **12**(5): p. 341-55.

90. Glover, D.J., H.J. Lipps, and D.A. Jans, *Towards safe, non-viral therapeutic gene expression in humans.* Nat Rev Genet, 2005. **6**(4): p. 299-310.

91. Seow, Y. and M.J. Wood, *Biological gene delivery vehicles: beyond viral vectors.* Mol Ther, 2009. **17**(5): p. 767-77.

92. Kay, M.A., *State-of-the-art gene-based therapies: the road ahead.* Nat Rev Genet, 2011. **12**(5): p. 316-28.

93. Rusconi S, C.M., *Chapter: Vector for Gene Delivery. Modern therapeutics in rheumatic diseases* G. Tsokos, Editor. 2002, Humana Press: Totowa, New Jersey.

94. Grimm, D., et al., *Novel tools for production and purification of recombinant adenoassociated virus vectors.* Hum Gene Ther, 1998. **9**(18): p. 2745-60.

95. Russell, D.W. and M.A. Kay, *Adeno-associated virus vectors and hematology.* Blood, 1999. **94**(3): p. 864-74.

96. Thomas, C.E., et al., *Rapid uncoating of vector genomes is the key to efficient liver transduction with pseudotyped adeno-associated virus vectors.* J Virol, 2004. **78**(6): p. 3110-22.

97. Duan, D., et al., *Structural analysis of adeno-associated virus transduction circular intermediates.* Virology, 1999. **261**(1): p. 8-14.

98. Yang, J., et al., *Concatamerization of adeno-associated virus circular genomes occurs through intermolecular recombination.* J Virol, 1999. **73**(11): p. 9468-77.

99. Acland, G.M., et al., *Long-term restoration of rod and cone vision by single dose rAAV-mediated gene transfer to the retina in a canine model of childhood blindness.* Mol Ther, 2005. **12**(6): p. 1072-82.

100. Yang, C.C., et al., *Cellular recombination pathways and viral terminal repeat hairpin structures are sufficient for adeno-associated virus integration in vivo and in vitro.* J Virol, 1997. **71**(12): p. 9231-47.

101. Inagaki, K., et al., *DNA palindromes with a modest arm length of greater, similar 20 base pairs are a significant target for recombinant adeno-associated virus vector integration in the liver, muscles, and heart in mice.* J Virol, 2007. **81**(20): p. 11290-303.

102. Miller, D.G., et al., *Large-scale analysis of adeno-associated virus vector integration sites in normal human cells.* J Virol, 2005. **79**(17): p. 11434-42.

103. Nakai, H., et al., *Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver.* J Virol, 2005. **79**(6): p. 3606-14.

104. Miller, D.G., L.M. Petek, and D.W. Russell, *Adeno-associated virus vectors integrate at chromosome breakage sites.* Nat Genet, 2004. **36**(7): p. 767-73.

105. Miller, D.G., E.A. Rutledge, and D.W. Russell, *Chromosomal effects of adeno-associated virus vector integration.* Nat Genet, 2002. **30**(2): p. 147-8.

106. Donsante, A., et al., *AAV vector integration sites in mouse hepatocellular carcinoma.* Science, 2007. **317**(5837): p. 477.

107. Coura Rdos, S. and N.B. Nardi, *The state of the art of adeno-associated virus-based vectors in gene therapy.* Virol J, 2007. **4**: p. 99.

108. McCarty, D.M., P.E. Monahan, and R.J. Samulski, *Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis.* Gene Ther, 2001. **8**(16): p. 1248-54.

109. Grimm, D. and M.A. Kay, *From virus evolution to vector revolution: use of naturally occurring serotypes of adeno-associated virus (AAV) as novel vectors for human gene therapy.* Curr Gene Ther, 2003. **3**(4): p. 281-304.

110. McFarland, N.R., et al., *Comparison of transduction efficiency of recombinant AAV serotypes 1, 2, 5, and 8 in the rat nigrostriatal system.* J Neurochem, 2009. **109**(3): p. 838-45.

111. Rabinowitz, J.E., et al., *Cross-dressing the virion: the transcapsidation of adeno-associated virus serotypes functionally defines subgroups.* J Virol, 2004. **78**(9): p. 4421-32.

112. Shen, X., T. Storm, and M.A. Kay, *Characterization of the relationship of AAV capsid domain swapping to liver transduction efficiency.* Mol Ther, 2007. **15**(11): p. 1955-62.

113. Hauck, B., et al., *Efficient AAV1-AAV2 hybrid vector for gene therapy of hemophilia.* Hum Gene Ther, 2006. **17**(1): p. 46-54.

114. Muzyczka, N. and K.H. Warrington, Jr., *Custom adeno-associated virus capsids: the next generation of recombinant vectors with novel tropism.* Hum Gene Ther, 2005. **16**(4): p. 408-16.

115. Kwon, I. and D.V. Schaffer, *Designer gene delivery vectors: molecular engineering and evolution of adeno-associated viral vectors for enhanced gene transfer.* Pharm Res, 2008. **25**(3): p. 489-99.

116. Bartlett, J.S., et al., *Targeted adeno-associated virus vector transduction of nonpermissive cells mediated by a bispecific F(ab'gamma)2 antibody.* Nat Biotechnol, 1999. **17**(2): p. 181-6.

117. Arnold, G.S., et al., *Metabolic biotinylation provides a unique platform for the purification and targeting of multiple AAV vector serotypes.* Mol Ther, 2006. **14**(1): p. 97-106.

118. Yang, Q., et al., *Development of novel cell surface CD34-targeted recombinant adenoassociated virus vectors for gene therapy.* Hum Gene Ther, 1998. **9**(13): p. 1929-37.

119. Wu, P., et al., *Mutational analysis of the adeno-associated virus type 2 (AAV2) capsid gene and construction of AAV2 vectors with altered tropism.* J Virol, 2000. **74**(18): p. 8635-47.

120. Girod, A., et al., *Genetic capsid modifications allow efficient re-targeting of adeno-associated virus type 2.* Nat Med, 1999. **5**(9): p. 1052-6.

121. Loiler, S.A., et al., *Targeting recombinant adeno-associated virus vectors to enhance gene transfer to pancreatic islets and liver.* Gene Ther, 2003. **10**(18): p. 1551-8.

122. Zincarelli, C., et al., *Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection.* Mol Ther, 2008. **16**(6): p. 1073-80.

123. Nathwani, A.C., et al., *Safe and efficient transduction of the liver after peripheral vein infusion of self-complementary AAV vector results in stable therapeutic expression of human FIX in nonhuman primates.* Blood, 2007. **109**(4): p. 1414-21.

124. Perabo, L., et al., *Heparan sulfate proteoglycan binding properties of adeno-associated virus retargeting mutants and consequences for their in vivo tropism.* J Virol, 2006. **80**(14): p. 7265-9.

125. Perabo, L., et al., *Combinatorial engineering of a gene therapy vector: directed evolution of adeno-associated virus.* J Gene Med, 2006. **8**(2): p. 155-62.

126. Mingozzi, F. and K.A. High, *Immune responses to AAV in clinical trials.* Curr Gene Ther, 2007. **7**(5): p. 316-24.

127. Lochrie, M.A., et al., *Mutations on the external surfaces of adeno-associated virus type 2 capsids that affect transduction and neutralization.* J Virol, 2006. **80**(2): p. 821-34.

128. Maheshri, N., et al., *Directed evolution of adeno-associated virus yields enhanced gene delivery vectors.* Nat Biotechnol, 2006. **24**(2): p. 198-204.

129. Duan, D., Y. Yue, and J.F. Engelhardt, *Expanding AAV packaging capacity with trans-splicing or overlapping vectors: a quantitative comparison.* Mol Ther, 2001. **4**(4): p. 383-91.

130. Mingozzi, F. and K.A. High, *Immune Responses to AAV in Clinical Trials.* Curr Gene Ther, 2011. **11**(4): p. 321-30.

131. Qiao, C., et al., *Hydrodynamic limb vein injection of adeno-associated virus serotype 8 vector carrying canine myostatin propeptide gene into normal dogs enhances muscle growth.* Hum Gene Ther, 2009. **20**(1): p. 1-10.

132. Hermonat, P.L. and N. Muzyczka, *Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells.* Proc Natl Acad Sci U S A, 1984. **81**(20): p. 6466-70.

133. Xiao, X., J. Li, and R.J. Samulski, *Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus.* J Virol, 1998. **72**(3): p. 2224-32.

134. Grimm, D., M.A. Kay, and J.A. Kleinschmidt, *Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6.* Mol Ther, 2003. **7**(6): p. 839-50.

135. Aucoin, M.G., M. Perrier, and A.A. Kamen, *Critical assessment of current adeno-associated viral vector production and quantification methods.* Biotechnol Adv, 2008. **26**(1): p. 73-88.

136. Wright, J.F., *Transient transfection methods for clinical adeno-associated viral vector production.* Hum Gene Ther, 2009. **20**(7): p. 698-706.

137. Booth, M.J., et al., *Transfection-free and scalable recombinant AAV vector production using HSV/AAV hybrids.* Gene Ther, 2004. **11**(10): p. 829-37.

138. Urabe, M., C. Ding, and R.M. Kotin, *Insect cells as a factory to produce adeno-associated virus type 2 vectors.* Hum Gene Ther, 2002. **13**(16): p. 1935-43.

139. Yang, Q., F. Chen, and J.P. Trempe, *Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins.* J Virol, 1994. **68**(8): p. 4847-56.

140. Yuan, Z., et al., *A versatile adeno-associated virus vector producer cell line method for scalable vector production of different serotypes.* Hum Gene Ther, 2011. **22**(5): p. 613-24.

141. Qiao, C., et al., *A novel gene expression control system and its use in stable, high-titer 293 cell-based adeno-associated virus packaging cell lines.* J Virol, 2002. **76**(24): p. 13015-27.

142. Zolotukhin, S., et al., *Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors.* Methods, 2002. **28**(2): p. 158-67.

143. Feng, L., et al., *Improvement in the suspension-culture production of recombinant adeno-associated virus-LacZ in HEK-293 cells using polyethyleneimine-DNA complexes in combination with hypothermic treatment.* Biotechnol Appl Biochem, 2008. **50**(Pt 2): p. 121-32.

144. Durocher, Y., et al., *Scalable serum-free production of recombinant adeno-associated virus type 2 by transfection of 293 suspension cells.* J Virol Methods, 2007. **144**(1-2): p. 32-40.

145. Negrete, A. and R.M. Kotin, *Strategies for manufacturing recombinant adeno-associated virus vectors for gene therapy applications exploiting baculovirus technology.* Brief Funct Genomic Proteomic, 2008. **7**(4): p. 303-11.

146. Smith, R.H., J.R. Levy, and R.M. Kotin, *A simplified baculovirus-AAV expression vector system coupled with one-step affinity purification yields high-titer rAAV stocks from insect cells.* Mol Ther, 2009. **17**(11): p. 1888-96.

147. Aucoin, M.G., J.A. Mena, and A.A. Kamen, *Bioprocessing of baculovirus vectors: a review.* Curr Gene Ther, 2010. **10**(3): p. 174-86.

148. Cecchini, S., T. Virag, and R.M. Kotin, *Reproducible high yields of recombinant adeno-associated virus produced using invertebrate cells in 0.02- to 200-liter cultures.* Hum Gene Ther, 2011. **22**(8): p. 1021-30.

149. Owsianka, A., et al., *Functional analysis of hepatitis C virus E2 glycoproteins and virus-like particles reveals structural dissimilarities between different forms of E2.* J Gen Virol, 2001. **82**(Pt 8): p. 1877-83.

150. Yamada, T., et al., *Nanoparticles for the delivery of genes and drugs to human hepatocytes.* Nat Biotechnol, 2003. **21**(8): p. 885-90.

151. Ferrer-Miralles, N., et al., *Microbial factories for recombinant pharmaceuticals.* Microb Cell Fact, 2009. **8**: p. 17.

152. Noad, R. and P. Roy, *Virus-like particles as immunogens.* Trends Microbiol, 2003. **11**(9): p. 438-44.

153. Jenkins, N., *Modifications of therapeutic proteins: challenges and prospects.* Cytotechnology, 2007. **53**(1-3): p. 121-5.

154. Greene, J.J., *Host cell compatibility in protein expression.* Methods Mol Biol, 2004. **267**: p. 3-14.

155. Palomares, L.A., S. Estrada-Mondaca, and O.T. Ramirez, *Production of recombinant proteins: challenges and solutions.* Methods Mol Biol, 2004. **267**: p. 15-52.

156. Quartley, E., et al., *Heterologous expression of L. major proteins in S. cerevisiae: a test of solubility, purity, and gene recoding.* J Struct Funct Genomics, 2009. **10**(3): p. 233-47.

157. Stockmann, C., et al., *Process development in Hansenula polymorpha and Arxula adeninivorans, a re-assessment.* Microb Cell Fact, 2009. **8**: p. 22.

158. Ostergaard, S., L. Olsson, and J. Nielsen, *Metabolic engineering of Saccharomyces cerevisiae.* Microbiol Mol Biol Rev, 2000. **64**(1): p. 34-50.

159. McAleer, W.J., et al., *Human hepatitis B vaccine from recombinant yeast.* Nature, 1984. **307**(5947): p. 178-80.

160. Goffeau, A., et al., *Life with 6000 genes.* Science, 1996. **274**(5287): p. 546, 563-7.

161. Hamilton, S.R. and T.U. Gerngross, *Glycosylation engineering in yeast: the advent of fully humanized yeast.* Curr Opin Biotechnol, 2007. **18**(5): p. 387-92.

162. Yin, J., et al., *Select what you need: a comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes.* J Biotechnol, 2007. **127**(3): p. 335-47.

163. Maya, D., et al., *Systems for applied gene control in Saccharomyces cerevisiae.* Biotechnol Lett, 2008. **30**(6): p. 979-87.

164. Nevoigt, E., et al., *Engineering of promoter replacement cassettes for fine-tuning of gene expression in Saccharomyces cerevisiae.* Appl Environ Microbiol, 2006. **72**(8): p. 5266-73.

165. Kim, M.D., et al., *Production of antithrombotic hirudin in GAL1-disrupted Saccharomyces cerevisiae.* Appl Microbiol Biotechnol, 2004. **65**(3): p. 259-62.

166. Valenzuela, P., et al., *Synthesis and assembly of hepatitis B virus surface antigen particles in yeast.* Nature, 1982. **298**(5872): p. 347-50.

167. Kuroda, S., et al., *Induction of protection level of anti-pre-S2 antibodies in humans immunized with a novel hepatitis B vaccine consisting of M (pre-S2 + S) protein particles (a third generation vaccine).* Vaccine, 1991. **9**(3): p. 163-9.

168. Poland, G.A. and R.M. Jacobson, *Clinical practice: prevention of hepatitis B with the hepatitis B vaccine.* N Engl J Med, 2004. **351**(27): p. 2832-8.

169. Lexchin, J., N. Arya, and S. Singh, *Gardasil(R) - The New HPV Vaccine: The Right Product, the Right Time? A Commentary.* Healthc Policy, 2010. **5**(4): p. 26-36.

170. Lowy, D.R. and J.T. Schiller, *Prophylactic human papillomavirus vaccines.* J Clin Invest, 2006. **116**(5): p. 1167-73.

171. Stanley, M., *Prophylactic HPV vaccines: prospects for eliminating ano-genital cancer.* Br J Cancer, 2007. **96**(9): p. 1320-3.

172. Zhao, K.N. and I.H. Frazer, *Replication of bovine papillomavirus type 1 (BPV-1) DNA in Saccharomyces cerevisiae following infection with BPV-1 virions.* J Virol, 2002. **76**(7): p. 3359-64.

173. Jansen, K.U., et al., *Vaccination with yeast-expressed cottontail rabbit papillomavirus (CRPV) virus-like particles protects rabbits from CRPV-induced papilloma formation.* Vaccine, 1995. **13**(16): p. 1509-14.

174. Hale, A.D., et al., *Expression and antigenic characterization of the major capsid proteins of human polyomaviruses BK and JC in Saccharomyces cerevisiae.* J Virol Methods, 2002. **104**(1): p. 93-8.

175. Lowin, T., et al., *Parvovirus B19 VP2-proteins produced in Saccharomyces cerevisiae: comparison with VP2-particles produced by baculovirus-derived vectors.* J Vet Med B Infect Dis Vet Public Health, 2005. **52**(7-8): p. 348-52.

176. Freivalds, J., et al., *Assembly of bacteriophage Qbeta virus-like particles in yeast Saccharomyces cerevisiae and Pichia pastoris.* J Biotechnol, 2006. **123**(3): p. 297-303.

177. Sun, S., et al., *A new RNA vaccine platform based on MS2 virus-like particles produced in Saccharomyces cerevisiae.* Biochem Biophys Res Commun, 2011. **407**(1): p. 124-8.

178. Tsunetsugu-Yokota, Y., et al., *Yeast-derived human immunodeficiency virus type 1 p55(gag) virus-like particles activate dendritic cells (DCs) and induce perforin expression in Gag-specific CD8(+) T cells by cross-presentation of DCs.* J Virol, 2003. **77**(19): p. 10250-9.

179. Rodriguez-Limas, W.A., et al., *Molecular and process design for rotavirus-like particle production in Saccharomyces cerevisiae.* Microb Cell Fact, 2011. **10**: p. 33.

180. Juozapaitis, M., et al., *Synthesis of recombinant human parainfluenza virus 1 and 3 nucleocapsid proteins in yeast Saccharomyces cerevisiae.* Virus Res, 2008. **133**(2): p. 178-86.

181. Juozapaitis, M., et al., *Generation of henipavirus nucleocapsid proteins in yeast Saccharomyces cerevisiae.* Virus Res, 2007. **124**(1-2): p. 95-102.

182. Juozapaitis, M., et al., *Generation of menangle virus nucleocapsid-like particles in yeast Saccharomyces cerevisiae.* J Biotechnol, 2007. **130**(4): p. 441-7.

183. Petraityte, R., et al., *Generation of Tioman virus nucleocapsid-like particles in yeast Saccharomyces cerevisiae.* Virus Res, 2009. **145**(1): p. 92-6.

184. Slibinskas, R., et al., *Synthesis of the measles virus nucleoprotein in yeast Pichia pastoris and Saccharomyces cerevisiae.* J Biotechnol, 2004. **107**(2): p. 115-24.

185. Samuel, D., et al., *High level expression of recombinant mumps nucleoprotein in Saccharomyces cerevisiae and its evaluation in mumps IgM serology.* J Med Virol, 2002. **66**(1): p. 123-30.

186. Juozapaitis, M., et al., *Generation of Sendai virus nucleocapsid-like particles in yeast.* Virus Res, 2005. **108**(1-2): p. 221-4.

187. Razanskiene, A., et al., *High yields of stable and highly pure nucleocapsid proteins of different hantaviruses can be generated in the yeast Saccharomyces cerevisiae.* J Biotechnol, 2004. **111**(3): p. 319-33.

188. Saelens, X., et al., *Protection of mice against a lethal influenza virus challenge after immunization with yeast-derived secreted influenza virus hemagglutinin.* Eur J Biochem, 1999. **260**(1): p. 166-75.

189. Sakamoto, S., et al., *Studies on the structures and antigenic properties of rabies virus glycoprotein analogues produced in yeast cells.* Vaccine, 1999. **17**(3): p. 205-18.

190. Klepfer, S.R., et al., *Characterization of rabies glycoprotein expressed in yeast.* Arch Virol, 1993. **128**(3-4): p. 269-86.

191. Nozaki, C., et al., *Expression of herpes simplex virus glycoprotein B gene in yeast.* Virus Res, 1985. **4**(1): p. 107-13.

192. Ulrich, R., et al., *Core particles of hepatitis B virus as carrier for foreign epitopes.* Adv Virus Res, 1998. **50**: p. 141-82.

193. Roy, P. and R. Noad, *Virus-like particles as a vaccine delivery system: myths and facts.* Adv Exp Med Biol, 2009. **655**: p. 145-58.

194. Clark, K.B., et al., *Expression and characterization of human group C rotavirus virus-like particles in insect cells.* Virology, 2009. **387**(2): p. 267-72.

195. Kotze L, S.J., den Haan, van Zyl W.H, Görgens J.F, *Expression of human papillomavirus type 16 (HPV16)*

L1 protein in Pichia pastoris. African Journal of Biotechnology Vol. 10 (2), pp. 214-219, 10 January, 2011, 2011. **10**(2): p. 5.

196. Zhao, K.N. and I.H. Frazer, *Saccharomyces cerevisiae is permissive for replication of bovine papillomavirus type 1.* J Virol, 2002. **76**(23): p. 12265-73.

197. Rosenfeld, A.B. and V.R. Racaniello, *Hepatitis C virus internal ribosome entry site-dependent translation in Saccharomyces cerevisiae is independent of polypyrimidine tract-binding protein, poly(rC)-binding protein 2, and La protein.* J Virol, 2005. **79**(16): p. 10126-37.

198. Blanco, R., L. Carrasco, and I. Ventoso, *Cell killing by HIV-1 protease.* J Biol Chem, 2003. **278**(2): p. 1086-93.

199. Alves-Rodrigues, I., et al., *Saccharomyces cerevisiae: a useful model host to study fundamental biology of viral replication.* Virus Res, 2006. **120**(1-2): p. 49-56.

200. Wickner, R.B., *Double-stranded RNA viruses of Saccharomyces cerevisiae.* Microbiol Rev, 1996. **60**(1): p. 250-65.

201. Nagy, P.D., *Yeast as a model host to explore plant virus-host interactions.* Annu Rev Phytopathol, 2008. **46**: p. 217-42.

202. Ishikawa, M., et al., *In vivo DNA expression of functional brome mosaic virus RNA replicons in Saccharomyces cerevisiae.* J Virol, 1997. **71**(10): p. 7781-90.

203. Chen, J. and P. Ahlquist, *Brome mosaic virus polymerase-like protein 2a is directed to the endoplasmic reticulum by helicase-like viral protein 1a.* J Virol, 2000. **74**(9): p. 4310-8.

204. Rubino, L., A. Di Franco, and M. Russo, *Expression of a plant virus non-structural protein in Saccharomyces cerevisiae causes membrane proliferation and altered mitochondrial morphology.* J Gen Virol, 2000. **81**(Pt 1): p. 279-86.

205. Panavas, T. and P.D. Nagy, *Mechanism of stimulation of plus-strand synthesis by an RNA replication enhancer in a tombusvirus.* J Virol, 2005. **79**(15): p. 9777-85.

206. Weeks, S.A. and D.J. Miller, *The heat shock protein 70 cochaperone YDJ1 is required for efficient membrane-specific flock house virus RNA replication complex assembly and function in Saccharomyces cerevisiae.* J Virol, 2008. **82**(4): p. 2004-12.

207. den Boon, J.A. and P. Ahlquist, *Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories.* Annu Rev Microbiol, 2010. **64**: p. 241-56.

208. Janda, M. and P. Ahlquist, *RNA-dependent replication, transcription, and persistence of brome mosaic virus RNA replicons in S. cerevisiae.* Cell, 1993. **72**(6): p. 961-70.

209. Pantaleo, V., L. Rubino, and M. Russo, *Replication of Carnation Italian ringspot virus defective interfering RNA in Saccharomyces cerevisiae.* J Virol, 2003. **77**(3): p. 2116-23.

210. Panavas, T. and P.D. Nagy, *Yeast as a model host to study replication and recombination of defective interfering RNA of Tomato bushy stunt virus.* Virology, 2003. **314**(1): p. 315-25.

211. Price, G.E., et al., *Perforin and Fas cytolytic pathways coordinately shape the selection and diversity of CD8+-T-cell escape variants of influenza virus.* J Virol, 2005. **79**(13): p. 8545-59.

212. Delan-Forino, C., M.C. Maurel, and C. Torchet, *Replication of avocado sunblotch viroid in the yeast Saccharomyces cerevisiae.* J Virol, 2011. **85**(7): p. 3229-38.

213. Angeletti, P.C., et al., *Stable replication of papillomavirus genomes in Saccharomyces cerevisiae.* J Virol, 2002. **76**(7): p. 3350-8.

214. Raghavan, V., et al., *The DNA-A component of a plant geminivirus (Indian mung bean yellow mosaic virus) replicates in budding yeast cells.* J Virol, 2004. **78**(5): p. 2405-13.

215. Ruffing, M., H. Zentgraf, and J.A. Kleinschmidt, *Assembly of viruslike particles by recombinant structural proteins of adeno-associated virus type 2 in insect cells.* J Virol, 1992. **66**(12): p. 6922-30.

216. Steinbach, S., et al., *Assembly of adeno-associated virus type 2 capsids in vitro.* J Gen Virol, 1997. **78 (Pt 6)**: p. 1453-62.

217. Hoque, M., et al., *Nuclear transport of the major capsid protein is essential for adeno-associated virus capsid formation.* J Virol, 1999. **73**(9): p. 7912-5.

218. Kronenberg, S., J.A. Kleinschmidt, and B. Bottcher, *Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids.* EMBO Rep, 2001. **2**(11): p. 997-1002.

219. Warrington, K.H., Jr., et al., *Adeno-associated virus type 2 VP2 capsid protein is nonessential and can tolerate large peptide insertions at its N terminus.* J Virol, 2004. **78**(12): p. 6595-609.

220. Farr, G.A., L.G. Zhang, and P. Tattersall, *Parvoviral virions deploy a capsid-tethered lipolytic enzyme to breach the endosomal membrane during cell entry.* Proc Natl Acad Sci U S A, 2005. **102**(47): p. 17148-53.

221. Grieger, J.C., et al., *Surface-exposed adeno-associated virus Vp1-NLS capsid fusion protein rescues infectivity of noninfectious wild-type Vp2/Vp3 and Vp3-only capsids but not that of fivefold pore mutant virions.* J Virol, 2007. **81**(15): p. 7833-43.

222. Vihinen-Ranta, M., et al., *The VP1 N-terminal sequence of canine parvovirus affects nuclear transport of capsids and efficient cell infection.* J Virol, 2002. **76**(4): p. 1884-91.

223. Qiu, J. and D.J. Pintel, *The adeno-associated virus type 2 Rep protein regulates RNA processing via interaction with the transcription template.* Mol Cell Biol, 2002. **22**(11): p. 3639-52.

224. Pereira, D.J., D.M. McCarty, and N. Muzyczka, *The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection.* J Virol, 1997. **71**(2): p. 1079-88.

225. McKune, K., et al., *Six human RNA polymerase subunits functionally substitute for their yeast counterparts.* Mol Cell Biol, 1995. **15**(12): p. 6895-900.

226. Romanos, M.A., C.A. Scorer, and J.J. Clare, *Foreign gene expression in yeast: a review.* Yeast, 1992. **8**(6): p. 423-88.

227. Qiu, J. and D. Pintel, *Processing of adeno-associated virus RNA.* Front Biosci, 2008. **13**: p. 3101-15.

228. Wistuba, A., et al., *Intermediates of adeno-associated virus type 2 assembly: identification of soluble complexes containing Rep and Cap proteins.* J Virol, 1995. **69**(9): p. 5311-9.

229. Cathomen, T., D. Collete, and M.D. Weitzman, *A chimeric protein containing the N terminus of the adeno-associated virus Rep protein recognizes its target site in an in vivo assay.* J Virol, 2000. **74**(5): p. 2372-82.

230. Kushnirov, V.V., *Rapid and reliable protein extraction from yeast.* Yeast, 2000. **16**(9): p. 857-60.

231. Watts, F., C. Castle, and J. Beggs, *Aberrant splicing of Drosophila alcohol dehydrogenase transcripts in Saccharomyces cerevisiae.* EMBO J, 1983. **2**(11): p. 2085-91.

232. Henikoff, S., *Cloning exons of mapping of transcription: characterization of the Drosophila melanogaster alcohol dehydrogenase gene.* Nucleic Acids Res, 1983. **11**(14): p. 4735-52.

233. Becerra, S.P., et al., *Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript.* J Virol, 1988. **62**(8): p. 2745-54.

234. Friedl, A.A., et al., *Ty1 integrase overexpression leads to integration of non-Ty1 DNA fragments into the genome of Saccharomyces cerevisiae.* Mol Genet Genomics, 2010. **284**(4): p. 231-42.

235. Meyer, M. and J. Vilardell, *The quest for a message: budding yeast, a model organism to study the control of pre-mRNA splicing.* Brief Funct Genomic Proteomic, 2009. **8**(1): p. 60-7.

236. Kochetov, A.V., N.A. Kolchanov, and A. Sarai, *Interrelations between the efficiency of translation start sites and other sequence features of yeast mRNAs.* Mol Genet Genomics, 2003. **270**(5): p. 442-7.

237. Chen, S.J., et al., *Translational efficiency of a non-AUG initiation codon is significantly affected by its sequence context in yeast.* J Biol Chem, 2008. **283**(6): p. 3173-80.

238. van den Brink, J., et al., *Energetic limits to metabolic flexibility: responses of Saccharomyces cerevisiae to glucose-galactose transitions.* Microbiology, 2009. **155**(Pt 4): p. 1340-50.

239. Demir, O. and I. Aksan Kurnaz, *An integrated model of glucose and galactose metabolism regulated by the GAL genetic switch.* Comput Biol Chem, 2006. **30**(3): p. 179-92.

240. Wright, J.F., et al., *Identification of factors that contribute to recombinant AAV2 particle aggregation and methods to prevent its occurrence during vector purification and formulation.* Mol Ther, 2005. **12**(1): p. 171-8.

241. Grimm, D., et al., *Titration of AAV-2 particles via a novel capsid ELISA: packaging of genomes can limit production of recombinant AAV-2.* Gene Ther, 1999. **6**(7): p. 1322-30.

242. Nony, P., et al., *Evidence for packaging of rep-cap sequences into adeno-associated virus (AAV) type 2 capsids in the absence of inverted terminal repeats: a model for generation of rep-positive AAV particles.* J Virol, 2003. **77**(1): p. 776-81.

243. Halbert, C.L., et al., *Capsid-expressing DNA in AAV vectors and its elimination by use of an oversize capsid gene for vector production.* Gene Ther, 2011. **18**(4): p. 411-7.

244. DiPrimio, N., et al., *Surface loop dynamics in adeno-associated virus capsid assembly.* J Virol, 2008. **82**(11): p. 5178-89.

245. Raj, A. and A. van Oudenaarden, *Nature, nurture, or chance: stochastic gene expression and its consequences.* Cell, 2008. **135**(2): p. 216-26.

246. Brauer, M.J., et al., *Coordination of growth rate, cell cycle, stress response, and metabolic activity in yeast.* Mol Biol Cell, 2008. **19**(1): p. 352-67.

247. Kim, E.J., et al., *Expression of hepatitis B surface antigen S domain in recombinant Saccharomyces cerevisiae using GAL1 promoter.* J Biotechnol, 2009. **141**(3-4): p. 155-9.

248. Kahali, B., S. Ahmad, and T.C. Ghosh, *Selective constraints in yeast genes with differential expressivity: codon pair usage and mRNA stability perspectives.* Gene, 2011. **481**(2): p. 76-82.

249. Smith, D.W., *Problems of translating heterologous genes in expression systems: the role of tRNA.* Biotechnol Prog, 1996. **12**(4): p. 417-22.

250. Fredrick, K. and M. Ibba, *How the sequence of a gene can tune its translation.* Cell, 2010. **141**(2): p. 227-9.

251. Zhou, J., et al., *Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability.* J Virol, 1999. **73**(6): p. 4972-82.

252. Bucarey, S.A., et al., *The optimized capsid gene of porcine circovirus type 2 expressed in yeast forms virus-like particles and elicits antibody responses*

in mice fed with recombinant yeast extracts. Vaccine, 2009. **27**(42): p. 5781-90.

253. Cervelli, T., A. Backovic, and A. Galli, *Formation of AAV single stranded DNA genome from a circular plasmid in Saccharomyces cerevisiae.* PLoS One, 2011. **6**(8): p. e23474.

254. Gietz, R.D. and R.H. Schiestl, *Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier.* Yeast, 1991. **7**(3): p. 253-63.

255. Galli, A. and R.H. Schiestl, *Effects of DNA double-strand and single-strand breaks on intrachromosomal recombination events in cell-cycle-arrested yeast cells.* Genetics, 1998. **149**(3): p. 1235-50.

256. Galli, A., T. Cervelli, and R.H. Schiestl, *Characterization of the hyperrecombination phenotype of the pol3-t mutation of Saccharomyces cerevisiae.* Genetics, 2003. **164**(1): p. 65-79.

257. Wang, X.S., S. Ponnazhagan, and A. Srivastava, *Rescue and replication of adeno-associated virus type 2 as well as vector DNA sequences from recombinant plasmids containing deletions in the viral inverted terminal repeats: selective encapsidation of viral genomes in progeny virions.* J Virol, 1996. **70**(3): p. 1668-77.

258. Blakesley, R.W., et al., *Duplex regions in "single-stranded" phiX174 DNA are cleaved by a restriction endonuclease from Haemophilus aegyptius.* J Biol Chem, 1977. **252**(20): p. 7300-6.

259. Nishigaki, K., et al., *Type II restriction endonucleases cleave single-stranded DNAs in general.* Nucleic Acids Res, 1985. **13**(16): p. 5747-60.

260. Nony, P., et al., *Novel cis-acting replication element in the adeno-associated virus type 2 genome is involved in amplification of integrated rep-cap sequences.* J Virol, 2001. **75**(20): p. 9991-4.

261. Tullis, G.E. and T. Shenk, *Efficient replication of adeno-associated virus type 2 vectors: a cis-acting element outside of the terminal repeats and a minimal size.* J Virol, 2000. **74**(24): p. 11511-21.

262. Baker, A., et al., *Adenovirus E4 34k and E1b 55k oncoproteins target host DNA ligase IV for proteasomal degradation.* J Virol, 2007. **81**(13): p. 7034-40.

263. Stracker, T.H., C.T. Carson, and M.D. Weitzman, *Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex.* Nature, 2002. **418**(6895): p. 348-52.

264. Querido, E., et al., *Regulation of p53 levels by the E1B 55-kilodalton protein and E4orf6 in adenovirus-infected cells.* J Virol, 1997. **71**(5): p. 3788-98.

265. Hong, G., P. Ward, and K.I. Berns, *In vitro replication of adeno-associated virus DNA.* Proc Natl Acad Sci U S A, 1992. **89**(10): p. 4673-7.

266. Wang, X.S., S. Ponnazhagan, and A. Srivastava, *Rescue and replication signals of the adeno-associated virus 2 genome.* J Mol Biol, 1995. **250**(5): p. 573-80.

267. Gottlieb, J. and N. Muzyczka, *In vitro excision of adeno-associated virus DNA from recombinant plasmids: isolation of an enzyme fraction from HeLa cells that cleaves DNA at poly(G) sequences.* Mol Cell Biol, 1988. **8**(6): p. 2513-22.

268. Ryan, J.H., S. Zolotukhin, and N. Muzyczka, *Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats.* J Virol, 1996. **70**(3): p. 1542-53.

269. Ward, P., P. Elias, and R.M. Linden, *Rescue of the adeno-associated virus genome from a plasmid vector: evidence for rescue by replication.* J Virol, 2003. **77**(21): p. 11480-90.

270. Ward, P. and K.I. Berns, *In vitro rescue of an integrated hybrid adeno-associated virus/simian virus 40 genome.* J Mol Biol, 1991. **218**(4): p. 791-804.

271. Xiao, X., et al., *A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle.* J Virol, 1997. **71**(2): p. 941-8.

272. Hauck, B., et al., *Undetectable transcription of cap in a clinical AAV vector: implications for preformed capsid in immune responses.* Mol Ther, 2009. **17**(1): p. 144-52.

273. Musatov, S., et al., *A cis-acting element that directs circular adeno-associated virus replication and packaging.* J Virol, 2002. **76**(24): p. 12792-802.

274. Ni, T.H., et al., *In vitro replication of adeno-associated virus DNA.* J Virol, 1994. **68**(2): p. 1128-38.

275. Yalkinoglu, A.O., H. Zentgraf, and U. Hubscher, *Origin of adeno-associated virus DNA replication is a target of carcinogen-inducible DNA amplification.* J Virol, 1991. **65**(6): p. 3175-84.

276. Alexander, I.E., D.W. Russell, and A.D. Miller, *DNA-damaging agents greatly increase the transduction of nondividing cells by adeno-associated virus vectors.* J Virol, 1994. **68**(12): p. 8282-7.

277. Alexander, I.E., et al., *Effects of gamma irradiation on the transduction of dividing and nondividing cells in brain and muscle of rats by adeno-associated virus vectors.* Hum Gene Ther, 1996. **7**(7): p. 841-50.

278. Zhou, C., Q. Yang, and J.P. Trempe, *Enhancement of UV-induced cytotoxicity by the adeno-associated virus replication proteins.* Biochim Biophys Acta, 1999. **1444**(3): p. 371-83.

279. Zentilin, L., A. Marcello, and M. Giacca, *Involvement of cellular double-stranded DNA break binding proteins in processing of the recombinant adeno-associated virus genome.* J Virol, 2001. **75**(24): p. 12279-87.

280. Cervelli, T., et al., *Processing of recombinant AAV genomes occurs in specific nuclear structures that overlap with foci of DNA-damage-response proteins.* J Cell Sci, 2008. **121**(Pt 3): p. 349-57.

281. Schiestl, R.H. and T.D. Petes, *Integration of DNA fragments by illegitimate recombination in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 1991. **88**(17): p. 7585-9.

282. Lipps, G., R. Fullkrug, and E. Beck, *Cathepsin B of Schistosoma mansoni. Purification and activation of the recombinant proenzyme secreted by Saccharomyces cerevisiae.* J Biol Chem, 1996. **271**(3): p. 1717-25.

283. Schiestl, R.H., M. Dominska, and T.D. Petes, *Transformation of Saccharomyces cerevisiae with nonhomologous DNA: illegitimate integration of transforming DNA into yeast chromosomes and in vivo ligation of transforming DNA to mitochondrial DNA sequences.* Mol Cell Biol, 1993. **13**(5): p. 2697-705.

284. Manivasakam, P. and R.H. Schiestl, *Nonhomologous end joining during restriction enzyme-mediated DNA integration in Saccharomyces cerevisiae.* Mol Cell Biol, 1998. **18**(3): p. 1736-45.

285. Schmitt, M.E., T.A. Brown, and B.L. Trumpower, *A rapid and simple method for preparation of RNA from Saccharomyces cerevisiae.* Nucleic Acids Res, 1990. **18**(10): p. 3091-2.

286. Angeletti, P.C., *Replication and encapsidation of papillomaviruses in Saccharomyces cerevisiae.* Methods Mol Med, 2005. **119**: p. 247-60.

287. Krysan, P.J., S.B. Haase, and M.P. Calos, *Isolation of human sequences that replicate autonomously in human cells.* Mol Cell Biol, 1989. **9**(3): p. 1026-33.

Acknowledgements

Ever since the formative stages of this project, to the final draft of my thesis, I owe an immense debt of gratitude to my supervisor, Dr. Alvaro Galli. His invaluable sound advice, insightful criticisms and careful guidance created simulative, productive and, above all, pleasant working environment for which I express my sincere gratefulness.

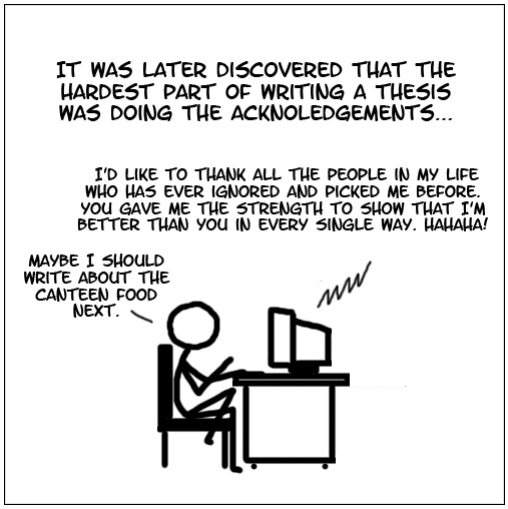
For her efforts and assistance throughout almost four years of co-working and exclusively for having kept my spirits up when the muses failed me, a special thanks to my colleague Tiziana Cervelli.

Special thanks I owe to Mauro Giacca and Lorena Zentilin who contributed to my research work in numerous ways; by letting me use their lab facilities whenever needed, and by generously giving countless, valuable recommendations and counsels.

I am also very grateful to Jurgen Kleinschmidt for his priceless suggestions and material sent to us by his laboratory; and to Matthew Weitzman for sending us plasmid constructs. Finally, I would like to thank Alessandra Salvetti for precious electron micrographs she made!

For being always profoundly analytic and a great critic that educate and challenge, I owe my gratitude to Monica Zopé.

I would be remiss without thanking to the remarkable institution of Scuola Normale Superiore and its always helpful staff. Here I would specifically mention Monacci Lucia and Michele Padricelli whose extreme generosity and cheerful faces will always be remembered.



Thanks to....

There are no rules on this page and there will be no revisions nor corrections....

This thesis wouldn't have been as enjoying as it was without my lab mates: Antonella, Barbara, Laura C. & S., Awatef, Cristina, Sara, Valentina, Vania, Silvia, Rudi, Serena, Francesco, Dani, Vero, Caro etc.. whose treasured advice, support, smiles and comforting words, leave me indebted for good... Durante questi anni di dottorato mi avete fatto imparare tante cose: a competere con la vostra eccellenza scientifica, a parare italiano (con un certo "strabismo" che ormai fa parte del mio profilo 'Bakova'), ridere alle battute italiane, litigare in italiano (contro la vostra volontà), cantare in italiano (questo pero' non lo sapete perché vi voglio bene).. e alla fine, sognare in italiano e l'Italia...

..E il sogno Italiano non si sarebbe completato senza Silvia, Claudia e Marco, Luca, Claudio, Stefano, Tommaso e Lorenzo... No words, in no language, can express the sense of thankfulness that I have for your unconditional love and support!

Sir Bobo, Antonio, Luca, Nicola e Sara, Teresa, Max etc... you may not be aware of how much you have contributed in the final four of this game, by defending me from stress and terrifying bad mood! A huge thank to my "Defense Line"! Comunque, sembra come se foste stati presenti sin dall'inizio...e questa vostra presenza la porterò ovunque andro'..

Saska i Marko, malo su dve reci 'Hvala Vam' pa sam pronasla bolji nacin da izrazim svoje osecaje..Dodeljujem vam Bakova Nagradu za najvecu moralnu i tehnicku podrsku pruzenu tokom izrade ovog doktorata u trajanju od skoro 6 meseci! Ova nagrada nema nikavu materijalnu vrednost jer ne bih da Vas uvredim..svaka cifra bi bila mala!!

To each of the above, I extend my deepest appreciation.

Ana B.

19.12.2011.

And last but not the least, I'll to express the overwhelming thankfulness to my parents and my sister. I fear I will have to write a new manuscript only for this purpose and still, it will remain unpublished...because it will never get to the end! ..For universe of smiles & for worming my heart, a final devotion to my adored Kiki..