UNIVERSITE BRETAGNE LOIRE

UNIVERSITÉ DE NANTES

# Thèse de Doctorat

Janina GERGEN

*Mémoire présenté en vue de l'obtention du*
***grade de Docteur de l'Université de Nantes***
*sous le sceau de l'Université Bretagne Loire*

**École doctorale :** Biologie-Santé

**Discipline :** Biologie, Médecine et Santé
**Spécialité :** Immunologie
**Unité de recherche :** CRTI UMR 1064, INSERM

**Soutenue le** 15.12.2017
**Thèse N° :**

# The CRISPR/Cas9 system as an anti-viral strategy against the human cytomegalovirus

## JURY

| | |
|---|---|
| **Président du jury** | **Julie DÉCHANET-MERVILLE**, Directeur de Recherche, CNRS UMR 5164 - Immuno ConcEpT, Université de Bordeaux, 33076 Bordeaux, France |
| Rapporteurs : | **Audrey ESCLATINE, Professeur**, I2BC, Université Paris-Sud, 92290 Châtenay-Malabry, France |
| | **Andrea CRISANTI, Professeur**, Faculty of Natural Science, Department of Life Sciences, Imperial College London, London SW7 2AZ, United Kingdom |
| Examinateurs : | **Julie DÉCHANET-MERVILLE**, Directeur de Recherche, CNRS UMR 5164 - Immuno ConcEpT, Université de Bordeaux, 33076 Bordeaux, France |
| | **Emiliano RICCI**, Chargé de Recherches, INSERM CIRI U1111 École normal supérieure de Lyon, 69007 Lyon, France |
| Directeur de Thèse : | **Franck HALARY**, Chargé de Recherche Titulaire d'une HDR, CRTI UMR 1064, INSERM, Université de Nantes, 44093 Nantes, France |
| Co-directeur de Thèse : | **Fabienne HASPOT**, Chargé de Recherche, CRTI UMR 1064, INSERM, Université de Nantes, 44093 Nantes, France |

# Acknowledgment

The past three years in France have been very exciting, challenging and sometimes difficult. I gathered a lot of personal and professional experiences, which will help me in my future. In my time here I had support and guidance and I would like to express my gratitude for that.

First of all, I want to thank my jury members Prof. Dr. Audrey Esclatine, Prof. Dr. Andrea Crisanti, Dr. Julie Déchanet-Merville and Dr. Emiliano Ricci for taking the time to review my work and attend to my thesis defence. I hope we will have an interesting scientific discussion about my work. Especially, I want to thank Dr. Julie Déchanet-Merville and Dr. Tuan Huy Nguyen, who both followed my work over the three years as my thesis committee and encouraged me with open discussions and critical questions.

I want to thank my supervisors, Fabienne Haspot and Franck Halary, for the opportunity to do my PhD at the Université de Nantes and INSERM on such an interesting project. Thank you very much for your confidence in me so that you gave me the chance to work independently and develop my project actively.

I also would like to thank all members of the Immuno-Virology team for the constructive critic and new ideas in our weekly meetings. Especially, I would like to thank Flora Coulon for her technical advice and patience with me. I enjoyed trying to speak French with you.

I would like to thank Labex IGO for the research grand financing my work.

I would like to thank Olaf Pinkenburg for our collaboration and your advice for the cloning design.

A big thanks goes to Isabel Vogel. We only met two years ago. You helped me to keep trying to succeed my life and work here in France. You were a mentor, a teacher and a friend. Thank you very much for everything and especially for the comments and corrections on this manuscript.

Aubin Mekeze, you were on my side for the last year. It was not always easy and not always happy, but at the end you gave me the strength to not give up. Thank you, for believing in me.

Finally, I want to thank my family, for the open door to come home whenever I wanted. Thank you for the emotional back up and spiritual support over the years.

# Table of content


**LIST OF FIGURES** ..... **1**

**LIST OF BOXES** ..... **3**

**LIST OF ABBREVIATIONS** ..... **4**

**I PREFACE** ..... **8**

**II GENERAL AIM OF THIS THESIS** ..... **9**

**III INTRODUCTION** ..... **11**

**1 Human Cytomegalovirus** ..... **11**
1.1 Epidemiology ..... 11
1.2 Transmission ..... 12
1.3 Polymorphism, genotypes and viral strains ..... 12
1.4 Pathology ..... 13
1.4.1 Congenital infection ..... 13
1.4.2 CMV in HIV patients ..... 14
1.4.3 CMV in intensive care unit patients ..... 15
1.4.4 CMV in glioblastoma ..... 15
1.4.5 CMV in transplantation ..... 16
1.4.5.1 Solid organ transplantation (SOT) ..... 16
1.4.5.2 Hematopoietic stem cell transplantation (HSCT) ..... 17
1.5 Molecular Characteristics of HCMV ..... 20
1.5.1 The viral particle ..... 20
1.5.2 Replication cycle ..... 21
1.5.3 Major Immediate Early proteins ..... 23
1.5.4 Latency ..... 26
1.6 Antiviral therapy ..... 29
1.6.1 Approved anti-viral chemotherapy ..... 29
1.6.2 Anti-HCMV drugs under development ..... 32
1.6.3 Adoptive T-cell therapy ..... 34
1.6.4 Vaccination ..... 36

**2 CRISPR/Cas9** ..... **41**
2.1 The origin of the CRISPR/Cas9 system ..... 42
2.2 Principle of genome editing via CRISPR/Cas9 ..... 43
2.3 Specificity of the CRISPR/Cas9 system ..... 47
2.3.1 Modification of the Cas9 and gRNA to improve the specificity ..... 49
2.3.2 Detection of off targets ..... 53
2.4 Delivery of the CRIPSR/Cas9 system ..... 54
2.4.1 Non-viral delivery ..... 55
2.4.2 Viral delivery ..... 56
2.4.3 Transgene organization of plasmid and viral vector delivery ..... 58

2.5 Cas9 variants and orthologues ..... 59
2.5.1 Cas9 variants ..... 59
2.5.2 Cas9 orthologues ..... 62
2.6 Applications of the CRISPR/Cas9 system for human diseases ..... 65
2.6.1 The CRIPSR/Cas9 system as an antiviral strategy ..... 68
2.6.1.1 HIV ..... 68
2.6.1.2 HBV ..... 70
2.6.1.3 Herpesviruses ..... 70
2.6.1.4 HPV ..... 71
2.6.1.5 RNA viruses ..... 71
2.6.1.6 Resistances ..... 72

**IV HYPOTHESIS 1 ..... 75**

**1 Design of the anti-*UL122/123* CRISPR/Cas9 strategies ..... 76**
1.1 General Principle ..... 76
1.2 Methods ..... 77
1.2.1 Lentiviral vector construction ..... 77
1.2.2 Lentiviral vector production ..... 78
1.2.3 RT-qPCR GAG/POL ..... 78
1.2.4 HCMV production ..... 78
1.2.5 Titration of HCMV ..... 78
1.2.6 FACS ..... 79
1.2.7 Cell lines ..... 79
1.2.8 HCMV infections ..... 79
1.2.9 PCR ..... 79
1.2.10 T7-assay ..... 80
1.2.11 PCR cloning ..... 80
1.2.12 qPCR *US8* ..... 80
1.2.13 Western Blot ..... 80
1.2.14 Virion release analysis by trans-infection plaque assay ..... 81
1.3 Results ..... 82
1.3.1 Design of the gRNAs ..... 82
1.3.1.1 Singleplex ..... 82
1.3.1.2 Validation of the most efficient gRNA for the singleplex strategy ..... 85
1.3.1.3 Multiplex ..... 89
1.3.1.4 Internal control ..... 91
1.3.2 Creation of the stable cell lines for analyzing the anti-viral activity of the singleplex and multiplex CRISPR/Cas9 strategies ..... 92
1.3.3 Multiplex CRISPR/Cas9 system impairs HCMV replication by excising an essential viral gene 95
1.3.3.1 Anti-IE CRISPR/Cas9 system reduces IE expression in primary fibroblasts ..... 96
1.3.3.2 Stable expression of the anti-IE CRISPR/Cas9 system induces mutations in the *UL122/123* gene in HCMV-infected U373-MG cells ..... 97
1.3.3.3 Dramatic decrease of IE protein expression in HCMV-infected U373-MG cells expressing gRNA/Cas9 ..... 99
1.3.3.4 The *multiplex* strategy is superior over the *singleplex* strategy to inhibit the viral genome replication and late protein expression ..... 101
1.3.3.5 The anti-HCMV multiplex strategy strongly impairs virion release from U373-MG cells 102
1.4 Discussion ..... 103

1.4.1 Our anti-HCMV CRISPR/Cas9 strategies are competitive with actual state-of-the-art anti-viral CRISPR/Cas9 designs ..... 103
1.4.2 Delivery of the singleplex and multiplex strategies into the target cells ..... 104
1.4.3 Cleavage efficiency of the anti-UL122/123 CRISPR/Cas9 approach on the viral genome ..... 105
1.4.4 Targeting the *UL122/123* gene at multiple target site is more efficient to inhibit viral replication than a single target site ..... 105
1.4.5 Comparison of the standard anti-HCMV treatments with our new anti-HCMV CRISPR/Cas9 approach ..... 106
1.4.6 Viral escape from anti-viral CRISPR/Cas9 systems ..... 106
1.4.7 Conclusion ..... 107

**V HYPOTHESIS 2 ..... 108**

**1 Validation of the duplex strategy on the lytic replicating virus ..... 109**
1.1 General principle ..... 109
1.2 Methods ..... 111
1.2.1 LV cloning ..... 111
1.2.2 PCR ..... 111
1.2.3 qPCR *US8* ..... 111
1.3 Results ..... 112
1.3.1 Design gRNA ..... 112
1.3.2 Validation of the duplex strategy on the lytic viral replication ..... 115
1.3.2.1 Selection of primary fibroblast cells stably expressing the duplex/Cas9 ..... 115
1.3.2.2 Cleavage analysis of the viral genome targeted by the duplex strategy ..... 116
1.3.2.3 Targeting the viral genome at multiple sites efficiently diminishes the viral genome copy number inside the cells ..... 117
1.3.2.4 Fragmentation of the viral genome prevents immediate early and late events of the lytic replication ..... 118
1.4 Discussion ..... 120

**2 Perspectives: *Prevention of HCMV reactivation from latently infected myeloid progenitor cells* 122**

**VI GENERAL DISCUSSION ..... 123**

**1 The CRISPR/Cas9 system as an anti-viral therapy ..... 123**

**2 Transmission to clinical application ..... 126**
2.1 Delivery ..... 126
2.1.1 Engineered graft for optimized CRISPR/Cas9 delivery ..... 126
2.1.2 Possible high efficient delivery systems to achieve prolonged gRNA/Cas9 expression ..... 127
2.2 Potential adverse effects and their prevention ..... 129
2.2.1 Off-target effects on the human genome ..... 129
2.2.2 Cas9-induced immune reaction ..... 130

**3 Further applications ..... 131**

**VII PERSPECTIVES ..... 132**

**VIII SUMMERY IN FRENCH ..... 133**

IX LIST OF REFERENCES ..... 136

X APPENDIX ..... 172

1 Publication ..... 172

2 Oral Presentations ..... 221

3 Poster presentations ..... 221

4 Awards ..... 221

# List of figures

Figure 1: Development of early *versus* late CMV disease in seropositive hematopoietic stem cell transplant recipients ..... 17

Figure 2: Overview of the relation of the HCMV serostatus and reactivation. ..... 19

Figure 3: HCMV genome structures ..... 20

Figure 4: Lytic replication cycle ..... 23

Figure 5: Gene and transcript structures of the major immediate early gene. ..... 25

Figure 6: Schema of the molecular mechanism of the MIEP regulation in latency and reactivation ..... 27

Figure 7: Molecular structure of the anti-HCMV nucleoside/nucleotide analogues ..... 30

Figure 8: Target sites of anti-HCMV drugs ..... 34

Figure 9: Representation of the different types and subtypes of the bacterial and archaea CRISPR/Cas systems ..... 42

Figure 10: Different types of non-homologous end joining (NHEJ) after a double strand break (DSB) ..... 44

Figure 11: Frequency and characterization of mutations after multiplex CRISPR/Cas9 cleavage. ..... 46

Figure 12: Homologous recombination (HR) DNA repair mechanism.47

Figure 13: Tolerance of mismatches in the target region ..... 48

Figure 14: Modifications on the CRIPSR/Cas9 system of *streptococcus pyogenes* to enhance the specificity ..... 50

Figure 15: Types of genome editing for therapeutic application ..... 66

Figure 16: Type of mutations induced by the CRISPR/Cas9 system over time ..... 72

Figure 17: Lentiviral vector constructs ..... 77

Figure 18: Schema of the anti-*UL122/123* CRISPR/Cas9 strategies to prevent IE1 and IE2 expression ..... 82

Figure 19: T7 assay of the *UL122/123* gene exon 2 ..... 85

Figure 20: Decrease of the IE expression by the singleplex CRISPR/Cas9 systems ..... 87

Figure 21: Off-target analysis of the singleplex strategies by T7 assay ..... 87

Figure 22: Schematic representation of the creation of cells stably expressing the CRISPR/Cas9 system ..... 92

Figure 23: Generation of U373-MG cell lines expressing the CRISPR/Cas9 strategies ..... 93

Figure 24: Cas9 expression in transduced MRC5 cells ..... 95

Figure 25: Anti-HCMV CRISPR/Cas9 system induces mutations resulting in the decrease of IE protein expression in primary fibroblasts ..... 97

Figure 26: Mutations in the *UL122/123* gene induced by the CRISPR/Cas9 system anti-HCMV in U373-MG cell line ..... 98
Figure 27: Decrease of IE expression by the HCMV targeting CRISPR/Cas9 systems. ..... 100
Figure 28: Reduced progression of the viral replication cycle by the multiplex strategy ..... 101
Figure 29: Inhibition of virion release by the CRISPR/Cas9 systems anti-HCMV ..... 102
Figure 30: Schema of the target regions of the duplex strategy. ..... 112
Figure 31: Cas9 expression in MRC5 cells after puromycin selection ..... 115
Figure 32: Cleavage analysis of the effect of the duplex strategy on the viral genome ..... 116
Figure 33: Titration of the viral genome inside the cells ..... 117
Figure 34: Inhibition of early and late events of the viral replication cycle ..... 118

# List of Boxes

*Box 1: HCMV in HSCT* ..... *19*
*Box 2: Importance of the MIE for the viral replication* ..... *25*
*Box 3: Development from latency to reactivation* ..... *29*
*Box 4: Comparison of the different therapy approaches* ..... *40*
*Box 5: General principal of the CRISPR/Cas9 system* ..... *41*
*Box 6: How to target viruses with the CRISPR/Cas9 system* ..... *74*
*Box 7: Graphical abstract of the hypothesis 1* ..... *75*
*Box 8: Graphical abstract of the hypothesis 2* ..... *108*

# List of abbreviations

**4-HT**: 4-hydroytarmoxyfen

**AAV**: adeno-associated vector

**Adv**: adenovirus

**AIDS**: acquired immune deficiency syndrome

**AML1b**: acute myeloid leukemia1b

**AP-1**: activator protein 1

**ATF**: activator of iron transcription

**AV**: adenovirus vector

**BAC**: bacterial artificial chromosome

**Bcl2**: B-cell lymphoma 2

**BLESS**: break labelling enrichment on streptavidin and next generation sequencing

**bp**: base pair

**BSD**: double strand break

**cas**: CRISPR associated

**Cas9-HF**: high fidelity Cas9

**Cas9n**: Cas9 nickase

**CASFISH**: Cas9-mediated fluorescence in situ hybridisation

**cccDNA**: covalently closed circular DNA

**CCL22**: CC-chemokine ligand 22

**CD**: Cluster of differentiation

**CDK2**: cyclin-dependent kinase 2

**CDV**: Cidofovir

**ChIP**: Chromatin immunoprecipitation

**CIK** cells cytokine induced killer cells

**CLT**: cytolytic T cell

**cmr**: chloramphenicol-resistance gene

**CMV**: Cytomegalovirus

**CMVD**: Cytomegalovirus disease

**CNS**: central nervous system

**Cpf1**: CRISPR from *Prevotella* and *Francisella*

**CPP**: cell-penetrating peptide

**CREB**: C-AMP response element-binding protein

**CRISPR**: Clustered Regularly Interspaced Short Palindromic Repeats

**CRISPRi**: CRISPR interference

**CRISPR-TF**: CRISPR transcription factor

**crRNA**: CRISPR RNA

**crs-acting element**: cis-acting repressive sequence

**DAXX**: death domain associated protein

**DBD**: DNA-binding domain

**DC**: dendritic cells

**dCas9**: dead Cas9

**DC-SGIN**: dendritic cell-specific intracellular adhesion molecule -3-grabbing non-integrins

**DD**: destabilization domain

**DE**: delay early

**DEA**: delay early antigen

**DISC**: transgenic disabled infectious single cycle vaccine strains

**DLI**: donor lymphocyte infusion

**DMD**: Duchenne muscular dystrophy

**DNA**: deoxyribonucleic acid

**ds**: double strand

**DSBR**: double strand break repair

**EBV**: Epstein-Barr-Virus

**EF1α short**: elongation factor-1 α short

**eIF2α:** eukaryotic translation initiating factor 2α

**ERF**: ethylene response factor

**eSpCas9**: enhanced SpCas9

**F**: Fusion protein (measles virus)

**F49A-FTP**: Fusion-toxin protein: variant F49A of CX3CL1 fused to the catalytic domain of *Pseudomonas exotoxin* A

**FAH**: fumarylacetoacteate hydrolase

**FKBP**: FK506 binding protein 12

**FOI:** force of infection

**FRB**: FKBP rapamycin binding domain

**GATA2**: GATA binding protein 2

**gB**: glycoprotein B

**gC**: glycoprotein complex

**G-CSF**: granulocyte-colony stimulating factor

**GCV**: Ganciclovir

**GFP**: green fluorescent protein

**gH**: glycoprotein H

**gL**: glycoprotein L

**GLN**: glutamine

**GLP-1**: glucagon-like-peptide 1

**GLU**: glutamic acid

**GLY**: glycine

**gO**: glycoprotein O

**gRNA**: guide RNA

**GSC**: glioblastoma stem cells

**GUIDE-Seq**: genome-wide unbiased identification of DSB enable by sequencing

**GVHD**: graft-versus-host-disease

**GVT**: graft-versus-tumour effect

**H**: hemagglutinin

**HAART**: highly active antiretroviral therapy

**Has-miR-200b**: homo sapiens microRNA 200b

**HBV**: hepatitis B virus

**HDAC**: histone-deacetlyase

**HDR**: homologous direct repair

**HF-Cas9**: high fidelity Cas9

**HHV6**: human herpes virus 6

**HIV**: human immunodeficiency virus

**HLA**: human leukocyte antigen

**HPV**: human papillomavirus

**HR**: homologous recombination

**HSC**: hematopoietic stem cell

**HSCT**: hematopoietic stem cell transplantation

**HSV**: herpes simplex virus

**HT1**: hereditary tyrosinemia type 1

**HTGTS**: high throughput genomic translocation sequencing

**HTT**: huntingtin

**ICU**: intensive care unit

**IDLV**: integrase deficient lentiviral vector

**IE**: immediate early

**IEA**: immediate early antigen

**IFN**: interferon

**IL**: interleukin

***Indels***: small insertions and deletions

**iPSC**: induced pluripotent stem cells

**IR**: internal region

**ISG**: interferon stimulated genes

**iTOP**: induced transduction by osmocytosis and propane betaine

**iv**: intravenous

**JC virus**: John Cunningham virus

**kb**: kilo base

**KRAB**: Krüppel associated box

**LA**: late antigen

**LAcmvIL-10**: latency-associated viral cellular IL-10 homologues

**Lin**: lineage

**LTR**: long terminal repeats

**LUNA**: latency unique natural antigen

**LV**: lentivirus

**MAPK**: mitogen-activated protein kinase

**MCMV**: murine Cytomegalovirus

**MCP-1**: monocyte chemoattractant protein 1

**mDC**: myeloid dendritic cells

**MDDC**: monocyte derived dendritic cell

**MDM**: monocyte derived macrophage

**MHC**: Major Histocompatibility Complex

**MIE**: major immediate early

**MIEP**: major immediate early promoter

**min**: minimal

**MIP-b**: macrophage inflammatory protein b

**miRNA**: micro RNA

**MLC**: mixed lymphocyte cultures

**MMEJ**: microhomology-mediated end joining

**MN**: meganucleases

**MOI**: multiplicity of infection

**MT**: methyltransferase

**mTOR**: mechanistic Target of Rapamycin

**MVA**: modified vaccinia ankara

**NES**: nuclear export signal

**NFkB**; nuclear factor κ of activated B-cells

**NGS**: next generation sequencing

**NHEJ**: non-homologous end joining

**NIEP**: non-infectious enveloped particle

**NK-cells**: natural killer cells

**NLS**: nuclear localization signal

**nt:** nucleotide

**ODN**: oligonucleotide

**$ori_{Lat}$**: latent origin of replication

**$ori_{Lyt}$**: lytic origin of replication

**PADRE**: pan DR epitope

**PAM**: protospacer adjacent motif

**PBMC**: peripher blood mononuclear cells

**PC**: pentameric glycoprotein complex

**PCR**: polymerase chain reaction

**PD1**: programmed cell death protein 1

**PEI**: polyethylenimine

**PFA**: phosphonoformate (Foscarnet)

**pi**: post infection

**PI(3)K**: phosphoinositide 3-kinase

**PKR :** protein kinase R

**PML**: promyelocytic leukaemia protein

**PML-NB**: promyelocytic leukaemia protein nuclear bodies

**POU2F2**: POU domain, class 2, transcription factor 2

**pp65**: phosphoprotein 65

**pp71**: phosphoprotein 71

**PPARγ**: Peroxisome Proliferator-activated receptor γ

**Pre-crRNA**: precursor CRISPR RNA

**R:** Purine

**RFLP**: fragment length polymorphism analysis

**RGR**: gRNA sequence by self-cleaving ribozyme

**RNA**: ribonucleic acid

**RNP**: ribonucleoprotein

**RT**: room temperature

**SCD**: sickle cell disease

**SDSA**: synthesis dependent stand annealing

**SES**: socio-economic status

**SeV**: Sendia virus

**SFFV**: spleen focus-forming virus

**SLiCES**: self-limiting Cas9 circuit for enhanced safety and specificity

**SNHL**: sensorineural hearing loss

**SNP**: single nucleotide polymorphism

**SOT**: solid organ transplantation

**ss**: single strand

**ssODN**: single strand oligonucleotide

**STAT**: signal transducer and activator of transcription

**SV40**: Simian virus 40

**TALEN**: transcription activator-like effector nuclease

**R0:** basic reproductive number

**RANTES**: Regulated on Activation, Normal T Cell Expressed and Secreted

**tDNA**: target DNA

**Thy-1**: Thymocyte antigen 1

**TMEJ**: theta mediated end joining

**TMP**: trimethoprim

**TNFa**: tumour necrosis factor α

**TOPOIIb**: Topoisomerase II beta

**TPD**: third party donor

**TR**: terminal region

**tra:crRNA**: trans-activating CRISPR RNA

**tRNA**: transfer-RNA

**tru-gRNA**: truncated guide RNA

**UL**: unique long

**US**: unique short

**vIL-10**: viral IL-10

**VLP**: virus like particle

**VP16**: activator protein 16 of the Herpes simplex virus

**WT**: wild type

**Y:** pyrimidine

**YY1**: yin yang 1

**ZAL**: zwitter ionic amino lipids

**ZFN**: zinc-finger nucleases

# I Preface

My thesis's project was to develop an anti-viral CRISPR/Cas9 system to target the human Cytomegalovirus.

In this manuscript, before presenting my results, I give an overview of different pathologies associated with the human Cytomegalovirus infection. Further, I present the molecular characteristics of the virus: particle structure, replication cycle and state of latency. I focus on the involvement of the major immediate early gene, coding for the most essential proteins of the virus replication cycle, because it was chosen as one of our target genes for the CRISPR/Cas9 system. Moreover, I introduce the different approaches to treat Cytomegalovirus infections by the standard therapy based on nucleotide analogues and new arising treatment strategies. Finally, I illustrate an inside view of the CRISPR/Cas9 technologies, their molecular mechanism, problems and improvements, and applications options.

The main body of this thesis is separated in two hypotheses based on two different designs of the anti-viral CRISPR/Cas9 systems and their different target sites in the viral genome. Most of the past three years, I worked on the first hypothesis. Read out methods were established and stable cell lines were selected. A publication was written based on those results (currently in submission) and the manuscript is provided in the appendix in this thesis. In the third year, the second hypothesis was build up based on the progress in the field of anti-viral CRISPR/Cas9 applications and on our results on the first hypothesis. Preliminary results regarding this second hypothesis are also presented.

In this thesis, I show the proof of principle of our new designed anti-viral CRISPR/Cas9 systems to prevent the Cytomegalovirus replication. As these strategies were envisioned to be used to treat a cell suspension in the context of hematopoietic stem cell transplantation, I discuss at the end possible obstructions for clinical applications and their potential solutions. Finally, I propose new experiments, which would be needed to bring these anti-viral strategies to the preclinical testing and one day to clinical application.

Throughout this thesis, I designed small Boxes (List of Boxes, p.3), which point out the most important information necessary to follow my project.

# II General aim of this thesis

The human Cytomegalovirus is a typical opportunistic infection in immunocompromised patients. Especially during transplantation the patient is weakened by the immunosuppressive treatments. HCMV can then either reactivate from the latent reservoir of a HCMV positive recipient or can be transferred from a HCMV positive donor organ. Consequences range from gastrointestinal disease to severe pneumonitis and are associated with a high mortality rate. Standard anti-CMV treatments by nucleotide analogues can reduce the viral replication in the recipient and alleviate the CMV associated diseases. However, they cannot prevent the reactivation from latency neither clear the infection, because nucleotide analogues only target the lytic replicating virus and have no effect on the latent virus pool.

**Therefore, this thesis aims to develop a new anti-HCMV approach, which targets the lytic and latent viral pool and prevents the reactivation of HCMV from latency.**

Genome editing tools have been tested against chronic and latent viral infections since the discovery of zinc finger nucleases (ZFN). The advantage of viral genome modification over standard chemotherapy to treat viral infections is that mutations in essential parts of the viral genome abrogate the viral replication permanently or silence the virus in latency. Therefore, only one efficient treatment period is theoretically necessary to prevent further reactivations or replication. However, the older genome editing tools (MN, ZFN and TALEN) have only a limited efficiency to prevent viral infections and require time-consuming protein engineering procedures for their development. The newly discovered CRISPR/Cas9 system revolutionised the genome editing field and enables fast, cost-efficient and feasible applications. Several other viral infections, like HIV and HBV, have been efficiently targeted by the CRISPR/Cas9 system showing a higher efficiency than TALEN.

Consequently, we choose the CRISPR/Cas9 system as the optimal tool to target HCMV. In this thesis, we developed three anti-HCMV CRISPR/Cas9 systems based on two hypothesises. **In the hypothesis 1, I designed two anti-HCMV CRISPR/Cas9 approaches, which are targeting the essential major immediate early gene UL122/123 to prevent the lytic viral replication.** The UL122/123 gene is the first viral gene expressed after infection of a cell and regulates the initiation of all further steps of the viral replication cycle. To prevent the UL122/123 expression the viral gene was either targeted by a single gRNA closed to the start codon to induce a frame-shift mutation or with three gRNA to excise the UL122/123 gene. **In the hypothesis 2, I designed a CRSIPR/Cas9 system to target the viral genome in latency considering the chromatin condensation.** Two gRNA were selected targeting the virus in the latency associated gene LUNA essential for viral reactivation and in two homolog structural regions. The cleavage at those three target sites simultaneously might even destroy the viral genome and not just silence it in latency. We analysed the mutations introduced in the viral genome and the subsequent effect on the different stages of the viral replication cycle: MIE protein expression, genome replication, late protein expression and virion release.

Even though, CRISPR/Cas9 delivery methods do not yet allow the delivery to all latently infected cells *in vivo*, the *ex vivo* treatment of HSC suspension is already applicable. **Therefore, the first approach with our new anti-HCMV CRISPR/Cas9 system would be the treatment of the donor hematopoietic stem cells to prevent the reactivation of HCMV in the recipient.**

Further applications will open up soon with the rapidly progressing research on optimized delivery methods.

# III Introduction

## 1 Human Cytomegalovirus

### 1.1 Epidemiology

The human Cytomegalovirus (CMV) infection is one of the most common viral infections in the world. Although, the infection is usually asymptomatic in healthy adults, it can cause morbidity and mortality in neonates and immune-compromised patients (transplantation and AIDS). It has a seroprevalence of up to 100% in the human population with big variations between the high-developed western countries and less developed regions. Western Europe has worldwide the lowest seroprevalence as for example in Germany with 27,4%[1]. In southern Europe as well as in the USA the seroprevalence is slightly higher and varies between 50% and 77%[2–4]. The highest occurrence of CMV is in Africa (77%[5]-97.2%[6]), south America (80%-90%[7]) and Asia (70%[8]-99%[9]). This is also represented in the differences of the basic reproductive number (R0) in the different regions. While there is a worldwide R0 of 2.4[10], it ranges from 1.7[11]/1.8[12] in the USA and Australia, to 5.64[12]/5.7[13] in Asia and Brazil, respectively.

In general, the seroprevalence is rising with increasing age[14]. In developed countries for example USA, the increase is nearly linear with a force of infection (FOI, risk of infection per year) of around 2%[15]. Interestingly, two main peaks can be observed by detailed analysis of different age groups. At early childhood (<5 years old) the force of infection is the highest (FOI = 8%) and decreases to less than 1% during adolescence. The second peak is during child bearing age (20-40 years, FOI = 5%), which is correlated with risk factors for CMV infection like parity[8,12,18] and day care attendance[1,19] due to the high risk of CMV transmission by young children shedding CMV.

In population with an overall very high seroprevalence, this effect is less important, because the seroconversion usually happens in early childhood[3,5,6,16]. For example in China, 60.37% of the CMV infection occurs in the first years of life[12]. The seroprevalence increases till the age of 25 (97.3%) and then stagnated from the age of 40. Also, the FOI for the population of 1-25 years old is very high of 12.6%.

Moreover, CMV seroprevalence strongly depends on the ethnic background and place of birth. In Europe, it was shown that immigrants have an up to 30% higher infection rate[1,14]. Similarly, in the USA, the difference between the ethnic groups is also very high. By the age of 40, just 50% of the white Americans are CMV positive in comparison to 90% of the Afro-Americans and Hispanic Americans[3], which can be also measured on the differences of the basic reproduction numbers. In average, the R0 in the USA is 1.7 but is significantly higher in the Black Americans and Hispanic Americans (R0 = 4.1 and 3.7)[11]. This is probably strongly influenced by differences in the socio-economic status (SES) (high vs. low: R=1.6 vs. 2.7[11]) and living habits[21]. The SES of a family is closely related with the living conditions, hygiene standards, health care/insurance and close contact with potential CMV positive persons. All those conditions lead to a higher risk of CMV infection[1,16,18,19]. Till today there is no genetic polymorphism known which is connected with a predisposition for CMV infections[22].

## 1.2 Transmission

All the risk factors for HCMV infections mentioned above (chapter II-1.1) can be traced back in one or another way to the different routes of CMV transmission.

A seropositive mother can infect the fetus, when she acquires a primary infection during pregnancy or experiences a reactivation. The transmission rate by a primary infection is much higher (30% up to 50%)[23,24] than during reactivation or reinfection, because the maternal immunity protects the fetus from the virus (transmission rate 1.4%)[23,25].Since the primary infection is much more dangerous, several studies tested an education program for pregnant women to learn about contamination sources and special hygiene practices in contact with young children[26,27]. Indeed, young children are the most probable source of infection by shedding the virus through saliva and urine in the first years of life[28]. These education programs have shown a significant decrease of primary infections and are right now the only way to protect the mother from the infection, because there is no vaccine available[26,27].

Beside the intrauterine transmission, the infant can acquire HCMV at birth, due to HCMV in the genital/urinal tract and cervix (26-57%)[24] or during the first weeks of life by breastfeeding (63%)[24,29]. 74.1% of the women have CMV in their breast milk and 13.8% of them transfer the virus to their babies[30].

During Childhood, the most common infection route is close contact with other infected individuals. The highest risk factor is a seropositive sibling, but also day care attendance can give rise to the HCMV infection[14,31].

In adolescence and adulthood, CMV is also transmitted by sexual contact. The virus has been detected in the semen and vaginal secretions[32]. Various studies have described an impact of the number of sexual partners and the age of the first sexual intercourse on the seroprevalence[14,31]. Furthermore, the sexual transmission of CMV has been also shown in animal models[33].

Lastly, CMV infection by transfusion of blood or blood products is especially of concern for high-risk groups like low-weight-born babies, pregnant woman and transplant patients. Earlier those high-risk groups were provided with blood products from CMV negative donors. Nowadays, the common procedure is the usage of leukocyte reduced blood products[7,34], but there is so far no international recommendation for blood transfusion products or procedures with regards to the CMV status of the donor and the type of recipient. Several hospitals use dual safe products (CMV- and leucocyte depleted) for neonates and pregnant woman and single safe products for transplant patients.

## 1.3 Polymorphism, genotypes and viral strains

In the human population, HCMV appears not as a limited number of defined viral genotypes as known for other viral infection like HBV (8 genotypes), but more as a high variability of viral strains. Polymorphisms or defined genotypes have been described in around 32 genes of HCMV[35]. None of this gene variations are linked to each other so that wide range of viral strains are possible[36]. Moreover, there is a general worldwide distribution of all genotypes/viral strains and no geographical areas for the occurrence of specific viral strains is documented[35]. Investigations of viral strains or specific genotypes and their relevance for transmission and clinical outcome of CMV infection or reactivation are very controversial and did not yet show a clear high virulent viral strain[35,36]. In the main focus of those studies are the surface glycoproteins, important for viral

entry, and viral proteins regulating cellular immune responses[37–41]. In addition to that, most of this studies describe the occurrences of infection with more than one HCMV strain[35,40,42]. Numbers of this mixed infection vary strongly (15-90%) depending on the studied population and the sensitivity of the read-out method[35]. Still, most of these studies do not detect a more severe outcome due to mixed infections. Although, in solid organ transplantations, the infection with different viral strains lead to a more severe pathology, but this is more correlated with the lack of neutralizing antibodies against the donor viral strain than to a high virulent viral genotype[38,43].

## 1.4 Pathology

The primary infection of HCMV is usually asymptomatic in healthy individuals. Rarely, it can cause a mononucleosis syndrome with fever, myalgia, adenopathy and hepatomegaly[44]. HCMV causes more severe diseases in immunodeficient individuals like neonates, AIDS patients and transplant patients. This will be further discussed in the chapters below.

### 1.4.1 Congenital infection

Congenital HCMV infection (cCMV) is one of the most common intrauterine transmitted diseases with partially fatal outcome. In average, around 0.65% (0.1-13.6%[45]) of all live births worldwide have a congenital CMV infection[46,47] with differences between developed and developing countries as explained by the epidemiology mentioned above. Whereas, the prevalence in the western world is relatively low (0.3-0.7%), it is much higher (1-2% [outliner:6.1% and 13.6%]) in developing countries[23,46,47]. Most of the congenital infection are observed in seropositive women (non-primary infection). Indeed, even in the low seroprevalence countries such as in the USA, only 25-29% of the cCMV cases are caused by the primary infection, because the FOI is relatively low (see chapter II-1.1)[14,15]. In high seroprevalence regions, the FOI for women in the childbearing age is much higher. Similar observations can be made between different ethnic groups. In the USA, around half of white women are seronegative and have only a risk of primary infection of 1.38% per pregnancy, while black and Hispanic women have a seroprevalence of already 83%, but the CMV negative woman are exposed to a higher risk of primary infection of 3.4-3.85% at child bearing age 20-49 years[11] as compared to the white woman. This difference is even stronger for adolescence females (12-19 years) with a risk up to 7.33% for black Americans (seroprevalence 58%) and a risk of only 0.15% for white Americans (seroprevlaence 39%). Finally, the same number of women in both groups will experience a primary infection during pregnancy. The influence of reinfection *vs* reactivation is not in details assessed so far, because most of the diagnostics do not distinguish between reactivation and reinfection. Probably, the reinfection is an important factor especially in high seroprevalence regions[13], because the maternal immunity cannot protect the foetus from the reinfection with a different viral strain.

The outcome of the congenital infection can be very dreadful and in the worst case it can lead to a stillborn child or abortion (1.1%)[48]. Usually, 85-90% of the children are born asymptomatic, but are prone to develop sequelae[49,50] such as sensorineural hearing loss (SNHL, 10%), microcephaly (5%) and chorioretinitis (2%)[51]. Children, which are symptomatic at birth, face a more severe outcome[47]. Their survival is lower (5% mortality) and the long-term effects are worse[51]. The clinical manifestations at birth can be clustered in four categories: cerebral abnormalities, non-cerebral abnormalities, hepatobiliary abnormalities and other physiological defects. Nearly all infants have cerebral abnormalities[52,53]. Further physiological signs are low birth weight[54], prematurity (34%) and petechia (79%).

Most of the symptoms are direct consequences of the viral replication in the fetus and placenta. Several non-structural viral proteins inhibit apoptosis, organogenesis, fetal development and growth[46,55]. For example, IE2 and US28 induce smooth muscle cell migration, which is leading to uncontrolled inflammation and proliferation, which narrows the blood vessels[46]. This causes vascular injuries and consequently hypoxia and brain damage. Other viral proteins decrease the protection against the maternal immune system[56]. Furthermore, HCMV-infected neuronal cells have an increased cellular trans-activator Peroxisome Proliferator-activated receptor γ (PPARγ) expression, which inhibits directly neuronal development, as in the brain of a congenital infected foetus[57]

Beside these severe symptoms at birth, 30-90% of the symptomatic congenital infected children will also later on develop sequelae[54,58], which are usually more severe than in asymptomatic children. The majority suffers from mental retardation sometimes associated with strong seizures and the inability to speak or walk[59]. Also very common is SNHL, ocular damage, microcephaly and motor defects[51,60]. In the development of SNHL, HCMV replicates in specific areas in the inner ear, like the stria vascularis and the reissner membrane, which collapses following the infection[61,62]. The spiral ganglions are also infected and destroyed[63]. The viral replication induces inflammation, which leads to tissue damage and cell loss [63]. It also causes an imbalance of the ions potential in the ear, which is important for the signal transduction. The cochlear transplant can improve the languages and speech perception again, but most of the children will still have a lower languages score due to the further CNS impairment as descript above[64].

However, not only the active viral replication in the inner ear can lead to the SNHL. The infection of the fetus early in development induces chromosome breakage at two specific loci, 1q42 and 1q21, which are both related to hearing loss and ocular damage[65].

In conclusion, a congenital HCMV infection is a severe issue for newborns especially when they are born with symptoms. Risk factors for symptomatic congenital HCMV[46] are high viral load in the amniotic fluid[51], gestation age at time of infection (first trimester)[51,64], primary maternal infection[51], the maternal immune status, HCMV induce placenta damage and the type of HCMV strain[24]. Till today there is no vaccine available and the standard anti-HCMV treatment (Ganciclovir) cannot be given to pregnant women due to its teratogenic adverse effects (Cymevene®, Roche, data sheet).

### 1.4.2 CMV in HIV patients

CMV is one of the most common opportunistic infection in human immunodeficiency virus (HIV) patients. Nearly all HIV patients are CMV positive due to similar transmission routes[66,67]. CMV usually reactivates in later stages of the acquired immune deficiency syndrome (AIDS)[68] and increases HIV viral loads. Indeed, the immunosuppression favors CMV replication and both viruses can transactivate each other[66]. The CMV reactivation further increases the immune suppression and is a risk factor for AIDS progression and death[66,69]. Furthermore, there is a broad range of CMV diseases (CMVD) in HIV patients. The most common ones are CMV retinitis (85% of the CMVD) followed by gastrointestinal disease (18% of CMVD), neurological diseases (1% of the CMVD), pneumonitis and adrenalitis[67,70–72]. CMV retinitis is a painless progressive pathology leading to vision impairment, but it is not fatal. However, it is correlated with HIV encephalitis[72] and increased mortality[73]. Most severe are the neurological CMV diseases, like necrotizing ventriculoencephalitis with strong neurological symptoms like seizures, confusion, dementia and an average survival of 42 days.

Since the introduction of the highly active anti-retroviral therapy (HAART) for HIV in the mid-1990s, the occurrence of CMV disease strongly declined by 80-90% in the western countries[67,74,75]. In developing countries, this effect is not observed[76,77]. This is probably due to the diagnosis of AIDS at later stages and less control on the development of HCMV resistances to the treatment. Moreover, due to the high cost of the HAART therapy less patients are treated efficiently to prevent AIDS progression. Usually, HAART leads to the immune recovery in around 50% of HIV patients[73] including CMV specific immunity. CMV retinitis heals spontaneous without specific CMV treatment[78,79]. The occurrence of other CMV diseases under HAART treatment is now very low[75]. Relapse of CMV is just observe by a failure of immune recovery or resistances against HAART[73].

### 1.4.3 CMV in intensive care unit patients

CMV infections are getting more and more attention in non-immunocompromised critically ill patients in intensive care units (ICU). Around 15-33% of the ICU patients have a CMV reactivation[80,81]. Even though, it was firstly described more than 20 years ago[82], it is still controversially discussed if CMV is a pathogen causing disease or only a bystander as a result of the underlying disease[83]. The main consensus is that the CMV reactivation is correlated with a prolonged hospitalization and ICU stay[80,84], as well as prolonged mechanical ventilation of the patients[80,84]. Several studies also observed an increased overall mortality[84,85] in CMV reactivated patients and a higher sequential organ failure assessment score[83,86].

The main site of reactivation is the lung, which leads to increased lung fibrosis[87,88] and pneumonitis[89] often associated with respiration failure and mechanical ventilation[80,83]. Further CMV-induced diseases in ICU patients are hepatitis and colitis[83,90]. Moreover, it increases the risk of a bacterial or fungal superinfection[80,82,83], which are often the cause of mortality in ICU patients.

Up to now, the mechanism of reactivation in ICU patients is not completely understood. Due to their critical illness, ICU patients have an altered immune response with impaired T- and NK-cell function[82,83]. This favors the HCMV reactivation[83]. In addition, there are several factors like mechanical ventilation[91], previous lung injury[88], blood transfusion[86,91], enteral feeding[80] and previous corticosteroid treatment[80,83], which increase the risk of reactivation. Moreover, bacterial superinfection and sepsis, which are common in ICU patients, are also connected with CMV reactivation [81,88].

The standard treatment of CMV is Ganciclovir (GCV), but there are no clinical trials based on the effect of GCV treatment on ICU patients. In a mouse model it was shown that only prophylactic treatments have a beneficial effect but not preemptive treatments after CMV diagnosis[88,92]. The benefice-risk balance cannot yet be estimated, but the side effects of GCV might be problematic in this patients group. Up to now, there is no recommendation for the treatment of HCMV infection/reactivation in ICU patients.

### 1.4.4 CMV in glioblastoma

In 2002, Cobbs *et al*. discovered that nearly 100% of glioblastomas, a very common and fatal brain tumor, are positive for the CMV antigens IE and pp65 as well as for the viral genomes[93]. In the following years, the presence of CMV in glioblastomas has been controversial[94,95], whereas more recent publications agree with Cobbs first findings[96–98]. CMV antigens and replication are mainly detected in grade IV glioblastomas[96] and can be also associated with a viremia[98]. CMV is only detected in the tumor itself, mainly in tumor/glioma stem cells (GCS)[99,100] and macrophages/microglia[101], but not in the healthy surrounding tissue[93,96].

Whereas the presence of CMV seems to be confirmed, the consequence of a CMV infection in this situation is still unclear. CMV is usually a non-oncogenic virus, so it is unlikely to cause the transformation of healthy tissue into tumor cells. Especially, since there is no correlation between the seroprevalence of CMV and the occurrence of glioblastoma[102]. It is more likely, that CMV has oncomodulatory functions and influences tumor progression and metastasis. Several viral gene products (like immediate early (IE) proteins, viral IL-10, glycoprotein B (gB) and the constitutive active G-protein coupled receptor US28) have been shown to influence different hallmarks of tumorigenesis[99,103–105]. They can induce the stem-like phenotype of tumor cells[100,105], influence tumor-suppressor genes[103,106], mitogenesis and invasiveness of tumor cells[107], and favor the immunosuppressive microenviroment[101].

Consequently, anti-CMV treatment could be beneficial for the survival of the patient and decrease the tumor progression. Indeed, the usage of GCV for glioblastoma patients increases the 2-year survival rate from 18% to 62-90% depending on the duration of the treatment with an overall survival prolongation of 33 months[108]. More recently, there are several clinical trials ongoing examining the possibility and efficiency of cell-therapies[109]. Adoptive T-cell therapies based on CAR-T-cells or CMV specific T-cells are currently evaluated and show an overall survival of 403 days[97,110]. Another option is the use of CMV antigen loaded autologous dendritic cells (DC) to induce a potent immune response against the CMV positive glioblastoma. This DC based therapy prolongs the overall survival to 40 months[105].

### 1.4.5 CMV in transplantation

Over 50 years ago, it has been observed for the first time that CMV reactivates in transplant patients, which leads to severe organ invasive diseases and non-transplant related mortality[111]. Therefore, CMV was described later as the « Troll of transplantation, who takes its toll » by Balfour[112]. He compared the old fairytale of “the three Billy goats Gruff" i.e. as the patients cross the bridge of transplantation; where the troll (HCMV) would take one out of three patients as toll while leaving the other crossing the bridge. This nickname of CMV catches on till today, not least because it did not lose its importance.

Approximately 50% of the transplant patients experience at least one CMV reactivation or primary infection in the first year after the transplantation[44]. HCMV has different features in solid organ transplantation (SOT) and hematopoietic stem cell transplantation (HSCT), which will be discussed separately. Since HSCT is more important for this thesis, I decided to focus more on this second part.

#### 1.4.5.1 Solid organ transplantation (SOT)

Briefly, the majority of SOT patients with CMV infection develops the so called CMV syndrome[113], characterized by at least two of the following symptoms: fever, malaise, 5% atypical lymphocytes, thrombocytopenia and elevated hepatic transaminase level[114]. Furthermore, 10-50% of those patients suffer from an organ-invasive CMV disease (end organ disease) in the early phase after transplantation (< 1 year)[113,115]. Around 20% of the patients show a late CMV disease such as gastrointestinal diseases (55%) and pneumonia (14%)[115,116]. Less common end organ diseases include hepatitis, chorioretinitis, nephritis, cysteitis, myocarditis and pancreatitis[44]. Beside those direct effects, HCMV induces some indirect effects like increased risk of graft rejection and ateriostenosis[44,116]. In addition, CMV infected SOT patients are more susceptible to bacterial and fungal superinfections[117]. All those factors lead to a higher mortality[118].

It is not exactly known what triggers CMV reactivation and its progression to the CMV disease, but the immunosuppressive treatment has a big impact on this. The balance of immunosuppressive drug to prevent graft rejection and preserve enough immune response to fight CMV is very difficult. The common immunosuppression with cyclosporine A or mycophenolate is correlated with a higher risk of CMV reactivation[119]. mTOR inhibitors, the new type of immunosuppression, can be more beneficial to optimize the balance[120]. They have direct anti-HCMV effects, which reduce viral replication, but also guide the T-cell response away from anti-graft rejection to anti-HCMV response[119]. Furthermore, several other risk factors are observed: previous graft rejection[121], diabetes mellitus[122], age[122] and HHV6 co-infection[123]. Moreover, heart transplantations[122] are connected with a higher risk of CMV disease. The most important factor is the serostatus of the recipient and donor. For example, CMV negative recipients are more prone to acquire a CMV infection from the donor organ than a CMV positive recipient[121]. Even in a seropositive recipient, it is much more probable (86%) that the donor organ causes the CMV infection than the virus reactivating from the recipient[124]. CMV positive recipients who receive a CMV negative donor organ have the lowest risk.

#### 1.4.5.2 Hematopoietic stem cell transplantation (HSCT)

In HSCT, the hematopoietic stem cells of the recipient will be replaced by the donor stem cells (or by their own stem cells = autologous). Therefore, the immune system of the recipient gets depleted and reconstituted by the donor cells. During the reconstitution, those patients are unprotected against pathogens and have an impaired immunological memory. The recipients are under special risk for viral reactivation (CMV, EBV, HSV and more) till the immune reconstitution by the donor stem cells is complete. This can take up to two years post transplantation[125].

In general, one third up to one half of the HSCT recipients experience CMV infection at some point after transplantation[126,127]. With the discovery of anti-CMV drugs in the 1980ies[128,129], the incidence of CMV disease of around 20-30% was halved[130]. Earlier, the highest occurrence of CMV reactivation and disease was in the early-onset after transplantation (<100 days) due to the lack of immune cells. The treatment with Ganciclovir and Foscarnet reduced the early CMV diseases to less than 5%, while there is an increase of late-onset CMV diseases (>100 days post transplantation) to around 15%[125] (Figure 1). The different treatment strategies are explained in more details in Chapter II-1.6.1 "*Approved anti-viral chemotherapy*".

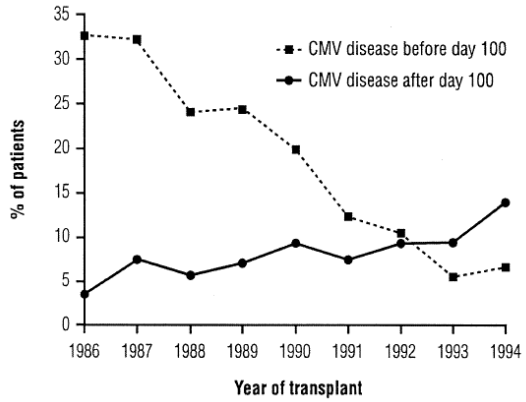



**Figure 1: Development of early *versus* late CMV disease in seropositive hematopoietic stem cell transplant recipients. The antiviral therapy with Gancicolvir or Foscarnet started at 1986 and this diagram represents the effect on all HSCT recipients regardless of antiviral treatment. The start of the administration of the GCV and Foscarnet strongly decreases the occurrence of early-onset CMV disease, but leads to a slight increase in late-onset CMV disease. Figure from: Boeckh et al. 2003, *Biol Blood Marrow Transplant*.**

Even with the treatment today, there is still a 5% risk to develop CMV disease. The early-onset of CMV diseases are mainly gastrointestinal diseases, which are usually not fatal[131,132]. Gastrointestinal diseases are also dominant in the late-onset. However, almost 17% of the recipients develop the more severe CMV pneumonia[44,133], which is related to a high mortality of 60-85% even under anti-viral therapy[134,135]. Less common CMV diseases are retinitis (0,2%)[136] and encephalitis[137].

The development from CMV reactivation to disease is associated with some risk factors. Most important is the viral load[126,138], which is therefore used as a main factor to start anti-viral therapy. Several other graft-related factors can increase the risk of CMV disease like allogeneic, related (haploidentical) or unrelated HLA-mismatch grafts[127,139] as well as graft versus host disease (GVHD)[127,131]. If the source of the transplant cells (cord blood, peripheral G-CSF mobilized blood or bonw marrow) has an influence on the occurrence of CMV disease is controversial[140]. Even though some groups described a higher incidence of CMV disease with stem cells from cord blood[141,142], which could be related to an delayed $CD8^+$ T-cell reconstitution from cord blood in comparison to other sources[143]. In general, the reconstitution of the immune system is strongly correlated with the control of CMV reactivation. Patients with lymphopenia are more likely to develop a symptomatic CMV infection[127,144]. This is due to the lack of CMV-specific T cells, which are essential to control the CMV reactivation[145]. CMV-specific T cells are present in the recipients in average of 5.5 weeks after transplantation. However, it takes up to two years to reach a level of a normal healthy person[146]. Patients with a delay of the start of T-cell reconstitution are more prone to develop a CMV disease[145,146]. It has been observed that the treatment with GCV and Foscarnet leads to a delay immune reconstitution[125,147]. To overcome the lack of T cells in the first weeks after transplantation, a non-myeloablative regimen can be applied before transplantation, where the recipient T cells are not depleted and are present till 80 days post transplantation. Those T cells provide protection for that early time frame after transplantation until the donor stem cell reconstitute the T cells in the recipient[125,148].

The serostatus of the recipient and the donor organ also influence the probability of CMV infection/reactivation and the development of a CMV disease[133,149,150] (Figure 2). Recipients, who are already CMV seropositive at the time of transplantation, are more likely to reactivate CMV and also have a higher risk of CMV disease. Only 19% of the stem cells from a CMV positive donor will transmit the virus to the recipient to induce a primary CMV infection[151]. Moreover, those patients infected by the donor cells are less prone to develop a CMV disease[149,152]. Beside this, there is an overall risk of 3% to acquire a transfusion-transmitted CMV infection by the blood products used during the transplantation procedure[153].

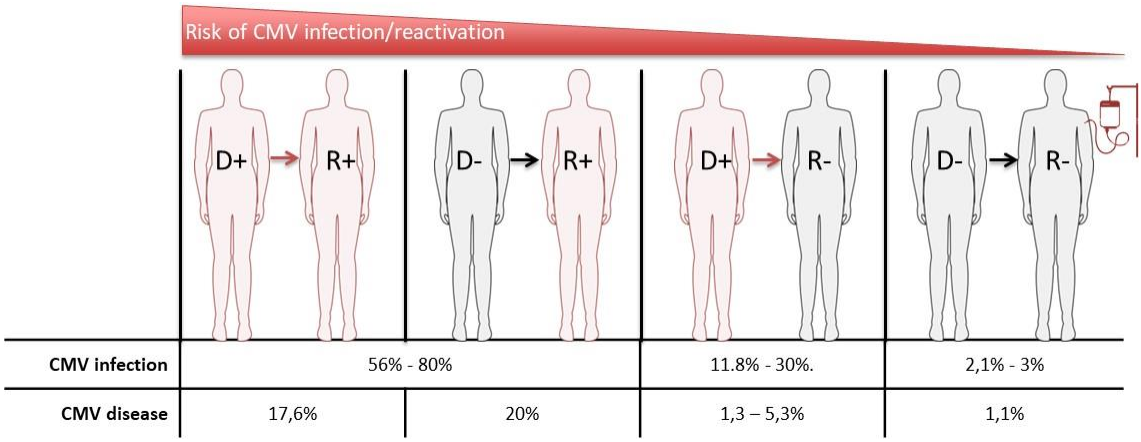



**Figure 2: Overview of the relation of the HCMV serostatus and reactivation. In HSCT, the serostatus is one of the major risk factors of HCMV reactivation. Seropositive recipients are at the highest risk that their own virus reactivates, because of the depletion of a protective immune system. The transmission of HCMV from the donor cells to the recipient is also possible but less likely. Finally, even in a complete naive donor and recipient setting there is still the risk to acquire HCMV by blood transfusion. CMV infection: viremia without symptoms (reactivation or primary infection), CMV disease: viremia with symptoms.**

Finally, the CMV disease is not the only problem induced by the CMV reactivation. CMV replication is correlated with graft failure[126] and makes the patients more susceptible to bacterial and fungal superinfection[154,155]. In conclusion, CMV reactivation/infection increases the non-relapse mortality by causing CMV pneumonia and enables bacterial and fungal infections[151,155]. On the other hand, it has to be mentioned that CMV reactivation can also have a positive effect: the so called virus-versus-leukemia effect, whereas CMV decreases the relapse of the underlying disease[155,156].

***Box 1: HCMV in HSCT***

| *Risk factors:* | *Outcome:* |
|---|---|
| • *HCMV seropositive recipient*<br>• *HLA-mismatch of donor and recipient*<br>• *Myeloablative regimen / T-cell depletion*<br>• *Immunosuppressive treatment* | • *Pneumonia with a high probability of mortality*<br>• *Gastrointestinal disease*<br>• *Graft failure* |

## 1.5 Molecular Characteristics of HCMV[1]

### 1.5.1 The viral particle

The human Cytomegalovirus is a typical Herpesvirus. It belongs to the subfamily of *β-herpesvirinae* and the genus of Cytomegalovirus. HCMV is also known as the Human Herpesvirus 5.

The viral particle has a size of around 150-200 nm with the typical structure of all herpesviruses: core, capsid, tegument and envelope.

The core contains the viral genome, which is around 230 kb of double-stranded DNA wrapped around a fibrillar protein. The viral genome consists of two unique regions, $U_L$ (long) and $U_S$ (short), which are separated by terminal or internal repeats (TR or IR). Those repeats consist of **a**, **b** and **c** sequences. The **b** sequences are flanking the $U_L$ region, while the **c** sequences are flanking the $U_S$ region. They are separated by several repeats of the **a** sequences, which are important for circularizing of the genome inside the cell and the cleavage of the viral genome for packing in the viral particle[157]. There are four isomers of the viral genome possible, which are based on rearrangements in the TR/IR regions with different orientation of the unique regions (Figure 3). All isomers are infectious[158]

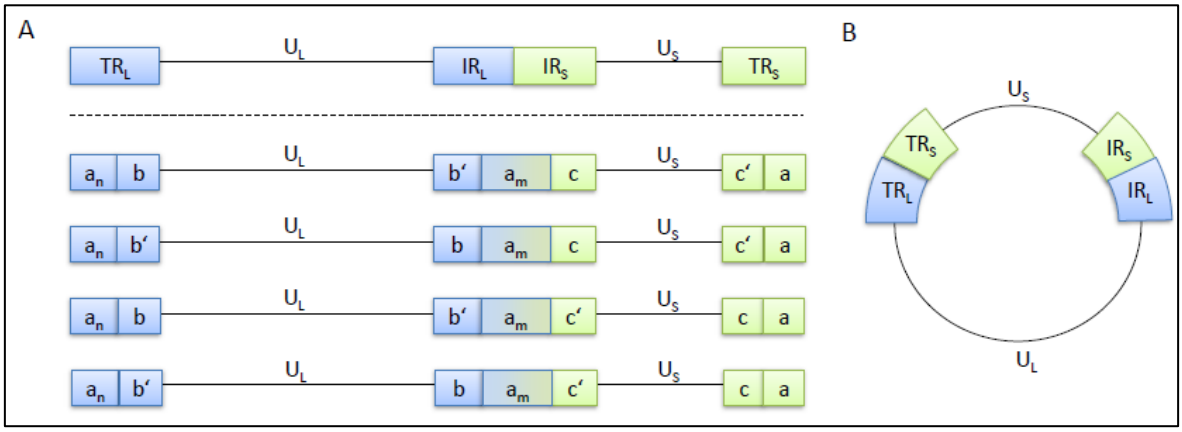



**Figure 3: HCMV genome structures. A) Schema of the linear viral genome inside the viral particle consisting of the unique regions, $U_L$ and $U_S$, separated by the terminal and internal repeats. Detailed representation of terminal and internal repeat structure and the four different isomers depending on the orientation of the $U_L$ and $U_S$ region. B) Schema of the episome as it is present inside the cells. It is circularized at the terminal region. $U_L$: unique long region; $U_S$: unique short region; TR: terminal repeat; IR: internal repeat**

The core is coated by the capsid. It is made out of five viral proteins. The capsid shell is icosahedral. The major capsid protein (UL86)[159] assembles into hexons (150) and pentons (12) giving this structure. It is hold together by a triplex structures consisting of the minor capsid protein (UL85) and the minor capsid binding protein (UL46)[160]. Finally, the smallest capsid protein (UL48-49)[161] binds the hexons. A protease (UL80a), located in the inside of the capsid, is important for the maturation of the capsid.

The capsid is surrounded by a layer of viral and cellular proteins, so called tegument. It is the biggest part of the viral particle containing around 60 viral and 70 cellular proteins[162]. Some of those proteins are essential for viral entry, gene expression, immune evasion, assembly and

[1] If not noted otherwise, the books: Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis[540] and Molekulare Viologie[541], are used as source.

egress[162]. To name some of the most essential tegument proteins: pp71 is essential to induce the immediate early gene expression; pp65, the most abundant tegument protein, prevents the presentation of the immediate early proteins through major histocompatibility complex I (MHC-I) and pUL32 is essential for the tegument formation and the capsid transport to the nucleus in the infected cell. The cellular proteins are mainly parts of the cytoskeleton, translational control, vesicle trafficking and signal transduction[163].

Finally, the capsid is enveloped by a membrane, which contains up to 9 glycoproteins. The three main complexes are gC-I, gC-II and gC-III[164]. gC-I, a heterodimer of gB, and gC-II, a heterodimer of gM/gN[165] are responsible for the binding on the cellular receptor to mediate viral entry[166,167]. The third glycoprotein complex is a heterotrimer consisting of gH/gL/gO, which induces the fusion of viral and host membrane[168,169]. Alternatively, gH/gL can built a complex with UL128-131, which then alters the receptor specificity and cell tropism (more detailed II-1.5.2).

During the replication of HCMV in cell culture, the infected cell releases two other types of particles: dense bodies and non-infectious enveloped particles (NIEP), which early assumed to decoy the immune system to protect the virions[170], but today researchers estimate is rather as an artifact of HCMV amplification in cell culture. Dense bodies are enveloped particles with some viral glycoproteins on the surface and contain some tegument proteins, dominantly pp65. They do not have a capsid or viral DNA. However, NIEPs have a normal envelope and tegument. They also contain a premature viral capsid, which still contains the scaffold protein (UL80.5) instead of the viral genome. Both particles are not infectious.

### 1.5.2 Replication cycle

The human Cytomegalovirus has a broad cell tropism for the usually highly specific herpesvirus family. It can infect fibroblasts, endothelial and epithelial cells, monocytes and macrophages, dendritic cells, hematopoietic progenitor cells, smooth muscle cells, stroma cells, neuronal cells, neutrophils and hepatocytes. The cell tropism is defined by the host cell surface receptors and some intracellular factors, but also by the viral glycoprotein complexes. For example, the ratio of gH/gL/gO trimer to the gH/gL/UL128-131 complex defines, if the virus can infect fibroblast or epithelial and endothelial cells[171].

In most of this cell types the virus enters the cell and initiate the lytic replication cycle, which can be divided in several steps: entry, immediate early phase, delay early phase, genome replication, late phase, assembly and egress (Figure 4). The first attachment to the cell is mediated by the interaction of gB to the common surface protein heparansulfate proteoglycans[166,167,172]. It is hypothesized that this is followed by the interaction with a cell type specific receptor. For example, DC-SIGN serves as the attachment receptor for the infection of dendritic cells by binding gB[173]. Another possible receptor is the platelet-derived growth factor-α, which is commonly expressed on fibroblast and epithelial cells, but also highly upregulated in glioblastoma cells[103,174]. PDGFR-α binding by gB and subsequent activation of the phosphoinositide-3-kinase pathway (PI(3)K) is essential for viral entry and stimulated angiogenesis and mitogenesis in glioblastoma. The receptor binding induces the internalization of the viral particle and the release of the capsid in the cytoplasm. The process is cell type dependent and mainly mediated by the gH/gL complexes[175]. Furthermore, gB has a fusogenic function, too, if it undergoes S-palmitoylation during the virus assemble and egress[176]. The viral envelope can fuse with the plasma membrane of fibroblast[177]. To enter endothelial and epithelial cells, it is internalized by endocytosis which

requires a low pH to fuse with the endosomal membrane[177]. Dendritic cells pick up the viral particle by macropinocytosis and release the capsid in a pH independent way[178].

Once the capsid enters the cytoplasm, it gets transported via microtubules to the nucleus. When the viral genome arrives in the nucleus it circularizes and forms the episome. This is followed by three stages of gene expression, the immediate early (IE), delay early (DE) and late phases[179]. In the immediate early phase, the tegument proteins of the virus transactivate the first set of genes. Those immediate early antigens (IEA) are mainly strong transactivators, which induce the expression of delay early antigens (DEA). They also push the host cell in the S-phase of the cell cycle to favor viral replication.

The DE phase starts around 8 h post infection (pi) and requires the IEA expression. The main DEA are viral proteins necessary to build the genome replication complex. It consists of the viral DNA-Polymerase UL54, the processivity factor UL44, the helicase-primase-complex UL105/UL70, DNA binding protein UL57 and the $ori_{Lyt}$ (origin of replication, lytic) binding protein UL84. The viral polymerase is similar to the human polymerase, it has a 5'-3' elongation function and a 3'-5' exonuclase function necessary for proof reading. The genome replication start at the $ori_{Lyt}$ [180]. The exact mechanism for the initiation of the genome replication is not known, but dependent on the interaction of UL84 with the major immediate early protein 2 (IE2)[181]. Once the replication complex is formed, the genome replication occurs by the rolling circle mechanism. This leads to the formation of a long concatemer, containing several copies of the viral genome in one long strand (Figure 4 IV).

After the genome replication, the late phase of gene expression is initiated. It includes the production of structure proteins, which induces directly the assembly of new virions (Figure 4, VI). The capsid proteins are transported to the nucleus, whereas the major capsid proteins need to be associated with the scaffold protein to enter the nucleus. This association also induces the generation of the hexon and penton structures, which then build with the triplex the premature capsid. DNA packing is induced by the viral protein UL57 and UL89, which interact with the **a** sequence in the TR region. They guide the viral genome through the portal UL104 in the capsid and cleave the concatemer at the end of each copy of the viral genome[182].

The viral egress is a process of primary envelopment, de-envelopment and second envelopment[182,183] (Figure 4,VII). The mature capsid leaves the nucleus by budding from the nuclear membrane. This premature virion fuses with the outer nuclear membrane and the mature capsid is released in the cytoplasm. There, it acquires the tegument. This structure is only partly defined. The capsid interacts directly with UL36 and there is a second layer of UL37. Further proteins are associated in an unknown way. Finally, the capsid/tegument buds in the trans-Golgi network by binding to the cytoplasmic tail of some viral glycoproteins. The mature infectious virion is released through the secretory pathway.

Another characteristic of HCMV is the formation of syncytia. The expression of viral glycoproteins, on the cell surface can induce cell-cell fusion, mainly mediated by the gH/gL complex[168], and transmit the mature capsid like that to the neighbor cells dependent on gB[184].

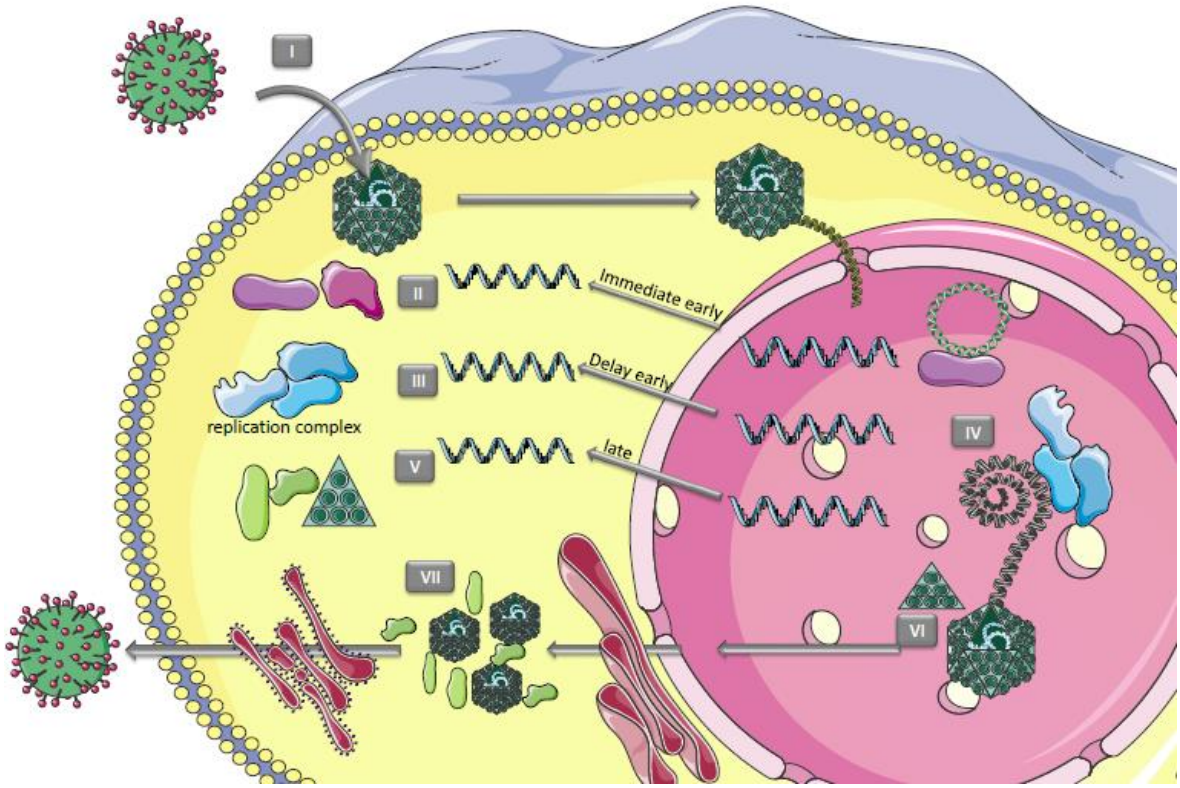



**Figure 4: Lytic replication cycle. I entry: The virus enters the host cell by binding to a receptor on the cell surface (e.g. heparanesulfate proteoglycans, integrin) and mediate the fusion of the cellular membrane with the viral envelop via the viral glycoprotein complex gH/gL. The capsid is released into the cytoplasm and transported to the nucleus. This enables the viral genome to enter the nucleus and circularize to the episome. II immediate early phase: The expression of immediate early proteins, which are strong transactivators regulating the further step of the viral replication cycle, is initiated by viral tegument proteins. III delay early phase: The expression of delay early proteins is induced by the major immediate early genes IE1 and IE2. IV genome replication: Those delay early proteins are viral kinases and polymerases forming the replication complex to perform viral genome replication by the rolling circle mechanism. V late phase: The expression of structure proteins (capsid, tegument, glycoproteins) is also regulated by the major immediate early proteins. VI assembly: The capsid proteins enter the nucleus and form of new viral capsids internalizing a copy of the viral genome. VII egress: The mature capsid is released into the cytoplasm by a process of budding and fusion with the nuclear membrane and endosomal compartment. In the cytoplasm the tegument proteins interact with the capsid and then induce a final budding in trans-Golgi network to be release through the secretory pathway.**

### 1.5.3 Major Immediate Early proteins

HCMV expresses several proteins in immediate early kinetics. Those genes are expressed without viral de novo protein synthesis necessary[185]. The two major immediate early proteins (MIE) are transcribed from a shared locus *UL122/123* and are called IE1 (UL123) and IE2 (UL122). Their expression is regulated by the major immediate early promoter (MIEP). It consists of a proximal and distal enhancer with several cis-acting elements[186]. This cis-acting element can be activated by several cellular factors like NFκB, CREB/ATF and AP-1 and the viral proteins IE1 and pp71. In fact, the first induction of the MIEP is ensured by the tegument protein pp71. pp71 induces the degradation of DAXX, a cellular factor, which recruits histone deacetylases (HDAC), and prevents the repression by deacetylation of the MIEP[187]. Later on, IE1 activates the MIEP in a similar manner[188,189]. At later time points of the replication cycle, IE2 can inhibit the MIEP by binding the a crs-acting element right in front of the transcription start[190,191]. The activation of the MIEP is the key step to the lytic replication cycle and depends on cell type and cell differentiation state.

Many studies have shown that IE1/IE2 are the most important transactivators of the virus and are indispensable for the viral replication[192–195] (Box 2). IE2 is the strongest transactivator. It binds on various cellular and viral promoters and induces their activation[196]. Its main function is the induction of the DEA expression. For example, the expression of the viral polymerase UL54[197,198] as well as UL44[199] and UL112[200] depends on the expression of IE2 usually in combination with some cellular factors. Without IE2 expression the viral replication cycle is abortive and no new virions will be produced[193,201]. On the other hand, IE1 is only essential at infections with a low multiplicity of infection (MOI)[202,203]. It positively autoregulates its own promoter and ensures the expression of IE2[194]. Therefore, IE1 works as an amplification system for the essential IE2 expression. Additionally, IE1 can inhibit HDACs, which results in the acetylation and activation of different IE and DE promoters[204].

IE1 and IE2 not only regulate the viral replication itself, they also have various effects on the cell cycle. Both MIEs are known to inhibit apoptosis[205] and to induce the progression to the S-phase of the cell cycle[206,207]. Moreover, IE1 and IE2 have a strong influence on the immune response of the cells[208]. IE1 blocks the induction of interferon-stimulated genes (ISG) by IFN-α. It complexes the mediators STAT1 and STAT2 and prevents the transactivation of ISGs[209]. Additionally, IE2 inhibits the expression of IFN-β, RANTES and other pro-inflammatory cytokines and chemokines[210]. On the other hand, IE1 induces the expression of interleukin 2 (IL-2) and IL-6[208]. Both cytokines are known to be associated with graft rejection. IE1 can even counteract the immunosuppressive drug cyclosporine A, which leads to the induction of IL-2[211]. IL-2 is involved in the expansion and activation of T cells and can increase their anti-graft activity.

The *UL122/123* consists of 5 exons. IE1 and IE2 have the first 3 exons in common and are alternatively spliced to either exon 4 or exon 5, respectively[212,213]. The start codon is localized in exon 2. Beside those two major IE proteins, there are several splice variants expressed (Figure 5). They are generated by different splicing sites and have partly distinct function from their full-length proteins. There are four main isomers of IE2 (IE55, IE60, IE40 and IE18[214]). IE55 is also expressed at immediate early time point during replications. It counteracts IE2 and activates the MIEP[215]. Moreover, at late time points during viral replication the isomers IE60 and IE40 are expressed. They lack the commonly used exons 1-3. Like IE2, they are also essential for the viral replication. IE60 interacts with UL84 and might influence the genome replication[216] and IE40 is important for late antigen expression[217,218]. For IE1, there are two main isomers known. IE19 counteracts its full-length protein and inhibits the MIEP[219]. Furthermore, there is a IE1x4 isomer, which is mainly expressed in latency and ensures the maintenance of the viral genomes[220], which will be closer discussed in the next chapter (II-1.5.4 latency).

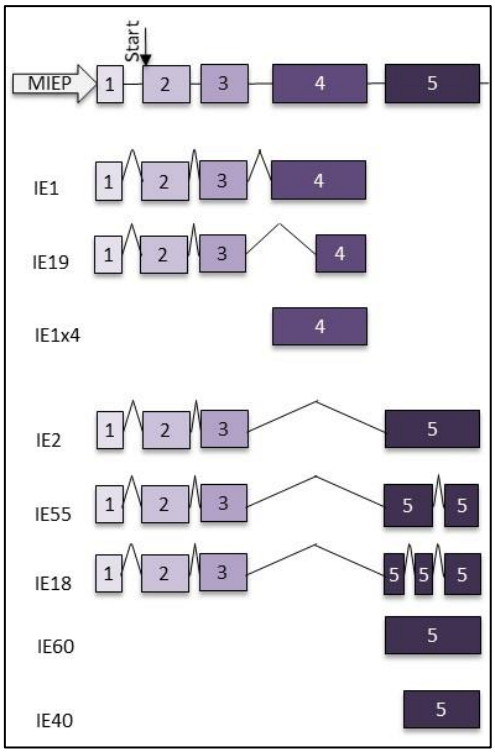
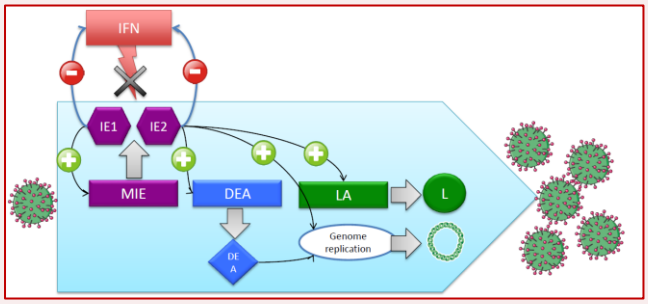



**Figure 5: Gene and transcript structures of the major immediate early gene.**

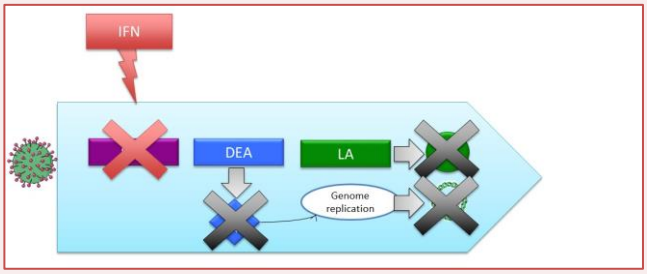
In conclusion, the MIE proteins are the strongest modulators for the viral replication cycle and ensure the expression of DEA and LA. But they also regulate important processes in the host cells to favor viral replication and prevent a full anti-viral immune response.

### *Box 2: Importance of the MIE for the viral replication*

***A**: The MIEs are involved in all major steps of the viral replication cycle. IE1 ensures its own expression against cellular repression. IE2 induces viral promoters to induce the expression of delay early and late antigens and is directly part of the genome replication. Moreover, the MIEs inhibit the interferon response of the host cell and prevent the activation of the innate immune system.*




***B**: If the expression of the MIEs is disturbed in any way, it will lead to an abortive viral replication. Also, the interferon response is not blocked anymore and an efficient induction of the innate anti-viral immune reaction is possible. Infected cells can be eliminated.*



### 1.5.4 Latency

After the primary infection, HCMV develops into latency to persist lifelong in the host. Latency describes a state, in which there is no active viral replication and no production of new infectious viral particles. Meanwhile, the viral genome is maintained inside the cells and only very few viral genes are transcribed. The major sites of latency are $CD14^+$ blood monocytes[221], granulocyte-monocyte progenitors[222] and $CD34^+$ myeloid progenitors[223]. Among the $CD34^+$ progenitors, the virus remains mainly in the highly primitive cells ($CD38^-$, $Lin^-$, $Thy\text{-}1^+$)[224]. The latent infection changes their ability to differentiate in all linage subsets and especially decreases the differentiation in myeloid colonies[224,225]. In a healthy latently infected person, only 0.01 - 0.004% of the mononuclear blood cells or bone marrow cells contain the viral genome and each latently infected cell holds only 2-13 viral genome copies[226].

HCMV is silenced in latency, however it is still not completely understood, how the virus establishes latency and what are the consequences for the infected cells. The establishment of latency is a process regulated by viral and cellular factors, which act together to inhibit the lytic viral replication. The entry process in these non-permissive cells is comparable to highly permissive cells, which support the lytic viral replication. After the membrane fusion, the viral genome enters the nucleus and circularizes to the episome. The MIE genes are directly silenced, which is the crucial step to establish latency. This is initiated by the cellular defense mechanism of the promyelocytic leukemia protein nuclear bodies (PML-NB), which is mainly conducted by a nuclear protein DAXX[227]. DAXX recruits histone deacetylases (HDAC), as well as other PML-NB components like Sp100 and PML, and deacetylates the histones at the MIEP leading to the silencing of the MIE genes (Figure 6 A). Further repressive factors, like YY1 and ERF, also support the inhibition of the IE expression[228]. The cellular expression levels of those factors (DAXX, YY1 and ERF) are very high in undifferentiated cells, but are also expressed in highly permissive cells[229]. The difference of highly permissive *versus* undifferentiated non-permissive cells is that the viral tegument protein pp71 enters in the nucleus of highly permissive cells, binds DAXX and induces its degradation. In undifferentiated cells, pp71 remains in the cytoplasm and cannot overcome the intrinsic defense mechanism[230]. Beside this major silencing event, there are a few further repressive mechanisms. For example, the host miRNA has-miR-200b inhibits the *UL122* and supports the latency[231]. The viral proteins derived from the *UL138* gene are essential to establish latency. Its two isoforms inhibit the IE expression, even though the molecular mechanism is not known yet, since the UL138 proteins are localized in the Golgi apparatus[232].

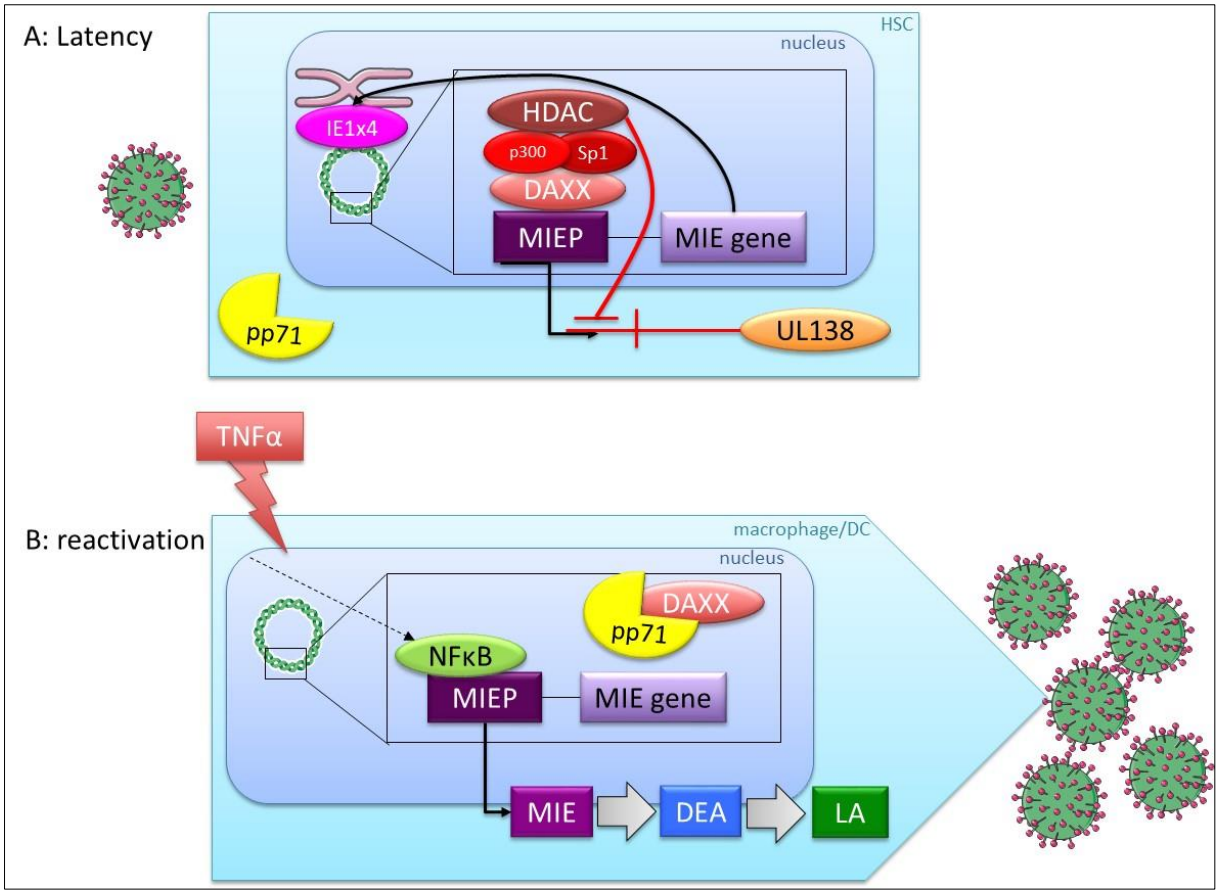



**Figure 6: Schema of the molecular mechanism of the MIEP regulation in latency and reactivation. A) In latency, the viral tegument protein pp71 remains in the nucleus. Cellular factors like DAXX interact with HDAC and inhibit the MIEP and prevent lytic viral replication. The MIEP independent IE1 splice variant IE1x4 attaches the viral genome with the host genome and ensures the viral genome maintenance. B) Viral reactivation occurs in differentiated cells, where pp71 translocates in the nucleus and induces the degradation of DAXX. Furthermore, NFκB, which is triggered by pro-inflammatory cytokines such as TNFα, is necessary to transactivate the chromatin-open MIEP and thus initiate viral replication. HSC: hematopoietic stem cell, DC: dendritic cell**

After silencing the lytic replication cycle, the virus needs to develop a way to ensure its maintenance in the cell and enable a minor genome replication. There are only a few regions in the HCMV genome with open chromatin during latency. One of these regions is the terminal repeat (TR) region. It is a repetitive region in the genome separating the $U_S$ and the $U_L$ region (see 1.5.1). The TR-region is essential for the genome maintenance and is suggested as the latent origin of replication ($ori_{lat}$)[233]. A few years ago, an IE1 isoform was found to be expressed in latency. It is lacking the first exons of the IE1 and only consists of exon 4 giving rise to the name IE1x4[234]. It is expressed independent of the MIEP. Its different protein domains interact with the cellular chromatin and the viral TR-region. Therefore, it is highly probable that IE1x4 tethers the HCMV genome to the human genome to ensure the viral genome maintenance in latently infected cells. Moreover, it also recruits the cellular Topoisomerase IIb (TOPOIIb), which is an important part of the cellular replication complex and could initiate the viral genome replication by the host enzyme complex. This would also be in line with the assumption that the TR-region could be the $ori_{lat}$.

IE1x4 is one of the rare viral transcripts during latency. Further ones are UL144, US28, UL111.5A and several viral miRNAs, to only name a few[235]. Most of the CMV latency-associated transcripts (CLT) are part of the immune evasion of HCMV to ensure that latently infected cells are not cleared by the immune system. For example, US28 is a G-protein couple receptor, which binds and internalizes CC chemokines and cytokines to reduce the concentration of pro-inflammatory

cytokines[236]. UL144 induces the production of CCL22, which induces Th2-cell differentiation and attenuates therefore a proper B- and T-cell response against HCMV. UL111.5A encodes for a viral interleukin 10 (vIL-10). During lytic replication the full-length vIL-10 is expressed, while in latency only a truncated isoform is produced, the LAcmvIL-10[237]. Both isoforms display similar biological functions to the cellular IL-10. They downregulate the major histocompatibility complex II (MHC II) expression to prevent cytotoxic T-cell activation and create an immunosuppressive environment[238,239].

Beside the expression of CLT, there are also changes in the host gene expression during HCMV latency. One of the most important effects is the down regulation of the host miRNA has-miR-92a, which normally inhibits GATA2[240]. GATA2 is one of the major regulators in $CD34^+$ progenitor cells. It regulates proliferation, lineage commitment and survival. Moreover, GATA2 induces the expression of cellular IL-10, which inhibits apoptosis through the increased expression of the anti-apoptotic protein Bcl2 and works synergistically with the LAcmvIL-10/vIL-10 to inhibit cytotoxic T cells. But GATA2 is also a major regulator of the latent viral gene expression. It has been shown that GATA2 induces the expression of most of the CLT[241,242].

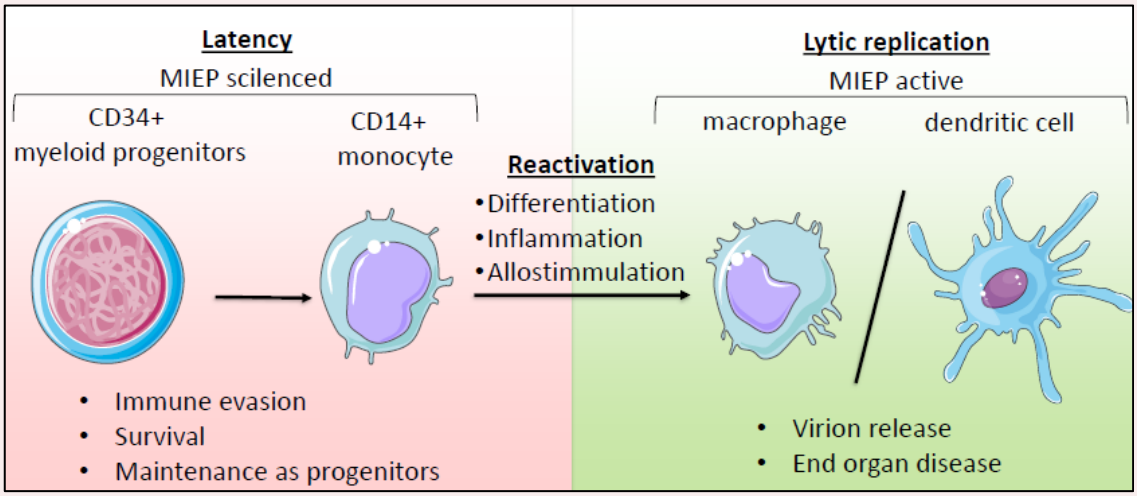
Most of the other changes in the host gene expression are related to host cell defense, immunity, cell growth, cell signaling and gene expression to favor the survival of latently infected cells[243]. Some examples are the up-regulation of POU2F2 and AML1b, which represses the IE expression and recruits HDAC, the up-regulation of MCP-1 and MIP1-β, which recruit leukocytes and facilitate the spread of HCMV to other permissive cells and the up-regulation of CD169, an adhesion molecule ensuring the persistence in the bone marrow or tissue and the contact to other permissive cells.

The latent HCMV infection itself would not be of concern, since there is nearly no harm to the host. But the virus can reactivate frequently. In a healthy individual, the reactivation and local lytic replication will be rapidly cleared by the immune system without any damage to the host. Only if the immune system fails or the person is already immunocompromised, there is a progression to CMV disease.

There are several hypotheses how HCMV reactivates from latency[244]. In general, HCMV reactivates in the process of cell differentiation from $CD34^+$ progenitors or $CD14^+$ monocytes to mature macrophages (MDM) and dendritic cells (mDC)[245]. This can happen quite frequently, but is normally quickly controlled by the immune system. In the setting of transplantation, the allogeneic stimulation and histoincompatibility can also reactivate the virus[246]. For this, cell-cell contact between alloreactive $CD8^+$ T cells and the infected monocytes is required, as well as the activation of $CD4^+$ T helper cells to produce IL-2, which co-stimulates the $CD8^+$ T cells[247,248]. The pro-inflammatory cytokines released by the $CD8^+$T cells and the direct contact lead then to the initiation of IE expression and lytic HCMV replication. The last possibility, which can promote reactivation, is an infection by another pathogen leading to a major inflammatory response or sepsis. It has been shown in the settings of transplantation that a bacterial superinfection is correlated with a higher risk of HCMV reactivation[81]. In this case, the resulting cytokine storm, including tumor necrosis factor α (TNF-α) and interferon γ (IFN-γ), activates the MIEP. Although theses cytokines are normally known as part of the anti-viral response, they have no anti-HCMV effect. They promote the activation of NFκB, which binds the MIEP and induces the transcription of the major IE genes initiating the lytic replication cycle[249] (Figure 6 B). Beside the external stimuli, some viral proteins are also essential to enable reactivation, like the latency unique natural antigen (LUNA)[250].

In conclusion, HCMV is well adapted to humans. Under normal circumstances they can coexist without harming each other for decades. Synergistically, the host cells and the virus ensure latency, maintenance and survival of latently infected cells. Only the virus developed a mechanism to escape the host under highly immunogenic conditions to ensure its survival in the human population, while the human immune system is not able to clear the infection.

***Box 3: Development from latency to reactivation***




*HCMV infects a broad spectrum of cells but it can only establish latency in progenitor cells. This mechanism allows HCMV to persist live long in the host. To ensure its maintenance, HCMV favours the survival of the infected cell and decreases its differentiation potential. In this state, the MIEP is total silenced, which abrogates the lytic viral replication cycle. Several circumstances can activate the MIEP, like the natural process of cell differentiation but also external stimuli like inflammation and allogenic stimulation. Especially in transplantation, all three stimuli are present and favour reactivation, which can then lead to severe end organ diseases.*

## 1.6 Antiviral therapy

This chapter will give an overview of the different available treatment options against HCMV; reviewing the already approved pharmaceutical drugs as gold standard treatment, new developed drugs in clinical or preclinical studies, approaches of adoptive cell therapy and the advances of vaccine development. The advantages and disadvantages of the different treatment options are summarized in Box 4 at the end of the chapter. The focus will be mainly on the treatment of transplant patients as it is the major interest of the thesis.

### 1.6.1 Approved anti-viral chemotherapy

Today, there are mainly three drugs available for the current standard treatment of transplant recipients: Ganciclovir (GCV), Foscarnet (phosphonoformate, PFA) and Cidofovir (CDV) [251,252]. In some countries, Acyclovir (ACV) is also approved. In particular in France, only GCV and PFA are

approved for HCMV treatment[2]. ACV is restricted as prophylaxis for transplant patients and CDV is since 2014 only available by an "Autorisation temporaires d'Utilisation" (ATU, temporary authorisation for use) under specific circumstances.

GCV, ACV and CDV are nucleoside and nucleotide analogues, respectively, (Figure 7), which inhibit the viral genome replication. The viral DNA polymerase UL54 incorporates the nucleotide analogue during elongation in the new synthetized DNA strand and the acyclic nature of the drug prevents further attachment of nucleotides, which leads to the premature chain termination. Unlike CDV, which requires two phosphorylations by a host kinase, GCV and ACV are highly specific to infected cells, because their phosphorylation to the active triphosphate is initiated by the viral kinase UL97 and completed by a host kinase.

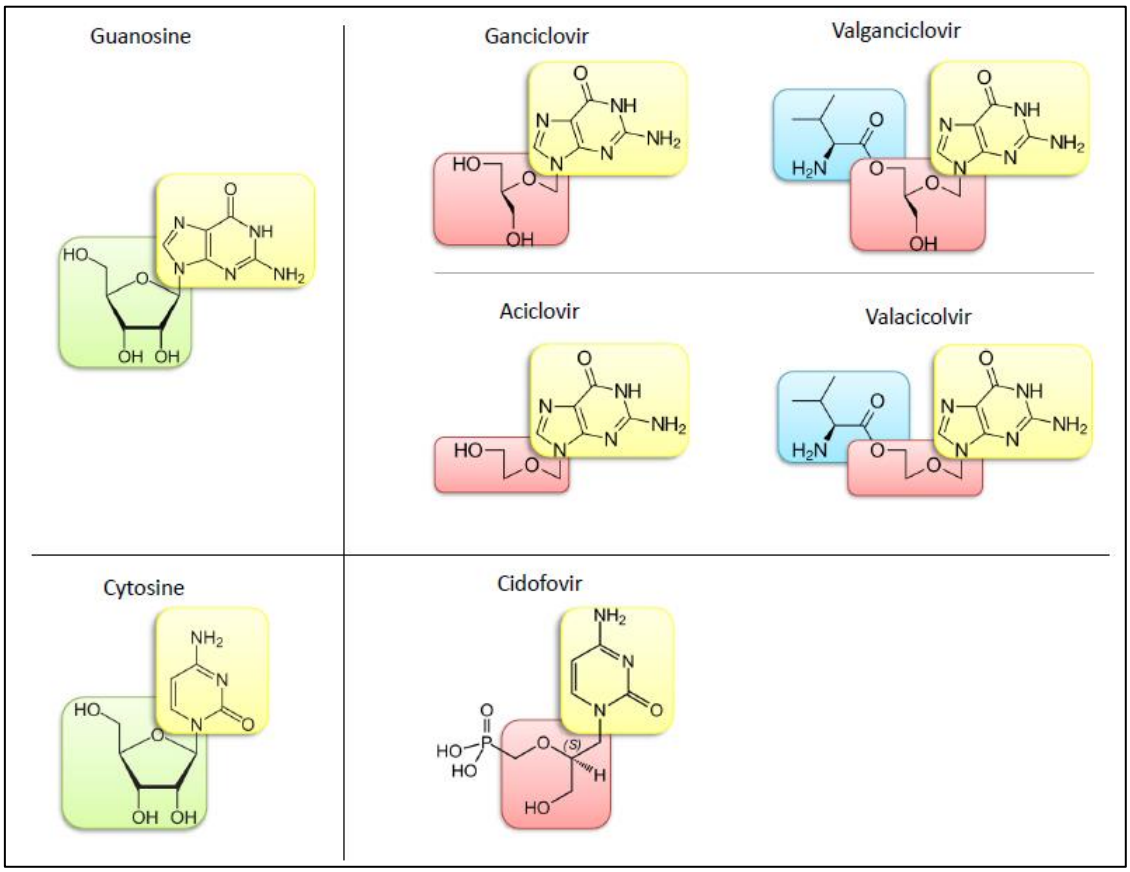



**Figure 7: Molecular structure of the anti-HCMV nucleoside/nucleotide analogues. Ganciclovir and Acyclovir are guanosine analogues and Cidofovir is a cytosine analogue. All anti-HCMV nucleoside/nucleotide analogues consist of the organic base, guanine or cytosine (yellow). However, the backbone of the nucleoside/nucleotide, the cyclic deoxyribose domain (green), is modified to an acyclic domain (red), which does not allow the elongation to other nucleotides. Ganciclovir and Acyclovir also exist with a valine ester modification (blue) to achieve a high bioavailability for oral administration.**

PFA is a pyrophosphate analogue, which does not require any activation. It is a non-competitive inhibitor of the viral DNA polymerase UL54 and targets the phosphate-binding site to prevent the cleavage of the triphosphate during elongation of the DNA strand.

GCV and PFA have the same efficiency to prevent CMV disease and improve the survival in transplant patients[253,254]. They decrease the early-onset CMV disease to less than 5% after HCST[130]. Moreover, there is a 50% reduction of CMV pneumonia and mortality under this anti-viral treatment[255]. Possible side effects of GCV are neutropenia, thrombocytopenia, tremor,

[2] Agence national de sécurite´du medicament et des produits de santé, www.ansm.sante.fr

seizure and mental confusion[253,256]. PFA is associated with renal dysfunction, abnormal electrolyte, decreased haemoglobin, white blood cells and platelets as well as hallucinations and tremor[254,257]. CDV is also highly efficient but the side effects are much more severe, with strong renal and kidney toxicity and severe neutropenia[251]. For this reason, CDV was used as a second or third line treatment for a while, but is no longer available in clinic. ACV shows much less side effects[130,251]. It has been proven efficient to decrease CMV infection, but has less impact on CMV disease progression and survival.

In transplantation, two treatments strategies are used in clinical practice : prophylaxis[258] and pre-emptive therapy[255]. Prophylaxis aims to prevent CMV infection/reactivation. All patients at risk are treated with the drug for the first 100 days post transplantation (in HSCT). This nearly abolishes the occurrence of early-onset CMV disease, but gives a slightly decreased risk of late-onset disease of 5-17% [256,259]. Due to the prolonged treatment period, side effects can be worse. The total prevention of CMV replication may also delay the reconstitution of CMV specific T cells, because T-cell priming is not possible without antigen presentation[259]. Moreover, during prolonged prophylactic treatment it is more likely to develop resistant viral strains[260,261]. In contrast to prophylaxis, pre-emptive therapy depends on the detection of CMV reactivation and prevents its progression to CMV disease. The difficulty for the pre-emptive therapy is the accurate detection of CMV reactivation. This includes strict surveillance (1-2 weekly) and sensitive CMV tests based on molecular assays[134,262]. The patients are treated for a minimum of 2-3 weeks or till the viral load decreases under the detection threshold. There is not significant difference in the overall efficiency of prophylaxis and pre-emptive therapy to prevent CMV disease and mortality[256,263], but most centres apply the pre-emptive therapy[264] as it reduces the risk of resistant viral strains and decreases possible side effects. In general, intravenous GCV (5 mg/kg) or its prodrug oral Valganciclovir (900mg) is the first choice for pre-emptive therapy[263]. It can also be used for prophylaxis, but its strong side effects usually leads to discontinued treatment[254]. ACV is less efficient than GCV, but also causes much less side effect and is therefore preferred in prophylaxis[264].

Drug resistance is a serious problem in the management of HCMV infection. Prolonged prophylactic treatment or inefficient clearance of viremia without decreasing the viral load can lead to the development of drug resistance[130,265]. Mutations in the *UL97* or *UL54* genes give rise to resistant viral strains. Usually, mutations in the *UL97* gene are acquired first as a resistance against GCV[266–268]. A switch to the second line drug PFA can rescue the patient and enable the repopulation of the wild type virus[261,269,270]. A small proportion of resistant viruses remains in the patients and might expand rapidly, if the patients is again treated with GCV. However, mutations in the *UL54* gene lead to resistant viruses against PFA with a possible cross resistance to GCV and make further treatment difficult[267].

Beside the nucleotide analogues, there is another type of drug approved to treat HCMV, but only for CMV retinitis in HIV patients[271]. The antisense oligonucleotide ISIS2922 (known as Fomivirsen) is complementary to the exon 5 of the *UL122/123* gene and therefore targets the IE2 expression[195,272]. The blockage of IE2 expression prevents the progression of the viral replication cycle and the release of infectious virions. This drug is only applied intravitreal and has never been tested for systemic CMV infection. Therefore, it is not suitable for the treatment of transplant patients. Even though Fomivirsen is a very effective treatment against HCMV retinitis in AIDS patients, it was withdrawn by Novartis in 2002 in Europe (and later on in the USA) due to rare requirements of the medicament. Indeed, since the introduction of HAART, the occurrence of CMV diseases and CMV retinitis in HIV patients has been reduced by 80-90%[73,75].

In conclusion, the introduction of GCV and PFA as highly potent anti-HCMV drugs improved the survival rate of transplant recipients significantly. Still, severe side effects and the occurrence of drug resistant viral strains prevent the total control of HCMV. New anti-viral strategies are required to manage CMV infection and reactivation after transplantation.

### 1.6.2 Anti-HCMV drugs under development

Several drugs are under development and are showing promising results in phase II and III clinical trials. The most promising new drug is Maribavir (1263W94)[273]. It is a benzimidazol nucleoside, which inhibits the viral kinase UL97 and therefore the viral genome replication (Figure 8). In a phase II clinical study, three oral doses (100 mg 2x/day, 400 mg 1x/day and 400 mg 2x/day) were tested against a placebo group as a prophylactic treatment of transplant patients[274]. All treatment groups had less CMV infections, with the high doses (400 mg 1-2x/day) being the most effective. Observed side effects were taste disturbance, nausea and vomiting, but no myelotoxicity or neutropenia (common problems with the GCV treatment). A follow up phase III clinical trial based on the lowest dose (100 mg 2x/day) failed to show a reduction of CMV infection and disease compared to the placebo group[275]. This clinical trial has been criticised due to a small study population and a low dose administration[276]. New phase III clinical trials are on-going: with an intermediate dose 200 mg 2x/days in HSCT as a pre-emptive treatment and in transplant patients with GCV/CDF/PFA resistant viral strains (NCT02931539, NCT02927067).

Terminase inhibitors are another novel and very promising approach to control CMV infection. Letermovir (AIC246) is a UL89 terminase inhibitor and prevents the cleavage of the viral genome[277]. This blocks the maturation of the viral particles and prevents the genome packing into the viral capsid. A prophylactic treatment of transplant patients prevented the development of CMV infections dose-dependently with no side effects in a phase II clinical trial[278]. A phase III study of Letermovir against a placebo control in HSCT with two different doses (240mg/days and 480 mg/day) has recently been closed, but results are not yet published (NCT20137772). Another terminase inhibitor is: BAY 38-4766[279], a non-nucleotide inhibitor, targeting the terminase complex UL56/UL89. It showed promising results in an *in vivo* infection study in guinea pigs[280]. Clinical development of BAY 38-4766 was started, but no follow up clinical studies were done. The benzimidazol nucleoside BDCRB[281] and its derivative GW275175X[282] also inhibit the concatemer cleavage by the terminase complex. However, their anti-viral efficiencies were below maribavir, so that clinical studies were not started.

Beside these drugs with new target mechanisms, there are also studies on a modified version of CDV. Brincidofovir (CMX001) is a lipid-conjugate of CDV with a much higher bioavailability, increased half-life and an improved safety profile[276]. A phase II clinical trial showed that the prophylaxis with 100 mg twice per week is efficient to reduced CMV infection in transplant patients[283]. Higher concentrations are needed, if the patient has already a low level CMV replication at the beginning of the treatment. Side effects are diarrhoea and nausea, which can be reduced by a dose-bodyweight adjustment to <3,5 mg/kg bodyweight. Recently, a phase III clinical trial was finished on CMV seropositive HSCT patients testing safety and efficiency, but results are not yet available (NCT01769170).

Maribavir (2013), Letermovir (2017) and Brincidofovir (2014) are already available by ATU in France[3]. With the start of the usage of Maribavir in clinic, first case reports descripted already the evolution of resistances by mutations in the UL97[284,285].

Further drugs are under development in pre-clinical studies. Some already established drugs for other diseases were shown to have anti-HCMV effects. Artesunate is already an approved treatment against malaria. Its side effects are very weak and it has a high selectivity index (toxic dose 50/effective dose 50) for HCMV, too[286]. *In vitro* tests could demonstrate an inhibition of viral replication at immediate early time points after infection by the downregulation of NFκB and Sp1, two important transactivators of the *MIE* genes[287]. It also causes cell cycle arrest in the late G1-phase[288]. Since artesunate targets host proteins, no resistant viral strains were observed so far. Valpromide and Valnoctomide, drugs approved for the treatment of epilepsy, mental disease, seizures and mania, inhibit HCMV *in vitro* and *in vivo*[289]. They prevent the attachment of the virus to the cell and the viral entry. In a mouse model of infection and placenta transmission, it could even increase the survival of the foetus and reduced the viral load. Therefore, it could not only be a treatment option for transplant patients but maybe even prevent congenital infections.

Furthermore, several high-throughput screens have identified potential anti-HCMV drugs. For example, Cohen *et al*. found concallatoxin, a cardiac glycoside, to inhibit HCMV IE transcription by the inhibition of methionine transport inside the cell[144]. Loregian and colleges detected 4 molecules (degulin, nitazoxaid, thioguanosin, alexidine dihydrochlorid), which inhibit the transactivator function of IE2 and prevent the progression of the viral replication cycle[290]. Two different studies analysed a host kinase inhibitor libraries and discovered molecules which inhibit the expression of IE2 and pp28[291,292]. For example, R00504985, which inhibits the kinases CDK2 (cyclin dependent kinase 2) and MAPK1 (mitogen-activated protein kinase 1). CDK2 usually phosphorylates IE2 and its inhibition decreases IE2 levels, while MAPK1 phosphorylates pp28, which is destabilized and degraded under MAPK1 inhibition.

Finally, two very different strategies were developed to target HCMV with modified host proteins. Spiess *et al.* designed a fusion protein (F49A-FTP) based on the chemokine $CX_3CL1$ and the *pseudomonas* exotoxin[293]. This F49A-FTP binds the viral G-protein coupled receptor US28 with a higher affinity than the natural receptor $CX_3CR1$. It gets internalized after binding and the toxin then kills the cell. Modification of the $CX_3CL1$ increased the specificity to US28 leading to a high selectivity index for infected cells[294]. This strategy is also the first to target latently infected cells, because US28 is expressed by lytic and latently infected cells. The other strategy is a custom-designed RNase P-based ribozyme to target HCMV mRNAs[295]. RNase P ribozyme is a ribonucleoprotein complex, which is guided by an RNA to identify the substrate RNA (by its tertiary structure), which further gets cleaved by the ribozyme. Naturally, it processes tRNAs. Yang *et al.* customized the RNA to target the viral mRNA of UL80 and UL80.5. Lentiviral delivery of the customized RNase P ribozyme to infected cells reduced the target mRNA by 99% and led to a concomitant inhibition of viral protein expression.

[3] Agence national de sécurite´du medicament et des produits de santé, www.ansm.sante.fr

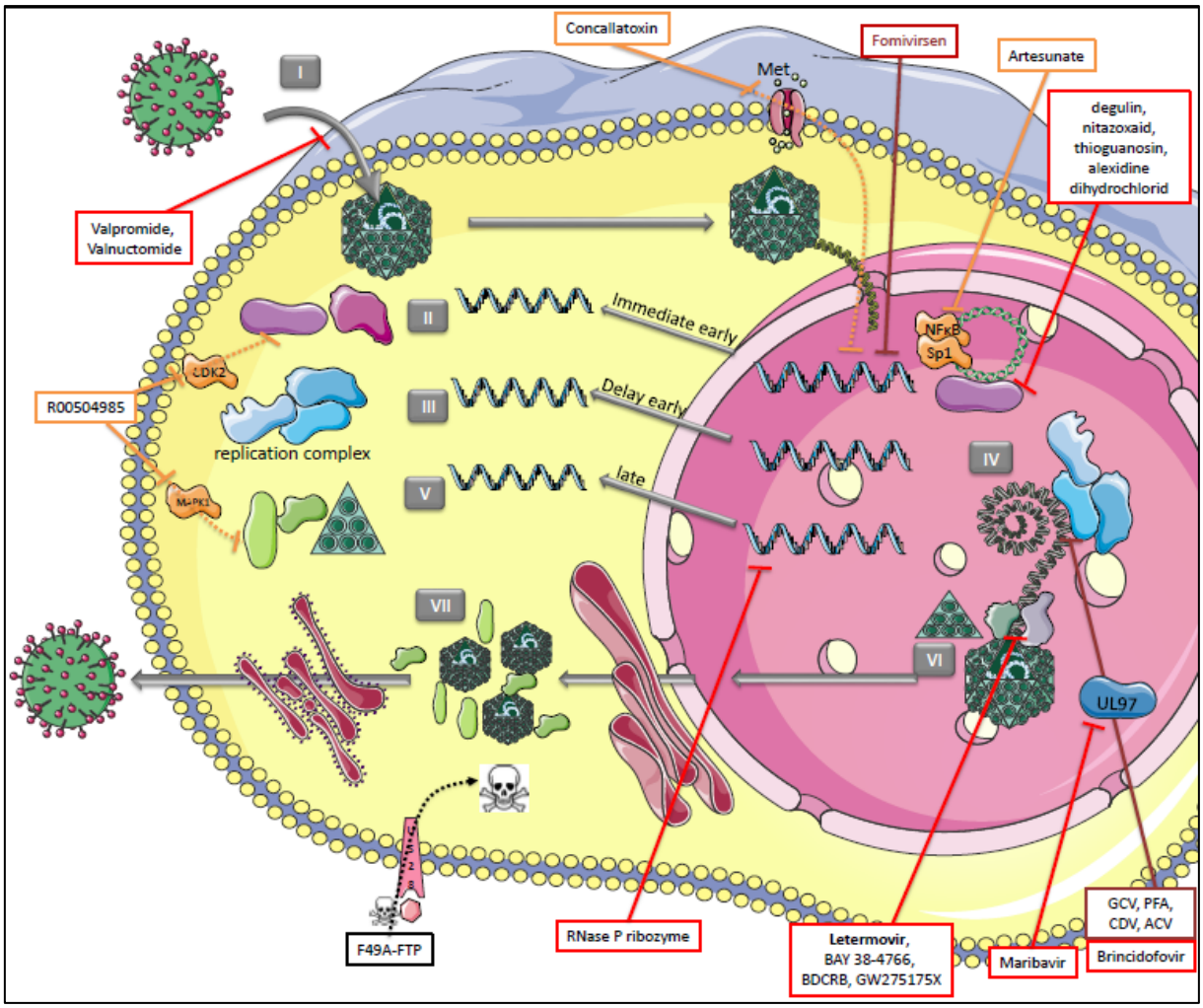



**Figure 8: Target sites of anti-HCMV drugs. Standard treatments (dark red) with GCV, PFA, CDV and ACV all inhibit the viral genome replication, beside the oligonucleotide ISIS2922, which target the IE mRNAs. Potential new drugs (light red) have various viral targets. Maribavir inhibits the viral kinase UL97 as part of the genome replication process. Others target the genome packing and capsid maturation, IE expression or viral entry. Furthermore, some new drugs are targeting cellular proteins (orange), which then indirectly block the viral replication. One completely new mechanism is the introduction of a toxin in the infected cell to induce cell death (black).**

All these novel strategies show innovative and promising ways to target HCMV to prevent replication or even clear the infection. Among those new treatment approaches, Maribavir and Letermovir already proved to be efficient in clinical studies and will probably soon be part of the standard treatment of transplant patients. Similar to GCV and PFA, they only target lytic replicating virus and cannot clear the infection. Therefore, the new host protein based strategies could be a promising approach to completely clear the infection in latently infected patients.

### 1.6.3 Adoptive T-cell therapy

One of the strongest risk factors to develop CMV disease is a delayed T-cell reconstitution (in HSCT)[296]. More than 20 years ago Riddell *et al.* performed the first adoptive T-cell transfer in HSCT patients to reconstitute the CMV specific immunity[296,297]. In those days, billions of cells and repeated infusions were needed to achieve a robust anti-CMV immune response and the cell preparation was expensive and very time consuming. However, he accomplished a T-cell reconstitution in nearly all his patients and the transferred T cells persisted for more than 12

weeks. Since then, a lot of research has been done to improve the cell preparation and to find the best cell type for adoptive cell therapy.

In general, adoptive cell therapy is defined as the infusion of effector cells to treat or prevent a disease. In transplantation, mostly T cells are used as effector cells for adoptive cell therapy. Since the reconstitution of $CD8^+$ T cells is essential to control CMV reactivation, therapy approaches were primary focused on infusing CMV-specific T cells. $CD8^+$ T-cell infusion has been shown to be protective in a mouse model of MCVM infection, while the infusion of $CD4^+$ T cells alone failed to protect mice from MCMV[298]. First studies in humans used mixed lymphocyte cultures (MLC) containing mainly $CD4^+$ but also $CD8^+$ T cells[297,299]. Cells were isolated from peripheral blood mononuclear cells (PBMC), co-cultured with a feeder cell line and supplemented with IL-2. To induce CMV reactivity, cell cultures were stimulated with infected autologous fibroblasts. After at least 4 weeks of *ex vivo* culture, the MLC could be used for infusion. This preliminary protocol was associated with a higher risk of GVHD and the infection of the recipient with HCMV from the infected fibroblasts. Analysis of the antigen repertoire of CMV specific T cells showed that 70-90% of those T cells recognise pp65 (major tegument protein of CMV), while only a minor proportion target IE1 or other CMV antigens[300,301]. Due to this finding, lymphocyte stimulation was subsequently done in the presence of pp65 peptides. These peptides span the whole length of the pp65 protein and were either added directly to the culture medium or pulsed on monocyte derived dendritic cells (MDDC). The stimulation by professional antigen presenting cells like MDDCs was much more efficient and is thus most commonly used today. Alternatively to the pulsed MDDC, MDDC can be nucleofected with the pp65 plasmid or transduced with an adenoviral vector coding for the pp65 protein[302]. This stimulation strategy allows also the generation of trivirus specific cell lines, which enable the clearance of CMV, but at the same time protect for Epstein-Bar-virus (EBV) or adenovirus (Adv), both also pathogens with severe pathologies in transplant patients[303].

Beside the usage of MLC, several groups isolated specific T-cell populations to increase the specificity against HCMV of the infused cells and increase their persistence in the recipient. The infusion of $CD8^+$ T-cell clones was able to prevent CMV infection[297], but they only persisted in the recipient, if the recovery of $CD4^+$ T cells from the HCST happened at the same time. Surprisingly, also the infusion of $CD4^+$ T-cell lines were protective against CMV[304,305]. It induced the recovery of $CD8^+$ T cells from the HCS. Other groups tried to only select activated T cells after stimulation with CMV antigens either by the selection of interferon-γ (IFNγ) positive cells or by CD137 expression, a T-cell activation marker[306–308]. Both strategies were successful in selecting CMV-specific cells.

One of the week points of adoptive T-cell therapy is the long culture and expansion period *ex vivo*. Most protocols culture the cells for at least 21 days. Cell preparations can fail because of insufficient expansion or contaminations during the culturing process. Moreover, large cell expansion is time consuming and expensive. For these reasons, minimal manufacturing procedures were tested. PBMC were stimulated over night with CMV pp65 peptides and selected on the next day based on IFNγ expression or by CMV-epitop streptomers[306,309]. Thus, only activated or CMV-specific cells were selected and the cell number required for efficient immune reconstitution was much lower.

Most adoptive T-cell therapies use immune cells from the HSC donor to reconstitute the immune system of the recipient. However, 30-50% of the donors are CMV negative and cannot provide CMV-specific T cells[3]. In this situation, there are two options available: PBMCs from the naïve donor can be primed by pp65 presenting MDDCs and further expanded and stimulated to gain

CMV-specific T cells[307]. In this case, it is beneficial to deplete regulatory T cells from the culture to boost the stimulation by CMV antigens. *In vitro* tests are promising, but it is not yet tested in HSCT recipients. The other possibility is the usage of third-party donor (TPD) cells. The difficulty is to find a HLA matching donor in the short time-frame after HCST. Two clinical studies tested the use of TPD cells with contrary results. Leen *et al*. were successful to reconstitute T cells in 17 out of 19 recipients[305], while Neuenhahn *et al.* failed to detect the donor cells in all but one patient[309]. Both studies used very different protocols for the cell preparation, which makes a conclusion of efficiency very difficult, but a high HLA match of the TPD cell and the HSC recipient and donor seems to be important to gain efficient T-cell reconstitution.

In addition to the infusion of commonly used $CD4^+$ or $CD8^+$ T cells, first studies are running on the usage of natural killer cells (NK-cell), cytokines induced killer cells (CIK-cells) or γδ T cells. The advantage of these strategies is the independency of MHC class I presentation. CMV developed several immune evasion mechanisms to prevent MHC-I presentation, like the phosphorylation of pp65 and IE1 to prevent the processing in the proteasome, the prevention of the antigen loading on MHC-I by US6 and the general downregulation of MHC-I expression[310]. All this can negatively influence the efficiency of adoptive T-cell therapy. NK-cells have cytolytic and proinflammatory functions based on an ubiquitous receptor, which is upregulated on infected cells[276]. They have been tested in a mice and protect against MCMV infection[310]. CIK-cells are a mixed culture of NKT cells and $CD8^+$ T cells, which are highly reactive against CMV and myeloid tumour cells *in vitro*[311]. γδ T cells are also not MHC-I restricted and response to a self-antigen upregulated on stressed/infected cells[312]. In mice, they clear the MCMV infection in the absence of αβ T cells.

In the last 20 years, the adoptive T-cell therapy was optimized in several ways and efficiently tested in clinical settings. But the lack of large controlled studies, comparing different protocols and infused cell types, makes a conclusion about the most efficient procedure difficult. In general, adoptive T-cell therapy can become a real alternative to prophylactic conventional therapy and might even give an economic advantage by reducing the time of hospitalization[313]. However, the highly intense immunosuppressive treatment after transplantation might counteract adoptive T-cell therapy, as it will affect the infused T cells and prevent their expansion *in vivo*.

### 1.6.4 Vaccination

A vaccine primes the immunes system for a specific pathogen to later protect the person from the infection by this pathogen. Most of the research on CMV vaccines aims to prevent congenital infections, since there is no treatment strategy available to prevent or decrease the pathology in this situation. Problematic is that even naturally infected mothers with a normal immune response to CMV can transmit the virus to the foetus leading to birth defects. It is not clearly known, which features of the immune response could prevent the transmission and would need to be induced by a vaccine. However, other patient groups could also profit from a vaccine. As mentioned above, transplant patients are at high risk of CMV disease due to their suppressed immune system. In fact, most of the clinical studies on CMV vaccination are done in HSCT patients[314], because clinical studies on pregnant woman are difficult and the occurrence of congenital CMV infection is relative low. Very large studies groups would be necessary to be able to show significant improvements. On the other hand, HSCT patient (donor or recipient) are often already seropositive for CMV. In this case, a vaccine can only boost the immune system to be more efficient to prevent CMV reactivation, but it cannot be tested for the prevention of a primary infection. Apart from specific patient groups, a generalized vaccine could be beneficial for the entire human population. Moreover, CMV is a species-specific pathogen, which only infects

humans. This makes CMV a perfect candidate to be eradicated by a generalized vaccination of the human population. A mathematic simulation has estimated that a vaccine with an efficiency of only 59-62% would be enough to eradicate CMV[10]. This is based on the worldwide average reproductive number (R0) of 2.4[10] (range 1.7[11] - 5.7[13]). However, taken into consideration that the R0 number is much higher in high seroprevalence regions, like in Brazil (seroprelavence: 87%[315], R0=5.7), more efficient vaccines would be needed in those regions. The success of a vaccine depends then also on the time of vaccination, the coverage in the population and the length of protection achieved by the vaccine. A mathematical model by Lanzieri et al.[13] assumed an efficiency of 70% and 90% coverage with a 5 year protection after injection, which then would lead to the eradication of HCMV, if the population is vaccinated within the first 12 months of life. Vaccinations later in life (>12 month) would still lead to a substantial reduction of the overall seroprelavence and would significantly decline the occurrence of congenital CMV infections, but an eradication is less probable. Moreover, for the vaccination of adolescences and adults, the impact of the vaccine on reactivation strongly influence the efficiency of the prevention of congenital infections.[15]

For over 40 years, researchers try to find a vaccine against HCMV. First studies tried to attenuate the virus by 50 till more than 125 passages in *in vitro* cell cultures in order to used theses attenuated viral strains as a source for the vaccine[316,317]. The long culture time of CMV in only one cell type leads to mutations in different parts of the viral genome. One of the most typical mutations is the disruption of the pentameric glycoprotein complex gH/gL/UL128/UL130/UL131 (PC)[314]. Most of the vaccine strains have a frame-shift mutation in the *UL130* gene. This abrogates the epithelia and endothelia phenotype and reduces infectivity and immunogenicity of the virus. The first live attenuated CMV vaccine tested, was the attenuated AD169 strain[316]. It induced neutralizing antibodies, which persisted for 10 years. Later on, the Towne CMV strain was more commonly used[317,318]. It also induces neutralizing antibodies as well as CMV-specific T cells[319]. At that time, human challenge studies were allowed, where the vaccinated study group gets infected with a clinical CMV isolate, like Toledo[317]. This study compared the immune response and viral replication of not vaccinated persons with naturally infected or vaccinated individuals. Low dose of the vaccine viral strain protected from a primary infection, but slightly higher doses increased the risk of infection by the vaccine strain and shedding of CMV. Today, this kind of virus challenge studies are ethically highly questionable and are not used in clinical research anymore. Since transplant patients have at significant high risk of CMV infection and enable the observation of efficiency of the vaccine in a relative small study population without introducing a virus challenge, most clinical studies are done in this patient group. Nowadays, two types of attenuated live vaccine are in the focus of research. The chimera viral strain of Towne and Toledo showed a good safety profile in a phase I clinical trial, but induced seroconversion only in 11 out of 36 vaccinated individuals[320]. This study tested several different chimeras and doses and concluded that a follow up study would be needed on one of the chimera with a broader dose range. The second type of attenuated live vaccines are the so called “DICS” (transgenic disabled infectious single cycle vaccine strains). Those viruses can be replicated *in vitro* in specific cell lines, but are replication incompetent *in vivo*[314]. The strain V160 is in a phase I clinical trial. It is genetically modified that the IE1 and IE2 as well as *UL51* genes are fused to the unstable FKB12 domain. FKB12 can be stabilized by the addition of Shield-1 in the culture medium *in vitro*, but *in vivo* it will be rapidly degraded leading to a replication incompetent virus. FKB12 fusion to IE1/2 has been used before to study the function of the IE genes[192].

Attenuated live vaccines have always the risk of adaptation of the vaccine virus to the host and loss of the attenuation, which could cause an infection of the host, shedding of the vaccine strain and infection of other individuals. Other safer vaccination strategies have been developed. The special feature of CMV to produce dense bodies was adjusted to deliver viral proteins in the host without an infectious particle[321]. Dense bodies are enveloped particle, which contain mainly the viral tegument protein pp65, but also express viral glycoproteins on the surface[163,170]. In different mouse models, dense bodies could deliver the viral proteins to the host cells[321,322]. Those proteins can be process and presented by MHC I and II molecules to induce a strong immune response *in vivo* including cytotoxic T cells (CTL) and neutralizing antibodies. A similar approach has been establish using the transfection of gB and the gag protein of the Moloney murin leukemia virus[314,323]. The coexpression of gag with gB leads to a capsid formation of gag, which is enveloped by a membrane containing the CMV gB. To increase the immunogenicity, gag can be fused to pp65 of CMV. These enveloped virus-like particles (VLP) is in a phase I clinical study on heathy seronegative volunteers (NCT02826798).

The use of other viral proteins or particles as delivery methods goes further than the construction of VLPs. The attenuated canarypox virus ALVAC can be used as a vector for some CMV antigens[324]. The first ALVAC vector expressed gB, but it failed to induce neutralizing antibodies. Those ALVACs has been tested efficiently only for priming the host before immunization with an attenuated live vaccine. Later, ALVAC was used to express pp65 resulting in a much better induction of neutralizing antibodies and CTL priming[325]. A phase II clinical trial in HSCT patients is ongoing. A different vector vaccine is based on the modified vaccinia Ankara (MVA) Triplex[326], expressing the CMV antigen pp65, IE1 exon 4 and IE2 exon 5. There is nearly no immunity against the vaccinia virus in the human population since the common pox vaccination was stopped in the 1970ies. The safety profile was good, even for vaccinia seropositive individuals, and a strong immune response specific to CMV was induced.

Another very common vaccination strategy is a so-called subunit vaccine, meaning the injection of a few proteins or peptides of the virus to prime the immune system against the main antigens. This is for example commonly used for the influenza virus vaccination. The subunit vaccine against CMV consists of gB, modified to delete the transmembrane domain and the internal cleavage site, and the adjuvant MF59, which has been proven more efficient than the common aluminium hydroxide gel[327]. In the phase I clinical study, it showed a good safety profile. Different follow up phase II clinical trials tested the efficiency of the vaccine[314]. While the vaccination of adolescents did not give significant results, the vaccination of women at least 1 year prior to a pregnancy showed a 50% protection and a boost of the immune system in seropositive women. Also, in SOT patients it reduced the reactivation of CMV and the transfer of CMV from the donor organ to the recipient. The disadvantage of the gB subunit vaccine is that the CTL induction is not optimal, since CMV-specific T cells mostly react against pp65 during the natural infection to control reactivation [310]. Based on this observation, a peptide vaccine was developed in which the immune dominant peptide of pp65 was either fused to the T-helper epitope PADRE (pan DR epitope) or a tetanus toxin peptide[328]. The fusion peptide has been shown to induce a strong T-cell response, which was even stronger in combination with the adjuvant PF03512676, a CpG single strand oligonucleotide and Toll-like-receptor 9 agonist. Under the name CMVPepVax it has been positively tested in a phase I clinical trial leading to a phase II study in HSCT, which is still ongoing[329].

A different approach, started in the late 1990ies, is based on the injection of a plasmid coding for 1-2 viral proteins (gB and/or pp65)[330]. This DNA vaccine was examined in a preclinical study in

Guinea pigs to analyse the effect on trans-placenta transmission to the foetus. First studies showed a high neutralizing antibody titre. As ASP0113, it went to clinical studies with a co-injection of two plasmids expressing pp65 and gB[331]. The good safety profile in the phase I tests led to a phase II trial in HSCT patients. A slight reduction of CMV viremia was shown, but no difference in antiviral therapy and CMV disease development. A more advanced model of the DNA vaccine include the usage of bacterial artificial chromosomes (BAC) coding for a replication incompetent CMV virus[332]. This was very promising in Guinea pigs and could significantly decrease mortality of the foetuses.

Overall the vaccine development made huge progress, but long-term studies of large study groups are needed to assess, if the prevention of CMV infection can be achieved. Moreover, even if an efficient vaccine is available, it will take decades to decrease CMV infection in the population due to the broad seroprevalence. Furthermore, the protection of the foetus from the maternal CMV infection will be very difficult to control, since the immune system of the mother often fails to prevent the transmission. Therefore, the vaccination of HSCT patients is in the focus of current research. It is difficult to immunize patients before the full immune reconstitution, which takes up to 2 years[125]. The priming and boosting of the immune response early after transplantation can be of advantage, as shown by several of the above discussed clinical studies. However, especially patients with a delay in immune reconstitution might not response to the vaccination and are still under high risk of CMV infection. Therefore, treatment with conventional therapies or novel treatment approaches targeting the virus are still needed.

***Box 4: Comparison of the different therapy approaches***

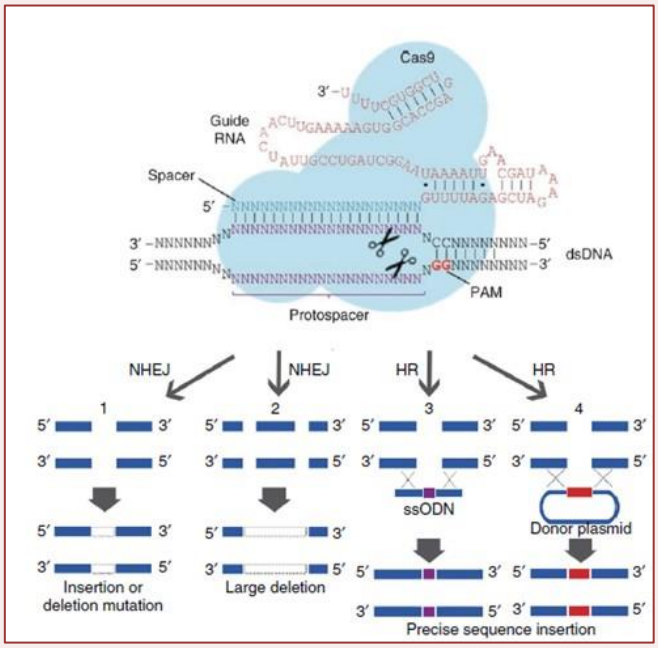
| THERAPY | ADVANTAGE | DISADVANTAGE |
|---|---|---|
| ***Gold standard (nucleotide analogues)*** | + Inhibition of lytic replication<br>+ Reduction of CMV viremia and disease in transplantation to less than 5% and 2%, respectively<br>+ Valine ester modification allows high bioavailability | - No clearance of the infection (do not target latently infected cells)<br>- No effect on established CMV disease<br>- No protection against congenital infection<br>- Side effects: myelosuppression, delayed immune reconstitution<br>- Drug-resistant viral strains |
| ***Anti-HCMV drugs under development*** | + No drug-resistant viral strains so far (exception Maribavir)<br>+ Multiple target sites allow the combination of the nucleotide analogues and the new drug to gain a synergistic effect<br>+ Several drugs target at IE time points and prevent the expression of all viral proteins preventing cell cycle control by CMV<br>+ Improved safety profile<br>+ *F49A-FTP has the potential to clear the infection by killing latently infected cells* | - No clearance of the infection (do not target latently infected cells) *(exception F49A-FTP)*<br>- Probably no protection against congenital infection |
| ***Adoptive T-cell therapy*** | + No side effects<br>+ Control CMV infection in the most natural way<br>+ Shorter hospitalization<br>+ Less use of pharmaceutical drugs | - Higher risk of GVHD<br>- Bad efficiency under intense immunosuppressive regimen<br>- Complicated protocol for cell generation<br>- High costs<br>- Probably no protection against congenital infection |
| ***Vaccination*** | + Prevents infection with CMV<br>+ Only possible prevention for congenital infection<br>+ Opens the way to eradicate CMV | - Difficulties in transplant patients:<br>  - It is not possible to induce an immune response before the reconstitution of the immune system<br>  - Patients are often already CMV positive |

# 2 CRISPR/Cas9

Genome editing has been used for decades to uncover gene functions, develop treatment for genetic disorders or to fight infections. First known were meganucleases (MN), which originate from yeast and are endonucleases with a >14 bp DNA recognition site[333]. Their adaption for research and programming for new targets is very difficult and includes new protein engineering with a low efficiency rate. Several years later, the discovery of Zink-finger-nucleases (ZFN) opened new possibilities. They are chimeric enzymes, which combine DNA-binding zinc-finger domains with the *Fok*I endonuclease domain. To target a DNA sequence a pair of ZFN needs to be designed, because the *FokI* domains needs to dimerise to induce a double strand break (DSB). Furthermore, the reprogramming still depends on protein engineering. More recently, transcription activator-like effector nucleases (TALEN) became an easier genome editing tool. They are engineered enzymes consisting of a DNA biding domain, found in a plant pathogen, fused to the *Fok*I endonuclease domain[334]. The DNA recognition is mediated by a repeat sequence in the DNA binding domain. Each repeat is specific for one base pair (bp). This gives a code to reprogram the DNA binding site more rapidly than for MN or ZFN. Still, complex cloning was necessary to achieve new programmed nucleases, especially because the repeats often lead to recombination. Moreover, the delivery of TALEN inside the target cell is very difficult, because as for ZFN, TALEN needs to dimerize to induce a DSB, and has therefore a very large size[333]. The field of genome editing was revolutionized when the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated nuclease 9 (Cas9) was introduced as a genome editing tool in

***Box 5: General principal of the CRISPR/Cas9 system***

*A) The CRIPSR/Cas9 system is a two component genome editing tool consisting of the Cas9 endonuclease and the gRNA. The Cas9 screens the target DNA for the protospacer adjacent motif (PAM) and the gRNA binds the protospacer by Watson-Crick-base pairing. If both components find the correct target region, the Cas9 induces a double strand break (DSB), which is subsequently repaired by the cellular repair mechanism: non-homologous end-joining (NHEJ) or homologous recombination (HR).*




*B) NHEJ is error prone and induces small indels at the target. Those indels can induce frameshift mutations and premature stop codons leading to the knock out of the target gene. Larger deletions can be created by using several Cas9/gRNAs. HR depends on a donor DNA template, either a single strand oligonucleotide (ssODN) to induce single nucleotide substitutions or large dsDNA (e.g. plasmids) to introduce whole genes.*

*Picture: modified* Mali et al. 2013; *Nat Methods*. and Shao et al. 2014;*Nat Protoc*.

2012/2013[335]. It is a much easier system of two components: the Cas9 endonuclease, which induces a DSB in the target DNA, and a small guide RNA (gRNA) guiding the Cas9 to the target site (***Box 5***). Since the recognition of the target site depends on RNA-DNA Watson-Crick base pairing, the reprogramming of the CRISPR/Cas9 is much more feasible and needs only the new design of the RNA compound. It gained rapidly recognition in the scientific community and was declared the scientific breakthrough of the year 2015. Since then the number of publication using this technology is exponentially rising.

## 2.1 The origin of the CRISPR/Cas9 system

Originally, the CRISPR/Cas9 system is an adaptive immune system of prokaryotes to protect against phages and plasmids[336]. 40% of all bacteria and up to 90% of archaea have some type of CRISPR/Cas system. Firstly described was the particular structure of the CRISPR locus in *Escherichia coli* in 1987[337]. However, the real functional properties of the CRISPR locus were only discovered in 2002 by Jansen *et al.*[338].

The CRISPR/Cas system is very diverse in the different species. It can be classified in three main types (I-III) and several subtypes (A-E) (Figure 9)[339]. In general, it consists of the CRISPR locus and different types of *cas* genes. The CRISPR locus starts with the leader sequence of several 100 bp and in average around 60 short repeats of 23-47 bp, which are separated by unique short spacer elements of 21-72 bp[336,338]. Those spacers originate from foreign DNA (phage, plasmid)[340]. There are up to 45 different cas families with distinct DNA modifying function[341]. They include nucleases, helicases, polymerases and polynucleotide binding proteins and are necessary to provide protection against the invading DNA[342].

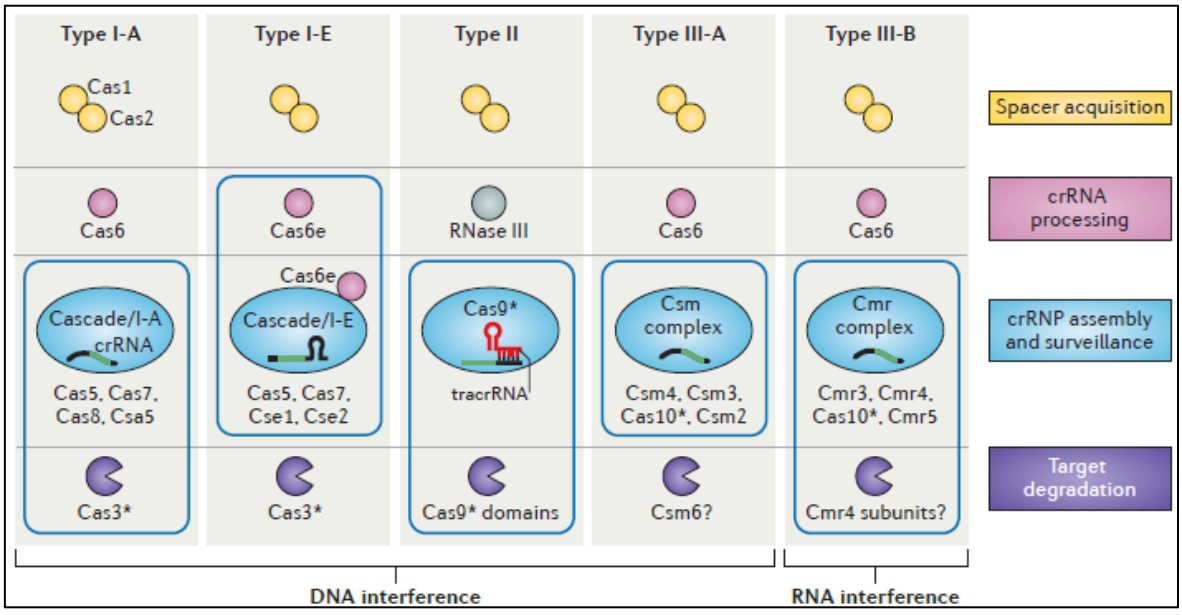



**Figure 9: Representation of the different types and subtypes of the bacterial and archaea CRISPR/Cas systems. They all include multiple enzymes, which are involved in the acquisition of new spacers. The main difference is the enzyme complex, which is responsible for the screening and cleavage of the DNA. CRISPR/Cas type I and III use multi protein complexes in connection with a CRISPR RNA (crRNA). In comparison, the CRISPR/Cas type II is more simple and only dependent on the Cas9 endonuclease in connection with the transactivating crRNA(tracrRNA):crRNA duplex. The simplicity of the system makes it a perfect candidate for the adaptation for genome editing. Figure from: van der Oost et al. 2014; *Nat Rev Microbiol*.**

The main mechanism to acquire protection and destroy foreign DNA is the same for the three CRISPR/Cas systems and can be divided in three steps[339]. Firstly, the bacterium or archaea is

infected with an invading pathogen, which releases its genetic material inside the plasma. This invading DNA is processed by the cas proteins and the small DNA fragment can be integrated in the CRISPR locus in bacterial/archaeal genome[343]. The integration of a new spacer element is regulated by the protospacer adjacent motif (PAM), which is species specific[344]. Depending on the PAM and further cas proteins the new spacer is inserted at the leader sequence and a new repeat sequence is generated. The integration of a new spacer element can be concomitant with the loss of an already acquired older spacer element to control the size of the CRISPR locus. In step two, the CRISPR locus gets transcribed to one long precursor CRISPR RNA (pre-crRNA) starting with the leader sequence followed by the repeat-spacer repetitions[339]. Subsequently, host RNase III or the Cas6 protein processes the pre-crRNA to the mature crRNA and allows the assembly with the nuclease complex. Finally, this ribonucleoprotein (RNP) complex degrades invading DNA sequence-specific dependent on the PAM and the crRNA.

## 2.2 Principle of genome editing via CRISPR/Cas9

The bacterial CRISPR/Cas type II served as a basis for the actual genome editing tool, because surveillance and degradation is only mediated by one protein, the Cas9[339]. The Cas9 consists of two major lobes: the NUC (nuclease lobe), which contains two nuclease domains (HNH and RuvC) and a small DNA binding domain for the PAM recognition (PI = PAM interaction domain), and the REC (recognition lobe) consisting of three domains (REC1-3) important for structural changes dependent on target-DNA binding of the gRNA for the nuclease activation[345]. Emmanuelle Charpentier and her team adapted the Cas9 of the *Streptococcus pyogenes* (SpCas9) for the usage in eukaryotic cells[335]. They inserted two nuclear localisation signals (NLS). Later on, the sequence of the Cas9 was human codon-optimized for better usage in human cells[346]. A further simplification was introduced on the RNA design. Originally, the bacterial Cas9 build a complex with the tracrRNA:crRNA duplex. The crRNA contains the spacer sequence and the constant repeat region and the tracrRNA base pairs with the repeat part of the crRNA to build secondary structure to interact with the Cas9. Charpentier constructed a chimeric guideRNA (gRNA) to replace the tracrRNA:crRNA duplex. It has a 21 nt single strand region at the 5' end to bind the target sequence (spacer) and a 3' end hairpin structure to connect with the Cas9. Today, there are several expression plasmids and vectors available, which only need the introduction of the 21 nt target sequence to complete a fully functional genome editing tool[346–348]. In comparison to the older techniques like ZFN or TALEN, the CRISPR/Cas9 is much more cost efficient and needs only a few weeks of planning and cloning.

Genome editing by the CRISPR/Cas9 system is meditated by the cleavage of the target gene leading to a DSB and its repair by the host cell DNA repair machinery to introduce changes in the target gene. The Cas9 itself is inactive as long as it is not interacting with the gRNA. The binding of the gRNA repositions the PI domain to enable its interaction with the DNA. Futhermore, the interaction of the gRNA/Cas9 pre-orders the gRNA spacer region proximal to the PAM (seed sequence) into an A-from to facilitate the interaction with potential target DNA sequences. The PAM distal gRNA region is unstructured but protected in a cavity of the Cas9 between the two nuclease domains[345]. The molecular mechanism depends then on the Cas9 screening the DNA/genome for the PAM sequence[334]. This first interaction is very short (<1 s). If the PAM recognition is successful, the Cas9 melts the DNA double helix and flips the first base of the target sequence towards the gRNA forming the R-loop[349]. The base-paring of the pre-ordered gRNA seed sequence with the target DNA goes in hand with conformational change in the Cas9 to facilitate

the gRNA distal base pairing with the target DNA. The RNA-DNA heteroduplex lies in a positive charged channel between the two lobes (REC and NUC). Finally, the REC3 domain detects sequence-independently the perfect base pairing of the gRNA/target DNA and induces the reorientation of the REC2 domain[350]. This reorientation allows the HNH domain to interact with the target DNA, which subsequently leads to the rearrangement of the linker domain between the HNH and RuvC nuclease domain positioning the RuvC domain closer to the target DNA to activate its nuclease function[345]. The cleavage of the DNA is then induced three base pairs upstream of the PAM sequence[335]. The DSB will be repaired by repair mechanisms of the host cell. The most common one is the template independent non-homologous end joining (NHEJ) known to be error-prone and to introduce small insertion and deletions (*indels*) at the cleavage site[17]. While the classical NHEJ is relatively accurate, the two alternative NHEJ mechanisms, microhomology-mediated end joining (MMEJ) and theta mediated end joining (TMEJ), are much more mutagenic[351] (Figure 10). The classical NHEJ re-joins the DNA unspecifically without further modifications. The blunt end DNA is captured by Ku70/80 to protect it from degradation. Further connection to the DNA by a protein kinase and other factors stabilizes the DNA and activates the ligation by host ligase IV. In comparison, the MMEJ is Ku independent[352]. It is a process starting with the resection of the DNA ends in 5'- 3' direction to generate a 3' overhang. The two ssDNA arms anneal depending on small microhomologies. Overhanging non-complementary 3' ends get removed and the open gap gets filled-in by a polymerase. Finally, the strands are ligated by the host ligase III/I. This repair mechanism nearly always generates small deletions, but hyper mutations are observed in an area of 10 kb around the DSB[353]. The TMEJ relays on a similar mechanism[351]. The major difference is the involvement of the translesion DNA polymerase θ, which extends the 3' overhang template independently by several base pairs. This leads to the formation of *indels* round the DSB. The mutagenic potential of the alternative NHEJ mechanisms are of advantage in genome editing. The introduction of *indels* can lead to a frame-shift mutation, which then usually results in a premature stop-codon or a malfunctional truncated protein. Often, it leads to a knock-out of the target gene.

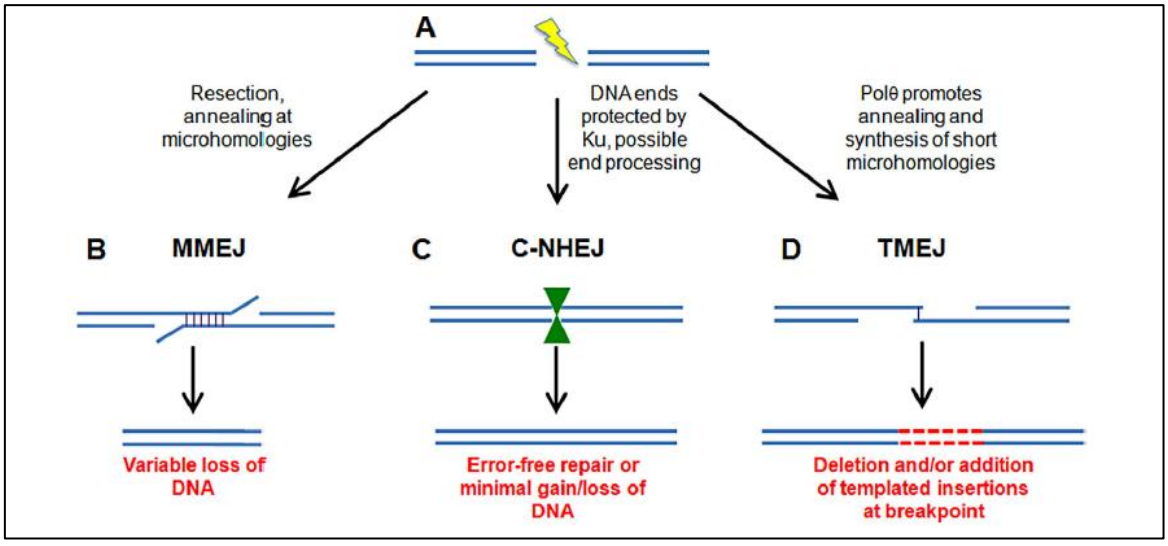



**Figure 10: Different types of non-homologous end joining (NHEJ) after a double strand break (DSB). A: The CRISPR/Cas9 system induced DSB by blunt end cleavage of the target site. Without a homologous region available or in a resting cell, the DSB gets repaired by NHEJ. B: Microhomology-mediated end joining (MMEJ) starts with the resection of the DNA blunt ends to generate single strand 3' overhangs, which can base pair at microhomologies. 3' non-complementary overhangs are removed and the reminding gaps get filled-in and re-ligated. MMEJ nearly always leads to small deletions. C: classical NHEJ is a mainly error-free direct ligation of the two blunt ends, depending on the protection of the free DNA ends by the proteins KU70/80. D: Theta-**

**mediated end joining (TMEJ) is named after the translesion polymerase θ, which can add bases at the 3' end template-independently. The mechanism is similar to the MMEJ, 3' overhangs, generated by resection and/or extension, anneal and the gap gets filled-in and re-ligated. TMEJ can induce small deletions and insertions. Figure from: Rodgers et al. 2016; *J Cell Physiol*.**

The efficiency of a gRNA/Cas9 to induce *indels* at the target region is most commonly accessed by a SURVEYOR assay or T7-endonulcease assay[354]. Therefore, the target region gets PCR amplified. Subsequently, the PCR products get denaturated and slowly reannealed. When a WT and a mutated or two differently mutated PCR fragments reanneal, they contain small mismatches at the mutation sites. Those reannealed PCR products will be subjected to an endonuclease (surveyor or T7), which identifies the small mismatches and induces a DSB. Separating the PCR fragments by gel-electrophoreses, an uncut PCR product band and two smaller bands will be detected. The percentage of mutated sequences can be calculated based on the following formula, where $\boldsymbol{a}$ is the DNA concentration of the uncut PCR product and $\boldsymbol{b}$ and $\boldsymbol{c}$ are the DNA concentrations of the digested PCR products:

$$f = \frac{(b + c)}{a + b + c}$$

$$indel\ (\%) = 100 \times (1 - \sqrt{(1 - f)})$$

The method is relatively accurate for low and intermediate mutation efficiency, but it is biased at high efficiencies, because it estimates the uncut band as not mutated. At very high mutation efficiency, the uncut band contains WT and mutated PCR fragments, since the probability that two fragments with the exact same mutation reanneal is much higher. Today, there is a new method available, which does not depend on the realignment of PCR fragments[355]. The restriction fragments length polymorphism analysis (RFLP) uses the same gRNA/Cas9, as used for inducing the mutation, to digest the PCR fragments. WT PCR products will be cut by the gRNA/Cas9, while mutated PCR products cannot be targeted by the same gRNA. The efficiency calculation is thus based on a simple ratio of not cut (=mutated) and cut (=WT) PCR products.

Finally, there is a third method for *indel*-detection, which is independent of the cleavage of the PCR product. TIDE (tracking *indels* by decomposition) is based on sequencing. The target region is PCR amplified and subsequently analysed by basic Sanger sequencing. The CRISPR/Cas9 treated sample will give a mixed sequencing result starting at the target region. The TIDE software will decompensate the sequence mix to identify the different sequences available in the mixture compared to the wild type sequence. This provides information about the length and frequency of small insertions and deletions and analyses the base composition of the 1 bp insertion. This method provides even more detailed information about the target region than the commonly used T7 assay[356].

In general, the mutation efficiency of a singleplex CRISPR/Cas9 system is between 16-79%[357,358]. Therefore, it is much more efficient than TALEN (0-34% efficiency). The high variability depends on the cell type, the target gene, the secondary structure of the gRNA and the chromatin condensation of the target region. Highly closed chromatin region cannot be targeted by the CRISPR/Cas9 system, because the gRNA/Cas9 cannot bind to the DNA[359].

The CRISPR/Cas9 system can also be used as a multiplex strategy, where several gRNAs are designed to target multiple sites in one or several genes simultaneously[348]. This leads to a large deletion between the target sites (Figure 11), but can also lead to the inversion of the target region with a probability of around 12.9%. Still, most commonly it leads to *indels* at both target sites. Deletions of up to 1 Mb have been generated like this[360]. The efficiency to induce the large

deletions is inversely correlated to the size of the deletion and needs gRNAs, which are highly efficient alone (singleplex).

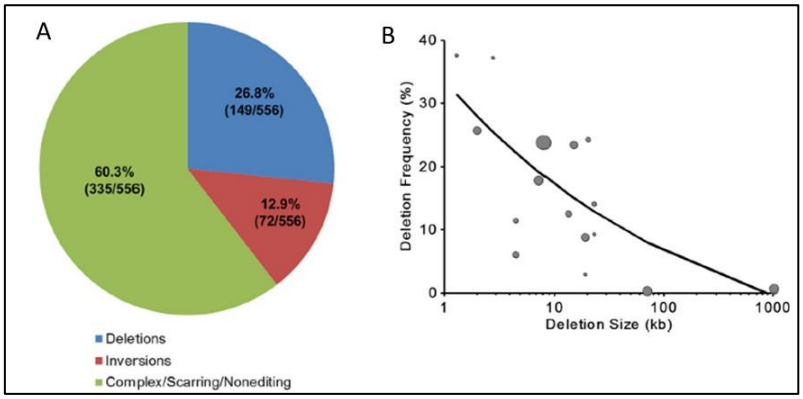



**Figure 11: Frequency and characterization of mutations after multiplex CRISPR/Cas9 cleavage. A) 278 clones were analyzed per allele by Sanger sequencing to determine the frequency of deletion, inversion and *indels*. B) The frequency of large deletions was analyzed in correlation with the deletion size in a non-linear regression model. Figure from: Canver et al. 2014; *J Biol Chem*.**

The introduction of specific point mutations is also possible with the CRISPR/Cas9 system. To achieve this, the DNA repair mechanism of homologous recombination (HR) is taken advantage of[333]. In general, all DSB could be repaired by HR. It is a template dependent mainly error free repair mechanism. However, it is only active in the S/G2-phase of the cell cycle to provide genome integrity during mitosis. There are several natural HR repair mechanisms (Figure 12), but the double strand break repair (DSBR) is the most common one[351]. Usually, the sister chromatid serves as a template to repair the DSB. After the DSB, a single strand (ss) 3' overhang is generated by the Sae2/CtIP nuclease. The single strand induces the invasion at the homologous region of the sister chromatid. A D-loop is generated, where the second single strand can base pair with the sister chromatid, too. The 3' end of the broken strand gets amplified till the 5' end of the other end is reached. Both ends are re-ligated, which can rarely lead to small mutations. The two sister chromatids are separated by the topoisomerase leading to the two original chromatids, or by inducing a crossover between the two chromatids leading to the translocation of the two arms of the chromatids. Alternatively, the synthesis dependent strand annealing (SDSA) relies on the same principle, with the difference that only one strand invades the sister chromatid, gets amplified and then reanneals with the original strand. A crossover between the two chromatids is here very unlikely. If the second strand is not available, the break induced replication (BIR) is initiated. Therefore, the invading strand gets replicated entirely. This can lead to the loss of heterozygosity and it is associated with mistakes and mutations[361].

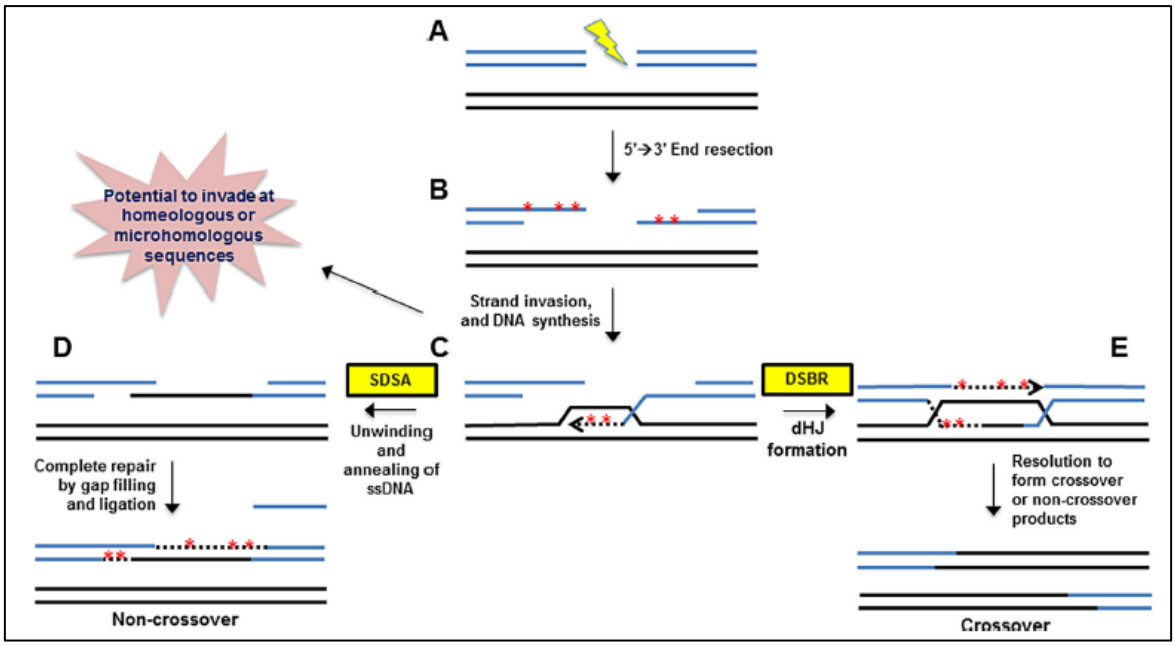



**Figure 12: Homologous recombination (HR) DNA repair mechanism. A) The introduction of the blunt end DNA double strand break (DSB) by the CRISPR/Cas9 system. B) HR depends on the base pairing of the broken DNA with a template DNA strand. Therefore, 3' overhangs are generated by the resection of the blunt end in 5'-3' direction. C) The 3' overhang subsequently invades the template DNA strand, which is most often the sister chromatid. D) During the synthesis dependent strand annealing (SDSA), the invading strand is template-dependently extended until it can dissolve and reanneal with the original strand. E) For the double strand break repair (DSBR) both 3' overhang anneal with the template DNA. The separation of the two DNA strands needs to be resolved by the topoisomerase to keep the integrity of both strands, or otherwise leads to a translocation (crossover) of the two chromatids. Figure from: Rodgers et al. 2016; *J Cell Physiol*.**

For genome editing and the introduction of a specific mutation, the researcher has to provide a donor DNA template. The insertion of whole genes or larger sequences depends on a template, which requires at least 400 bp long homology arms on both site of the insert sequence to induce the HR[362]. The generation of a point mutation is easier. It has been shown that it is highly efficient to provide a single strand oligonucleotide fragment containing the point mutations complementary to the non-target strand, because after the cleavage of the DNA, the Cas9 releases at first the non-target strand, allowing the invasion of the single stranded donor DNA. The general efficiency of specific mutagenesis dependent on HR is around 15%[358]. However, HR is only active in dividing cells, which narrows this application to specific highly proliferative cell types.

To conclude, the CRISPR/Cas9 is an easily feasible gene editing tool, which outrange its predecessors like TALEN and ZFN, in cost and time efficiency as well as in target efficiency. It is rapidly programmable by the generation of a new gRNA sequence of 21 nucleotides. Several web tools are available today, which helps to design and clone the desired genome editing tool. Open challenges are the efficient delivery and the avoidance of off-target effects, which will be discuss in the following chapters.

## 2.3 Specificity of the CRISPR/Cas9 system

The CRISPR/Cas9 system, as the easiest of the above-mentioned genome editing tools, has a great potential for the usage in clinical applications. In this context, a very high specificity is necessary to avoid severe side effects, which can be caused by mutations induced by the CRISPR/Cas9 system cutting at an off-target site. Off-target sites are sequences with a high homology to the original on-target site, which can be identified by the Cas9/gRNA in a less frequent manner. Those

mutations can have different consequences depending on the affected gene or region[363]. In the worst case, it might lead to the transformation of the cell and the development of a tumour. Other outcomes could be chromosome translocations and the dysregulation or disruption of a gene.

Several factors influence the specificity of the CRISPR/Cas9 system, which have to be taken in consideration for the design of a gRNA. In this chapter, I will focus on the most commonly used CRISPR/Cas9 system of *Streptococcus pyogenes*. The target sequence consists of a 20 nt protospacer, which is detected by the gRNA, and the PAM, in this case NGG[354]. The optimal gRNA would target a unique sequence with minimal homologies in the target genome. Detailed analysis of the binding specificity of the gRNA and Cas9 on-target regions discovered that the Cas9 undergoes a conformational change, when the gRNA matches perfectly with the target region[364]. Homologues regions with a few mismatches in the target region could be cut by the gRNA/Cas9, if the homology is still strong enough to induce the conformational change. Several studies tested the tolerance of the gRNA/Cas9 complex to mismatches between the target region and the gRNA. A, so called, *seed sequence* has been identified, which does not allow any mismatches[365]. It includes the first 8-14 nt proximal to the PAM (Figure 13 A). In the PAM distal region, mismatches are better tolerated and can lead to off-target activity. The off-target activity is not only correlated to the position in the target region but also to the type of mismatch, the number of mismatches and if they are separated or adjacent. Therefore, mismatches of the RNA:DNA base pairing by rC:dC have a stronger impact on the Cas9 activity than other nucleotide mismatches. Up to 7 mismatches can be tolerated[364], but the increasing number of mismatches is inversely correlated to the cleavage activity of the Cas9, whereas adjacent mismatches reduce the cleavage efficiency much stronger than several single mismatches[366].

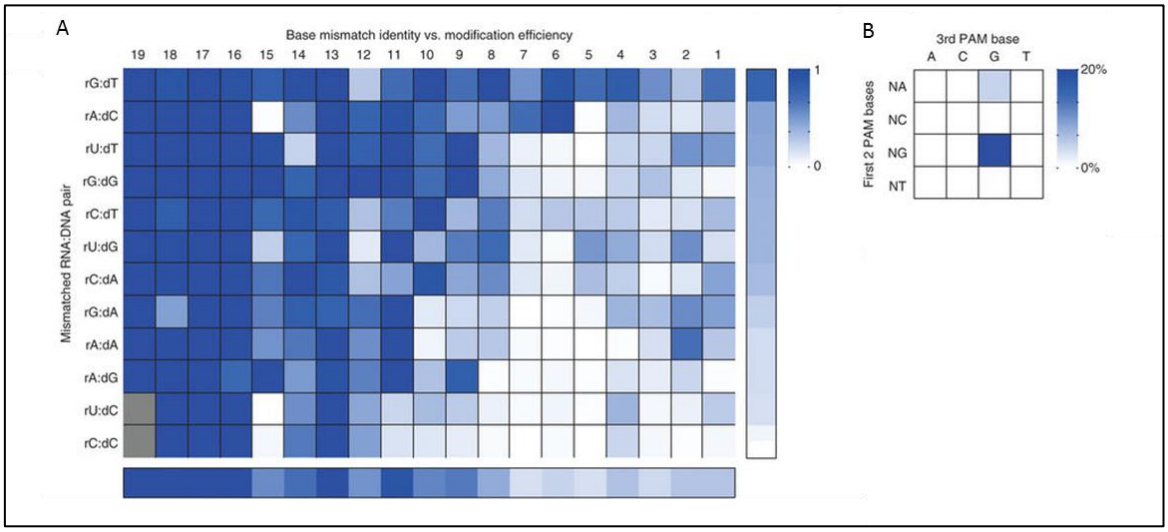



**Figure 13: Tolerance of mismatches in the target region. A) Heat map of the T7 assay analysis of gRNA/Cas9 with single mismatches at different position in the gRNA. PAM proximal mismatches abolish the cleavage efficiency nearly entirely. RNA:DNA mismatches of rC:dC have the strongest impact on the cleavage efficiency (increasing cleavage efficiency normalized to the original gRNA from white: no cleavage to blue: maximal cleavage). B) T7 assay analysis of alternative PAMs. Natural PAM of NGG is correlated with high cleavage efficiency. Low efficiency with the alternative PAM NAG is also detected (increasing *indel* frequency from white 0% to blue 20%). Figure from: Hsu et al. 2013; *Nat Biotechnol*.**

Furthermore, the gRNA or target DNA strand can build bulges (nucleotide insertions), which can be tolerated proximal and distal of the PAM[366]. The Cas9 is more sensitive to the DNA bulge (insertion of a nucleotide at the target strand) and only tolerates 1 nt bulge. In comparison, RNA bugles of up to 5 nucleotides still allow Cas9 activity.

The spacer region is not the only one being associated with off-target activity, the PAM region can be also associated with off-target activity of the Cas9. The first interaction with the DNA is mediated by the Cas9 detecting the PAM, usually NGG. But it has been discovered that the Cas9 can bind to alternative PAMs, NGA (and NAG) with a reduced probability[363,365] (Figure 13 B).

Today a lot of different web tools are available to design gRNAs and predict their off- and on-target potential[367]. They use an input DNA sequence of the target gene to design a gRNA and analyse this gRNA for off-target effects on a reference genome. Many factors should be taken into account to gain a highly specific gRNA. Obviously, general factors like CpG content and secondary structure of the gRNA are important[366]. Moreover, the off-target prediction should consider the number, position, type and nature (single/adjacent) of mismatches, the possibility of bulges and the alternative PAMs[363,365–367]. For off-target, and more importantly for on-target activity, the chromatin condensation can strongly interfere with the target efficiency and should be taken into account, if possible[359,368]. Finally, on-target activity can be affected by single nucleotide polymorphism between different individuals[367].

The choice of one of the different web tools strongly depends on the project and organism to work with. Simple knock-out experiments need different features than multiplex gene deletion experiments. To make a valid choice, addgene offers an interactive comparison table including 27 web tools available at the moment (CRISPR Software Matchmaker).

In this thesis, I worked with two different web tools. CRISPR design (http://crispr.mit.edu/) from Zhang lab, MIT was one of the first web tools online. After providing a target sequence of several hundred bp, it calculates possible gRNAs. The gRNAs are ranked by their off-target potential based on number, position and nature of mismatches. Later on, new web tools were released with more extended features, thus I continued working with the CRISPOR web tool (http://crispor.tefor.net/). It has a more detailed specificity analysis. gRNA selection is done considering microhomologies to increase the probability of frame-shift mutations. Off targets are calculated based on mismatches, alternative PAMs and the position in the genome (exon, intron, intergenic). The different scores lead to a more substantial gRNA selection.

### 2.3.1 Modification of the Cas9 and gRNA to improve the specificity

With the progression of knowledge about the action of the Cas9 and gRNA to interact with the target DNA and induce cleavage, modifications on both components were applied to increase the specificity. There are three major ideas to improve the specificity (Figure 14): (I) adjusting the gRNA/target DNA (tDNA) interaction to be more sensitive to mismatches, (II) engineering a Cas9 with reduced interaction with the tDNA to increase the sensitivity of the gRNA/tDNA interaction and (III) modifying the Cas9 cleavage activity to function as a dimer. Furthermore, there are some other Cas9 variants, which also improve specificity (IV).

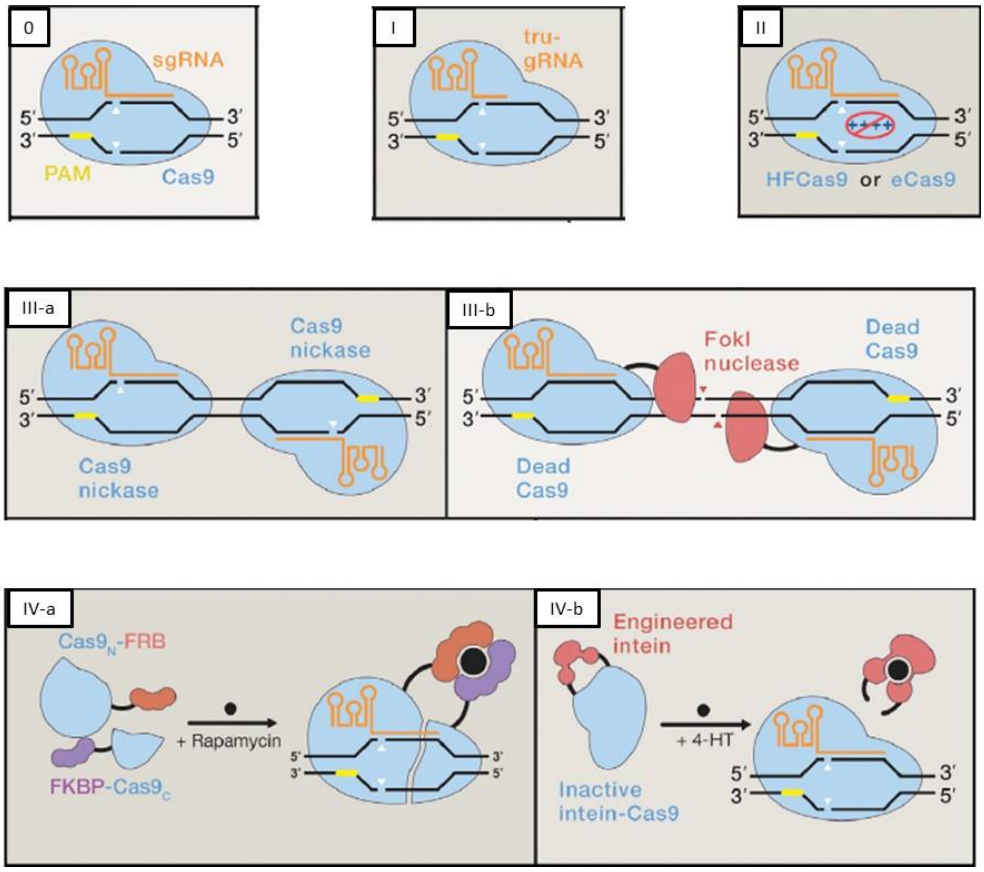



**Figure 14: Modifications on the CRIPSR/Cas9 system of *streptococcus pyogenes* to enhance the specificity. 0: Wild type gRNA/Cas9. I: Truncation of the gRNA structure by 2-3 nt reduces the total gRNA/tDNA binding energy, which leads to a higher sensitivity to mismatches. II: Attenuation of the Cas9 binding to the target DNA strands increases the importance of the perfect base pairing of the gRNA/tDNA to generate a stable R-loop for the induction of the DSB. III: Dimerization dependent approaches of the CRIPSR/Cas9 system increase the specificity, because the monomer alone cannot cause significant damage to an off-target site. III-a) The duplex strategy of the Cas9 nickase generates the single strand breaks leading in close proximity to a DSB with overhangs. III-b) The cleavage deficient Cas9 fused to the nuclease domain *Fok*I. *Fok*I is only a functional nuclease after dimerization, the same principle as TALEN and ZFN. IV: Inducible CRISPR/Cas9 systems have reduced off-target effects, because the exposure of the functional gRNA/Cas9 complex is limited. IV-a) The split-Cas9 is separated in two subunits, which are ligated to the domains FRB and FKBP, which dimerize in the presence of rapamycin. The dimerization reconstructs the functional gRNA/Cas9 complex. IV-b) The Cas9 is fused to intein, which inhibits the nuclease activity of the Cas9. 4-hydroxtamoxifen induced the removal of intein and releases a fully functional gRNA/Cas9. Figure from: Komor et al. 2017; *Cell*.**

(I) The advantage of modifying the gRNA is that it is easy and fast to apply. In a first attempt, the length of the protospacer region was changed. Surprisingly, the prolongation of the region failed to improve binding specificity, because it was processed *in vivo* to the typical 20 nt length[347]. However, the truncation of the protospacer domain successfully reduced the off-target activity by 5000 fold[369]. The truncation reduces the binding energy of the gRNA/tDNA and therefore increases the sensitivity to mismatches. A minimal length of 17 nt is necessary to mediate normal on-target activity (Figure 14 I). gRNAs shorter than 15 nt lead to a dead gRNA and cannot induce the conformational change to induce DNA cleavage[370]. Furthermore, the scaffold domain of the gRNA can also be adjusted[371]. The extension of the gRNA scaffold improves on-target efficiency significantly. Optimal is an extension of 5 bp at the RNA duplex for the interaction with the Cas9. Synergistic with the extension is the reduction of the T stretch between the scaffold region and the protospacer domain.

The four Ts are an RNA Polymerase III pause signal. The substitution of only one T to G/C efficiently enhances the on-target activity. Stronger on-target activity increases the specificity (on-target/off-target ratio) of the CRISPR/Cas9 system and allows then a shorter exposure of the gRNA/Cas9 to the cells or a reduced dose, which would further reduce the off-target activity. Therefore, the delivery of the gRNA/Cas9 can have impact on the specificity of the system as well. The transfection of an expression plasmid leads to Cas9 and gRNA expression usually under a highly active promotor for a few days. But, direct transfection of the Cas9 mRNA and the gRNA reduces the exposure time of the CRISPR/Cas9 system to the target genome and prevents off-target activity. Unfortunately, the on-target efficiency is also reduced, because the gRNA stability is very short and is already partly degraded before the Cas9 protein can be synthesised in the cell. Hendel *et al.* chemically modified the gRNA to ameliorate its stability and gained much higher on-target efficiencies[372]. Moreover, he validated that the direct delivery of the Cas9/gRNA ribonucleoprotein complex to the cells has an even higher on-target activity. The chemical modification of the gRNA is only possible, if a direct delivery of the *in situ* synthesised gRNA is chosen.

(II) Beside the gRNA enhancements, two groups have developed off-target reduced Cas9s. In 2015, the eSpCas9 was developed from the Zhang lab[373]. They validated their hypothesis, that the cleavage efficiency of the Cas9 depends on the separation of the target DNA strands. This is dependent on the Cas9 helicase activity by the interaction with the non-target strand and the gRNA/tDNA base paring. The attenuation of the helicase domain should reduce the force of the helicase to separate the DNA double strand and make the interaction of the gRNA/tDNA more important for this action. Zhang and his team remove some of the positive charged amino acids for the non-target DNA binding groove and discovered two eSpCas9 (1.0: K810A / K1003A / R1060A and 1.1 K848A / K1003A / K1060A) with an improved specificity profile (Figure 14 II). They have the same on-target efficiency, but their off-target probability is reduced to less than 0.2% and is therefore more specific than the truncated gRNA. The other Cas9 variant was published in 2016 by Kleinstiver *et al.*[374]. His modification idea, similar to the eSpCas9, reduced the Cas9 force to open the DNA double strand and refined the relevance of the gRNA/tDNA interaction. Therefore, the interaction domain of the Cas9 with the target DNA strand was weakened (Figure 14 III-a). The Cas9-HF (N497A / R661A / Q695A / Q926A) has a similar on-target efficiency than the wild type (WT) Cas9 but nearly no off-target activity. Both improved Cas9s are not compatible with the truncated gRNA, because such a complex strongly reduces the binding energy between the target DNA and the gRNA/Cas9 complex and the interaction is not stable enough to induce the cleavage.

Further detailed analysis of the two off-target improved Cas9s disproved the hypothesis that the mutations reduce the binding force of the Cas9 to the target DNA. In contrary, it has been discovered that both Cas9s interact with the same binding energy with the DNA as the WT Cas9 [350]. However, the mutated Cas9s are resting in the inactive state after gRNA/tDNA base pairing at the off-target position unable to undergo the conformational changes necessary to activate the nuclease domains HNH and RuvC. Moreover, with the increasing knowledge that the REC3 domain is responsible for the gRNA/tDNA base pair recognition and the induction of activating conformational change in the REC2 domain, a third off-target improved Cas9 has been developed. The so-called hyper accurate Cas9 (hypaCas9) has 4 substitutions in the REC3 domain increasing the threshold for the REC3 conformational change[350]. Its on-target efficiency is comparable to the WT Cas9 and its off-target activity is even lower than the eSpCas9 and the Cas9-HF1.

One further Cas9 modification was discovered to enhance the specificity. The PAM recognition of the Cas9 to NGG is leaky and allows the binding to NGA with reduced frequency (Figure 13 B). The amino acid at position 1135 defines the binding to the third position of the PAM. A D1135E mutation significantly improved the binding to NGG and abolished the tolerance to the alternative PAM NGA[375]. This approach can be combined with the Cas9-HF to further optimise the specificity.

(III) The off-target potential of TALEN or ZFN is much lower than for the CRISPR/Cas9 system, because they act as a dimer, while the WT gRNA/Cas9 is active as a monomer. The Cas9 has two nuclease domains, RuvC and HNH, cutting the target and non-target strand, respectively[335]. Shortly after the discovery of the CRISPR/Cas9 system as a genome editing tool, Ran *et al.* created a Cas9 nickase (Cas9n), which only cut one of the DNA strands, or a dead Cas9 (dCas9), which is cleavage deficient[347]. The Cas9 D10A with a knock out of the RuvC domain is the most used one. The Cas9n can reduce the off-target activity of the CRISPR/Cas9 system, when used as a duplex strategy (Figure 14 III-b). If the two gRNAs have target regions in close proximity, the two nicks induced by the Cas9n lead to a double strand break with an overhang and can be repaired as the WT Cas9 by NHEJ. For the efficient induction of *indels* the distance between the two protospacer should be between -4 and 20 nt[347,376]. Moreover, a 5' overhang is much more prone to induce *indels* than a 3' overhang. Therefore, the two Cas9n should be oriented tail-to-tail[376]. Depending on the target gene, the duplex Cas9n strategy can reduce the off-target to 0.1-1% with an on-target efficiency of 14-91%, so that the specificity is 200-1500 fold increased[347,377]. This strategy is inefficient to induce large deletion by multiple target sites[377].

An alternative dimer strategy of the CRIPSR/Cas9 system uses the dCas9 fused to a *Fok*I nuclease domain, similar to TALEN. The *Fok*I/dCas9 can only induce a DSB, if the two *Fok*I domains dimerise (Figure 14 E)[378]. The two target sites in the DNA should be optimally separated by 13-17 nt to allow a perfect dimerization of the *Fok*I domains. The cleavage efficiency is slightly lower than the WT gRNA/Cas9 (3-40%), but the off-target activity is decreased to 0.15%. This can be further reduced by 40% by the combination with a truncated gRNA[379].

(IV) Beside those major Cas9 variants, there are some new ideas to increase the target specificity, which have, however, not yet received much attention in the scientific community. There are two types of inducible Cas9. The intein-Cas9 is fused to the intein domain, which can be spliced of the Cas9 in the presents of 4-hydroxytramoxifen (4-HT) and releases an active Cas9 (Figure 14 IV-b)[380]. The intein domain naturally also reacts with β-estradiol, but a simple point mutation abolished this interaction. Due to the fact that 4-HT has to be present to create the active Cas9, the exposure time of the active gRNA/Cas9 complex to the target genome is reduced giving a 25 fold higher specificity in long term experiments. However, the removal of 4-HT cannot inactivate the gRNA/Cas9 complex, which will be active for several days until its natural degradation in the cell. The other inducible variant is the split-Cas9[381]. The Cas9 was split into two subunits, whereas the C-terminal part is fused to FK506 binding protein 12 (FKBP) and the N-terminal domain combined with FKBP rapamycin binding domain (FRB). In the presence of rapamycin, the two subunits reunite to a fully functional Cas9 (Figure 14 IV-a). To prevent a background activity of spontaneously assembled split-Cas9, the two subunits are localized in different cellular compartments. The Cas9(N)-FRB is retained in the cytoplasm by a nuclease export signal (NES). In contrast, the Cas9(C)-FKBP contains two NLS inducing the transport in the nucleus. After dimerization, the split-Cas9 is localized in the nucleus,

because the two NLS are stronger than the NES. These approaches can be useful, if the gRNA/Cas9 sequence is constantly expressed in a stable cell line by, for example, by lentiviral vector delivery.

Finally, there is the Cas9-DBD[370]. It has an attenuated PAM binding domain and is fused to a programmable DNA binding domain (DBD) from ZFN or TALEN. Their longer DNA recognition domains can provide an up to 150 fold higher specificity. However, this modification makes the CRISPR/Cas9 system less flexible and again dependent on protein engineering or complex cloning. It is not feasible for most of the applications.

These above presented advanced CRISPR/Cas9 systems have all pros and cons and the perfect selection of one of those modifications also depends on the type of application. Personally, I prefer the modifications on the gRNA. They have a strong impact on the specificity and are easily and rapidly added to already validated CRISPR/Cas9 systems. I expect, that the optimised gRNA in combination with the WT Cas9 or Cas9 D1135E (improved PAM specificity) are suitable to provide a highly specific genome editing tool.

### 2.3.2 Detection of off targets

For clinical application is the unbiased, highly sensitive detection of off-target sites very important. Before injecting modified cells into patients, it should be ensured that the off-target potential is as low as possible and that there is no risk of tumour development. Also in research applications, it can be of high interest to identify off targets to verify that the observed effect is due to on-target gene modifications and not an artefact based on an off-target effect.

On-target and off-target activities rank from 3-91% to less than 0.1%, respectively[377,378]. While the on-target effects are usually very easily detected by T7 assay (see chapter II-2.2), the sensitivity of the T7 assay to detect off-target effects is often too low (1%)[367]. Therefore, highly sensitive and less biased methods need to be applied to reliably detect off targets in the whole target genome. In fundamental research or in proof of concept applications, targeted sequencing is often the primary choice. It depends on the prediction of the off-target regions based on a web tool and the amplification of the few most probable off-target sides[363]. The amplicons are either analysed by basic Sanger sequencing or by the more sensitive next generation sequencing (NGS). Those methods are biased, because they only screen a limited number of sites in the whole genome and do not give information on the entire integrity of the genome. Moreover, the PCR amplification is a potential source of error and could lead to an overestimation of off-target effects[363].

There are several methods available, which are not biased by off-target prediction software, and screen the whole genome, which are in detail compared in the several reviews[363,364,367,370]. Whole genome sequencing is the most accurate method. It simply analyses the whole genome in comparison to the wild type cell and can identify the smallest changes. Detected mutations can then be analysed for homologies with the gRNA to ensure that it is a real off-target site and not a sequencing mistake or a natural single nucleotide polymorphism (SNP). This method is perfectly suitable for the analysis of clonal cell populations but not for bulk analysis, because each cell needs to be analyse separately. Moreover, it cannot detect off-target sites, which are perfectly repaired[367]. The assessment of off-target sites independent of mutations can be done by ChIP-seq (chromatin immunoprecipitation and NGS). It uses the interaction of the Cas9 with the genome and co-precipitate the DNA by the Cas9. To be able to analyse the bound DNA region, usually the

dCas9 is used for the off-target detection. Nevertheless, this method tends to overestimate the number of off-target sites, because the Cas9 also binds to regions with low homologies, which cannot induce the conformational change to induce a DSB[363]. In contrast to that, there are two methods, which detect off targets based on the induced DSB. Integrase-deficient lentivirus integration identifies off targets by the integration of a lentiviral genome in a DSB based on NHEJ. However, the integration efficiency is relatively low so that the sensitivity of this method is only 1% leading to the underestimation of the off-target potential[363]. GUIDE-Seq (genome-wide unbiased identification of DSB enable by sequencing) is highly sensitive (0.03 – 0.1%)[364,382]. It uses double stranded oligonucleotides (ODN) to captures DSB, which are subsequently analysed by deep sequencing. Another ODN based method is BLESS (break labelling enrichment on streptavidin and next generation sequencing). However, it only works on fixed cells limiting this method to one specific time point and so underestimates the amount of off targets[370]. In the fixed cells, the DSB are labelled by biotinylated ODN, which can be purified, enriched and sequenced. Those three methods based on DSB capture are biased by Cas9 independent DSB and require intensive analysis of all identified DSB to find homologies with the gRNA to distinguish off-target effects from natural DSB in the cell. Alternatively, the *in vitro* method Digenome-Seq is completely Cas9 dependent. Extracted genomes are subjected to the gRNA/Cas9 complex to fragment the genome. Those fragments are then analysed by NGS. Digenome-Seq has a high sensitivity of 0.1%, but overestimates off-target effects, because the chromatin condensation, which strongly interferes with the Cas9 binding, cannot be considered[367]. Recently, a more sensitive *in vitro* off-target screening method has been described. CIRCLE-Seq (circularization for in vitro reporting at cleavage effect by sequencing) is based on extracted genomic DNA, which is sheared and subsequently circularized[383]. The circular DNA fragments are then subjected to the gRNA/Cas9 complex linearizing DNA fragments containing on-target or off-target sides. Those linearized fragments are then PCR amplified after an adapter ligation and sequenced. This method is even more sensitive than GUIDE-Seq. Moreover, it was used to compare off-target activities in different individuals (2504 persons) and discovered in 50% of the individuals at least one unique variation and in general difference of 2.5% from one individual to another[383].

Finally, there are two less sensitive methods, which only identify a specific type of off-target effects. HTGTS (high throughput genomic translocation sequencing) detect translocations by catching the off-target DSB by the on-target DSB[364]. And transcription profiling recognizes phenotypically changed cells[367].

All those methods have certain disadvantages and most of them are also expensive and time consuming. Off-target analysis should be performed with those highly sensitive methods only in the final physiological target cells. For pre-screenings and tests, T7 assay analysis provides already a basis of information and safety for the most probable off-target sites, but is not sufficient to conclude about the safety of the CRISPR/Cas9 approach. For clinical application, for example adoptive cell therapy, a combination of CRICLE-seq to screen the individual patient for potential off-target site considering SNP and post Cas9/gRNA treatment analysis by GUIDE-Seq on patients cells considering chromatin condensation and epigenetic factors would be very accurate to estimate potential adverse effects.

## 2.4 Delivery of the CRIPSR/Cas9 system

The CRISPR/Cas9 system is a very powerful and highly efficient genome editing tool. The strongest limitation is the efficiency of delivery of the Cas9/gRNA and, if required, donor DNA for

HR. The choice of the right delivery system depends on the target cell type, *in vitro* or *in vivo* applicability and desired time of expression. A large variety of different delivery tools has been adapted for the optimal CRISPR/Cas9 delivery in the past few years, ranging from classical transfection to more complex viral vector systems. In the following, an overview of the most commonly used delivery methods and some new advancements of these systems is given.

In general, delivery strategies are divided in two groups: viral vector delivery and non-viral delivery. The non-viral delivery methods can then be further divided in two categories: chemical and mechanical. For both delivery methods, the Cas9 and gRNA can be transferred as either an expression plasmid, mRNA and gRNA, or as the mature ribonucleoprotein complex (RNP). Those non-viral deliveries have the advantages that they are rarely immunogenic and easy to scale up for large productions[384]. However, most of those methods are only specialized for *in vitro* applications of cell cultures or single cells (zygote, oocyte)[385]. Viral vectors are more suitable for systemic applications and can be optimized to cell or tissue tropism. They provide a prolonged Cas9/gRNA expression, which increases on-target efficiency, but also off-target effects. Moreover, the prolonged Cas9 expression *in vivo* can induce an immune response against the Cas9 protein, which leads to inflammation, tissue damage and the loss of transduced cells[386,387].

### 2.4.1 Non-viral delivery

The chemical methods include different types of transfection based on cationic lipids (lipofection) or polymer based nanoparticles[384]. Lipofection is mainly used to deliver expression plasmids. The cationic lipid interacts with its cationic head with the DNA and then self-assembles to a liposome by its hydrophobic tails. The liposome can penetrate the cell by endocytosis or micropinocytosis[388]. Lipofection is very efficient on some well-established cell lines, but has only a very low transfection efficiency on primary cells[384]. Moreover, it can induce a general inflammation and cytokine release in the cell culture. Extended research on lipid structure gives rise to a big variety of different lipid based transfection reagents, which are adapted to specific cell types and hard-to-transfect cells. So far, only one reagent has been shown to be efficient in *in vivo* delivery. Zuris *et al.* used RNAiMAX to deliver the RNP to the inner ear of mice and detected a 13% knock-out efficiency without tissue damage or immune cell infiltration[389]. A systemic application of lipid based systems is not possible because of the high interaction with serum proteins in the blood. Therefore, polymer based systems are theoretically more suitable. The polymers are positively charged to bind the DNA[384]. Furthermore, they act as a proton-buffer to cover the charge to cross the plasma membrane. The most common polymer is PEI (polyethylenimine), which is already used for decades to transfect cells. Over the time, it has been improved with several modifications, like PEGylation to increase the stability. Other modifications enhance the internalization (folate, Arg-GLy-Asp peptides, and galactose) or mediate receptor dependent entry (antibodies, sugar, peptides). A specific adaptation for the CRISPR/Cas9 system is the so-called nanoclews, which are DNA nanoparticle coated with PEI[390]. The DNA is partially complementary to the gRNA and can pack the RNP in the nanoparticle. But so far, this nanoclews are highly immunogenic. The problem with the administration of polymers is the bad biodegradability, which limits the usage *in vivo*. New nanomaterial, specially chosen for their high biodegradability, has been developed[391]. The amino-ester derived lipid-like nanoparticles can be digested by the esterase with a half-life of 24 h. After intra venous (iv) injection, it is cleared from the blood within 1 h and from the liver in 6-24 h without causing organ damage while still being efficient enough to transduce 18% of the cells in a xenograft tumour mouse model. The field of *in vivo* applicable nanoparticles is further advancing with ZAL (zwitterionic amino lipids) or iTOP

(induced transduction by osmocytosis and propane betaine)[388]. Beside the addition of lipids or polymers to deliver the Cas9/gRNA, it is possible to directly modify the RNP complex itself. So-called cell penetrating peptides (CPP) are known to be able to penetrate the plasma membrane[392]. They are short cationic peptides of less than 30 amino acids and have a high affinity to the lipid bilayer. Fused to a protein, CPPs can delivery proteins much bigger than themselves inside the cell. The Cas9 has been fused to a poly arginine (Cas9-mR9) and the gRNA has been complexed to a poly arginine peptide to be transferred to hard-to-transfect cells like embryonic stem cells[385]. Furthermore, the fusion of four further nuclear localisation signals of the simian virus 40 (SV40) to the N-terminus of the Cas9 (4x NLS-Cas9-2x NLS) also serves as a CPP[393]. It enhances the direct cell delivery ten-fold. This Cas9 has been tested *in vivo* for a local application in the brain of mice. Transfection efficiency was dose dependent. Nearly all different neural cell types were transfected, while astrocytes were not sensitive for the delivery. Systemic administration has not been tested.

Contrary to those chemical transport systems are the mechanical delivery methods. Those are not based on charge interactions or endocytosis mechanisms. Basically, the plasma membrane is disrupted, so that small transient holes arise, through which the CRISPR/Cas9 approach can enter the cell. The most widely used method is electroporation. Exposure of the cells to high voltage induces the formation of transient holes, which allows the entry of negatively charged material into the cell[388]. Classically, expression plasmids are delivered by this method and enable a transgene expression for 3-4 days. However, with increasing knowledge of the off-target activity of the CRISPR/Cas9 system, the administration of the mature RNP is becoming more in focus. RNP delivery enables the same *indel* efficiency, but the shorter Cas9/gRNA expression (<24 h) in the cell decreases significantly the off-target effects[358]. Moreover, RNPs are less toxic than plasmids. Electroporation can be adapted to a lot of different cell types by adjusting the voltage and time of exposure. It has also been tested that, if the RNP is delivered for homologous recombination (HR), a sequential delivery of two electroporations (one for the RNP and one for the donor DNA) is more efficient[394]. The disadvantage is that the high voltage exposure can induce substantial cell death or phenotypical changes. To avoid the high voltage cell disruption, a new micro fluid device allows the creation of transient holes in the plasma membrane by nano silicon blades[395]. This device was especially designed to transfect hematopoietic stem cells (HSC) with a very high transfection efficiency of 70% without extensive cell death or changes of the phenotype or differentiation potential. Electroporation and the nano silicon blades are adapted to cell suspension delivery. For single cell delivery, like to zygotes or oocytes, it is more common to use microinjection[385]. It is mainly used for the creation of transgenic animals. Finally, hydrodynamic injection is the only mechanical delivery method for *in vivo* genome editing[385]. It is only used for small rodents and probably not applicable for bigger animals. During the hydrodynamic injections, a large volume of DNA suspension is rapidly administered in the tail vein. This induces liver damage and concomitant DNA uptake. It only allows the delivery to the liver with an efficiency of 5-27% of *indels* or 1/250 cells for HR. However, the hydrodynamic injection is very rough on the animal and might not be the optimal delivery method *in vivo*. Viral delivery methods are gentler for the animals.

### 2.4.2 Viral delivery

Viral delivery in general is a delivery method, which gives rise to a prolonged transgene expression. There are three major viral vector systems: Adenoviral vectors (AV), Adeno-associated viral vectors (AAV) and lentiviral vectors (LV). Adenoviral vectors originate from the

Adenovirus, a linear dsDNA virus. It can deliver the transgene to dividing and non-dividing cells without integrating into the host genome[390]. Its package capacity is relatively large (~ 8 kb) and allows the delivery of the Cas9 and the gRNA, or a donor DNA for HR. AV delivery has been tested *in vivo* for the local muscular application in a mouse model of muscle dystrophy. Problematic is that AVs are immunogenic and can cause severe inflammation *in vivo* leading to the destruction of the transduced cells[386]. In contrast, AAV are nearly not immunogenic[388]. They are small non-enveloped particles with a high infection rate and low oncological risk. The different serotypes of AAVs allow a broad tissue tropism: muscles (serotype 9)[396], brain (serotype 1/2)[385], liver (serotype 2/8)[397] or systemic (serotype 5)[398]. The biggest disadvantage is the small transgene packing capacity of maximal 4.7 kb[388]. This limits the applicability of the CRISPR/Cas9 system dramatically to either dual vector systems, alternative Cas9 orthologues or combinational approach with chemical delivery. First applied was a dual AAV system with gRNA and Cas9 in two separated vectors each with a high transduction efficiency of 90 % and 95 %, respectively[398]. Later on, Chew *et al*. developed a split Cas9[387]. Each subunit can then be delivered by an AAV and leaves enough space for a reporter gene and the gRNA. More detailed analysis of the different Cas9 orthologues revealed the Cas9 of the *Staphylococcus aureus* (SaCas9) as highly efficient, while it is much smaller than the commonly used SpCas9[399]. The SaCas9 is optimal for the usage of AAV delivery, because it enables the co-delivery with one gRNA in the same vector. For HR, the AAV delivery has been tested in combination with lipofection *in vivo*[397]. The AAV was used to deliver the gRNA and the donor DNA, while the Cas9 was transfected by C12-200. The so-called nano.Cas9 particles are stable for 18 days and enable an expression of the Cas9 for 4-14 h after iv injection. The delivery of the Cas9 by nanoparticles and the shortened Cas9 expression reduced off-target risks, while still inducing 6.2 % gene correction in hepatocytes with the AAV delivered gRNA and donor DNA.

Finally, lentiviral vectors (LV) are commonly used for transgene delivery *in vivo* or to generate stable cell lines. Due to its origin from retroviruses, it integrates its genome and transgene in the host genome and induces a stable transgene expression. This feature has the disadvantage that the integration is not directed and can occur everywhere in the host genome. Consequently, LV can be carcinogenic[390]. LVs are produced from three plasmids: two helper plasmids coding for the gag/pol/rev (capsid proteins, reverse transcriptase/integrase and regulatory protein for nuclear export and packing) and the surface glycoproteins, and one lentiviral vector backbone plasmid containing the long terminal repeats (for integration) and the transgene. This last plasmid has a packing capacity of max. 10 kb transgene leading to a very large plasmid[388]. During the production of the LV it can be sensitive to recombination, if the transgene contains repetitive regions. That is in particular a problem for the multiplex strategy with several gRNAs. Modification of the helper plasmid can broaden the cell tropism and give rise to the several LV pseudotypes[400]. Most commonly used is the VSV-G pseudotype, which expresses the glycoprotein of the vesicular stomatitis virus (VSV) at the surface. It has a very general and broad cell tropism, but is less efficient for myeloid cell types. The usage of the hemagglutinin (H) and the fusion protein (F) of the measles virus increases the transduction efficiency of HSC. Furthermore, the mutation of the H protein, responsible for cell attachment, and fusion to a cytokine domain[401] or single chain antibody domain[400,402,403], specifies the cell tropism of the LV to different surface molecules, for example: CD4, CD8, CD133 (HSC) and a lot more. A relatively new LV pseudotype is based on the nipah virus G-protein and F-protein[404]. It transduced ephrin-B2 positive cells, but has a relatively low transduction efficiency for endothelial cells and HSC, which makes it ideal for *in vivo* applications. The LV particle will not be used up by the endothelial cells so that they reach their

target organ in higher quantity and have no potential dangerous side effects on HSC. Taken together, LVs represent a very suitable and adaptable delivery system. Even for the problematic integration, there is an integrase deficient LV (IDLV), which has all the advantages of the LV, but does not interfere with the integrity of the host genome[405]. It has the same transduction efficiency as the WT LV, but less off-target effects, because the transgene expression time is limited. Even if a DSB is induced by the CRISPR/Cas9 system, the integration potential of IDLV in the host genome is not increased (0.8%). Another way to decrease off-target effects is the so-called SLiCES (self-limiting Cas9 circuit for enhanced safety and specificity)[406]. It is a CRISPR/Cas9 system consisting of the Cas9, the target gRNA and a gRNA anti-Cas9. After lentiviral transduction, it allows an active Cas9 expression for two days, which is usually enough to achieve a high gene editing efficiency. The self-targeting gRNA prevents prolonged Cas9 expression and subsequently off-target activity. This system is ideal for *in vivo* genome editing, because it uses the high efficiency of LV delivery and limits its activity to a timeframe comparable to non-viral delivery systems.

Beside these well-established viral vectors, there are some new vectors developed. The baculoviral vector, originates from a dsDNA virus and has the big advantage of a very high packing capacity of up to 38 kb[407]. It is non-integrative and has a delivery efficiency of 18-100% depending on the cell type. A very different type of viral delivery is the sendaiviral vector (SeV)[408]. The Sendai virus is a RNA virus of the family of *Paramyxoviridae*. It delivers only RNA to the cell, which eliminates the unwanted and potentially carcinogenic integration of the transgene/vector in the host genome. The SeV allows the co-delivery of the Cas9-P2A-GFP and the gRNA in a single vector. Furthermore, it is very efficient to transduce myeloid, hard-to-transfect cells, like monocytes (up to 99%).

### 2.4.3 Transgene organization of plasmid and viral vector delivery

All the above introduced delivery methods have their advantages and disadvantages as well as some limitations, like cell tropism, toxicity, large scale production (especially for the viral vectors) and packing capacity. A standard expression plasmid has no real precise transgene limit, but it is commonly known that large constructs are produced in lower quantities and are less efficiently transferred to cells. Viral vectors are more sensitive to a specific limit of packing capacity to maintain their genome and particle structure and allow high titre productions. For these reasons, several advances have been made to improve the gene and promoter organisation of the CRISPR/Cas9 system. Basically, the Cas9 is regulated by a highly active RNA polymerase II promoter (commonly CMV or SFFV) and the gRNAs are transcripts from a RNA polymerase III promoter to avoid poly-A tailing and potential splicing mechanisms[409]. The first multiplex approach contained individual promoters for each gRNA. Kadabi *et al.* verified the efficiency of four different promoters for gRNA expression (human and murinU6, 7SK and UbC). Problematic for the multi-promoter system is the large final construct, which limits the delivery efficiency. To minimize the size of one gRNA cassette, Mefferd *et al*. proposed a small tRNA promoter system (70 bp promoter and 101 bp tRNA-gRNA transcript)[410]. He used an original viral mechanism, where a tRNA was excised from a transcript by the tRNAseZ. Fusing the tRNA and gRNA sequences together led to a transcript, which was processed by the cellular tRNAseZ and released a functional gRNA and tRNA. The efficiency depends on the tRNA (best: GLN, GLY, GLU, PRO) and the protospacer sequence of the gRNA (avoid high C:G contents). An alternative approach has been developed by Yoshioka *et al*., which used a mono-promoter system to express several gRNA and the Cas9[411]. To allow the gRNA expression from a RNA polymerase II promoter, as required for the Cas9 expression, the gRNA has to be excised out of the transcript. Therefore, Yoshika

flanked the gRNA sequence by self-cleaving ribozymes (RGR), which released a full functional gRNA. The fusion of the Cas9 sequence to an IRES (internal ribosome entry site) allowed the expression of the Cas9 for the same construct as the gRNAs. Those optimisations of the gene structure of the Cas9/gRNA are very helpful to create an efficient viral vector, with a smaller size to allow a high titre production.

The choice of the correct delivery system for the CRISPR/Cas9 system depends on the type of application. To test if a given gRNA achieves a knock-out or knock-in in a cell line (proof-of-principle), non-viral delivery methods are usually chosen. They are cheap and rapidly constructed. For primary cells, the right choice of delivery is more difficult. Most chemical transfection reagents are not efficient enough or have toxic side effects. Electroporation is the better choice in that case. Low toxic and highly efficient conditions can be found with some effort. The new micro fluid device with nano silicon blades is an interesting alternative. For *in vivo* delivery, viral delivery systems are the most efficient. The choice in this case depend on the desired target tissue, the Cas9 orthologue and the desired time of Cas9/gRNA expression.

## 2.5 Cas9 variants and orthologues

### 2.5.1 Cas9 variants

Since the adaptation of the CRISPR/Cas9 system to a genome editing tool, several modified Cas9 proteins have been designed to increase the specificity, alter the target sequence, improve the delivery or change the function. All these Cas9 variants have been summarized in Table 1: Cas9 variants. Most of these modifications are done with the SpCas9, but some exceptions are highlighted in the Table 1. Cas9s for improved target specificity (eSpCas9, Cas9-HF, inducible split Cas9, Cas9n, Cas9-FokI) or smaller Cas9 variants for more efficient delivery (split Cas9) have been discussed in more detail in the previous two chapters (II-2.3.1 off target and II-2.4 delivery). Here, I focus on the advanced function of the CRISPR/Cas9 system by Cas9 fusion proteins.

Early after discovery of the genome editing tool CRISPR/Cas9, Jinek *et al*. discovered that two substitutions in the nuclease domains (D10A= RuvC, H840A=HNH) lead to a nuclease deficient Cas9[335]. Further on, this inactive Cas9 was termed dead Cas9 (dCas9) and is the origin of all the Cas9-fusion variants, because the dCas9 has the possibility to guide fused effector proteins to a site-specific target sequence. This feature has been widely used for gene regulation systems. Firstly, the dCas9 itself downregulates the expression of a target region by sterical blockage of transcription factor (TF) binding and the initiation of transcription or by the abortion of elongation[412]. This feature is called CRISPR interference (CRISPRi) and leads to similar effects as RNA interference (RNAi)[413]. While CRISPRi regulates the gene expression at the transcriptional level, RNAi targets the mRNA transcripts and prevents the translation. As for RNAi, the inhibition by the dCas9 alone only reduces the gene expression, but does not result in a full blockage. Therefore, the dCas9 was fused to the Krüppel-associate box (KRAB), a well-known repressor[414]. If the dCas9-KRAB fusion is guided by a site-specific gRNA to an enhancer or a promoter region, it silences the concomitant gene expression. dCas9-KRAB is a much larger complex and can prevent even more efficient the TF binding to the enhancer/promoter sites. Moreover, it induces the methylation of the target region by recruiting the methyltransferase SEDBD1. Finally, the chromatin accessibility is decreased in a region of around 25 kb, enabling the inhibition of whole gene complexes. Off-target methylations can be detected in proximal regions, but the total transcription is only target-site-specifically downregulated. A more precise gene regulation can

be achieved by the dCas9 fusion to a methyltransferase. The first one was dCas9-DNMT3a, which enables the methylation of a much smaller range of 150 bp compared to the dCas9-KRAB[415]. However, the methylation of the target region took up to 7 days to reach its maximum. Two further dCas9-methyltransferases have been developed and induced a more site-specific methylation. The dCas9-MC/MN is a split version of a methyltransferase (MT) originating from *E.coli,* which was earlier used for the fusion to TAL[416]. One of the MT subunits is fused to the C-terminus of the Cas9 (MC), while the second subunit (MN) is unbound. The co-expression of both subunits is necessary to gain a site-specific methylation. The binding of the dCas9 to the target region induces a conformational change and enables the dimerization of MC/MN to a functional MT. This mature complex methylates site-specifically the target gene at a methylation signal 8-20 nt downstream of the PAM. The dCas9-MQ1 is an already active MT[415]. A substitution Q147L in the MT reduces off-target methylation of nearby sequences leading to a significantly smaller methylation range of 30 bp. Moreover, the dCas9-MQ1 is much faster than the dCas9-DNMT3a and reaches its maximal methylation already 24 h after transfection. All three dCas9-MT can be used for gene regulation studies and the influence on cell differentiation and early development of a foetus (mouse).

Beside gene repression, the CRISPR/Cas9 system can also be used as a TF for gene activation. Therefore, the dCas9 is fused to several copies of the activator protein 16 of the Herpes virus simplex (VP16)[417]. More copies of the VP16 induce a stronger activation. Commonly used is the dCas9-VP64[412,418] with 4 copies of the VP16, but up to 12 copies have been tested[417]. The programmed site-specific gene activation can be used for controlled cell differentiation or induced pluripotent stem cell (iPSC) generation. For better regulation of this system, Balboa *et al*. developed a chemically inducible system, where the dCas9 was fused to VP48 and a destabilization domain (dihydrofolate reductase derived destabilization domain = DD). The DD-dCas9-VP48 depends on the presence of Trimethoprim (TMP). In combination with the commonly used TetON system he created an elegant two-step differentiation system to generate pancreatic progenitor cells. Even stronger transactivation can be achieved by the dCas9-VPR which not only contains the VP64 but also 2 additional transcription activators (p65 and EBV R). Alternatively, the dCas9-p300 can acetylate the histone 3 and increases gene expression similar efficiently as the dCas9-VPR[419].

Finally, that dCas9 can be used for genome loci imaging. Firstly used was the dCas9-GFP fusion protein to detect repetitive regions with one gRNA in the human genome, like telomers[420,421]. To gain a clear signal of a genome loci at least 36 dCas9-GFP proteins need to be bind to the target region. Consequently, a multiplex design of at least 36 gRNA to target non-repetitive genome regions is required. The disadvantage compared to the dCas9-GFP is the limit to only one locus at the same time and the fixation to GFP. Later on, the development of dCas9-HALO broadened the possibilities[422]. HALO is an adaptor domain for a specific type of fluorochromes, which can be assembled *in vitro*, opening opportunities for multicolour applications with different dCas9-HALO/gRNA complexes for different gene loci. The authors promoted their discovery CASFISH (Cas9-mediated fluorescence in situ hybridization) as a cheaper and faster variant of the commonly used FISH technique. The genome loci imaging CRISPR/Cas9 approaches can be used for the analysis of telomere length, telomere and chromosome dynamics by live cell imaging or the detection of gene organizations.

**Table 1: Cas9 variants**

| NAME | | FEATURE | MODIFICATION | EFFECT/USAGE |
|---|---|---|---|---|
| **Cas9** | | Alternative PAM | Substitutions | Higher variability for target sequence in human genome |
| | VQR | NGAN / NGCG | D1135V/R1335Q/T1337R | |
| | EQR | NGAG | D1135E/R1335Q/T1337R | |
| | VRER | NGGG | D1135V/G1218R/R1335E/T1337R | |
| **SpCas9-HF1** | | Reduced interaction with Cas9 and target strand | N497A/R661A/Q695A/Q926A | Less off target |
| **SpCas9-HF2** | | Increased PAM specificity to NGG | D1135E | Less off target |
| **eSpCas9** | 1.0 | Attenuation of helicase function, reduced interaction with Cas9 and non-target strand | K810A/K1003A/R1060A | Less off target |
| | 1.1 | | K848A/K1003A/R1060A | |
| **Cas9-pDBD** | | More stringent DNA binding by Cas9 | Attenuated PAM interaction domain Fused to DNA binding domain (ZFN; TAL…) | Less off target, Independency of PAM |
| **Split-Cas9** | | Two subunits of Cas9 | Split Cas9 in subunits and fused the subunit the intein | Better delivery of small subunits by AAV |
| **Split-Cas9** | | Rapamycin dependent assembly | Split Cas9 in subunits and fused to FKBP and FRB | Regulation of Cas9 activity reduce off target |
| **dCas9-FokI** | | DSB | Fused to Fok-I endonuclease domain (functional as dimer) | Less off target |
| **Cas9n (nickase)** | | Single strand break = nick | Cas9-D10A: RuvC deficicent Cas9-H840A : HNH deficient | Reduce off target effects (50-1500x) |
| **dCas9** | | CRISPRinterference | Substitution D10A and H840A (nuclease deficient) | Gene downregulations (1000x) |
| **dCas9-KRAB** | | Repressive histone modifications of enhancer and promotor | Fused to Krüppel-associated-box | Gene downregulation |
| **dCas9-DNMT3a** | | Programmed methylation | Fused to catalytic domain of DNMT3a methyltransferase | Gene regulation, differentiation and early development regulation |
| **dCas9-MC /MN** | | | Fused to split subunit of methyltransferase and co-delivery of second subunit of methyltransferase | |
| **dCas9-MQ1, dCas9-MQ1 (Q147L)** | | | Fused to MQ methyltransferase Q147L substitution for less off target | |
| **dCas9-VP16x VP48 VP64 VP96 VP192** | | Transcriptional activation | Fused to HSV protein 16 domain (3x, 6x or 12x) | Gene activation iPSC generation |
| **DDdCas9-VP48** | | Inducible transcriptional activation | Fused to 4x HSV protein 16 domain Fused to DHFR-DD (dihyrdofolate reductase derived destabilization domain) | Controlled cell differentiation (example: pancreatic progenitor cells) |
| **dCas9-VPR** | | Transcriptional activation | Fused to 4x HSV protein 16 domain, p65 (NF-κB trans-activating subunit) and EBV R trans-activator | Strong gene transactivation, HIV reactivation |
| **dCas9-p300** | | Histone 3 actetylation = Transcriptional activation | Fused to p300 histon acetyltransferase | Strong gene transactivation, HIV reactivation |

| | | | |
|---|---|---|---|
| **dCas9-GFP** | Genome loci imaging | Fused to GFP | Gene copy number detection, Telomer length analysis, Telomer and chromosome dynamics (time laps), Organisation of genes |
| **dCas9-HALO** | CASFISH (Cas9-mediated fluorescence in situ hybridization) | Fused to adapterdomain HALO to couple fluorescence adapter | Multicolour gene localisation |
| **NmdCas9-LSD1** | Downregulation proximal enhancers | Fused to LSD1(histon-demethylase) | Gene regulation |
| **SaCas9n** | Single strand break = nick | Cas9-D10A: RuvC deficient<br>Cas9-N580A : HNH deficient | Reduce off target effects (50-1500x) |
| **SaCas9 KKH** | Alternative PAM: NNNRRT | E782K/N968R/R1015H | More target sites available |

### 2.5.2 Cas9 orthologues

The first Cas9 developed for genome editing originates from the bacterium *Streptococcus pyogenes*[335]. However, over 40% of all bacteria and 90% of all archaea have a CRISPR defence system[336]. Considering the feasibility of the system for genome editing, only the CRISPR type II subtypes are adaptable to genome editing, since they consist of only one nuclease and a RNA duplex[339]. Other systems contain large multiprotein complexes and complex processing of the RNA components. Advanced search on bacterial genomes discovered further CRISPR/Cas9 systems. They are listed in Table 2. Some of the new Cas9 orthologues were adapted to a genome editing tool, while other cannot be adapted to mammalian cells or are not efficient enough. The main principle of genome editing for all Cas9 orthologues is the same. The gRNA leads the Cas9 to the target region to induce a DSB 3 bp upstream of the PAM, which is repaired by host DNA repair mechanisms.

Most promising and best accepted in the scientific community is the Cas9 orthologues from the bacterium *Staphylococcus aureus* (SaCas9). It is much smaller than the commonly used SpCas9, but has characteristics similar to the SpCas9 i.e. a 21-23 nt protospacer target sequence, a 7-8 nt long seed sequence, but a more complex PAM (NNGRRT)[423]. In comparison to several other Cas9 orthologues, Ran *et al*. described this one as the only Cas9 with comparable on-target activity to the SpCas9. The features (small size, high activity) made the SaCas9 very interesting for *in vivo* applications (several times used in mouse models by AAV delivery). A couple of improvements used for the SpCas9 are now also available for the SaCas9, like nickase[424], alternative PAM[425] or split SaCas9. Also, the usage of truncated gRNA to minimal 19 nt to improve the off-target activity is possible with the SaCas9. Several web tools already included the SaCas9 for gRNA design, too.

Another Cas9 orthologue studied in great detail, originates from the *Streptococcus thermophilus*[426]. The first step to understand the CRISPR locus as a defence system has been made in this bacterium[343]. It actually has two CRISPR/Cas9 systems, St1Cas9 and St3Cas9[427]. Both have been tested for human genome editing and the St3Cas9 was found to be as efficient as the SpCas9, while the St1Cas9 was much less active. However, both StCas9s show much lower off-target activity due to their longer PAM NNAGAAW and GGG/NGGNG for the St1Cas9 and St3Cas9, respectively.

The NmCas9 (*Neisseria meningitis*) has also been tested for genome editing with similar results as the StCas9[428]. In comparison to the SpCas9, it has a slightly lower on-target efficiency, while it has also a reduced off-target activity. This does not make the NmCas9 the perfect candidate for

genome editing, but the discovery of a natural inhibitor of the NmCas9 by three proteins provides the opportunity of a tuneable CRISPR/Cas9 system without protein engineering, important for long term Cas9/gRNA expression[429].

Two further Cas9 proteins are worth mentioning. The CjCa9 (Camplyobacter jejuni) is the smallest Cas9 found so far[430]. Surprisingly, it has a comparable on-target efficiency with SpCas9 and SaCas9 but the lowest off-target potential. It has been successfully tested *in vivo* by AAV delivery in the muscle or the retina by local injection. It was only published in the beginning of this year and I think, it has the highest potential to become the new routinely used Cas9. It is especially suitable for therapeutic applications, since it has better features for delivery and off-target activity than the SpCas9 and SaCas9 and no new disadvantages. Finally, the Cas9 from the bacterium *Francisella novicidea* (FnCas9) is interesting, because it targets not only dsDNA but also mRNA[429]. The target specificity is regulated by two different types of crRNAs (crRNA and small crRNA).

However, even in this bacterium there was another subtype of CRISPR type II found. The Cpf1 belongs to the CRISPR type II but subtype V (distinct from the Cas9 subtype II)[431]. It is a two-component system of the Cpf1 nuclease and the crRNA. It does not require a tracrRNA as the original Cas9. Cpf1 has only one endonuclease domain (RuvC) and induces a DSB with a 5 nt 5' overhang. Furthermore, its target region is differently arranged than the one of the Cas9. Cpf1 detects a PAM (TTN, YTN) at the 5' end of the protospacer of 23-25 nt, which only has a small seed sequence of 5-8 nt, but is also highly sensitive to mismatches at the cleavage positions 18 to 23 proximal of the PAM. Additionally, Cpf1 has an endoribonuclease domain and processes the crRNA itself without further RNases[432]. In the context of genome editing, this is ideal for multiplex application, because it excises the crRNA simply from a transcript without the need of multi-promoter systems or ribozyme flanked gRNA. Variable numbers of crRNA can be fused together in one transcript[433]. The FnCpf1 does not have any activity in human cells, but two Cpf1 orthologue have been found and adapted to human cells: AsCpf1 (*Aciddaminicoccus sp*) and LbCpf1 (*Lachnospiraceae bacterium*)[431]. They have a lower overall on-target but also off-target activity than the SpCas9. However, the creation of a 5' overhang is ideal for application with HR.

**Table 2: Cas9 orthologues**

| | *Origin* | Size | PAM | Alternative PAM | Protospacer | Seed sequence | DSB Type | Position | On target | Off target |
|---|---|---|---|---|---|---|---|---|---|---|
| SpCas9 | *Streptococcus pyogenes* | 1368 aa | NGG | NGA NAG | 20 nt min. 18 nt | 8-10 nt | Blunt end | 3' upstream PAM | ++++ | SpCas9 |
| SaCas9 | *Staphylococcus aureus* | 1053 aa | NNGRRT | NNGRR | 21-23 nt min.19 nt | 7-8 nt | Blunt end | 3' upstream PAM | ++++ | SaCas9 |
| NmCas9 | *Neisseria meningitis* | 1083 aa | NNNGATT | NNNGCTT | 24 nt | 9 nt | Blunt end | 3' upstream PAM | ++ | NmCas9 |
| St1Cas9 | *Streptococcus thermophiles* | 1121 aa | NNAGAAW | | 20 nt min. 18 | 5 nt | Blunt end | | ++ | St1Cas9 |
| St3Cas9 | *Streptococcus thermophiles* | | GGG NGGNG | | 20 nt min. 18 | | Blunt end | | ++++ | St3Cas9 |
| CjCas9 | *Cymplyobacter jejuni* | 984 aa | NNNNACAC | NNNNRYAC | 22 nt min. 20 | | | | ++++ | CjCas9 |
| FnCas9 | *Franciselle novicida* | 1629 aa | NGG | | | | DNA: blunt end, ssRNA | | +/- | FnCas9 |
| Cas9 | *Streptococcus pasteurious* | 1130aa | NNGTGA | | | | | | + | Cas9 |
| Cas9 | *Neisseria cinereae* | 1083 aa | NNNNGTA | | | | | | + | Cas9 |
| Cas9 | *Camplyobacter lori* | 1003 aa | NNGGG | | | | | | + | Cas9 |
| Cas9 | *Parvibaculum lavenentiourous* | 1037 aa | NNNNCATN | | | | | | + | Cas9 |
| Cas9 | *Corynebacterium diphtheria* | 1084 aa | NGG | | | | | | + | Cas9 |
| Cpf1 =Cas12a | *Franciselle novicida* | | TTN, YTN (5' end of protospacer, non-target strand) | CTA | 23-25 nt | 5-8 nt, (Position 18-24 nt) | 5' overhang | 18 (non-target strand) and 23 (target strand) downstream PAM | - (Mammalian cells) | Cpf1 =Cas12a |
| AsCpf1 | *Aciddaminicoccus sp* | | | | | | | | ++ | AsCpf1 |
| LbCpf1 | *Lachnospiraceae bacterium* | | | | | | | | ++ | LbCpf1 |

aa= amino acid, R = Purin, Y=pyrimidin, nt= nucleotides, min. =minimal, PAM = protospacer adjacent motif

## 2.6 Applications of the CRISPR/Cas9 system for human diseases

Over the previous chapters, the broad variability, high efficiency and specificity of the CRISPR/Cas9 system has been illustrated. Its easy feasibility to different targets and cell types, and the fast and cost-efficient design makes this system ideal for the application to cure or alleviate human diseases, including genetic disorders, acquired protein expression deficiencies and infectious diseases. Even only after 5 years of the discovery of the CRISPR/Cas9 system a huge variety of human diseases have been targeted by this multifunctional genome editing tool. The basic strategies to apply the CRISPR/Cas9 system can be classified in 4 major categories: gene disruption, gene correction by NHEJ, gene correction by HDR and gene addition by HDR (Figure 15)[333].

Gene disruption is the classical usage of a genome editing tool. It has been used over decades with ZFN and TALEN to create knock-out animals to study gene functions and generate rodent models for human diseases. However, some hereditary human diseases are based on malfunctioning and toxic proteins expressed from mutated genes[434]. One example is the neurodegenerative disorder, Huntington's disease. In the patient's brain, toxic aggregates of the mutated huntingtin (HTT) disturb the cellular function and lead to the classical defect of motor function. The HTT gene has an insertion of several repeats of CAG leading to a poly-glutamine in the HTT and a dysfunctional protein. The frame-shift mutation induced by the Cas9/gRNA targeting the CAG repeats downregulates the mutated HTT expression[435]. This decreases the accumulation of toxic aggregates in the brain and alleviates the motor function.

Another approach based on gene disruption is based on knock-out of the PD1 expression on cancer-specific T-cells to prevent the inhibition of the T cells by the immunosuppressive tumour microenvironment. A Chinese group, led by Lu You, started the first clinical trial based on the CRISPR/Cas9 system with those Cas9/gRNA modified T cells in the end of 2016[436].

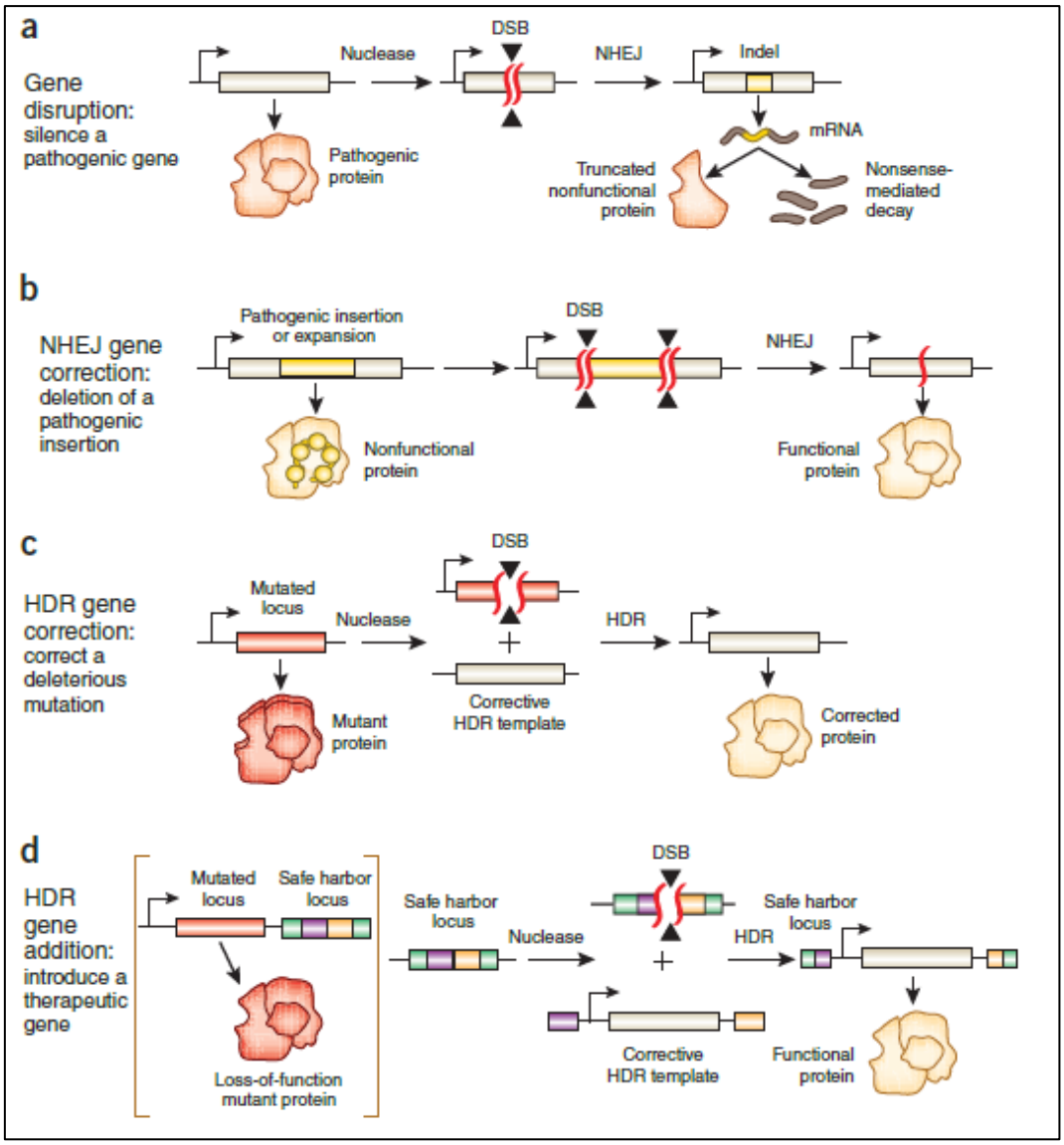



**Figure 15: Types of genome editing for therapeutic applications: A) DSB repaired by NHEJ and the introduction of small *indels* can lead to a knock-out of the target gene or a non-functional truncated protein expression. These induced loss-of-function mutations are applicable to diseases based on toxic protein intermediated from mutated genes. B) The multiplex CRIPSR/Cas9 strategy targeting multiple sites of a mutated gene can remove malfunctioning genes or the mutated gene regions leading to the loss-of-function or non-toxic truncated protein intermediates. C) The multiplex CRISPR/Cas9 strategy combined with donor DNA can be used to exchange mutated genes with the WT sequence and to restore the normal WT protein expression. D) Similarly, the multiplex strategy can be used with a donor DNA to safely insert the WT gene at a safe harbours locus in the host. Figure from: Turitz Cox et al. 2015; *Nat Med*.**

Gene correction by NHEJ is based on a multiplex gRNA/Cas9 approach to excise partially or completely a malfunctioned gene. This is one of the strategies, which can be applied to alleviate the Duchenne muscular dystrophy (DMD)[434]. It is a degenerative muscular disease based on a frameshift mutation induced by intragenic deletions of one or more exons leading to a dystrophin deficiency. The large excision of the mutated part of the gene (from the intron in front of exon 45 till the intron after exon 55) leads to a truncated but partly functional dystrophin expression. It changes the phenotype of the disease to the less severe Becker muscular dystrophy. This has been successfully tested in the mdx mouse model (SpCas9 knock-in and gRNA delivery by local muscular electroporation)[437].

Gene correction by HDR is a more classical approach. It aims to restore the original gene by replacing the mutated region with the WT sequence[333]. This strategy has been tested to cure β-thalassemia, sickle cell disease (SCD) and hereditary tyrosinemia type I (HT1)[434,438]. SCD and HT1 both are based on a single base pair substitution. Using the CRISPR/Cas9 system to induce a DSB close to the mutations and providing a donor temple can restore the WT gene. If the gRNA targets directly the mutations, it can be even applied to heterozygote genotypes without the danger of disrupting the WT gene. The treatment of HT1 is validated in a fah-/- mouse model, carrying a substitution in the gene for the fumarylacetoacetate hydrolase (FAH). Its deficiency causes the accumulation of toxic metabolites of the tyrosine catabolism in the liver. The CRISPR/Cas9 system delivered by AAV was able to significantly reduce liver damage and partially reinstate FAH expression[397].

Finally, gene addition by HDR can insert a WT gene to overcome a gene knock-out or increase the expression of a missregulated gene under a different promoter control. This is used for an innovative idea to treat high-fat diet induced diabetes. Yue *et al*. inserted the GLP-1 (glucagon-like-peptide 1) in skin epidermal progenitor cells[439]. Tested with mouse keratinocytes, he created a skin graft containing the GLP-1 expressing cells on an acelluarized mouse dermis and engrafted obese mice. The rise of GLP-1 in the serum of the engrafted mice led to an increased insulin production and the prevention of weight gain under high-fat diet. Importantly, the uncontrolled gene integration in a genome can be carcinogenic by influencing proto-oncogenes. Therefore, the insertion should be directed in safe harbours[333]. These are genetic locations with low or non-oncogenic potential for gene integration. Alternatively, only several exons can be inserted or can replace a mutated gene region. This is the second approach to cure DMD[434]. The deletion of exon 44 can be restored by inducing a DSB in exon 45 and inserting the WT sequence of exon 44. This could be achieved so far in iPSC with an efficiency of up to 75%[440].

The efficiency of the different treatment strategies depend on several factors: the efficiency of the gRNA, the delivery system, the type of genome editing (NHEJ>HDR) and the fitness of the cells[333]. Most of those factors has been discussed in the previous chapters, beside the fitness of the modified cells. For example, if a CRISPR-induced modification gives rise to a better fitness in the organism, it will overgrow the unmodified cells. In this case, the delivery and modification efficiency do not need to be very high. This is the case for HT1, where the modified hepatocytes have a survival advantage over fah-/- cells, because they do not acquire toxic metabolites[397]. In contrast, if the CRIPSR-modified cells have no survival advantage, they expand parallel to the unmodified cells. Depending on the disease, a higher number of modified cells is needed to ensure an alleviating effect. A typical example is haemophilia B, a clotting factor IX deficiency. The restoration of only 1% of factor IX expression improves strongly the clinical condition of the patient[333]. Finally, if the CRISPR-modified cells have a growth or survival disadvantage, then the CRISPR/Cas9 system needs to edit 100% of the target cells. This is the case for tumour cells to restore the transformations or the immunosuppressive action. Targeted cells will die or be removed by the immune system, while unmodified cells can proliferate uncontrolled.

As discussed here, the CRISPR/Cas9 system can be a potentially very efficient tool to improve human diseases. The most common application is in hereditary disorders with protein deficiencies or malfunctioning proteins. However, it can be also applied to infectious disease, which is of special interest in this thesis. Applications and strategies are discussed in more details in the next chapter (II-2.6.1).

### 2.6.1 The CRIPSR/Cas9 system as an antiviral strategy

There are several viruses, which can cause chronic and persistent infections in humans. Usually, these viral infections cannot be cleared by actual anti-viral chemotherapy, because the viral genome rests as a very stable circular intermediate in the nucleus or even integrates in the host genome. The usage of a genome editing tool can assess this persistent viral reservoir and destroy or silence the viral genome. The main targets are viruses, which have a dsDNA viral genome stage during their viral replication cycle (like HIV, HBV, Herpesviruses, HPV and JC), because the Cas9/gRNA is specialized to targeted DNA.

#### 2.6.1.1 HIV

The human immunodeficiency virus (HIV) causes the acquired immune deficiency syndrome (AIDS) by infecting and destroying $CD4^+$ T cells. After infection, it integrates into the human genome and persists lifelong in the host. The main reservoir of HIV are $CD4^+$ memory T cells, which are located all over the body (blood, brain, lymph tissue, gastrointestinal tract and more)[441]. Standard highly active anti-retroviral therapy (HAART) can only control active replicating virus, but cannot disrupt the latent reservoir. Therefore, targeting HIV with a genome editing tool could assess the integrated provirus and potentially clear the infection. This has been already tried with previous genome editing tools like MN, ZFN and TALEN[442,443]. HIV is also the first virus targeted by the CRIPSR/Cas9 system[444]. There are basically three strategies against HIV infections: functional cure, sterilizing cure and vaccination[445].

The functional cure aims to inactivate the integrated virus to prevent its reactivation or the expression of toxic viral intermediates. Ebina *et al*. in 2013 followed that idea by targeting the main promoter of HIV, the long terminal repeats (LTR)[444]. Induced DSB in the LTR region led to *indel* formations and silencing of viral protein expression. It was also observed that the HIV genome was excised in some cells.

For the sterilizing cure the complete reservoir of the integrated HIV genomes or infected cells needs to be removed from the patient, so that the patient is completely HIV negative again. Singleplex targeting of the LTR region of the integrated provirus (on both end of the viral genome inside the host genome) enables the excision of the entire viral genome from the host genome[446]. Moreover, targeting the LTR alone or in combination with essential viral genes by a duplex strategy is much more efficient[446–448]. The most efficient approach is a multiplex strategy. Yin *et al.* used a quadruplex strategy (4 gRNAs) targeting the LTR and two essential viral genes (gag and pol) and successfully fragmented the viral genome[449]. The excision of the HIV genome *in vivo* has been tested in several mouse models; the most common is the Tg26 mouse, which has several copies for a truncated HIV genome integrated into the host genome[448]. Duplex and multiplex approaches can significantly reduce the amount of integrated HIV. Multiple tissues were assessed by the delivery of the SaCas9 and two or four gRNAs with an AAV[449]. The liver and peripheral blood cells were the most efficiently targeted. Even in the brain HIV excision was detected, which is very important for HIV CNS diseases (dementia). In NCr mice, a HIV permissive mouse model, the simultaneous infection and injection of AAV Cas9/gRNAs reduced the amount of infected cells with an efficacy of up to 96% depending on the analysed tissue. Finally, the multiplex CRISPR/Cas9 system significantly cuts out the viral genome of human cells in humanized mice after the HIV infection through the natural route of the vaginal mucosa. In all three mice strains, the multiplex Cas9/gRNA treatment significantly reduced the reservoir of HIV positive cells, but total clearance of the infection has not been yet observed. There are two possible explanations for

this. On the one hand, actual delivery systems are not sufficient enough to assess all HIV positive cells, especially resting cells in the tissue. On the other hand, HIV has extremely high mutation rates, so that one infected individual has a lot of different quasi-species in the infected cells. By next generation sequencing (NGS) of HIV patient peripheral blood samples, it was detected that the total HIV reservoir acquire 10-20 mutations per year only in the LTR region[450]. It reaches a plateau around six years after the primary infection. This feature makes it very difficult to target the complete reservoir. The correct gRNA design depends on NGS to compare the different quasi-species. But still, only 50% of the analysed samples were able to be targeted with a gRNA set comprising less than 10 gRNAs. Moreover, due to the high mutation rate, the gRNA set is only valid for less than one year. Under the actual technical conditions a sterilizing cure is not yet possible.

A very different strategy can also lead to the clearance of HIV from the patients. The so-called "shock and kill" strategy is based on the reactivation of all latent infected cells, which then will be destroyed by the immune system. This approach can only be considered before the progression to AIDS. A clinical study using an HDAC inhibitor (SAHA) failed so far to decrease the HIV reservoir in patients[445]. However, the actual latency-reversing agents are unspecific and have side effects like T-cell activation. Moreover, the reactivation of HIV is not strong enough to lead to a complete and strong viral protein expression. Currently, researchers are testing the dCas9-transactivators to reactivate HIV by targeting the LTR region. The standard dCas9-VP64 is not sufficient to reactivate HIV[451], but two related CRISPR transactivators were very efficient: dCas9-VPR and SAM[451–453]. The dCas9-VPR is fused to three different transactivator domains, the VP12 of HSV, the p65 subunit of NFκB and the EBV R transactivator domain. The SAM system is still based on the dCas9-VP64, but uses a modified gRNA which can recruit a bacteriophage coat protein fused to transactivation domains (p65, HSF1). Besides those CRISPR-TF, the dCas9-p300 (histone acetyltransferase) also induces reactivation very efficient[453]. The advantage using the CRISPR/Cas9 system as an activator is, that it only activates HIV and has no adverse effect on non-infected cells. Moreover, it can also be used in combination with SAHA, which has a synergistic effect[451]. *In vivo* studies showing the reduction of the HIV reservoir are still missing.

Finally, the vaccination strategy treats uninfected cells to prevent the infection by HIV. It has been shown that cells transduced with the Cas9/gRNA against the LTR cannot just target the integrated viral genome, but also cleaves the dsDNA intermediate of HIV during primary infection[454]. Primary T cells and iPSC generated monocytes can be protected from HIV infection by a stable Cas9/gRNA expression[454,455]. Alternatively, the CRISPR/Cas9 system can be used to destroy the entry receptor of HIV. However, CD4 is essential for T cell signalling and the co-receptor CXCR4 is an essential homing receptor for HSC, it is not possible to knock-out this receptors[442]. But, the CCR5 co-receptor can be targeted, because individuals with a CCR5Δ32 mutation (small internal deletion) have no phenotype and are completely healthy. Moreover, as shown by the "Berlin Patient" the CCR5Δ32 cells cannot be infected by HIV leading to the cure of the patients by HSCT from a CCR5Δ32 donor. Since then, the knock-out or induction of the specific internal deletion is considered a treatment for HIV by cell therapy and is in a clinical trial with ZFN modified cells. Efficient knock-out of the CCR5 has been tested in T cell lines[456], primary T cells[455], iPSC generated macrophages[457] and HSC[458] without interfering with the phenotype, cytotoxicity or changes of the differentiation potential. All $CCR5^-$ cells were resistant to HIV infection by R5-tropical strains. These cells have a selective advantage towards WT cells and overgrow the WT cells under HIV infection conditions, which is very promising for cell therapy[456]. $CCR5^-$ HSCs are tested *in vivo* to

reconstitute humanized mice in the first and second generation[459]. Moreover, mice with CCR5$^-$ HSC have a significantly reduced viremia and enrichment of CCR5$^-$ cells.

HIV cure is still far away, but adoptive autologous or allogenic cell therapy shows promise in improving the T cell count and therefore prevents or slows down the progression to AIDS.

#### 2.6.1.2 HBV

The hepatitis B virus causes chronic infections in the liver leading to liver cirrhosis and carcinoma. In infected patients, it infects nearly all hepatocytes and remains in the cells as a very stable cccDNA (circular covalently closed DNA). More rarely, it can also integrate into the host genome.

Similar as for HIV, CRISPR/Cas9 targeting strategies aim to silence or destroy the HBV reservoir in the hepatocytes. Several singleplex and duplex strategies have been tested, showing that the duplex strategy is always more efficient[460]. Furthermore, to reduce off-target effects on the human genome, the Cas9n and the dCas9 have been tested for efficiency to cut or destroy the viral genome[461–463]. Surprisingly, both Cas9 variants led to an inhibition of viral protein expression. However, the efficiency was reduced in comparison to the WT Cas9 and might not be sufficient to be used as an anti-viral treatment[462,463]. The correct target gene selection is more variable in HBV than in HIV. Basically, all major ORFs have been tested as an efficient target for the Cas9/gRNA. The targeting of the reverse transcriptase (RT) directly in the catalytically motif YMDD shows the most promising results[464]. In this case, *indels* leading to frameshift or in-frame mutations are both sufficient to destroy the functional protein and to prevent viral replication. Therefore, it is more efficient to target this motif. Moreover, duplex strategies often lead directly to the loss of the cccDNA, which is obviously more sensitive to cleavage than integrated viral genomes or other viral genome intermediates[464,465]. The complete degradation of the HBV genome can be ensured by a multiplex strategy based on eight gRNAs[466]. It very efficiently eliminated nearly all viral genome intermediates and reduced HBV antigen expression *in vivo*. For most of the strategies, a prolonged exposure of the Cas9/gRNA against the HBV genome increases the on-target efficiency and reduces viral replication for over 300 days[465,467]. *In vivo* mouse models validated the reduction of the viral replication by the delivery of CRISPR/Cas9 system by hydrodynamic tail vein injection or AAV (SaCas9), but never achieved a complete HBV clearance[466,468–471]. The delivery is the limiting factor. A mathematical model calculated that multiple doses are necessary to target all hepatocytes. For example, with a compound allowing a delivery efficiency of 99.99%, it would require only two to three doses of the compound to target 100% of the hepatocytes. On the other hand, using a less efficient compound leading to a 50% delivery efficiency, it would need at least 38 doses to reach 100% of the hepatocytes[472]. However, multiple doses increase the risk of an immune reaction against the Cas9 or the delivery agent and might lead to massive cell loss of transduced cells. A combination of the CRISPR/Cas9 system with the standard nucleotide analogue based anti-viral therapy could overcome this problem. Additive or synergistic effect have been proven already[464].

#### 2.6.1.3 Herpesviruses

Herpesviruses are a family of very large dsDNA viruses. Primary infections are usually not severe, but the virus persists in latency lifelong in the host. The latent reservoir depends on the virus type, either locally in ganglion nervous cells (HSV) or more systemic in different resting cell types (EBV, HCMV). The viral genome remains in the nucleus of infected cells as an episome and does not integrate in the host genome.

EBV (Epstein-Barr Virus) causes mononucleosis as a primary infection, but is more known as an oncovirus leading to Burkitt lymphomas or nasopharyngeal carcinomas, usually in immune deficient patients. The major aim is to reverse the transformation or loss of the viral genome maintenance. Targeting the oncogenes EBNA3C and LMP-1 can remove the transformations and induce apoptosis in Burkitt lymphoma cells[473]. However, the loss or destruction of the entire viral genome is more efficient. DSB in the internal repeat regions can lead to the fragmentation of the genome. The most efficient strategy used a multiplex approach with seven gRNAs targeting repeats, oncogenes and genes essential for viral genome maintenance.

HSV (Herpes simplex virus) causes cold sores (HSV1) or genital herpes (HSV2). While a singleplex strategy failed to protect against the HSV infection, a duplex strategy targeting essential viral genes for genome replication prevented the virion release completely[474]. Surprisingly, the Cas9/gRNA was not able to target the latent viral genome. The HSV genome is highly chromatin-condensed and associated with nucleosomes during latency, which prevents Cas9 binding to the target site. This problem can probably be overcome by choosing a target site in viral genes, which are expressed during latency.

HCMV has also been targeted by the CRISPR/Cas9 system. Singleplex strategies to knock-out Delay Early antigens of the HCMV only prevented viral replication over a short period of time but the virus escaped rapidly leading to virion release and resistance to the CRISPR/Cas9 system in primary cells[474].

#### 2.6.1.4 HPV

The human papillomavirus (HPV) is a commonly known oncovirus mainly causing cervical cancer. Two main viral genes induce the transformation of infected cells. E6 inhibits the tumour-suppressor gene p53 and E7 leads to the degradation of Rb. Treatment strategies against HPV aim more the destruction of these oncogenes than the clearance of the infection. Already the destruction of a single viral oncogene (E6 or E7) leads to cell cycle arrest and apoptosis[475–477]. However, the duplex strategy is more efficient targeting either both antigens or one oncogene and the E6/E7 promoter[478,479]. This CRISPR/Cas9 strategy can be enhanced by the co-treatment with a chemotherapy agent CDDP, which inhibits E6[479]. The synergistic effect could also be confirmed in a xenograft mouse model and led to a reduced tumour size and less metastasis.

#### 2.6.1.5 RNA viruses

RNA viruses are not a classical target of the CRISPR/Cas9 system, but there are two possibilities using the CRISPR/Cas9 system to target RNA. Using a PAMmer can change the substrate of the Cas9 from dsDNA to ssRNA[480]. The PAMmer is a ssDNA containing the PAM sequence and is partly complementary to the target RNA sequence, which does not need to contain the PAM. Binding to the PAM enables the gRNA to bind ssRNA and the nuclease domain HNH of the Cas9 cuts the RNA.

Alternatively, the FnCas9 from the *Francisella novidica* can cut dsDNA with the classical tracrRNA:crRNA complex. However, if the FnCas9 binds a small crRNA (scrRNA) it digest ssRNA[429]. The FnCas9 has been successfully used against the RNA Hepatitis C virus[481]. Unexpectedly, the viral RNA targeting was PAM independent and the gRNA sequence at the 5' end was more important for efficient HCV suppression than the 3' seed sequence for DNA targeting.

The advantage to target RNA viruses is the reduced probability of off-target effect on the host genome. Target sites are PAM independent and should reduce the possibility of homologue

regions in the host genome. Moreover, most RNA viruses stay during their replication cycle in the cytoplasm. Keeping the Cas9 in the cytoplasm abrogates the possibility of DSB in the host genome.

#### 2.6.1.6 Resistances

The biggest challenge besides a highly efficient delivery method for *in vivo* applications, is the development of CRISPR/Cas9 resistant viruses. The usage of the singleplex strategy targeting the viral genome leads statistically in 33% of all DSB to an in-frame mutations (+/- 3nt). This virus mutants are mostly replication-competent, but have a disrupted target site, which abrogates the repetitive cutting by the Cas9/gRNA. So far, HIV, HSV, HCMV and EBV have been reported to escape from the CRISPR/Cas9 system[474,482,483]. Extent analysis on the viral escape mutants and the prevention of breakthrough replication has been mainly done on HIV. First of all, the type of escape mutation strongly depends on the target region in the viral genome[482]. Highly conserved genes usually acquire nucleotide substitutions or small in-frame mutations to maintain the ORF and functional protein expression. The substitutions often even cause no amino acid change in the protein. Less conserved regions or non-essential genes gain normal *indels*, because of the missing selection pressure. Promoter and regulatory regions can obtain normal *indels*, too, but some conserved regions like secondary structures or transcription factor binding sites may only allow small *indels* or just substitutions[484].

To prevent the viral escape, a minimum of two genes/regions needs to be targeted[474,485,486]. Larger multiplex strategies probably decrease the viral escape probability further. Still, the right selection of target regions can be further optimized. It has been observed, that a duplex strategy with target sites with more distance to each other provide better protection[485]. Moreover, both gRNAs need to be highly efficient (tested as a singleplex strategy) to ensure that the duplex strategy leads to large deletions and not just small *indels* at both target sites[486]. Finally, to ensure no viral rescue by non-treated cells, a prolonged Cas9/gRNA treatment is necessary[485]. The prolonged Cas9/gRNA exposure pushes the mutations from nucleotide substitution to larger deletions over time (Figure 16). For HIV, it has been shown that one to three months of Cas9/gRNA exposure is necessary to prevent breakthrough replication in a mixed culture of treated and untreated cells. This is very important for the right choice of the delivery system. Viral delivery systems, which ensure transgene expression for at least one month will therefore be more efficient. In general, there is a higher risk of off-target effects on the human genome by prolonged CRISPR/Cas9 expression, but most of the studies mentioned here could not detect any off-target effects at all[446,467,487].

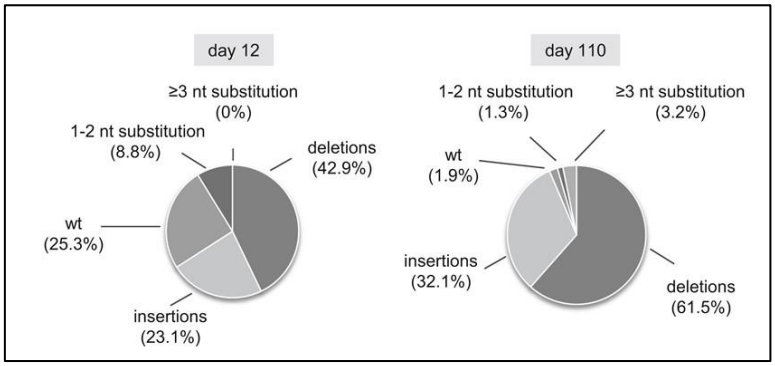



**Figure 16: Type of mutations induced by the CRISPR/Cas9 system over time. The CD4+ T cell line (SupT1) was transduced with a duplex Cas9/gRNA lentiviral vector and subsequently infected with HIV. Twelve and 110 days pi, DNA was extracted from the cells, the HIV provirus was PCR-amplified and sequenced based on TA cloning. The type of mutations changed over time from minor mutations to large deletions. Figure from: Wang et al. 2016; *Cell Reports*.**

In conclusion, the usage of CRISPR/Cas9 system to target chronic or persistent viral infections is a very promising tool for the future, but several factors have to be carefully managed. The right target gene selection is the key to a safe treatment strategy. The destruction of the entire viral genome prevents viral escape and avoids the expression of toxic viral intermediates. Most viral infections cannot yet be cleared by the CRIPSR/Cas9 system because of insufficient delivery methods. However, EBV and HCMV are typical opportunistic infection for transplant patients giving the opportunity to treat tissues or cell suspension *ex vivo* with a higher delivery efficiency and the chance to select Cas9/gRNA positive cells.

***Box 6: How to target viruses with the CRISPR/Cas9 system***

***Target selection:***

- *Essential viral gene: disruption of the target gene has to lead to a replication-incompetent virus*
- *Highly conserved: less variability between different viral strains and only small mutations will be tolerated*
- *Multiplex > duplex > singleplex: multiple target sites prevent viral escape mutations and may even allow the degradation of the viral genome*
- *Non chromatin compensated regions: Cas9 cannot target highly chromatin condense regions, which gives the virus the chance to escape the treatment*

***Strategies:***

I. *Destruction of essential viral genes or fragmentation of the viral genome/gene and subsequent loss of the viral genome = abolish viral replication*
II. *Inactivation or excision of integrated viral genome copies = prevention of reactivation or loss of latent viral reservoir*
III. *Reactivation of the latent virus and subsequent destruction of infected cells by the immune system = reduction of viral reservoir*
IV. *Knock-out of essential viral host factors like entry receptors = prevent the infection of modified host cells (vaccination)*

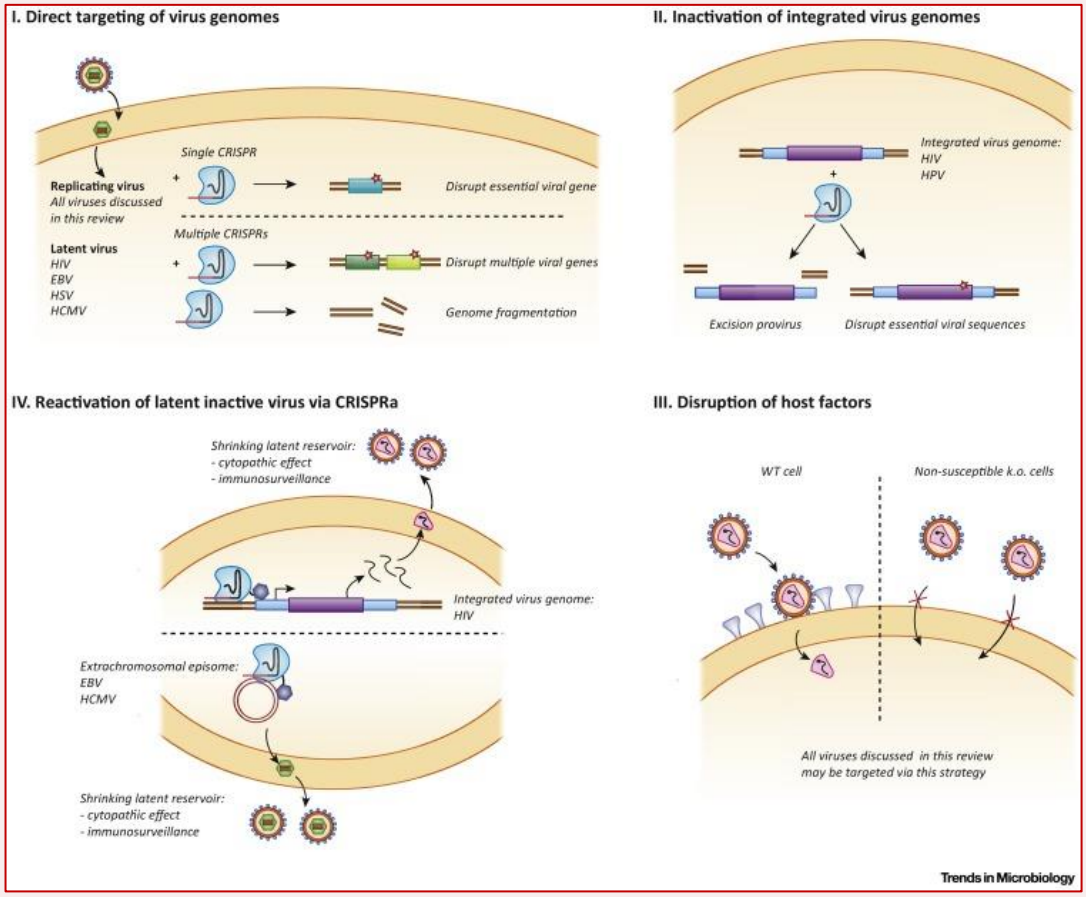


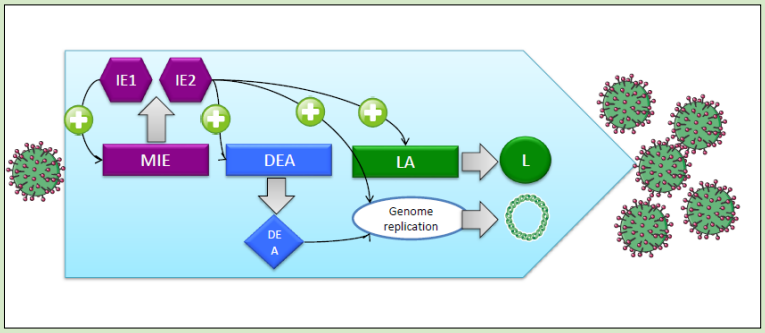

*Figure from:* Soppe et al. 2017; *Trends Microbiol.*

# IV Hypothesis 1

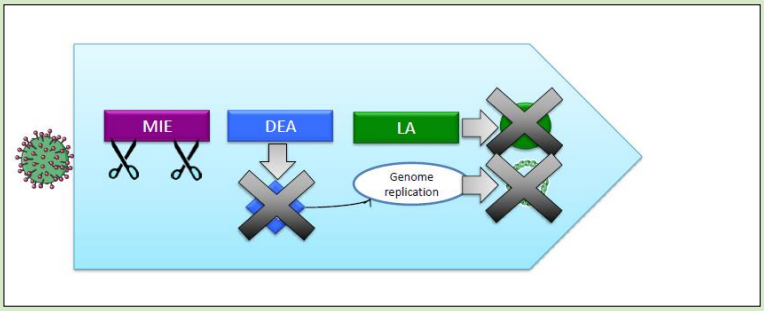
## *The knock-out of the major Immediate Early gene UL122/123 by the CRISPR/Cas9 system inhibits lytic viral replication*

***Box 7: Graphical abstract of the hypothesis 1***

***A**: The MIE proteins are the key regulators of the CMV replication cycle and are absolutely essential for genome replication and virion release.*




***B**: Disruption of the MIE genes by the CRISPR/Cas9 system abrogates the viral replication. No new virions are released.*



# 1 Design of the anti-*UL122/123* CRISPR/Cas9 strategies

## 1.1 General Principle

As shown in previous studies, the MIE gene *UL122/123* encodes the key regulators of the viral replication cycle: IE1 and IE2[194,488]. IE2 is the major transactivator responsible for the expression of delay early and late antigens[193,196]. IE1 is responsible for the positive autoregulation of the *UL122/123* expression and is especially necessary for infection with a low MOI (≤ 0.1)[203]. The disruption of the MIE gene expression leads to a replication-incompetent virus. The viral infection is abortive directly after cell entry[194]. Neither viral genome replication nor newly produced virions can be detected[488]. Moreover, a splice variant of IE1, expressed during latency, is responsible for the maintenance of the viral genome in the host cell[234]. Therefore, Azad *et al*. developed in the early 90's an antisense oligonucleotide targeting the exon 5 of the MIE gene, which prevents the viral replication very efficiently[195]. Later on, this approach was developed to a treatment for CMV retinitis in HIV patients.

**This led subsequently to the hypothesis that the knock-out of the *U122/123* gene is an appropriate target for a new anti-viral approach based on the CRISPR/Cas9 system**. We assume that the permanent disruption of the *UL122/123* ORF will completely inhibit viral replication. The usage of a single gRNA should lead to *indels*, which can induce a frameshift mutation and a premature stop codon. A multiplex approach based on three gRNAs is expected to excise the *UL122/123* gene between the first and third gRNA. The second gRNA should prevent the inversion of the gene and should lead to the fragmentation of the target gene.

In order to analyse the anti-viral efficiency of the CRISPR/Cas9 system targeting the *UL122/123* gene we designed stable cell lines expressing the Cas9 and one or three gRNAs. Subsequently, those cell lines were infected with HCMV and the progression of the viral infection and replication cycle was monitored. An abortive viral replication cycle at the immediate early time point was expected.

## 1.2 Methods

### 1.2.1 Lentiviral vector construction

For the singleplex strategy and the unsp. gRNA/Cas9, gRNAs were design based on the consensus motif $N_{20}$NGG. Concomitant oligonucleotides were designed and inserted in the CRISPR/Cas9 expression plasmid pX330 (addgene) following the manual instructions. Later on, the gRNA cassettes were PCR amplified using the following primers F: 5'ATATGAATTCTT TTGCTCACATGTGAGGGC3', R: 5'ATATGAATTCCGCGCTAAAAACGGACTAGC3' and cloned by EcoRI digestion in the first LV backbone containing the SpCas9 fused to GFP[489] (Figure 17 A). The multiplex gRNA cassette was synthetized (GeneScript, Piscataway, USA) and also inserted in the lentiviral vector[489] by EcoRI digestion.

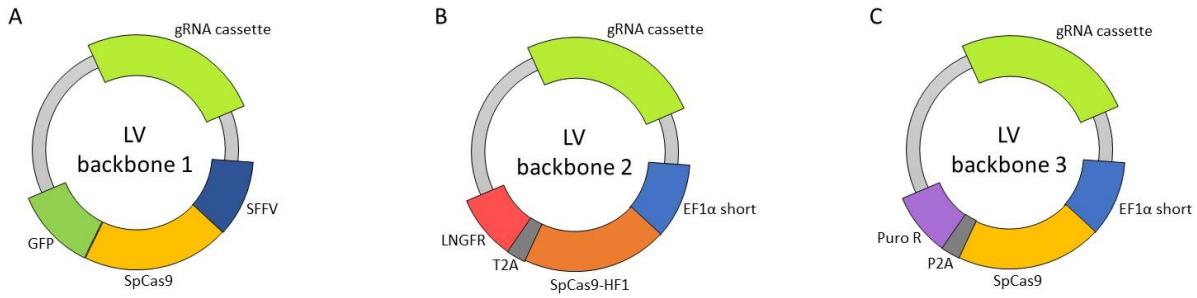



**Figure 17: Lentiviral vector constructs. All constructs contained one of the three gRNA cassettes (unsp. gRNA, singleplex, multiplex), which are in detail explained in the results section. A) The LV type 1 was based on an LV previously published in our lab containing the SpCas9-GFP under the control of the SFFV promoter[489]. B) The LV type 2 was designed and constructed in collaboration with Olaf Pinkenburg to gain an LV optimized for high titer production. Subsequently, the SpCas9HF-T2A-LNGFR under the control of the EF1α short promoter was inserted. C) The LV type 3 is based on an optimized CRISPR/Cas9 LV backbone from the Zhang lab, MIT containing the SpCas9-P2A-puromycin resistance (addgene plasmid # 52961). LV: lentiviral vector, SFFV: spleen focus-forming virus promoter, LNGFR: low-affinity nerve growth factor receptor, EF1α short: elongation factor-1 α short**

The second LV backbone contained SpCas9HF1-T2A-LNGFR (Figure 17 B). It was designed and constructed in cooperation with Olaf Pinkenburg. He built an LV backbone for high titre production from the pWPT-GFP (a gift from Didier Trono (Addgene # 12255)) as followed: A replacement of the fragment XbaI - PvuI (1347 bp) with the SpeI – PvuI fragment (462 bp) from pGL4.13 (PROMEGA) was done in order to remove unwanted restriction sites. Also, the BamHI-KpnI fragment was replaced with BamHI-KpnI fragment from pWPXL (a gift from Didier Trono (Addgene # 12257) to add an XmaI restriction site. To create the final pWPTΔΔ., the restriction sites for MluI - XbaI and NotI - EcoRI up- and downstream of the promoter EF1α, respectively, were inserted by PCR amplification with the following primers: EF1α forward: TATAACGCGTCCCCTCTAGAGGCTCCGGTGCCCGTCAGTG and EF1α reverse: TTGAATTCTGCGGCCGCGCGTCACGACACCTGTGTTCTGGC. The Cas9-HF1 (VP12, addgene #72247) was fused to the T2A-LNGFR construct synthesized by GeneScript (Piscataway, USA) and cloned into this new empty LV backbone (pWPTΔΔ) by NotI and XmaI digestion. Lastly, the multiplex cassette, excised from the first LV, and the gRNA2 or unsp. gRNA cassettes, PCR amplified from the pX330 (Primers: Forward: ATATACGCGTTTTTGCTCACATGTGAGGGC and reverse: ATATTCTAGACGCGCTAAAAACGGACTAGC), were inserted by MluI and XbaI digestion.

The third LV backbone (Addgene plasmid # 52961) contains the SpCas9-P2A-Puromycin resistance and a gRNA cassette under the control of the U6 promoter (Figure 17 C). For the singleplex and the internal control, the gRNA2 and unsp. gRNA were synthesized as

oligonucleotides and inserted in the gRNA cassette following the manufacturer's instructions. The multiplex gRNA cassette was PCR amplified from the first LV backbone SpCas9-GFP by the following primers: F: TATAttaattaaacgcgtGAGGGCCTATTTCC, R: TATAgaattccgtacgaAAAAAAGCACCGA and inserted via PacI and EcoRI.

### 1.2.2 Lentiviral vector production

6*10^6 HEK293T cells were seeded in a 150 $mm^2$ cell culture dish in complete medium (DMEM, 10% fetal calf serum (FCS), 1% penicillin/streptomycin (P/S), 1% L-gulatmine (L-glu) and 1% HEPES) and incubated overnight. The LV backbone and the two helper plasmids psPAX2 (for gag, pol, rev and tat expression) and pMD2G (VSV-G pseudotype) were diluted in 1.5 mL 2x HBSS (0.28 M NaCl, 10 mM HEPES, 1.5 mM $Na_2HPO_4$ in $H_2O$) and 750 µL $H_2O$/ 2.5 mM HEPES. The DNA solution was slowly vortexed and dropwise mixed with 750 µL $CaCl_2$ (0.5 M). The final mix was vortexed for 5 s with maximum speed and further incubated for 20 min at RT. Subsequently, the transfection mix was vortexed again at maximal speed for 5 s and then dropwise added over the HEK293T cells. 16 h post transfection the medium was changed to fresh complete medium and incubated for another 24 h at 37 °C. Finally, the supernatant was collected and centrifuged for 5 min at 3000 rpm. The LV stock was further purified by filtration through a 0.45 µm PVDF filter and concentrated by ultracentrifugation (26000 rpm, 4 °C, 90 min). The LVs were resuspended in complete medium and stored at small aliquots at -80 °C.

### 1.2.3 RT-qPCR GAG/POL

To be able to sort LV transduced cells, they need to be declared "HIV negative" to leave the biosafety level 3 lab. Around three weeks post transduction, a part of the cells was harvested and the RNA was extracted from the cells by the NucleoSpin RNA Kit (Macherey Nagel, Düren, Germany). RNA was transcribed to cDNA by the TaqMan™ Reverse Transcription Reagents (Invitrogen, Thermo Fisher, USA).

To titre the amount of HIV transcripts, the cDNA samples were analysed by qPCR using the 2x sybr green reaction mix (Invitrogen, Thermo Fisher, USA) for the expression of the two retroviral genes *gag* and *pol* and in internal cellular control *HB2* (housekeeping gene). Samples were analysed in the 7900 qPCR machine with the following program: 10 min 95 °C, 40 cycle of 15 s 95 °C and 60 s 60 °C.

### 1.2.4 HCMV production

MRC5 primary fibroblasts were seeded in a 5-layer cell stack (5*10^7 cells) and cultured until they reached a confluency of 70%. Then, the medium was replaced by the inoculum to infect the MRC5 cells with an MOI of 0.02 in minimal medium (DMEM, 2% FCS, 1% P/S, 1% L-glu, 1% HEPES). MRC5 cells were incubated until 90% of the cells showed a cytopathic effect (7-8 days). Finally, the supernatant was harvested, centrifuged for 15 min at 3000 rpm and either directly aliquoted or concentrated *via* ultracentrifugation through a 20% sucrose cushion (2.5 h, 24000 rpm, 4 °C). The virus stock was store at -80 °C.

### 1.2.5 Titration of HCMV

MRC5 cells were plated in a 12-well-plate at a concentration of 1*10^5 cells/well and incubated overnight. Subsequently, the medium was removed and the virus suspension was serial diluted

1:10. MRC5 cells were infected in duplicates by a 2 h inoculation. Two days pi, the cells were trypsined and analysed by FACS for IE expression.

### 1.2.6 FACS

Cells were harvested by trypsin, added in a 96-well-plate (V-bottom) and washed one time (1.5 min, 2500 rpm, 4 °C) with PBS. For the viability stain, cells were resuspended in 50 µL of the Live/Dead Fixable Dead Cell Stain Kit (Invitrogen- Thermo Fisher Scientific, USA) and incubated for 30 min at room temperature (RT) in the dark. Cells were washed three times with PBS and fixed in 3.2% paraformaldehyde/PBS for 10 min at 4°C. After one washing step, cells were permeabilized by PBS/3% BSA/0.2% Triton for 30 min on ice. Intracellular IE was detected by either an anti-HCMV mAb (clone MAB810R; Millipore, Germany) or an anti-IE/E CMV antibody (Argene Biomérieux, France). HCMV glycoprotein B (gB) was detected intracellularly by a mouse anti-CMV gB antibody (1-M-12, Santa Cruz, USA). An anti-mouse IgG antibody conjugated to Alexa® 647 (BD Biosciences, USA) was used as a secondary antibody in all staining experiments presented in this study. After three further washing steps, the samples were analysed on the LSR II (BD Biosciences, California, USA).

### 1.2.7 Cell lines

U373-MG cells (astrocytoma cell line) and MRC5 cells (primary fibroblasts) were cultured in complete medium (DMEM, 10% FCS, 1% P/S, 1% L-Gut, 1% HEPES).

For the generation of stable U373-MG cell lines expressing the CRISPR/Cas9 strategies, U373-MG cells were transduced with the different LVs (first LV backbone) with an MOI of 5 to 10. After a three-week expansion phase, the U373-MG cells were declared HIV negative by a *gag* and *pol* RT-qPCR and subsequently FACS sorted based on the Cas9-GFP expression.

For the LV Cas9-GFP, MRC5 cells were transduced with an MOI of 5-10 and cultured for three weeks. After a negative HIV RT-qPCR, cells were FACS sorted based on the Cas9-GFP expression.

For the LV Puro, MRC5 cells were transduced with a dilution of concentrated LV of 1:500-1:1000 in the presence of 4 ng/µL polybrene and spinoculated at 1000 g for 90 min at 33 °C. After 8 h, the inoculum was replaced by fresh complete medium and cells were further cultured for two to three days. Subsequently, transduced MRC5 cells were treated for two days with 2 µg/mL puromycin or for nine days with 0.5 µg/mL puromycin (refreshing medium every two days).

### 1.2.8 HCMV infections

U373-MG cell lines and MRC5 cells were seeded in a 12-well plate at a concentration of 3*10^5 cells/mL and incubated overnight at 37 °C. Subsequently, cells were infected with the HCMV (strains: TB40GFP, Toledo, VR1814) at an MOI of 1 - 0.01 in complete medium. After 2 h inoculation, the medium was replaced by fresh complete medium and the cells were further incubated at 37 °C for two to eight days.

### 1.2.9 PCR

Total DNA was isolated from the infected cells with the NucleoSpin TriPrep Kit (Macherey-Nagel, Düren, Germany). The target region of the singleplex strategy was amplified by standard PCR with the following primers: F: 5' GTTCTCGTTGCAATCCTCGGTCAC 3' R: 5' CGTGGCGGTAGGGTATGTGT 3' following the Herculase II Fusion Enzyme with dNTPs Combo kit (Stratagene, Agilent

Technologies, CA, USA). A bigger PCR amplifies the whole *UL122/123* region (3862 bp amplicon; Primers F: ACATGAGGGGGAGAAGGACA; R: CGTGGCGGTAGGGTATGTGT) to analyse the multiplex target region. Samples were analysed by the Caliper LabChip GX device (PerkinElmer).

### 1.2.10 T7-assay

For the T7-assay, the PCR products were purified by the NucleoSpin column (Macherey Nagel, Düren, Germany). 200 ng of DNA was denatured for 5 min at 95 °C and slowly re-annealed in three steps consisting of 15 s at 95 °C, 15 s at 85 °C and 30 s at 25 °C. Small mismatches were detected by the digestion with the T7 endonuclease (New England Biolab Inc., UK) for 30 min at 37 °C. The addition of 2 µL 0.25 M EDTA stopped the reaction and samples were analysed by standard gel-electrophoresis or capillary electrophoresis on a Caliper LabChip GX device (PerkinElmer). The Caliper device calculated the quantity and concentration of each band and therefore allowed the calculation of the percentage of *indels* based on the formula from Hsu et al[365].

### 1.2.11 PCR cloning

The PCR products were inserted in an empty ampicillin/kanamycin vector *via* TA cloning with the StrataClone PCR cloning Kit (Agilent Technologie Devision, USA) according to the manufacturer's instructions. This expression vector was subsequently transformed by heat shock into competent cells. The competent cells were plated on LB medium with ampicillin and 2% X-gal and incubated overnight at 37 °C. Positive clones were selected the next morning by white/blue selection and send to MWG (Eurofins Genomics GmbH, Ebersfeld, Germany) for sequencing.

### 1.2.12 qPCR *US8*

Cellular and viral DNA were isolated with the NucleoSpin TriPrep Kit (Macherey-Nagel, Düren, Germany) and the viral genome was quantified by a *US8* qPCR with the following primers F: GGCACCAAATGCAGAGTGAG, R: AAGCCGTATTCCGTTTGCG, and probe: TGGTCCAAGTCCGTGGGCACC (FAM-BHQ-1, Eurofins Genomics GmbH, Ebersberg, Germany). The absolute quantification was performed based on a plasmid standard (10^6 to 10 copies/well) with the TaqMan™ Universal PCR Master Mix (Thermo Fisher Scientific, Waltham, MA USA) at 45 cycles of 94°C 20 s, 57°C 20s and 72°C 20s. Because the percentage of infected cells varied from one experiment to another, the genome copy numbers were first normalized to the amount of total cellular DNA and then expressed as an index of the amount of viral genomes in HCMV-infected untransduced U373-MG cells.

### 1.2.13 Western Blot

MRC5 cells(two days pi) or U373-MG cells (eight days pi) were harvested and the proteins were extracted by the NucleoSpin TriPrep Kit (Macherey-Nagel, Germany) according to manufacturer's protocol. 15 µL of the samples were separated by SDS-PAGE and transferred via semi-dry western blot on a nitrocellulose membrane (GE Healthcare life science, UK) for the U373-MG lysate or via liquid transfer on a PVDF membrane (Millipore) for the MRC5 cell-lysates. The membrane was blocked for 2 h in 5 % milk in TBST. IE were detected by the mouse anti-CMV antibody (MAB810R, Millipore, Germany) and a donkey anti-mouse HRP antibody (Jackson Immuno Research Labs, USA). After being washed, the membrane was incubated 5 min with the SuperSignal™ West Pico Chemiluminescent Substrate (Thermo Fisher Scientific, USA) and the signal revealed by the Luminescent Image analyzer LAS-4000 (FujiFilm, Japan). Following a short wash in TBST,

antibodies were removed from the membrane by Restore Western Blot Stripping Buffer (Thermo Fisher Scientific, USA) for 30 min at room temperature and were used for subsequent detections. GAPDH or actin detection was used as a housekeeping protein with mouse anti-GAPDH (6C5, Santa Cruz, USA) or the mouse anti-actin (C4, Santa Cruz, USA) antibodies and a secondary antibody donkey anti-mouse HRP. The Cas9-GFP from the U373-MG lysate was detected with a rabbit anti-GFP antibody (Invitrogen- Thermo Fisher Scientific, USA) and revealed by a goat anti-rabbit HRP secondary antibody (Jackson Immuno Research Labs, USA). The Cas9 from the MRC5 cell-lysate was detected with the anti-Cas9-A647 (clone 7A9-3A3 Alexa 488, Cell signaling, The Nederland) and revealed with a secondary antibody donkey anti-mouse HRP. For the quantification of the proteins on the membrane, the pictures were analyzed by the software GIMP 2. The background signal of the membrane was subtracted and the signal intensity of each band was calculated in an arbitrary unit/$mm^2$. The Cas9 expression was normalized to the actin expression.

### 1.2.14 Virion release analysis by trans-infection plaque assay

MRC5 cells were plated in a 24-well plate (Falcon, Corning Incorporation, USA) at a density of 2*10^5 cells/mL to be used the day after for trans-infection plaque assay. Eight days after HCMV-infection, U373-MG cells were harvested, counted, serially diluted (from 10^5 to 1 cell per well) and seeded over the MRC5 cell-monolayer in duplicates. After overnight incubation, liquid media was replaced by 0.8% agarose (Sigma, USA) in MEM (Gibco life technology, USA) (10% FCS, 100 U/mL Penicillin/Streptomycin, 2 mM L-Glutamine, 10 mM HEPES). After 7-14 days, plaques were observed by phase-contrast microscopy and counted.

## 1.3 Results

### 1.3.1 Design of the gRNAs

The *UL122/123* gene consists of 5 exons. The start codon is located in exon 2. Two main splice variants are expressed from the *UL122/123* gene, both containing the exons 1 to 3 and are alternatively spliced to either exon 4 (IE1) or exon 5 (IE2)[213]. Both splice variants are essential for the viral replication cycle. Therefore, we designed five gRNAs targeting the *UL122/123* gene and created two strategies to knock-out the IE1/2 expression. The singleplex strategy consists of one gRNA and the Cas9. To target IE1 and IE2 with only one gRNA, we chose a gRNA targeting close to the start codon, in order to induce a frameshift mutation and a concomitant premature stop codon or to disrupt the start codon itself (Figure 18 A). Three gRNAs were designed to be tested as a singleplex strategy to find the most efficient one.

For the multiplex strategy, we decided to excise the entire coding region of the *U122/123* gene. Therefore, we chose the most efficient gRNA used for the singleplex strategy and designed a second gRNA targeting the end of exon 5, but still inside the coding region (Figure 18 B). Several groups have reported that the excision by a duplex strategy leads to an inversion but not to the loss of the target region in around 12% of the events[490]. The inversion of the target region from exon 2-5 would nearly maintain the complete ORF for the UL122/123 gene. Even though, a low leaky expression of IE1/IE2 from the inverted region would be unlikely, some of the minor splice variants would be unaffected and could possibly be expressed normally. For example, the IE1x4, which is expressed MIEP-independently and is only encoded by exon 4. It has an essential role for genome maintenance during viral latency and therefore it would be beneficial to prevent its expression, too. To avoid the inversion and expression of IE splice variants, a third target site was chosen at the beginning of exon 5. This should fragmentise the *UL122/123* gene and lead to the deletion of the entire target region.

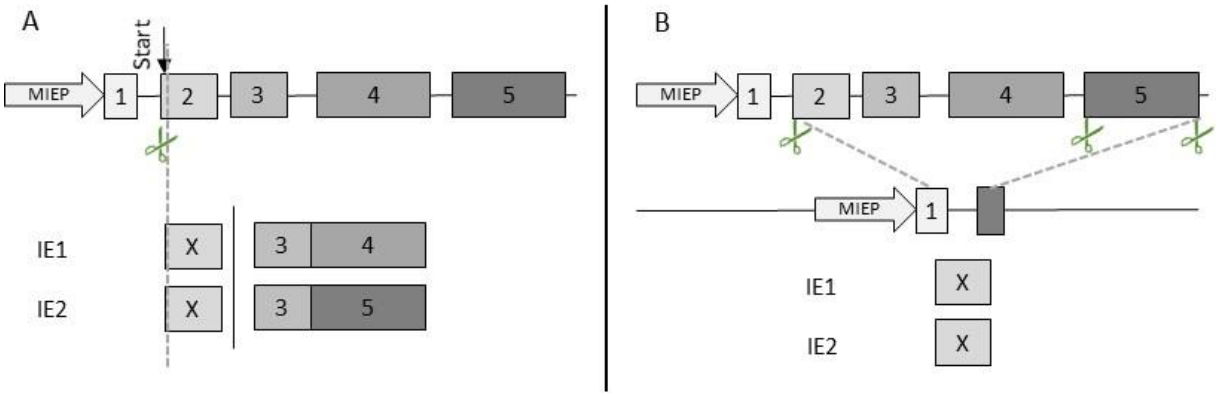



**Figure 18: Schema of the anti-*UL122/123* CRISPR/Cas9 strategies to prevent IE1 and IE2 expression. A) Singleplex: One gRNA targets the exon 2 of the *UL122/123* gene close to the start codon in order to induce a frameshift mutation or remove of the ATG. B) Multiplex: Three gRNAs target the *UL122/123* gene close to the start codon in exon 2 and at the beginning and end of exon 5, which should mainly lead to the deletion of the exons 2-5.**

#### 1.3.1.1 Singleplex

The singleplex gRNAs have been selected manually for the motif $N_{20}$NGG in the exon 2 of the *UL122/123* gene, because, at this time, the only web tool available for gRNA design did not allow the insertion of customized DNA sequences. Therefore, all available complete genomes of HCMV strains on the NCBI database were aligned for exon 2 of the *UL122/123* gene to choose the gRNAs

in a highly conserved region (Table 4). With the further progress of the CRISPR design web tools we validated our gRNAs later on with the Zhang lab, MIT[491] and CRISPOR software to screen for off-target sites. Finally, three gRNAs were selected (see Table 3).

**Table 3: gRNAs for singleplex**

| | Sequence | Strand | length | Number of off targets |
|---|---|---|---|---|
| **gRNA 1** | GTCCATCTTTCTCTTGGCAG**AGG** | + | 20 nt | 190 (min. 1 mismatch) |
| **gRNA 2** | GGACTCCATCGTGTCAAGGA**CGG** | + | 20 nt | 53 (min. 2 mismatches) |
| **gRNA 3** | **CCA**TCGTGTCAAGGACGGTGAC | - | 19 nt | 57 (min. 2 mismatches) |

Those gRNAs were cloned into an LV containing a Cas9-GFP under the control of the standard U6 promoter. Subsequently, these LVs were used to transduce the HCMV-permissive cell line U373-MG and transduced cells were sorted by flow cytometry in order to gain three stable cell lines each expressing one gRNA and the Cas9-GFP.

**Table 4: Alignment of different HCMV strains for the target region of the singleplex strategy**

| TB40/E clone Lisa | CGTGGCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAAAAGA<br>CGTGGCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAAAAGA<br>CGTGGCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAAAAGA |
|---|---|
| 3157 | **********A**A**********_*_ |
| 3301 | **********A**A********** |
| AF1 | **********A**A********** |
| HAN13 | **********A**A********** |
| HAN38 | **********A**A********** |
| JHC | \-********** |
| JP | **********_*_ |
| Toledo | ********** |
| Towne | ********** |
| U8 | ********** |
| U11 | **********A**A**********_*_ |
| VR1814 | ********** |

*Green : gRNA-PAM, yellow: nucleotide substitution, grey: deletion, purple: start codon*

### 1.3.1.2 Validation of the most efficient gRNA for the singleplex strategy

To test the on-target efficiency of the three gRNAs, each stable U373-MG cell line was infected with HCMV at three different MOIs (1, 0.1 and 0.01). Two, four and eight days post infection (pi), the infected cells were collected and the total DNA was extracted. Subsequently, the viral genome was analysed for small *indels* in the *UL122/123* gene in exon 2 by PCR/T7 assay. The T7 assay revealed *indels* at the target region for all three gRNAs. Figure 19 A shows representative PCR and T7 assay results at four days pi. The control U373-MG cells (non-transduced) had only the WT *UL122/123* gene, no *indels* could be detected in the T7 assay. At an MOI of 1 or 0.1 there was a high frequency of mutated *UL122/123* genes for all three gRNAs. Even though, the detection of *indels* at very low MOIs (0.01) was more difficult because of the low viral genome copy number in the DNA samples, we still detected some *indels*. Subsequently, the frequency of *indels* at all time points was calculated based on the formula from Hsu *et al.*[354] (Figure 19B).

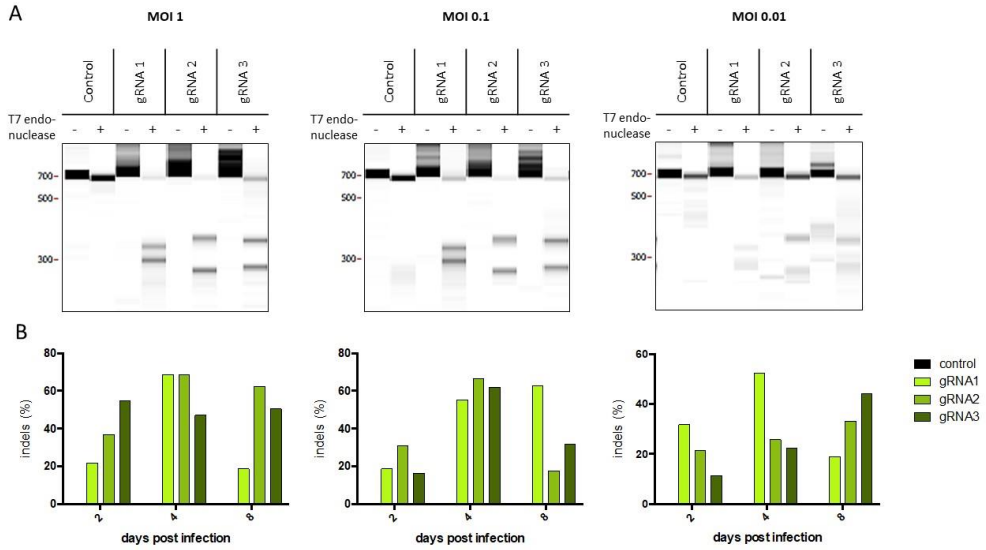



**Figure 19: T7 assay of the *UL122/123* gene exon 2. U373-MG cell lines and untransduced U373-MG control cells were infected with HCMV TB40GFP at different MOIs (1, 0.1 and 0.01). Two, four and eight days pi, cells were harvested, total DNA was extracted and exon 2 of the *UL122/123* gene was PCR amplified and analysed by T7 assay for *indels*. A) A representative electrophoresis picture of the Caliper device for the PCR and T7 assay four days pi is shown. B) Quantitative analysis of the *indel* frequency based on the formula by Hsu *et al.*[491] (n=1)**

Two days pi, *indel* frequencies of 10-50% were detected. The highest amount of mutated *UL122/123* was detected at four days pi, especially for gRNA 1 and 2. At eight days pi, the *indel* frequency decreased slightly, depending on gRNA and MOI. In general, lower frequencies were expected at the late time point, because the remaining WT viral genomes entered the replication cycle leading to the acquisition of more viral genome copies in the cell. Moreover, the viability of the cells is lower at day eight (observation by light-microscopy).

To validate that the observed *indels* were caused by the Cas9/gRNA DSB, the PCR products from day four pi were analysed by TA-cloning and Sanger sequencing (Table 5, Table 6 and Table 7). Small deletions of 1—12 bp were the most frequent mutations. For gRNA 1, around 60% of the mutations caused a frameshift mutation, because the gRNA targets inside the ORF of *UL122/123* gene. gRNA 2 and 3 were designed to target close to the start codon of the *UL122/123* gene, but they both induce the DSB in front of the ATG and therefore outside the ORF. Even though, the gRNA 2 induced a DSB further away from the ATG (8 bp) than the gRNA 3 (3 bp), it still was more efficient to disrupt the start codon (Table 6) and to cause *indels* (Figure 19).

**Table 5: Sequences of exon 2 *UL122/123* of the U373-MG gRNA1 cell lines**

| WT sequence | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTCCAT | effect |
|---|---|---|
| Sample 1 | TCAGGGTTGTCAGGGTCCATCTTTCTCT---CAGAGGACTCcAT | In-frame |
| Sample 2 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGG-AGAGGACTCCAT | Frameshift |
| Sample 3 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGG-AGAGGACTCCAT | Frameshift |
| Sample 4 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGG------ACTCCAT | In-frame |
| Sample 5 | TCAGGGTTGTCAGGGTCC----------CAGAGGACTCCAT | Frameshift |
| Sample 6 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTG-CAGAGGACTCCAT | Frameshift |
| Sample 7 | TCAGGGTTGTCAGGGTCCATCTTTCTCT---CAGAGGACTCCAT | In-frame |
| Sample 8 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGG-AGAGGACTCCAT | Frameshift |
| Sample 9 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGG-AGAGGACTCCAT | Frameshift |
| Sample 10 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTGG------ACTCCAT | In-frame |
| Sample 11 | TCAGGGTTGTCAGGGTCC----------CAGAGGACTCCAT | Frameshift |
| Sample 12 | TCAGGGTTGTCAGGGTCCATCTTTCTCTTG-CAGAGGACTCCAT | Frameshift |

*Green: protospacer + PAM; purple: start codon; blue: deletions*

**Table 6: Sequences of exon 2 *UL122/123* of the U373-MG gRNA2 cell lines**

| WT sequence | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAA | effect |
|---|---|---|
| Sample 1 | CTTGGCAGAGG----------GGACGGTGACTGCAGAA | Loss ATG |
| Sample 2 | CTTGGCAGAGGACTCCATCGTGTCA-GGACGGTGACTGCAGAA | NON |
| Sample 3 | CTTGGCAGAGGACTCCATCGTGTCAA--ACGGTGACTGCAGAA | NON |
| Sample 4 | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAA | WT |
| Sample 5 | CTTGGCAGAGGACTCCATCGTGTCAAAGGACGGTGACTGCAGAA | NON |
| Sample 6 | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTACAGAA | NON |
| Sample 7 | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAA | WT |
| Sample 8 | CTTGGCAGAGGACTCCATCGTGTCAA---CGGTGACTGCAGAA | NON |
| Sample 9 | CTTGGCAGAGGACTCCATCGTGTCAAAGGACGGTGACTGCAGAA | NON |
| Sample 10 | CTTGGCAGAGGACTCCATCGTGTCAAAAGGACGGTGACTGCAGAA | NON |
| Sample 11 | CTTGGCAGAGGACTCCATCGTGTCA-GGACGGTGACTGCAGAA | NON |
| Sample 12 | CTTGGCAGAGGACT----------GTGACTGCAGAA | Loss ATG |
| Sample 13 | CTTGGCAGAGGACTCCATCGTGTCAAAGGACGGTGACTGCAGAA | NON |

*Green: protospacer + PAM; purple: start codon; blue: deletions; yellow: insertions; dark blue: substitution*

**Table 7: Sequences of exon 2 *UL122/123* of the U373-MG gRNA3 cell lines**

| WT sequence | TTTCTCTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGC | effect |
|---|---|---|
| Sample 1 | TTTCTCTTGGCAGAGGACTCCATCGT--CAAGGACGGTGACTGC | Non |
| Sample 2 | TTTCTCTTGGCAGAGGACTCCATCG-GTCAAGGACGGTGACTGC | WT |
| Sample 3 | TTTCTCTTGGCAGAGGACTCCATCGTTGTCAAGGACGGTGACTGC | Non |
| Sample 4 | TTTCTCTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGC | WT |
| Sample 5 | TTTCTCTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGC | WT |
| Sample 6 | TTTCTCTTGGCAGAGGAC----------*GGTGACTGC | Loss ATG |
| Sample 7 | TTTCTCTTGGCAGAGGACTCCATCG-GTCAAGGACGGTGACTGC | NON |
| Sample 8 | TTTCTCTTGGCAGAGGACTCCATCGTTGTCAAGGACGGTGACTGC | NON |
| Sample 9 | TTTCTCTTGGCAGAGGACTCCATCGTTGTCAAGGACGGTGACTGC | NON |

*Green: protospacer + PAM; purple: start codon; blue: deletions; yellow: insertions*

Finally, the effect of the mutations was analysed by the reduction of IE1 and IE2 expression. Two, four and eight days pi, cells were harvested and stained for intracellular IE1/2 and analysed by flow cytometry. In control U373-MG cells, the fraction of $IE^+$ cells was decreased at eight days pi probably due to the progression of the viral replication cycle (IE expression is downregulated at late time points) and decreased viability of the cells due to prolonged culture without medium refreshments. The IE expression was reduced by all three gRNA to the same extend. At a high MOI, the IE reduction was 40% at all time points (Figure 20 A). The effect was much stronger for lower MOIs (0.1 and 0.01) with an IE inhibition of 60% in comparison to the non-transduced control cells (Figure 20 B and C). The CRISPR/Cas9 system is probably more efficient at lower MOIs, because the viral load per cell is much lower, so that the Cas9 has to target only a few HCMV genome copies per cell.

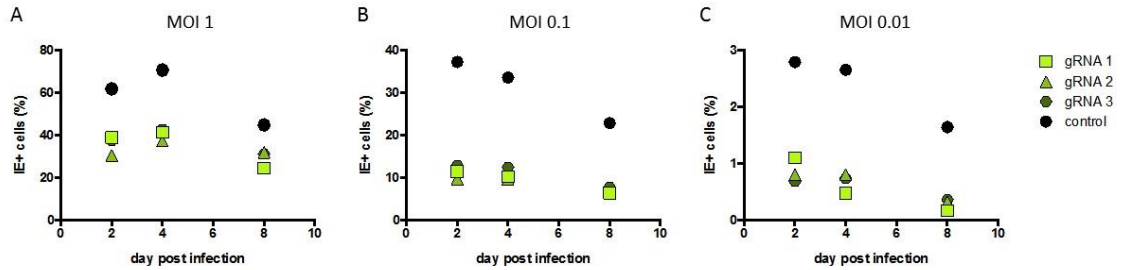



**Figure 20: Decrease of the IE expression by the singleplex CRISPR/Cas9 systems. The U373-MG cell lines were infected with HCMV TB40GFP at different MOIs. Cells were harvested and analysed by flow cytometry for intracellular IE expression two, four and eight days pi (n=1).**

In conclusion, all three gRNAs inhibit the IE expression to the same extent, even though we have shown that only gRNA1 directly targets the coding region. gRNA 2 and 3 only target close to the start codon without directly disrupting the ORF, but may still interfere with regulatory regions.

Since these gRNAs were only selected by hand without a proper off-target analysis program, it is necessary to validate that the gRNAs do not target the human genome. With the advancement of the CRISPR design web tool from Zhangs lab, MIT[491], we finally had the opportunity to screen for potential off-target sites in the human genome. The three most probable off-target sites for each gRNA were selected, amplified by PCR and subjected to the T7 endonuclease. Figure 21 shows the results of the T7 assay for the transduced U373-MG stable cell lines and the non-transduced control to ensure that detected cleavage products are due to mutations caused by the CRISPR/Cas9 system and not naturally occurring nucleotide polymorphisms. For gRNA1, *indels* were detected for the off-target site 1 (Figure 21 A, second line), but not for off-target sites 2 and 3. gRNA2 and 3 did not cause any *indels* detectable by the T7 assay.

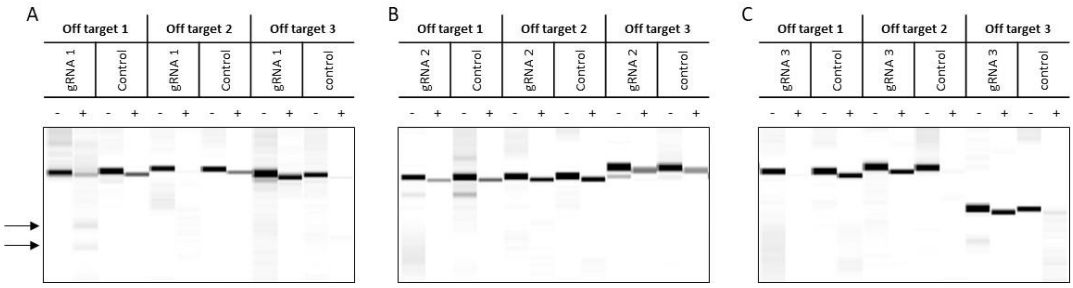



**Figure 21: Off-target analysis of the singleplex strategies by T7 assay. The total DNA of the three U373-MG cell lines and non-transduced U373-MG cells were extracted and the three most probable off-target sites for each gRNA (A: gRNA1; B: gRNA2 and C: gRNA3) were PCR amplified and subsequently subjected to the T7 endonuclease.**

Therefore, gRNA 1 needed to be excluded from further testing, because the risk of off-target mutations in the human genome is too high so that it cannot be used as an antiviral therapy. gRNA 2 and 3 both showed the same safety profile, but in the functional tests (T7 assay and IE expression) gRNA2 showed a slightly higher efficiency. For further analyses, gRNA2 was used for the singleplex strategy testing.

#### 1.3.1.3 Multiplex

These pre-tests revealed gRNA2 as the most efficient with the lowest off-target risks. Consequently, we used this gRNA2 for the multiplex design. gRNA4 and 5 were designed by the CRISPOR web tool on the exon 5 of the TB40GFP HCMV strain. Subsequently, the exon 5 of 14 different HCMV strains were aligned. Finally, highly specific gRNAs at the beginning and end of exon 5, designed by the CRISPOR software, were selected when matching a conserved region for all HCMV strains (Table 9 and Table 10). The new designed gRNAs have minimal homology with the human genome with minimally three mismatches for the most probable off-target site (Table 8). Moreover, the most probable off-target sites were mainly in intergenic regions or introns, and only rarely targeting exons.

**Table 8: gRNAs for the multiplex strategy**

| | Sequence | Strand | length | Number of off targets |
|---|---|---|---|---|
| **gRNA 2** | GGACTCCATCGTGTCAAGGA**CGG** | + | 20 nt | 53 (min. 2 mismatches) |
| **gRNA 4** | GGTGCTACTGGAATCGATAC**CGG** | + | 20 nt | 192 (min. 3 mismatches) |
| **gRNA 5** | GTCCTGGATGGCTGCCTCGA**TGG** | + | 20 nt | 68 (min. 3 mismatches) |

We aimed to use a single vector delivery for the multiplex strategy. Therefore, we designed a multiplex gRNA cassette containing the three gRNA each under a different RNA Pol III promoter. We chose the commonly used U6 promoter and two further promoters tested for the gRNA expression by Kabadi *et al.* (H1 and 7SK)[348]. This multiplex gRNA cassette was subsequently cloned into an LV for the transduction of the target cells (chapter IV-1.3.2, Figure 22)

**Table 9: Alignment of the target region of gRNA 4 in exon 5 of the *UL122/123* gene**

| TB40/E | CGCGACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
|---|---|
| 3157 | ********** |
| 3301 | **********T********** |
| AD169 | ********** |
| AF1 | ********** |
| HAN13 | ********** |
| HAN38 | ********** |
| JHC | ********** |
| JP | ********** |
| Toledo | ********** |
| Towne | ********** |
| U8 | ********** |
| U11 | ********** |
| VR1814 | ********** |

*Green: gRNA-PAM, yellow: nucleotide substitution*

**Table 10: Alignment of the target region of gRNA 5 in exon 5 of the *UL122/123* gene**

| TB40/E | CTATGTACAAGAGTCCATGTCTCT TTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
|---|---|
| 3157 | **********CT********** |
| 3301 | **********CT********** |
| AD169 | **********CT********** |
| AF1 | ********** ********** |
| HAN13 | ********** ********** |
| HAN38 | **********CT**********T********** |
| JHC | **********CT********** |
| JP | ********** ********** |
| Toledo | **********CT**********T******** |
| Towne | **********CT********** |
| U8 | **********G********* **********A** |
| U11 | **********CT********** |
| VR1814 | **********CT********** |

*Green: gRNA-PAM, yellow: nucleotide substitution, light blue: insertion*

#### 1.3.1.4 Internal control

For unbiased testing of the singleplex and multiplex anti-HCMV CRISPR/Cas9 strategies, we designed an unspecific gRNA (unsp. gRNA) to exclude the sequence independent Cas9 induced side effects on the viral replication. Therefore, a gRNA targeting a different DNA virus (BK polyomavirus) was designed by the CRISPOR web tool. A gRNA with minimal off-target potential was selected.

**Table 11: Unspecific gRNA as an internal control**

| | Sequence | Length | Number of off-targets |
|---|---|---|---|
| **Unsp. gRNA** | GAATTTCACCCTGACAAAGG**GGG** | 20 nt | 142 (min. 2 mismatches) |

### 1.3.2 Creation of the stable cell lines for analyzing the anti-viral activity of the singleplex and multiplex CRISPR/Cas9 strategies

To analyse the effect of the different CRISPR/Cas9 strategies on the HCMV lytic replication cycle, stable HCMV-permissive cells were needed. We decided to work with two different cell types: The well-established astrocytoma cell line U373-MG, which is permissive for HCMV, but only supports the viral replication very slowly and does not produce significant amounts of new virions. As a second cell type, we chose the primary fibroblasts MRC5 to assess the effect of the Cas9/gRNAs on the virion release. MRC5 cells are commonly used for HCMV production and are known to allow a fast viral replication. We tested three strategies to create CRISPR/Cas9 stable cell lines by transducing the cells with an LV and selecting the transduced cells by: **A**: FACS sort of transduced cells based on the Cas9-GFP expression, **B**: magnetic microbeads selection based on the expression of the LNGFR surface marker or **C**: Puromycin resistance (Figure 22).

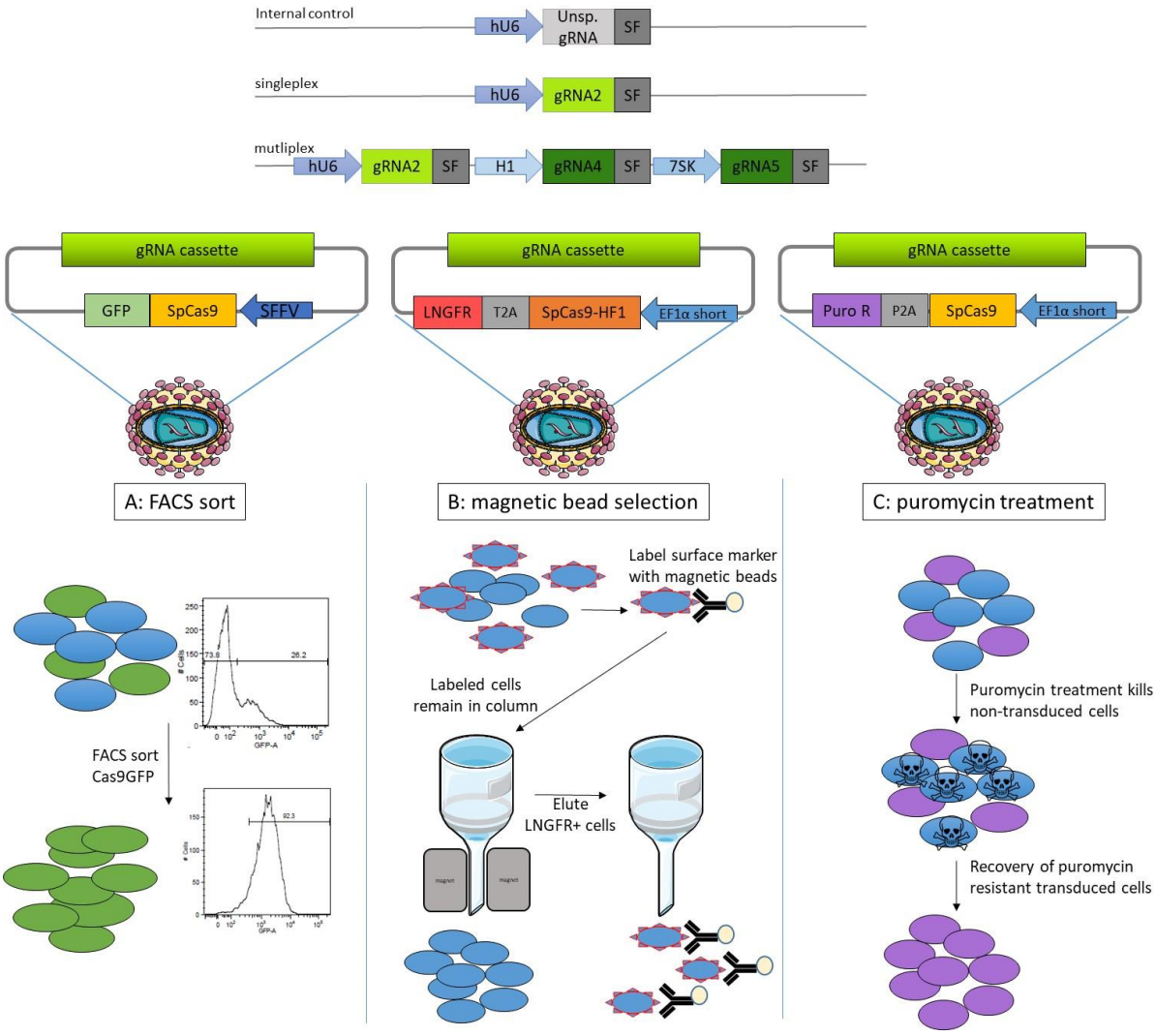



**Figure 22: Schematic representation of the creation of cells stably expressing the CRISPR/Cas9 system. The different gRNA cassettes or the internal control and anti-HCMV strategies (singleplex and multiplex) were cloned into the different types of LVs containing the Cas9 and either a GFP, a truncated LNGFR or a puromycin resistance as a marker gene. A) Demonstration of the selection of transduced cells by a FACS sort based on the Cas9-GFP expression. B) Anti-LNGFR-magnetic beads were used to positively select the transduced cells expressing the surface marker LNGFR. C) A short period of puromycin treatment killed non-transduced cells, while transduced cells were protected against puromycin and recovered in normal complete medium.**

For strategy **A**, the different gRNA cassettes were cloned into an LV containing the SpCas9 fused to GFP. U373-MG were transduced with the three LVs type 1 with an MOI of five to ten and FACS-sorted based on their Cas9-GFP$^{high}$ expression. All three cell lines had a similar mean of fluorescence (MFI = 1985-2203) for Cas9-GFP expression, which represents a similar expression level of the CRISPR/Cas9 system in the different cell lines (Figure 23 A). Throughout all experiments performed on these cell lines, growth kinetics were similar (same frequency of subculturing steps) and the viability of the U373-MG cells were high (Figure 23 B). Neither the Cas9 expression nor the anti-viral gRNAs were toxic. Cas9 expression was also monitored by fluorescence microscopy for GFP.

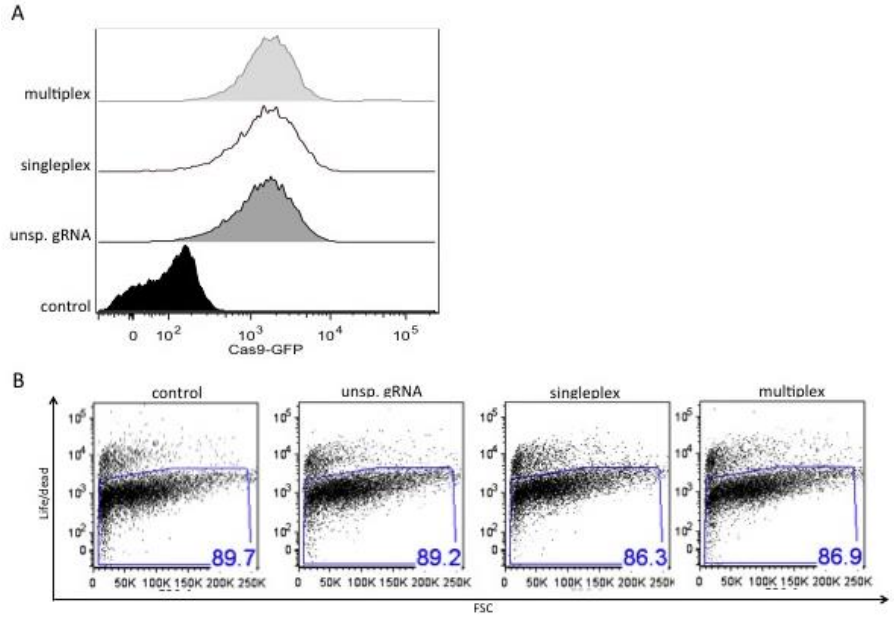



**Figure 23: Generation of U373-MG cell lines expressing the CRISPR/Cas9 strategies. U373-MG cells transduced with one of the three lentiviral vectors type 1 were FACS-sorted based on the Cas9-GFP$^{high}$ (fusion protein) expression. A) Post-sort analysis of each U373-MG cell line is presented. Mean of fluorescence is indicated above each histogram. B) Cell viability analysis by fixable Life/dead stain of the U373-MG stable cell lines and untransduced control cells by FACS.**

MRC5 cells are much harder to transduce and are more sensitive to the LV infection. Moreover, they are primary cells and can only be hold in cell culture for one to two months before they go into senescence. Following the protocol of LV transduction mentioned before, three weeks expansion and subsequent FACS sort failed in three independent experiments. Due to the low amount of GFP$^{+}$ cells, the low purity of the FACS sort and the prolonged culture period, MRC5 cells did not expand after sorting. To increase the purity of the FACS sort, different nozzles were tested, because MRC5 cells are significantly bigger than most other cell types. They sediment rapidly in the FACS tube, which makes it difficult to achieve the optimal cell concentration for the FACS sort. To optimize the concentration, the tubes were vortexed several times during the sort to prevent the settlement of the cells. However, none of the improvements led to a higher purity. Furthermore, to avoid the long expansion phase, MRC5 cells were FACS sorted only eight days post transduction, but only a very small number of cells could be selected at this moment and they did not start growing in cell culture again.

To avoid the rough and time consuming selection by FACS, a new LV vector was designed containing the SpCas9-HF1[374] and a truncated LNGFR (low-affinity nerve growth factor receptor) connected by a T2A sequence. This truncated LNGFR is supposed to be expressed on the cell

surface of the transduced i.e. Cas9 expressing cells. The transduction of MRC5 cells with this LV should enable the selection of transduced cells only a few days post transduction (pt). However, preliminary tests on easy-to-transduce HEK293T cells failed to show LNGFR expression. The transfection of the same constructs (LV backbone 2) gave rise to an intermediate LNGFR signal. We concluded that the LNGFR expression was not high enough by LV transduction to allow a clean selection of transduced cells, even though the constructs itself allowed LNGFR expression.

Finally, a third LV vector was constructed containing the SpCas9-P2A-puromycin resistance. Different transduction and puromycin treatment models were tested. Transduced cells were treated with puromycin (2 µg/mL, 2 days[474]) two to three days pt. This allowed the selection of nearly 90% of the MRC5 cells in the well. However, subsequent splitting and expansion of the cells led to a high mortality and the cells required a prolonged culturing phase to recover. This in turn was a problem, because the long culture time led the MRC5 cells to enter senescence. Consequently, only a few experiments could be performed on those cells. This treatment model was very time consuming and had strong impact on cell viability. To shorten the experimental design, MRC5 cells were transduced, puromycin treated and infected in the same well without a sub-culturing step in between. Unfortunately, in this model, cells showed a high susceptibility to the HCMV infection and not the expected protection. The transduction and puromycin treatment had a strong impact on the transgene expression and viability of the cells so that the protective effects could not be assessed anymore. Since high (2 µg/mL) and rapid puromycin treatment had a toxic effect even on transduced cells, we reduced the puromycin concentration to 0.5 µg/mL and extended the treatment phase from two to nine days. Subsequent sub-culturing still resulted in a high mortality of the MCR5 cells and a two to three weeks recovery phase was necessary before starting the infection with HCMV. Finally, we analyzed the Cas9 expression by western blot and detected only a very low amount of Cas9 expression in the singleplex and multiplex cells as compared to unsp. gRNA cells (Figure 24 A). These differences are either due to a toxicity induced by the anti-viral gRNAs or by different copy numbers of LV transgene integration. The viability of the transduced MRC5 cells were assessed after the recovery phase and expansion of the transduced cells by FACS (Figure 24 B). All MRC5 cells have the same viability independently of their transduction and the LV constructs. Thus, we monitored the gRNA/Cas9 expression in transduced MRC5 cells over time to analyse, if the Cas9 expression level decreases stronger with anti-viral gRNA than with the unsp. gRNA. Consequently, the Cas9 expression from the time after transduction till the end of the recovery phase was analysed by western blot. As shown in Figure 24 C and D the Cas9 expression level was reduced during consecutive subculture steps. After three passages, the Cas9 expression was decreased by 60% for the singleplex and multiplex MRC5 cells to 87% for the unsp. gRNA MRC5 cells. Therefore, the decrease of the Cas9 expression was even stronger in unsp. gRNA MRC5 cells than in singleplex or multiplex MRC5 cells, showing that the anti-HCMV gRNAs were not more toxic for the cells than the unsp. gRNA. The differences of the Cas9 expression level are more likely due to varying LV titres used for transduction, which were not titrated for this experiment. However, all LV used on MRC5 cells were produced on the same days.

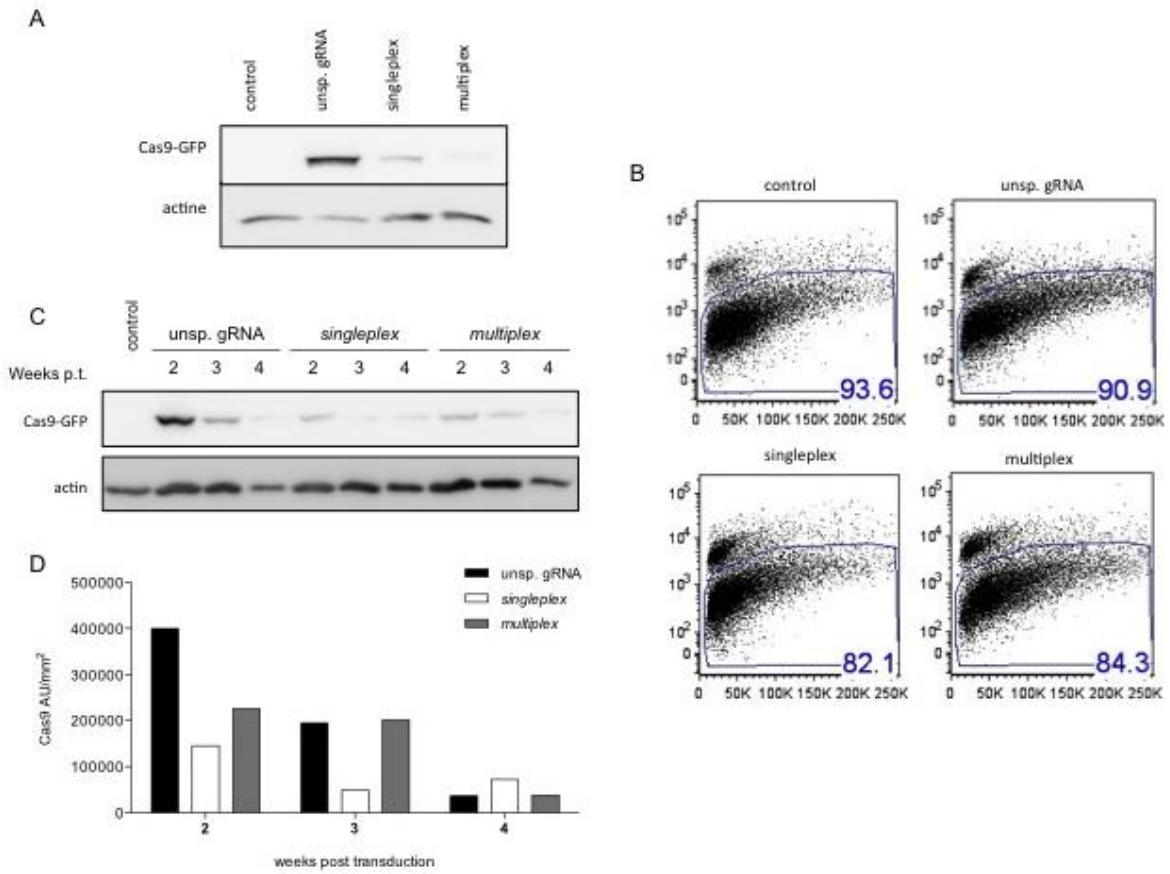



**Figure 24: Cas9 expression in transduced MRC5 cells. MRC5 cells were transduced with one of the three LV type 3 and selected by puromycin treatment (2µg/mL) for 2 days. A) After the selection and a three-week recovery phase, proteins were extracted with the TriPrep Kit and the Cas9 expression was assessed by Western Blot. B) Control and transduced MRC5 cells were stained with the fixable Life/Dead stain Kit to analyse cell viability by FACS. C) In a second transduction and selection experiment, a part of the transduced and selected MRC5 cells was collected during the recovery phase at each passage and analysed by western blot for the Cas9 expression. D) Relative quantification of Cas9 expression based on the Western Blot (C) normalized by housekeeping protein expression of actin. pt: post transduction.**

In conclusion, U373-MG cells were easy to transduce and could be FACS sorted based on a marker gene. They tolerated the LV transduction with nearly no mortality and allowed very long culturing periods without senescence or phenotypic changes. In contrast, MRC5 cells were more fragile. They are primary cells and thus only expand for several weeks before entering into senescence. Transduction with an LV is possible without high cytotoxicity, but the selection of the transduced cells was very challenging. These large cells could not be FACS-sorted with a high purity. Surface marker selections failed due to too low transgene expression. The only partly efficient selection was by puromycin resistance, even though prolonged culture periods for recovery were necessary and the Cas9 expression decreased over time.

### 1.3.3 Multiplex CRISPR/Cas9 system impairs HCMV replication by excising an essential viral gene

The following results are part of our publication: *Multiplex CRISPR/Cas9 system impairs HCMV replication by excising an essential viral gene, Gergen et al.* (in submission). We analysed the effect of the two anti-HCMV CRISPR/Cas9 systems (singleplex and multiplex) on the lytic viral replication cycle.

#### 1.3.3.1 Anti-IE CRISPR/Cas9 system reduces IE expression in primary fibroblasts

MRC5 primary fibroblast cells were transduced and selected via puromycin treatment as descripted before (chapter IV-1.3.2). Subsequently, cells were infected with Toledo (MOI 0.1) to assess the effect of the anti-IE strategies on the viral replication. Two days pi, the viral genome was analyzed for *indels* by T7-endonuclease assay (Figure 25 A). The efficiency of the singleplex cutting was calculated as described elsewhere[354,491,492] (Table 12). Twenty-nine percent of the viral genome had *indels* at the target site. To analyze the efficiency of the multiplex strategy, a PCR spanning exons 2 to 5 was performed, which gave a 3862 bp amplicon. Interestingly, while this WT amplicon was strongly detected in control MRC5 cells and in MRC5 cells expressing the unsp. gRNA; multiplex MRC5 cells showed a weaker WT band. Moreover, a weak amplification of 500 bp was also observed (Figure 25 B), which probably represents a deletion of 3300 bp in the *UL122/123* gene between the target sites of gRNA2 and gRNA5. When quantifying the weaker band (500 bp) in comparison with the wild type band (3862 bp), around 5% of the viral genome copies showed this deletion in the *UL122/123* gene. In PCR, the amplification of small products are favored in comparison to longer fragments, so that these percentages might not represent the exact quantity of mutations in the viral genome extracted from HCMV-infected MRC5 cells. However, the presence of the small PCR fragments still proofs that a part of the viral genome has a larger deletion in the *UL122/123* gene.

**Table 12: Relative quantification of CRISPR-induced mutations in *UL122/123* gene in MRC5 cells**

| | ***Mean in % ±SD*** | | | |
|---|---|---|---|---|
| ***HCMV strain*** | *control* | *unsp. gRNA* | *singleplex* [a] | *multiplex* [b] |
| *Toledo* | 0 ±0 | 0 ±0 | 29.02 ±2.31 | 5.36 ±0.86 |

*Mean percentages of indels are presented.* [a] *The percentage of mutation for the singleplex strategy is analysed based on the T7 assay and quantification of the PCR products and cleavage products by the Caliper microfluidic bioanalyzer.* [b] *Bigger deletions induced by the multiplex strategy are analysed by PCR and quantified by the Caliper microfluidic bioanalyzer (n=2 or 3 independent experiments for each transduced MRC5 cells).*

The effect of those mutations in the *UL122/123* gene on the IE expression was then analyzed by western blot. The singleplex and multiplex strategies enabled a strong reduction of the IE1 expression and nearly abrogated the IE2 expression (Figure 25 C). The unexpected strong decrease of IE protein expression by the multiplex strategy is probably due to a combination of large deletions (5%) as analyzed by PCR and *indels* at each target site itself. Furthermore, the cut by the gRNA/Cas9 and the subsequent repair takes at least 5 hours[362], which delay the IE expression, even if the gene is correctly repaired. The viability of the cells could not be assessed by western blot, but pretests of viral growth kinetics on MRC5 cells have shown that there is no viral induced cytopathic effect on the cells three days pi at an MOI of 0.1 (Figure 25 D). Moreover, the slightly higher expression of IE in the unsp. gRNA MRC5 cells in comparison to the control cells shows that the transduced cells are equal or even more susceptible to the HCMV infection than non-transduced MRC5 cells. Together, this confirms that the observed inhibition of IE expression is specific to the anti-HCMV CRISPR/Cas9 approaches.

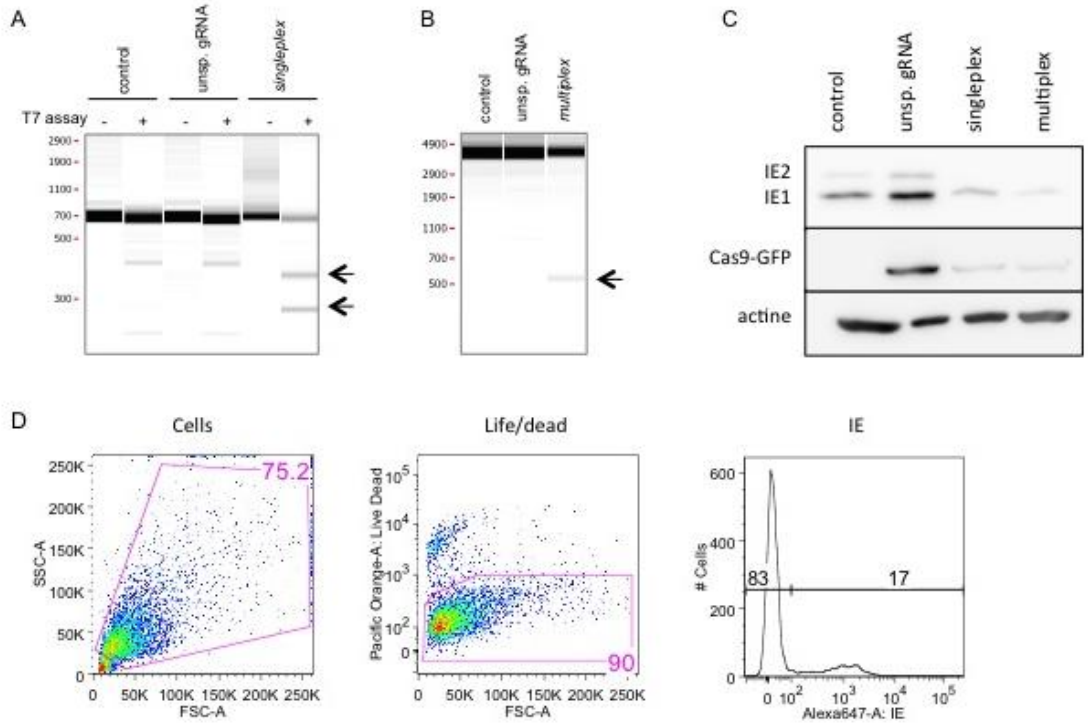



**Figure 25: Anti-HCMV CRISPR/Cas9 system induces mutations resulting in the decrease of IE protein expression in primary fibroblasts: MRC5 cells were transduced with one of the three LV type 3 and selected by puromycin treatment (2µg/mL) for two days. Control (untransduced) and puromycin-resistant MRC5 cells were subcultured prior to be infected with Toledo (MOI 0.1). Two days pi, proteins and DNA were extracted from the infected cells via the TriPrep Kit. A) Viral DNA extracts were PCR-amplified at the target region. Amplicons were subsequently subjected to the T7 endonuclease to detect *indels* induced by the singleplex strategy. B) PCR amplicons of the whole IE gene were analyzed to detect bigger deletions induced by the multiplex strategy. The arrows highlight the *indels* (singleplex) and bigger deletions (multiplex) induced by the anti-HCMV CRISPR/Cas9 strategies. (one out of three independent experiments is shown). C) Western Blot analysis of the IE and Cas9 expression two days pi (one representative western blot out of 3 independent experiments is shown). D) MRC5 cells infected with HCMV (MOI 0.1) were harvest 3 days pi, stained with the fixable Life/Dead stain Kit and intranuclear IE expression and analysed by FACS (n=1)**

Here we show that anti-HCMV CRISPR/Cas9 strategies disturb the viral genome at the target site, which results in a strong decrease of IE protein expression.

### 1.3.3.2 Stable expression of the anti-IE CRISPR/Cas9 system induces mutations in the *UL122/123* gene in HCMV-infected U373-MG cells

Since the Cas9 expression was not stable in MRC5 cells, we decided to perform a more detailed analysis of the efficiency of the singleplex and multiplex strategies in a HCMV-permissive astrocytoma cell line (U373-MG). These cells support a full lytic replication cycle of HCMV. As descripted before, U373-MG were transduced with the LV type 1 and FASC-sorted based on the Cas9-GFP expression (chapter IV-1.3.2, Figure 23).

The three transduced U373-MG cell lines were infected with three different low passage HCMV strains: TB40GFP, Toledo and VR1814[493]. The HCMV viral genome was extracted eight days pi to analyze mutations induced by the gRNA/Cas9. The efficiency of both, the singleplex and multiplex strategies, was assessed as described before with the MRC5 cells. The singleplex strategy yielded 30–50% *indels* (Figure 26 A and B, and Table 13) thus slightly higher than in the MRC5 cells. For the multiplex strategy, a major amplification of 500 bp (Figure 26 C and D,) representing the deletion of the target region between gRNA1 and gRNA3, and a smear above, representing smaller deletions, were detected. The quantification of this band and the smear above in comparison to

the WT band revealed that up to 95% of the viral genome was affected by the multiplex strategy (Table 13). Importantly, all three viral strains tested were similarly efficiently targeted showing the universal usage of our anti-*UL122/123* gRNAs (Table 13).

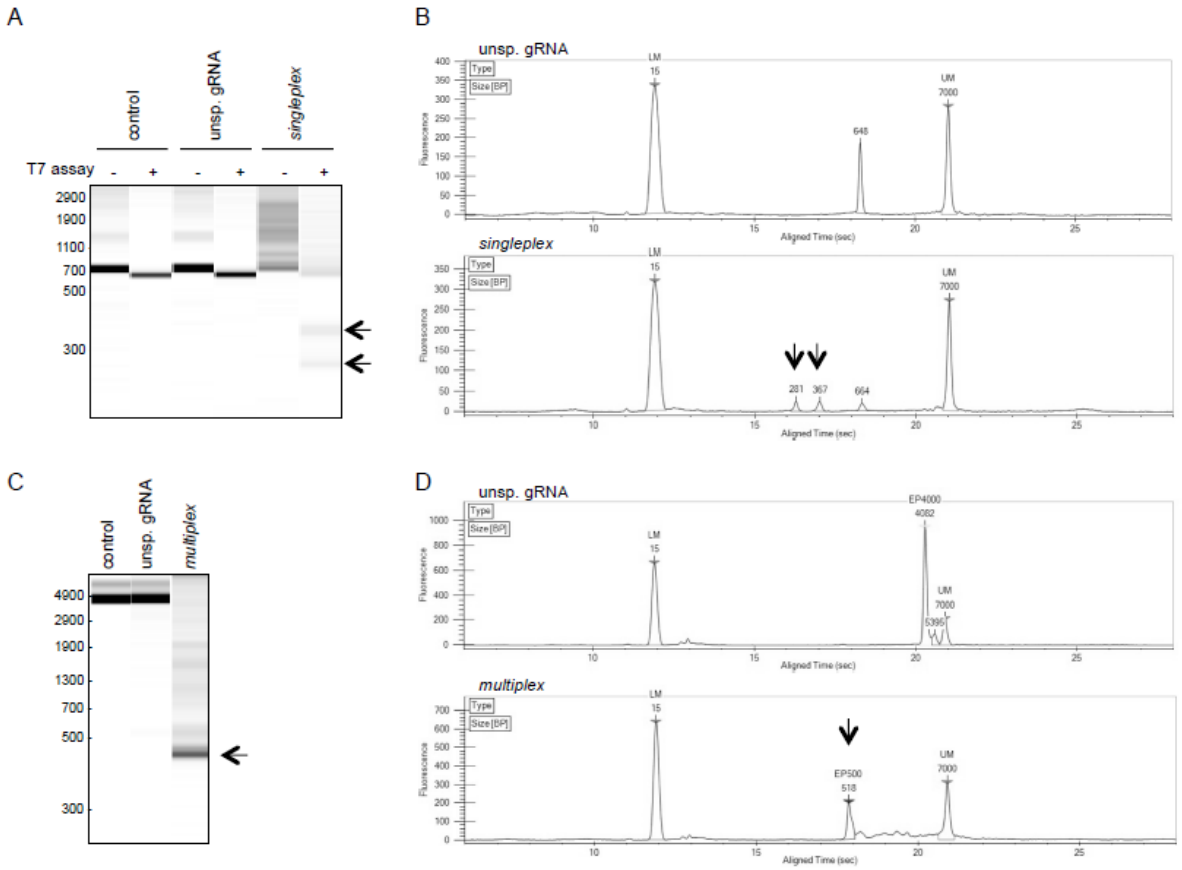



**Figure 26: Mutations in the *UL122/123* gene induced by the CRISPR/Cas9 system anti-HCMV in U373-MG cell line: Control and transduced U373-MG cells were infected with HCMV (Toledo, MOI 1) and cultured for eight days. Viral DNA was extracted and PCR amplified. A) T7-assay was performed on the PCR amplicon of exon 2 to detect *indels* induced by the singleplex strategy. B) Electrogramm of the T7 assay by the Caliper LabChip analysis for the in the unsp. gRNA and singleplex strategies. C) Large deletions induced by the multiplex strategy were highlighted by analyzing the whole *UL122/123* gene amplicon. D) Electrogramm of the PCR by the Caliper LabChip analysis identify a major amplicon of 500 bp and a smear above for the multiplex strategy. Arrows highlight the *indels* (singleplex) and bigger deletions (multiplex) induced by the anti-HCMV CRISPR/Cas9 strategies. One representative experiment out of three is shown for Toledo, similar data were found with TB40GFP and VR1814 (n=3 independent experiments per virus strain) LM, lower marker; UM, upper marker.**

**Table 13: Relative quantification of CRISPR-induced mutations in *UL122/123* gene in U373-MG cells**

| | ***Mean in % ±SD*** | | | |
|---|---|---|---|---|
| ***HCMV strain*** | *control* | *unsp. gRNA* | *singleplex* [a] | *multiplex* [b] |
| *TB40-GFP* | 0 | 0 | 50.63 ±9.25 | 95.18 ±5.47 |
| *Toledo* | 0.28 ±0.69 | 0.09 ±0.22 | 31.18 ±5.18 | 92.14 ±4.69 |
| *VR1814* | 0.86 ±1.62 | 1.83 ±3.09 | 46.46 ±11.78 | 80.00±8.58 |

*Mean percentages of indels are presented. [a] The efficiency of the singleplex strategy to induce mutations is analysed based on the T7 assay and quantified by the Caliper microfluidic bioanalyzer. [b] Detection of deletions induced by the multiplex strategy are analyzed by PCR and quantified by the Caliper microfluidic bioanalyzer (n=3 independent experiments per virus strain).*

Overall, the multiplex strategy was more efficient than the singleplex strategy and showed significant higher yields of mutations in the viral genome.

#### 1.3.3.3 Dramatic decrease of IE protein expression in HCMV-infected U373-MG cells expressing gRNA/Cas9

We analyzed, if the induction of mutations in the IE gene led to a concomitant reduction of IE expression in the different U373-MG cell lines, two and eight days pi, with three different viral strains. The unsp. gRNA cell line was equally permissive to HCMV infection compared to untransduced control cells for the three HCMV strains tested suggesting that there was no effect of the Cas9/unsp. gRNA on the viral infection (Figure 27 A and B). HCMV-infected singleplex U373-MG cells showed a reduction of IE positive cells of up to 50 % with TB40GFP or Toledo at a high MOI (1). The multiplex strategy was significantly more efficient than the singleplex and reduced the amount of IE positive cells by 75-85% (Figure 27 A and B). The endotheliotropic HCMV strain VR1814 could only be used at a low MOI (0.1). In this condition, the singleplex strategy was already significant to reduce the IE expression by up to 75 % (Figure 27 A and C). The decrease of IE positive cells by the multiplex strategy reached up to 95% for VR1814. In comparison to the strains, TB40GFP and Toledo, at a low MOI, the effect on IE reduction was also stronger than with an MOI of 1 (Figure 27 C). Subsequent analyses for TB40GFP and Toledo were done with an MOI of 1 to strongly challenge the anti-viral CRISPR/Cas9 system. Comparing the effect of both strategies between day two and day eight pi, it appears that the decrease of IE expression was stable over time when cells were infected with TB40-GFP. The IE expression decreased significantly over time in cells harboring the multiplex strategy infected with Toledo (MOI 1). In U373-MG cells, HCMV did not induce a significant cytopathic effect as shown in Figure 27 D comparing the viability of infected to non-infected cells. Still, on day 8, cell viability was reduced. However, this mortality is due to prolonged culture without media refreshment (to not interfere with the amount of released virions and the readout of the assay). In conclusion, the observed IE inhibition was indeed induced by the anti-viral CRISPR/Cas9 approaches and are not an artifact due to virus-induced mortality.

A western blot analysis was performed to analyze both major IE splice variants (IE1 and IE2). The expression of both IE variants was impaired by the anti-*UL122/123* CRISPR/Cas9 system with a higher effect on IE2 than on IE1. Importantly, the IE2 expression was undetectable with the multiplex strategy for Toledo and VR1814 (Figure 27 E), which could point out a possible knock-out of IE2. We also analyzed the expression of the Cas9 by western blot and could confirm a stable and comparable expression level of the Cas9 in all three U373-MG cell lines. Mutations induced by the singleplex and multiplex strategies led to a significant and stable decrease of the number of IE positive cells over time and to almost undetectable levels of the IE2 protein.

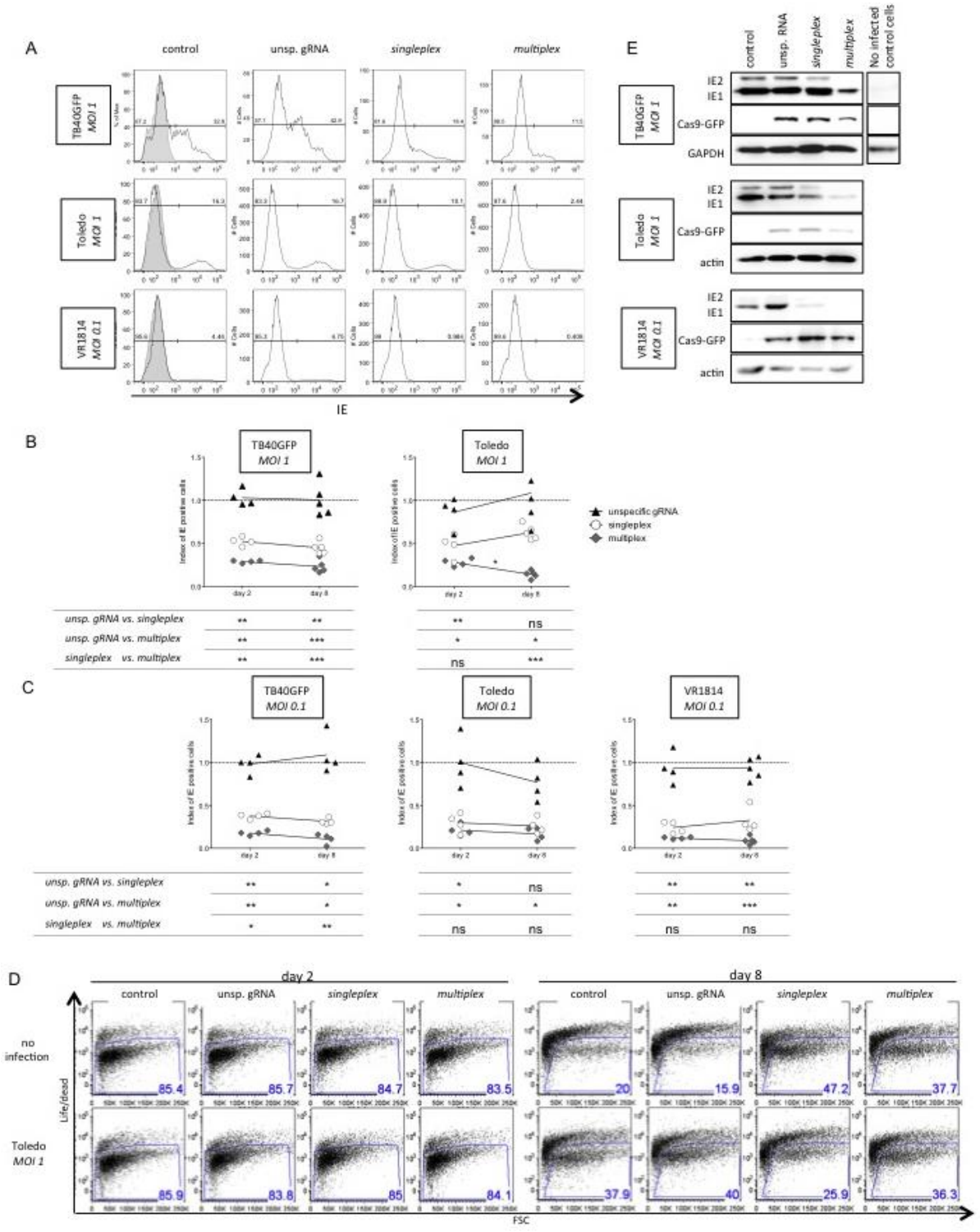



**Figure 27: Decrease of IE expression by the HCMV targeting CRISPR/Cas9 systems: Control and transduced U373-MG were infected with HCMV and harvested at two or eight days pi. A) Representative FACS histograms of intranuclear IE expression eight days pi are shown for all U373-MG cell lines and three different viral strains. IE stain was analysed on live cells. The grey histogram represents uninfected U373-MG cells. B and C) IE expression on the different U373-MG cell lines normalized to HCMV-infected control U373-MG cells (dash line) (n=3 to 5 independent experiments). One-way ANOVA, multiple comparison tests, were performed to compare the results within the different cell lines and are presented in the table under each graph. Mann Whitney tests were performed to analyze each cell lines over time (day 2 pi *vs* day 8 pi). The only statistical difference is noted in the graph. D) Representative FACS blots of cell viability analysis of non-infected *versus* HCMV (MOI 1) infected MRC5 cells stained with the fixable Life/Dead stain kit at day 2 and 8 pi. E) Western Blot analysis of protein extraction obtained using TriPrep kit eight days pi (one representative western blot out of 3 independent is shown for each virus strain as well as for the uninfected control U373-MG cells).**

#### 1.3.3.4 The *multiplex* strategy is superior over the *singleplex* strategy to inhibit the viral genome replication and late protein expression

IE proteins are transactivators and induce the production of delayed early proteins, essentials for the genome replication, and for the production of structure proteins, needed for the assembly of new virions[196,494]. Thus, the disruption of IE expression abrogates the progression of the viral replication cycle. We analyzed eight days pi, the effect of the anti-*IE* CRISPR/Cas9 strategies on the genome replication using a qPCR on the *US8* gene. While the singleplex strategy was only effective for VR1814 at a low MOI with a decrease of 80% of viral genome copies, we detected 70 to 90 % less viral genome in U373-MG containing the multiplex strategy compared to the untransduced cells for all viral strains (MOI of 1 or 0.1) (Figure 28 A).

Furthermore, we analyzed the expression of the viral envelope glycoprotein B (gB), by intracellular FACS eight days pi. The untransduced control U373-MG cells infected with Toledo harbored around 6.5% of gB positive cells (Figure 28 B). The use of gRNA2 alone only slightly decreased the percentage of gB positive cells, while the multiplex strategy nearly abrogated the gB expression (Figure 28 B and C). The expression level of gB for TB40GFP (MOI 1) and VR1814 (MOI 0.1) was not high enough to be detected by FACS analysis.

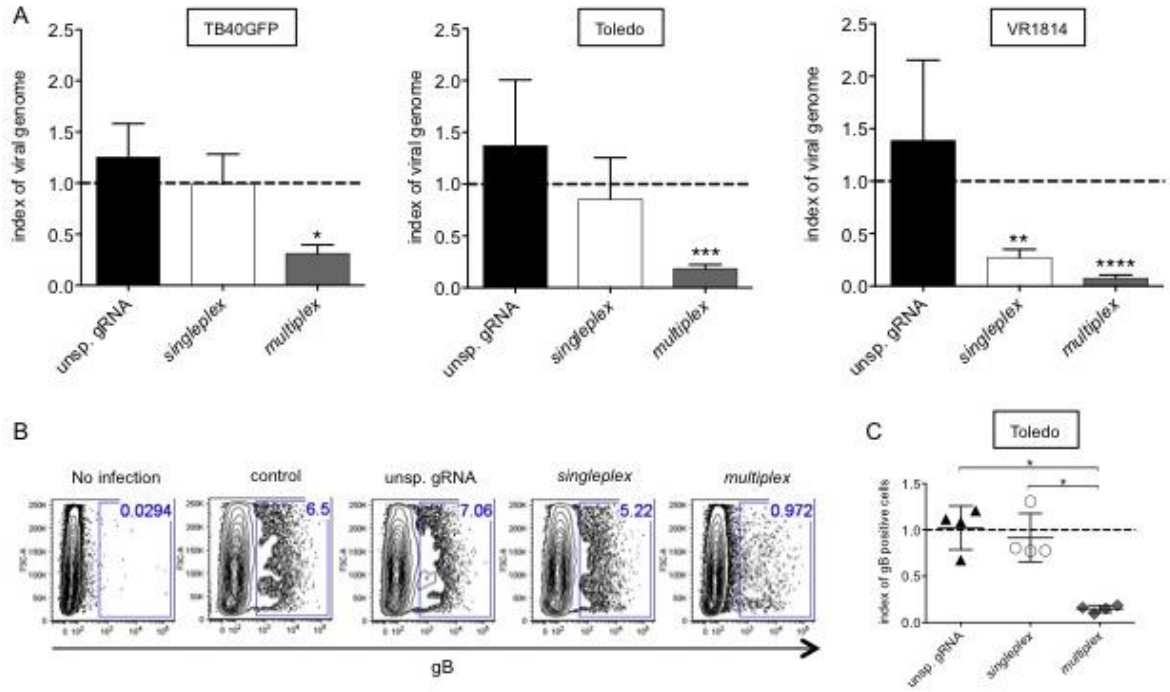



**Figure 28: Reduced progression of the viral replication cycle by the multiplex strategy: Transduced and control U373-MG cells infected with HCMV were harvested eight days pi. A) Relative viral genome quantification normalized to HCMV-infected control U373-MG cells (dash line) (n=3 independent experiments, +/- SD) One-way ANOVA, multiple comparison test, was performed and significant differences in comparison to the control are mentioned. B) Cells were FACS stained for total gB expression. Representative dot plots of total gB expression after infection with Toledo. gB stain was analysed on live cells. C) gB expression normalized to HCMV-infected control U373-MG cells (dash line) (triangles: unsp. gRNA; dots: singleplex; diamonds: multiplex) (n=4 independent experiments). One-way ANOVA, multiple comparison test, was performed.**

Overall the progression of the viral replication cycle was dramatically impaired by the multiplex strategy, as shown by a strong reduction of the genome replication and by the decreased expression of the late envelope glycoprotein B.

#### 1.3.3.5 The anti-HCMV multiplex strategy strongly impairs virion release from U373-MG cells

The multiplex anti-*UL122/123* CRISPR/Cas9 system efficiently decreased gB expression and genome replication. To assessed the production of infectious viral particles, we established a trans-infection plaque assay based on direct cell-to-cell transfer of the virus, because U373-MG cells only poorly secrete HCMV particles in the extracellular space[495]. Control U373-MG cells and the unsp. gRNA cell line, infected with TB40GFP reached a trans-infection plaque-titer of around 2500 plaques/10^5 cells (Figure 29 A). Singleplex U373-MG cells released 32 % less infectious virions (1600 plaques/10^5 cells). Importantly, targeting the *UL122/123* gene with the multiplex strategy decreased virion release by 80 % in average (Figure 29 A) (436 plaques/10^5 cells). In comparison, Toledo was produced in much higher amounts by the control and unsp. gRNA cells (7287 plaques/10^5 cells) (Figure 29 B). While the singleplex U373-MG cells released 67 % less infectious virions (4925 plaques/10^5 cells), multiplex U373-MG cells showed a remarkably 98 % inhibition of virion release (156 plaques/10^5 cells, Figure 6b).

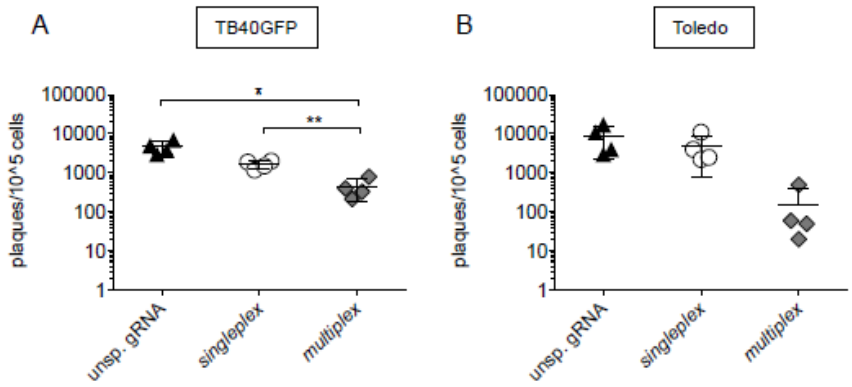



**Figure 29: Inhibition of virion release by the CRISPR/Cas9 systems anti-HCMV. Trans-infection plaque assay of infected U373-MG cells over MRC5 cells incubated in solid media. Plaque formation was observed seven to 14 days post trans-infection with TB40GFP-infected U373-MG cell lines (A) or Toledo-infected U373-MG cell lines (B). Each symbol represents medians of duplicates obtained in independent experiments. One-way ANOVA, multiple comparison test, was performed and only significant differences are mentioned in the figure.**

In conclusion, the multiplex anti-HCMV CRISPR/Cas9 system strongly inhibits the production of infectious viral particles and efficiently prevents viral spreading *in vitro.*

## 1.4 Discussion

HCMV is a widely spread infection in the human population and can cause severe end organ diseases in immunosuppressed patients such as solid-organ or HSC transplanted patients. Treatments exist, but are only efficient on replicative virus and have no effect on the latent virus pool. Here, we proposed an anti-viral strategy expected to be applicable for both lytic and latent viral infections. Using CRISPR/Cas9, we excised several exons from the essential *UL122/123* gene and further blocked IE-dependent steps of the viral replication life cycle.

### 1.4.1 Our anti-HCMV CRISPR/Cas9 strategies are competitive with actual state-of-the-art anti-viral CRISPR/Cas9 designs

We designed two CRISPR/Cas9 strategies to knock-out the *UL122/123* viral gene based on one or three gRNAs. The *UL122/UL123* gene encodes the Immediate Early molecules, IE1 and IE2, which are the first molecules expressed during the replication cycle and are essentials for the end of latency. Whereas IE2 is known to be essential for the viral replication[193,201,196] and is expressed at first in the lytic replication cycle[496]; IE1 is more responsible for transcriptional activation of immediate early and delay early promoters by the inhibition of HDACs [188,189] and is only essential for infections with a very low MOI[192,194,203]. Furthermore, the IE molecules are necessary for the initiation of reactivation from latency and a splice variant of IE1 is essential for the viral genome maintenance during latency[234]. The destruction of the *UL122/123* gene would therefore not only be efficient to inhibit the lytic replicating virus, but would also prevent the reactivation from latency and the persistence in the host cell. Moreover, Formivirsen, an approved anti-CMV-retinitis drug, is based on an antisense oligonucleotide targeting the *UL122/123* gene and is efficiently blocking local HCMV replication[271]. Consequently, the *UL122/123* gene is a suitable target for an anti-HCMV CRISPR/Cas9 system.

Our two anti-*UL122/123* CRISPR/Cas9 strategies are competitive with the actual state-of-the-art anti-viral CRISPR/Cas9 design. Classically, singleplex strategies target one essential viral gene somewhere inside the ORF. We designed here a singleplex strategy targeting close to the start codon in order to not just depend on frameshift mutation, but also disrupt the start codon itself. The multiplex design is an advancement of the singleplex strategy and aims to excise the entire ORF or the *UL122/123* gene. It is expected to enhance the efficiency of the anti-viral effect as shown before for EBV, HSV and HBV[460,474]. The novelty of our multiplex strategy is the focus on only one essential gene, while other groups build their multiplex strategies usually against several genes. Since the *UL122/123* gene is the absolute essential key regulator directly in the beginning of the viral replication cycle, the inhibition of this gene leads to an abrogative replication cycle[193,201]. Three gRNAs targeting one gene are expected to excise the *UL122/123* gene. However, in the case of repair in stet of excision, the ORF will still be disrupted at three independent positions and therefore would also prevent the MIE expression. Moreover, recently researchers have reported that the virus can escape the CRISPR/Cas9 system, if only one gRNA is used or if the cleavage efficiency of the gRNA/Cas9 complex is low[474,482,486]. Using three gRNAs is supposed to prevent viral escape, because three simultaneous in-frame mutations needs to occur in the viral genome to escape our multiplex strategy.

Van Diemen *et al*. tested the CRIPSR/Cas9 system against HCMV with a singleplex strategy[474]. They focused their main target sites in the delay early antigens, mostly proteins necessary for the genome replication. This singleplex strategies achieved short time inhibition of viral replication. However, the immediate early genes are still expressed and can cause side effects on the cell

viability in terms of cell cycle regulation and cytokine release[206,210]. Moreover, in case of clinical settings, the MIE are directly involved in some pathologies. For example, in congenital infections, IE2 is involved in the inhibition of apoptosis and interferes with the embryogenesis. It can also induce smooth muscle cell migration, which is leading to uncontrolled inflammation and proliferation and narrows the blood vessels[46]. Consequently, this leads to vascular injuries with hypoxia and brain damage of the foetus. In glioblastoma patients, the MIE are associated with immortalization and stem-like phenotype of the cells[100] and decreased tumour-suppressor gene expression[103,106]. Even though, van Diemen did not claim any future application options, the chosen target sites at delay early genes might be not optimal for all HCMV-associated pathologies.

Here, we improved the HCMV-targeting by choosing an earlier target gene (immediate early) and three gRNAs on the same gene to block viral protein expression and to prevent the further steps of replication cycle. Our anti-IE CRISPR/Cas9 systems are supposed to have a broader potential application field for most of the HCMV-caused diseases.

### 1.4.2 Delivery of the singleplex and multiplex strategies into the target cells

As shown for a duplex strategy against HIV[485] and a multiplex strategy against HBV[467], only a prolonged anti-viral Cas9/gRNA exposure can abolish the viral replication and reactivation entirely. Therefore, we chose LVs to delivery our CRISPR/Cas9 approaches. Furthermore, looking forward to possible clinical applications in the future, LVs are already an approved delivery method for *ex vivo* cell modification before injection into a patient[497,498]. One can imagine the treatment of HCMV-infected HSC with an LV containing the anti-viral CRISPR/Cas9 system to prevent HCMV reactivation and replication in the recipient.

For the proof of concept of our anti-*UL122/123* CRISPR/Cas9 strategies, we transduced HCMV-permissive cells with the LVs. To assess the full effect of the new anti-viral strategy, the transduced cells needed to be selected to exclude HCMV background replication in non-transduced cells. U373-MG cells allowed LV transduction and selection without complications. In contrast, the primary fibroblasts (MRC5) were more sensitive to transduction and different selection procedures. FACS-sort of transduced MRC5 cells was of low purity, due to their large shape was interfering with the optimal sort conditions. Prospectively, a recently acquired new FACS-sorter (Moflo) may allow the sort of large and dense cells for future applications as an easier and faster way to select transduced MRC5 cells. Surface expression marker selection failed, because of insufficient transgene expression. Finally, antibiotic resistance was used to select $Cas9^{+}$ MRC5 cells. Unfortunately, the LV transduction itself was very rough on the MRC5 cells and induced a high mortality. In general, LV transduction is possible for MRC5 cells, as shown by others[474]. However, the possible problem could be the purity of the LV production. We produced the LVs by our self. The production protocol only included to removal of bigger fragments by filtration (45 µm) and centrifugation. Potential toxic secrets from the producer cells still remained inside the LV stock and might have influenced the viability of the MRC5 cells. A three-week recovery phase finally give rise to Cas9/gRNA expressing MRC5 cells. However, the long procedure to gain the transduce cells only allowed a limited number of short experiments before cells enter into senescence.

### 1.4.3 Cleavage efficiency of the anti-UL122/123 CRISPR/Cas9 approach on the viral genome

In general, the efficiency of the CRISPR/Cas9 system depend on the amount of gRNA/Cas9 complexes available inside the cells, on the chromatin condensation and epigenetic factors, and on the copy number of the target sequence. In our experimental design, the cells were transduced with an LV containing the gRNAs/Cas9. We did not subclone the cells or analysed the copy number of the transgene integration. The idea was that in clinical application, the overall aim of the development of such an anti-viral CRISPR/Cas9 system, primary cells could not be subcloned. Analysing the stability and efficiency of the cell bulk with slight variability inside the culture is more accurate considering the major aim. However, the expression level of the Cas9 was comparable in the three different U373-MG cell lines allowing the comparison of the different strategies.

The efficiency of the singleplex strategy was analysed by T7 assay, a standard assay used to assess the frequency of *indels* at the target site. It is based on a PCR amplification of the target region, denaturation and reannealing of the PCR fragment, and digestions of the reannealed PCR products by the T7 endonuclease detecting small mismatches. We detected 30-50% *indels* at the target site by an infection with an MOI of 1 for TB40GFP and Toledo and 45% *indels* for infection with VR1814 (MOI 0.1). This suggests that the viral load have only minor influence on the cleavage efficiency. However, the T7 assay only allows the detection of *indels* in general and does not give information of the type of mutations induced. To have a more detailed analysis of the type of mutations induced, TIDE would be a method of interest giving the information about the size of each insertion/deletion. Therefore, the frequency of in-frame mutations could be assessed easily, as a first information about possible viral escape mutants. Testing the viral genome with TIDE would be an interesting perspective to analyse the singleplex strategy, especially in a time course experiment to see the possible acquisition of escape variants. Finally, to assess in detail the type of mutations induced, the total pool of viral genomes from treated cells would need to be analysed by next generation sequencing allowing then the exact identification of possible escape mutations.

For the analysis of the multiplex strategy, we presented here the simple ratio of WT PCR products *versus* mutated PCR products and detected 80-95% deletions at the target region. Unfortunately, the PCR amplification is biased here, because smaller products (with large deletion cause be the multiplex strategy) are favoured over the long WT sequence. Nearly all standard assays (sequencing, TIDE, RFLP) are influences by this. The only possible method could be the use of two quantitative PCR, one amplifying a sequence in the target region e.g. exon 5, which would be deleted by the multiplex strategy, and on amplifying a sequence elsewhere in the viral genome. The ratio of both qPCRs would give information about the frequency of large deletions, but would not consider *indels* at each target region itself. Since nearly all read out methods for the real frequency of mutations (large deletions and *indels*) are biased in some way, we decided to focus on the functional read out of the anti-*UL122/123* CRISPR/Cas9 system: the target gene expression and the effect on the viral replication cycle itself.

### 1.4.4 Targeting the *UL122/123* gene at multiple target site is more efficient to inhibit viral replication than a single target site

In this study, we challenged cells, pretreated with the singleplex or the multiplex strategy, with three different HCMV viral strains. When cells were infected at a low MOI (0.1), the singleplex strategy was efficient at decreasing the expression of IE molecules. Interestingly, the analysis of

the splice variant expression showed a possible knock-out of IE2 and a very strong decrease of IE1. Both variant are essential at this MOI, which explains the subsequent inhibition of the replication cycle. As expected, the reduction of IE molecules by the singleplex strategy for the viral strains Toledo and TB40GFP at a higher MOI (1) is not sufficient to prevent viral replication. The expression of a low amount of IE2 under this conditions, was probably sufficient to start the replication cycle and could be boosted by the positive autoregulation of IE1 as already proposed by others[194]. Van Diemen and colleagues also used a singleplex strategy anti-HCMV to target delay early genes with different efficacies at impairing the viral replication, even when a very low MOI (0.05) was used[474]. Furthermore, simultaneously targeting of the viral genome with several gRNAs completely abolishes the viral cycle as shown so far for HSV-1, HIV and HBV[466,474,485]. In line with this, we confirmed that targeting HCMV with a multiplex strategy was more efficient than a single gRNA. It abrogated IE expression at low and high MOI and led to a subsequent blockage of the viral replication cycle. Importantly, our *multiplex* strategy was effective on the three viral strains tested thus opening perspectives for its use for clinical application.

### 1.4.5 Comparison of the standard anti-HCMV treatments with our new anti-HCMV CRISPR/Cas9 approach

The use of RNA-guided endonucleases offers advantages over the actual standard treatment for HCMV infections, which are Ganciclovir and Foscarnet. They block the productive infection of HCMV by targeting the viral polymerase UL54[499]. This improves the health of patients facing HCMV diseases significantly, but side effects such as nephrotoxicity and myelosuppression are essential problems for the patient. Developing CRISPR/Cas9 strategies targeting the viral genome with low/no homology to the human genome should be less toxic and have no proven myelosuppressive effects.[458] Moreover, nowadays several high-fidelity Cas9s[373,374] have been shown to significantly reduce off-targets. As previously mentioned, controlling IE proteins expression was already used in clinic for CMV retinitis in HIV-1 patients before the development of highly active anti-retroviral therapy. The limit of such a strategy involving oligonucleotides is that the effect is only transient and usually does not completely inhibit the protein expression. As shown by Hamilton and colleagues[500], the knock-down of HCMV by siRNA targeting *UL122/123* mRNAs reduces viral replication and virion release. However, the application of siRNA is very transient and would not prevent HCMV replication over a longer time course. In contrast, the mutations or deletions induced by the CRISPR/Cas9 system are permanent and can provide a long-term protection, if all viral genome copies are efficiently targeted.

### 1.4.6 Viral escape from anti-viral CRISPR/Cas9 systems

Drug-resistances to Ganciclovir and Foscarnet are due to mutations in the kinase *UL97* or polymerase *UL54* genes[499,266]. Escape mutations against an anti-viral CRISPR/Cas9 singleplex system has been shown before on HIV[482,483], and on HCMV[474]. For example, the proposed anti-HCMV CRISPR/Cas9 systems by Van Diemen *et al.* using one gRNA targeting essential viral genes involved in viral genome replication gave rise to viral escape mutations. Those viral genomes harbored in-frame mutations, after being targeted by the anti-viral gRNA/Cas9. The probability of an escape mutation would be significantly lower with a multiplex strategy, because the cut within several targets leads nearly always to the deletion of parts or the complete targeted region and not just to small *indels*. Moreover, two studies on HIV have also shown that the duplex strategy can prevent escape mutations and viral breakthrough replication[485,486]. They have proven that the combination of several gRNAs diminishes the probability of in-frame mutations and that a longer

exposure to Cas9/gRNAs increases the frequency of bigger deletions in the viral genome. Furthermore, a more extensive multiplex strategy was successfully used against EBV, whose genome of 170 kbp could be destroyed entirely in Raji cells by the use of seven gRNAs simultaneously[473]. Our multiplex strategy is expected to prevent the viral escape as described for the other viruses. It induced mainly large deletions (80-95%) after an exposure of only eight days. Moreover, to reach this goal, high Cas9 expression is needed to target all copies of the viral genome before the expression of the IE molecules, which occurs as early as three hours after HMCV infection [496,501]. It has been shown by Richardson *et al*. that the Cas9 stays for around 5.5 h attached to the DNA after cleavage and therefore is not available to cut further target sites[362]. During lytic replication, the viral genome copy number increases rapidly exponential and it might not be possible for the Cas9/gRNA complex to target all copies. During natural latency in mononuclear cells from G-CSF mobilized blood or bone marrow no more than 13 viral genome copies are present per cell [502]. As a consequence, low Cas9/gRNA expression is expected to target all viral genome copies in a manageable exposure time.

### 1.4.7 Conclusion

In conclusion, we gave the proof-of-concept that targeting the *UL122/123* gene of HCMV genome with a *multiplex* strategy is efficient to affect the viral genome and to inhibit the virion release by up to 98%. In this study, we showed that even a singleplex strategy is efficient at inhibiting viral replication, if a low MOI is used. The multiplex strategy is superior over a single gRNA at low and high MOIs. Thus, these results pave the way to the development of a promising new therapeutic strategy that could be applicable to treat a hematopoietic stem cell suspension. Challenges for such a pre-emptive CRISPR/Cas9 therapy involves an optimized *ex vivo* delivery system, the selection of the targeted cells and the use of a high-fidelity Cas9[373,374].

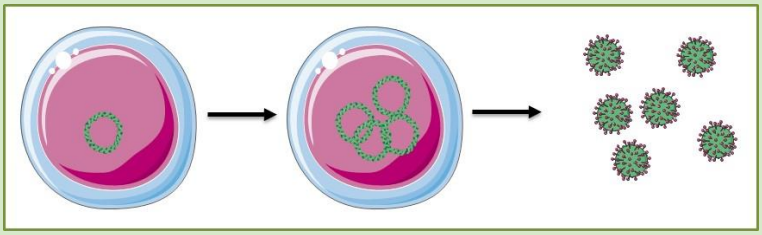
*Please find our submitted paper in the Appendix (X1).*

# V Hypothesis 2

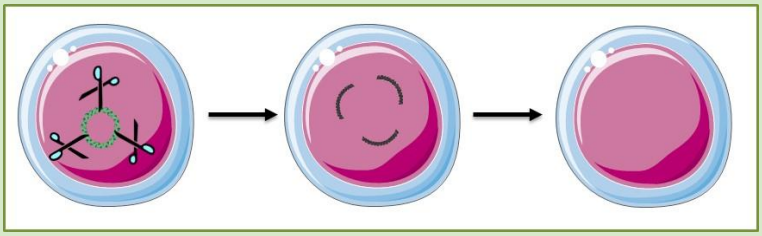
*Targeting the viral genome at multiple sites, which are chromatin open during latency, enable the fragmentation of the viral genome and may clear the HCMV infection from the host cell.*

***Box 8: Graphical abstract of the hypothesis 2***

***A**: Reactivation of HCMV leads to genome replication and the production of new virions.*



***B**: Cutting the viral genome at multiple target sites leads to the fragmentation of the viral genome and the loss of the viral genome form the infected cell. To destroy the viral genome in latency only chromatin open genome regions are targeted.*



# 1 Validation of the duplex strategy on the lytic replicating virus

## 1.1 General principle

Another aim of this thesis was to prevent HCMV reactivation from latency. One important feature to consider is that the HCMV genome is chromatin condensed during latency, which interferes with the ability of the Cas9 to target the viral genome. For HSV, it has been shown that there is no cleavage activity in viral genes, which are silenced in latency[474]. In collaboration with Benjamin Rauwel, we analysed the chromatin association of the *UL122/123* gene of HCMV and detected a wide chromatin condensation of this region. Consequently, this might be a problem for our singleplex and multiplex strategies, which were introduced in chapter IV. Therefore, it is necessary to validate our strategies in a latency model, because all of the gRNA target regions are silenced during latency.

The difficulty for the gRNA design against latent viral genomes is that most latency associated genes are non-essential for the viral replication or genome maintenance. So far, one MIE splice variant, IE1x4, is known to be necessary for genome maintenance[234]. The multiplex strategy excised the entire *UL122/123* gene and therefore would also prevent the IE1x4 expression, but the accessibility of the multiplex target regions is unknown. **Therefore, we created a duplex strategy, which only targets regions of the viral genome, which have been shown to be chromatin open during latency.** Most of the latency associated transcripts (LAT) are responsible for immune evasion[503]. UL138, the major LAT, promotes latency by inhibiting the MIE expression[232]. Those genes are not suitable as target genes during latency. Their knock-out would either have no effect on the reactivation and lytic replication cycle or would even induce the reactivation directly. *LUNA* is the only LAT, which is essential for reactivation, even though the molecular mechanism is still unknown[250]. As shown by us in the first hypothesis and by other[482,486], a singleplex strategy is not sufficient to permanently block viral replication and promote viral escape. Either a duplex strategy targeting two essential viral genes or a multiplex strategy fragmenting the viral genome is necessary to give rise to a full protection. Due to the lack of further essential genes, which are active during latency, we decided to follow the model from Wang *et al*., who created several multiplex strategies to destroy the viral genome of EBV[473]. He showed that three and seven gRNA are sufficient to destroy the viral genome. However, the delivery of extended multiplex strategies is difficult and often inefficient. **Therefore, we designed a duplex strategy consisting of two gRNAs to fragmentize the viral genome. One gRNA targets *LUNA*, which is essential for reactivation, and the second gRNA targets the two homologous TR regions.** The TR regions are also chromatin open, as shown by Rossetto *et al*.[504] and partly associated with the viral genome maintenance inside the cell[234]. They also encode two proteins, TRS1 and IRS1, which are essential for immune evasion during lytic viral replication. We assumed that targeting the viral genome at these three positions distributed over the entire viral genome would lead to the destruction of the viral genome and a HCMV-free cell. Moreover, in case of an imperfect destruction of the viral genome, the disruption of the ORF of LUNA should disable the reactivation from latency.

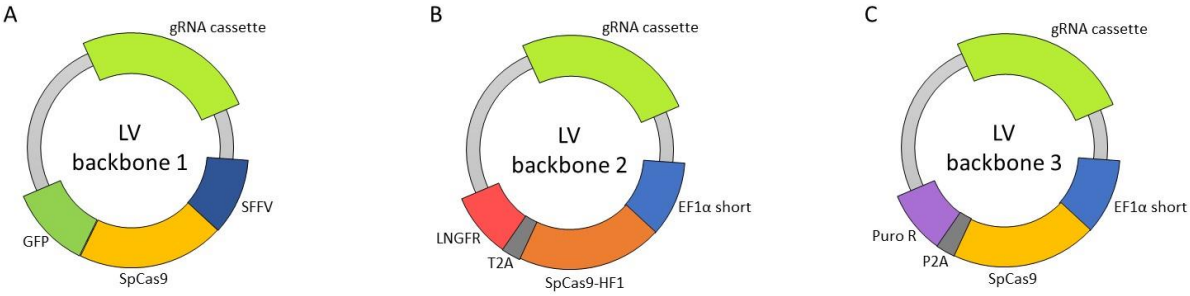
In order to test our hypothesis, we first validated the duplex strategy during lytic viral replication. Subsequently, we wanted to use a hematopoietic stem cell progenitor cell line, Kasumi-3. These cells are permissive for HCMV and allow the establishment of a latent HCMV infection and

reactivation after stimulation with TPA or TNF-α[505–507]. **We aimed to prove that this duplex strategy targets the HCMV genome in latency, prevents reactivation and reduces the amount of viral genomes in the cells.**

## 1.2 Methods

### 1.2.1 LV cloning

The gRNAs for the duplex strategy were designed by the CRISPOR web tool (TEFOR/CNRS network (http://tefor.net/crispor/crispor.cgi)). A gRNA cassette with two promoters (U6 and 7SK) each controlling one gRNA was designed and synthesised (GeneScript, Piscataway, USA). Subsequently, the duplex gRNA cassette was inserted in the third LV backbone SpCas9-P2A-Puromycin resistance by EcoRI and PacI (chapter IV1.2.1, Figure 17 C).




### 1.2.2 PCR

Cellular and viral DNA were extracted from infected MRC5 cells by the NucleoSpin TriPrep Kit (Macherey-Nagel, Düren, Germany). The target region of the gRNA6 (*LUNA*) was PCR amplified by the Herculase II Fusion Enzyme with dNTPs Combo kit (Stratagene, Agilent Technologies, CA, USA) with the following primers: F: GTGAGAGCCCCGTTGTTACC, R1: CATCCACTTTTTCCGCCGTG. The re-ligation site of *LUNA* and the TR2 region was amplified by the primers: F (see above) and R2: CGTGGCTGAGGGTGTAGAAG.

### 1.2.3 qPCR *US8*

Total DNA from infected MRC5 cells was extracted by the NucleoSpin TriPrep Kit (Macherey-Nagel, Düren, Germany) or viral DNA from the supernatant was extracted by the NucleoSpin® RNA Virus (Macherey-Nagel, Düren, Germany) and further diluted in water (1:10-1:10000). *US8* region was amplified with the following primers F: GGCACCAAATGCAGAGTGAG R: AAGCCGTATTCCGTTTGCG and the probe: BHQ-1 TGGTCCAAGTCCGTGGGCACC 6-FAM.46 and analysed on the StepOne plus qPCR machine.

*Further methods were explained in the previous chapter: Hypothesis 1 – methods (IV1.2).*

## 1.3 Results

### 1.3.1 Design gRNA

The duplex strategy consists of two gRNAs targeting the two TR regions and *LUNA* (*UL82-82* anti-sense-transcript).

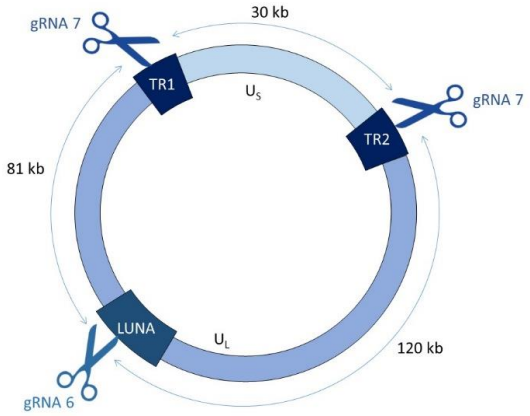



**Figure 30: Schema of the target regions of the duplex strategy. gRNA6 targets *LUNA*, which is located in the $U_L$ region. gRNA7 targets both homologous TR regions. Together these gRNAs cut the viral genome into three parts of 30 kb, 81 kb and 120 kb.**

The gRNA targeting *LUNA* was designed by using the CRISPOR web tool. The sequence of the *UL82-82* anti-sense-transcript was provided as the target region and the human genome was used as a reference for the calculation of off-target potentials. The most efficient gRNA options were aligned with 14 different HCMV strains (Table 15) and one gRNA targeting a conserved region was selected with a high selectivity score of 86 (Table 14). The second gRNA design to target the TR region was more complex. The TR region is a structural, mainly non-coding repeat region. This means, the sequence is not as highly conserved as the essential viral genes. The optimal target would have been the TRL region, which is bound by IE1x4 to maintain the viral genome inside the cell[234]. Unfortunately, the alignment of different viral strains did not show any conserved region. Extended alignments of the entire TR region detected the TRS region as partly conserved, because it overlaps with the ORF for *TRS1/IRS1* (Table 16 and Table 17). Those genes are immediate early genes and at least one needs to be expressed by the virus to assure viral replication (both proteins have the same function)[508]. They are responsible for the immune evasion from the innate immune response induced by dsRNA intermediates of the virus during viral replication. They inhibit the protein kinase R (PKR), which usually would activate the eukaryotic translation initiating factor 2α (eIF2α) leading to the shutdown of total protein expression[509]. More important, all the target regions are distributed throughout the entire viral genome so that a re-ligation of the different fragments is less likely. The three target sites by our two gRNAs digest the viral genome in three fragments with the size of 30 kb, 81 kb and 120 kb (Figure 30). This should be sufficient to destroy the viral genome as shown for EBV[473]. Duplex strategies targeting the host cells genome also have proven that the excision of up to 100 kb is possible[510].

**Table 14: gRNA for the duplex strategy**

| | Sequence | Length | Specificity score | Number of off targets |
|---|---|---|---|---|
| **gRNA 6** | AATCCGTACGGCCGTCCGAG**CGG** | 20 nt | 98 | 4 (min. 3 mismatches) |
| **gRNA 7** | GGACTCCATCGTGTCAAGGA**CGG** | 20 nt | 86 | 76 (min. 3 mismatches) |

**Table 15: Alignment of the gRNA 6 target region in the gene UL82-83 anti-sense-transcript (LUNA)**

| TB40/E | GATGAGAGGTCATCTTCGTCGTCCTCTTCCTCTTCTT | CCTCCTCTTCCTCGGTGGGTGGTAATCCGGGGGACTGCGGGAGAAACTCGGA |
|---|---|---|
| 3157 | ************************************ | ************************************************** |
| 3301 | ************************************ | ************************************************** |
| AD169 | **G********************************* | ************************************************** |
| AF1 | **G********************************* | ************************************************** |
| HAN13 | ************************************ | ************************************************** |
| HAN38 | **G********************************* | ************************************************** |
| JHC | **G********************************* | ************************************************** |
| JP | **G********************************* | ****************************T**C**T***************** |
| Toledo | **G**********G*********************CTT | ************************************************** |
| Towne | **G********************************* | ****************************T**C**T***************** |
| U8 | **G********************************* | ************************************************** |
| U11 | ************************************ | ************************************************** |
| VR1814 | ************************************ | ************************************************** |

*Green: gRNA-PAM, yellow: nucleotide substitution, light blue: insertion*

**Table 16: Alignment of the target region of gRNA7 in the TRS1 region**

| TB40/E | CCAACAGCACGGGCCGCGCCATGCGCAAGTGGTCGCAGCGCGACGCGGGCACGCTGCTGCCGCTCGGACGGCCGTACGGATTCTACGCGCGG |
|---|---|
| 3157 | ***********C********************************************************************************** |
| 3301 | ********************************************************************************T****A* |
| AD169 | ********************************************************************************** |
| AF1 | ************************************************G********************************* |
| HAN13 | ********************************************************************************** |
| HAN38 | ********************************************************************************** |
| JHC | ********************************************************************************** |
| JP | ********************************************************************************** |
| Toledo | ************************************************A********************************* |
| Towne | ********************************************************************************** |
| U8 | ********************************************************************************** |
| U11 | ********************************************************************************** |
| VR1814 | ********************************************************************************** |

*Green: gRNA-PAM, yellow: nucleotide substitution*

**Table 17: Alignment of the target region of gRNA7 in the IRS1 region**

| **TB40/E** | CCGCGCGTAGAATCCGTACGGCCGTCCGAGCGGCAGCAGCGTGCCCGCGTCGCGCTGCGACCACTTGCGCATGGCGCGGCCCGTGCTGTTGG |
|---|---|
| 3157 | **********G********** |
| 3301 | *T****A********** |
| AD169 | ********** |
| AF1 | **********C********** |
| HAN13 | ********** |
| HAN38 | ********** |
| JHC | **********G********** |
| JP | ********** |
| Toledo | **********T********** |
| Towne | ********** |
| U8 | ********** |
| U11 | ********** |
| VR1814 | ********** |

*Green: gRNA-PAM, yellow: nucleotide substitution*

### 1.3.2 Validation of the duplex strategy on the lytic viral replication

#### 1.3.2.1 Selection of primary fibroblast cells stably expressing the duplex/Cas9

For the validation of the duplex strategy against HCMV, MRC5 cells were transduced with LVs (SpCas9-P2A-Puromycin resistance) as described in chapter IV1.3.3. In brief, MRC5 cells were spinoculated with the LVs for 90 min in the presence of polybrene. Afterwards, the cells were cultured for two to three days in fresh complete medium and subsequently treated with puromycin for two days. MRC5 cells were further cultured and expanded in fresh medium. After a recovery phase of three weeks, remaining cells were treated a second time with puroymycin (0.5 µg/mL for 9 days) and were used for infection experiments before they enter senescence after another two to three weeks of culture.

To validate the Cas9 expression, a small portion of MRC5 cells was collected at each sub-culturing step and proteins were extracted by the TriPrep Kit (Machinery Nagel, Düren, Germany). As shown in Figure 31, the Cas9 expression decreased over time. As described in chapter IV for the singleplex and multiplex strategies, the decrease of the Cas9 expression level was not stronger for the MRC5 cells transduced with gRNA6 and 7 as compared to the unsp. gRNA leading to the conclusion that the duplex gRNAs are not toxic for the cells. However, four weeks after transduction only 10% of the original Cas9 expression was left which might interfere with the efficiency of the CRIPSR/Cas9 system. (Figure 31 B).

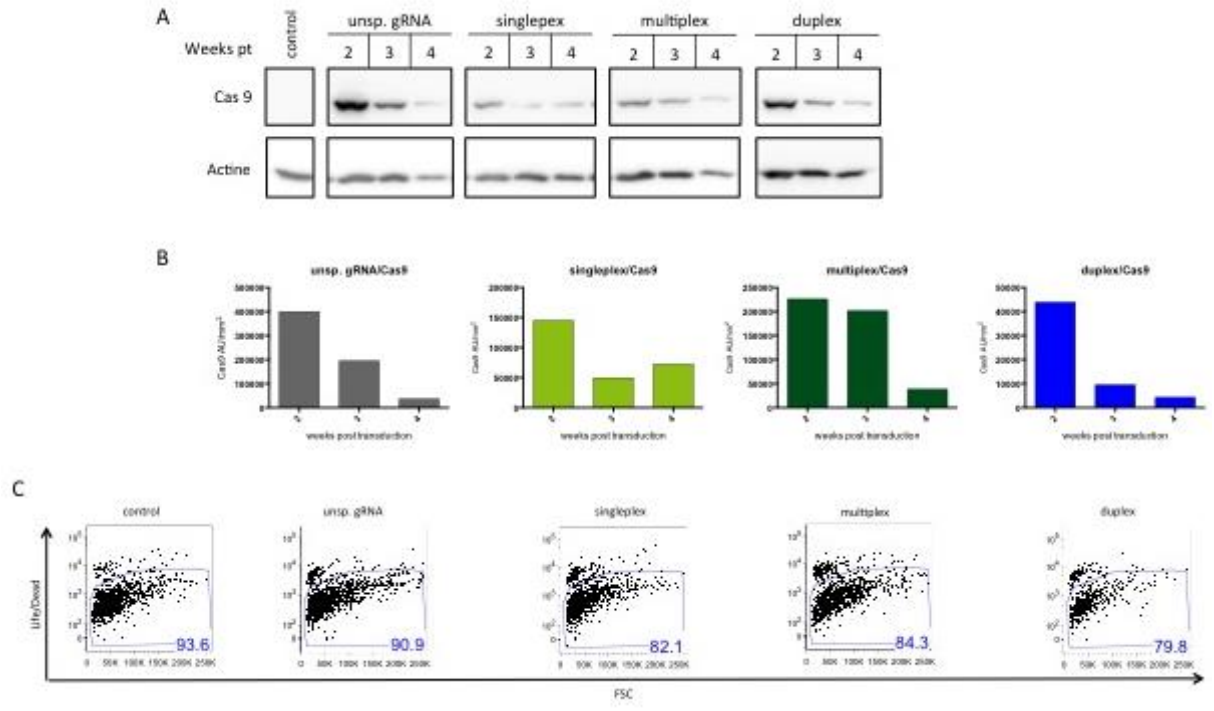



**Figure 31: Cas9 expression in MRC5 cells after puromycin selection. MRC5 cells were transduced with the different LVs containing a gRNA cassette and SpCas9-P2A-Puromycin. Transduced cells were selected by a two-day treatment with 2 µg/mL puromycin. At each following sub-culturing step, a part of the transduced MRC5 cells were collected and total protein was extracted. A) Cas9 expression and loading control actin expression were analysed by western blot. B) Relative quantification of the Cas9 expression based on the western blot, normalized by the actin expression level (n=1). C) Cell viability was analysed by fixable Life/Dead stain kit by FACS.**

### 1.3.2.2 Cleavage analysis of the viral genome targeted by the duplex strategy

The duplex strategy aims to destroy the viral genome by cutting it at the TR regions and in the *LUNA* gene. To analyse the cleavage efficiency, two pairs of primers were designed to either amplify the target region of gRNA6 (*LUNA*) (Figure 32 A) or to amplify a re-ligation of the biggest genome fragment, the $U_L$ region between TR2 and *LUNA* (Figure 32 C). The detection of *LUNA*-PCR amplicons confirms that not all viral genome copies are entirely fragmented as hypothesised and either are based on uncut or repaired viral genomes. Those amplicons of *LUNA* were analysed by T7 assay to assess the amount of *indels* and detected 7.8% (±0.38) *indels* for the MRC5 cells infected with Toledo and 9.7 % (±0.02) *indels* for the infection with TB40GFP. No *indels* were detected in the control MRC5 cells or the MRC5 cells expressing the unsp. gRNA (internal control). The results indicate that a proportion of the viral genome is correctly repaired or not cut at the *LUNA*-target site. Yet we cannot conclude, if the TR regions were targeted or if the $U_S$ region was excised for those viral genomes containing the WT *LUNA*.

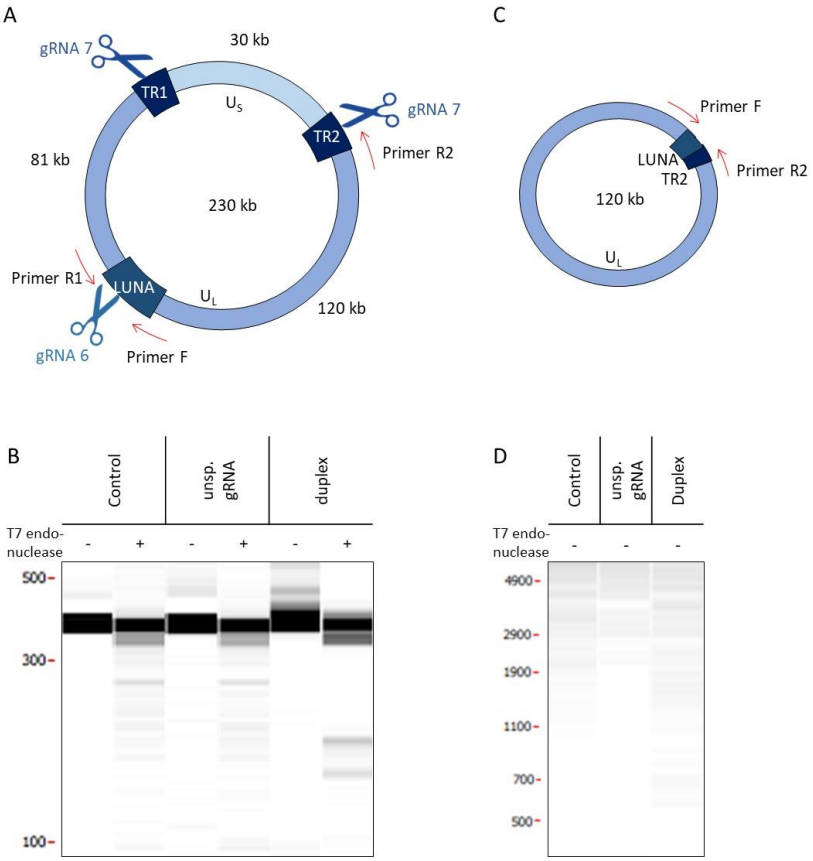



**Figure 32: Cleavage analysis of the effect of the duplex strategy on the viral genome. Control and transduced MRC5 cells were infected with HCMV Toledo at an MOI of 0.1. Four days pi, cells were harvested and total DNA was extracted by the TirPrep Kit. A) Schematic representation of the viral genome and the positions of the target region of the gRNA and the primers for PCR and T7 assay. B) The target region of gRNA6 (*LUNA*) was PCR amplified by Primer F and R1 and subsequently subjected to the T7 endonuclease. Cleavage products around 190 and 170 bp represent *indels* at the target site. C) Schematic representation of the potential re-ligation product of the $U_L$ region fragment from TR2 to *LUNA* after gRNA/Cas9 digestion. D) PCR amplification of the $U_L$ region over the potential re-ligation site by primers F and R2 with the expected size of 4300 bp. One representative experiment of three is shown.**

To test if the re-ligation of the biggest fragment of the viral genome after the cleavage of all three target sites is possible, a PCR spanning the re-ligation site between *LUNA* and TR2 was performed (Figure 32 A and C). The corresponding primer pair cannot amplify anything from the WT viral

genome, because the primer binding sites are 110 kb away from each other. A re-ligation would bring the primer binding sites closer together and would lead to a 4300 bp PCR product. No PCR product was amplified from the viral genome in the control nor in the duplex MRC5 cells (Figure 32 D). This was expected for the control and unsp. gRNA MRC5 cells, because they contain only WT viral genome copies. For the duplex strategy, no PCR amplification represents either that only WT viral genome copies remained inside the cells or that the cleavage with these two gRNAs led to the destruction of the viral genome and no re-ligation was possible.

In conclusion, only a minor part of the viral genome copies contained *indels* at the *LUNA* target site and no re-ligation product of the biggest fragment of the $U_L$ region could be detected.

#### 1.3.2.3 Targeting the viral genome at multiple sites efficiently diminishes the viral genome copy number inside the cells

As shown for EBV, targeting a herpesvirus genome at least at three target sites destroys the viral genome and leads to a reduced viral genome copy number in the cells[473]. Our duplex strategy was designed to have the same effect. The destruction of the viral genome would lead to a strong decrease or loss of the viral genome content in the cells, while *indels* at the three targets regions would only moderately decrease the genome replication. Therefore, the copy number of the viral genomes inside the cells was determined by qPCR of the *US8* gene. Figure 33 shows a comparable level of viral genomes in the non-transduced cells and the unsp. gRNA transduced cells. This shows that the expression of the Cas9 did not interfere with the viral genome replication. Comparison of the duplex strategy with the before characterized singleplex and multiplex strategies (see chapter IV1.3.3) showed that the duplex strategy was much more efficient than the singleplex strategy. It also reduced the viral genome copy number much stronger than the multiplex strategy, especially for TB40GFP. The hypothesis that the viral genome is fragmentised by the duplex strategy seems therefore possible. Simple *TRS1* and *IRS1* knock-out by gRNA7 would be less efficient as shown before for gRNA2 targeting *UL122/123*..

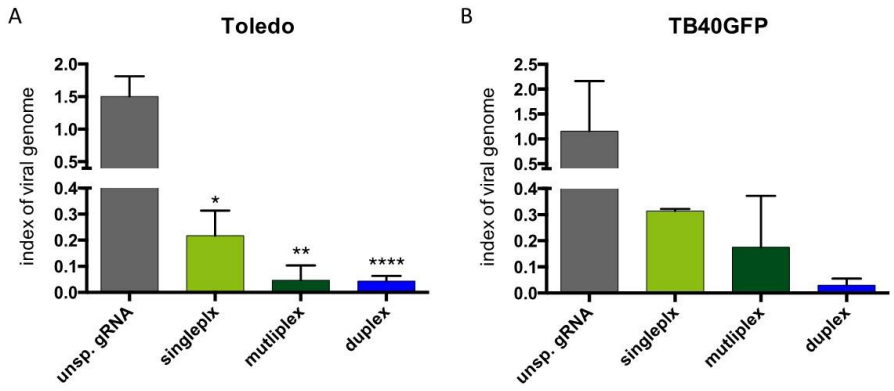



**Figure 33: Titration of the viral genome inside the cells. Transduced and control MRC5 cells were infected with A) Toledo or B) TB40GFP at an MOI of 0.1. two days pi, total DNA was extracted and subsequently analysed by qPCR of the *US8* gene. All values were normalized to non-transduced control cells. Toledo: n=3; analysis was done by ANOVA with multiple comparison; TB40GFP: n=2**

In conclusion, the digestion of the viral genome at three positions during the lytic replication cycle probably destroyed the majority of the viral genome copies and prevented the induction of the viral genome replication.

#### 1.3.2.4 Fragmentation of the viral genome prevents immediate early and late events of the lytic replication

To further investigate, if the duplex strategy destroys the viral genome or only downregulates the expression of *TRS1/IRS*, the MIE gene expression was analysed. It has been previously shown that *TRS1/IRS1* knock-out viruses are replication incompetent, because total protein expression is blocked at delay early time points. However, IE1 and IE2 expressions are not affected by the *TRS1/IRS1* knock out[508]. On the other hand, if the viral genome is disrupted directly after the entry, all steps of the viral replication cycle should be impaired including the MIE expression. Surprisingly, the unsp. gRNA transduced MRC5 cells were more permissive to the infection than the control non-transduced MRC5 cells. This might be due to the selection process of the transduce cells, which were associated with strong mortality. Comparing the three treatment conditions to the unsp. gRNA internal control, all CRISPR/Cas9 strategies showed a reduced MIE expression (Figure 33 A). As before, the multiplex strategy was more efficient than the singleplex strategy. Also, the duplex strategy reduced the MIE expression as efficiently as the multiplex strategy, even though the duplex strategy did not target the MIE gene directly. Therefore, we conclude that the duplex strategy destroys most of the viral genomes directly after cell entry and prevents the expression of immediate early proteins.

Finally, the progression of the viral replication cycle was analysed. Control MRC5 cells started to release new virions four days pi and reached higher titres around eight days pi. To assess the strength of the duplex strategy, we titrated the viral genome in the supernatant of infected MRC5 cells eight days pi. Figure 34 B shows that the control and the cells transduced with the unsp. gRNA released similar amount of virions. At eight days pi, the virion release reached a plateau and the control and unsp. gRNA cells were dying due to the viral infection. The singleplex strategy inhibits the virion release insufficiently by only 75%. The multiplex strategy, before proven to be very efficient to inhibit the viral replication in U373-MG cells, inhibited virion release by 95%. The destruction of the viral genome itself by the duplex strategy nearly completely abrogated the virion release by 98%. The MRC5 cells treated with the multiplex or duplex strategy were also more viable at eight days pi.

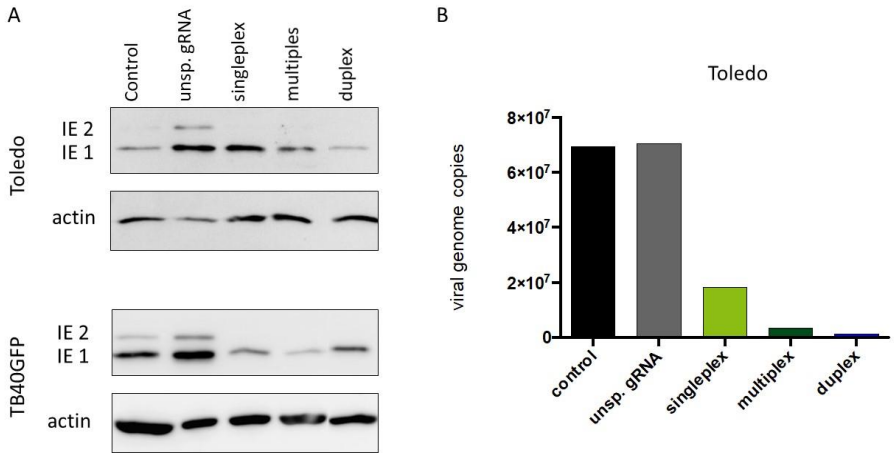



**Figure 34: Inhibition of early and late events of the viral replication cycle. A) MRC5 (transduced and control) cells were infected with HCMV at an MOI of 0.1. two days pi, proteins were extracted by the TriPerp kit and analysed by western blot to detect IE and actin (one representative western blot of three is shown). B) All MRC5 cells were infected with HCMV Toledo at an MOI of 0.05 for eight days. The supernatant was collected and the amount of viral genome copies was analysed by *US8* qPCR. (n=1)**

The destruction of the entire viral genome was more efficient than targeting one essential gene. The viral genome was cut directly after the virus entered the nucleus, which subsequently inhibited MIE expression and finally prevented the production of new virions.

## 1.4 Discussion

HCMV is a very widespread viral infection. After the primary infection, usually during childhood, it develops into a persistent latent infection in the host. Especially in transplant patients with suppressed immune system, HCMV can reactivate and is then associated with severe pathologies like HCMV pneumonitis. Anti-viral therapies against HCMV are based on nucleotide analogues, but they only inhibit the lytic replication virus and cannot assess the latent viral pool. Consequently, this treatment can neither prevent reactivation nor clear the infection. Therefore, we aim to develop a new therapeutically approach to target the latent viral genome to prevent reactivation and protect the patient from severe consequences.

Several viruses have been targeted by the CRISPR/Cas9 system in order to prevent replication and reactivation or to restore the transformation of cells. A special difficulty has to be overcome when targeting herpesviruses: All herpesviruses enter a phase of latency in the host, where the viral genome is silenced and chromatin condensed. In this state of latency, the Cas9 might not access chromatin condense regions of the viral genome and would be therefore inefficient to destroy or inactivate the viral genome. Van Diemen *et al*. showed on the example of HSV that the viral genome was unmodified during latency, but *indels* were detected after reactivation of the virus[474]. Therefore, we designed a duplex strategy, which is especially selected to target chromatin open region during latency to allow the cleavage of the silenced viral genome before reactivation and to prevent the potential dangerous entry into the lytic replication cycle. The target of only one gene by one gRNA has been proven several time as insufficient to prevent viral replication[474,482]. Rapidly, the virus gains in-frame mutations and escapes the anti-viral CRISPR/Cas9 system. Extended duplex and multiplex systems significantly decrease the risk of viral breakthrough replication[485,486]. However, Wang *et al.* designed a multiplex CRISPR/Cas9 system against EBV, which not just prevented the expression of the viral oncogene EBNA, but also destroyed the viral genome completely[473]. A similar multiplex approach with eight gRNAs targeting HBV removed the viral genome from the infected cells and decreased the viral genome load *in vivo*[466]. Both authors did not analysed the viral escape from their multiplex CRISPR/Cas9 systems, but the loss of the viral genome theoretically abolishes the risk of viral escape. Our duplex strategy was designed to fulfil both target criteria. It targets *LUNA*, an essential viral gene for the reactivation from latency[250], and the two immediate early genes *TRS1*/*IRS1* in the TR regions, to degrade the viral genome. Together, the cuts at those target sites should render the virus unable to reactivate (*LUNA* knock-out) and the fragmented viral genome should be removed from the infected cell.

We validated the anti-viral activity during lytic replication of the virus. The analysis of the fate of the viral genome after the cleavage at the three targets sites is challenging. Three outcomes would be possible. Either the DSB will be repaired by NHEJ leading to *indels* at all target sites, which then might induce frameshift mutations in the target gene. Alternatively, parts of the viral genome could be lost and the bigger fragments could be re-ligated. Those re-ligation products would probably not be able to produce new virions. Finally, the fragmentation of the viral genome in three parts would lead to the removal of the viral genome fragments from the cells and lead to the clearance to the infection. *Indels* were analysed at the target site for gRNA6 validating less than 10% of the viral genomes containing mutations in the *LUNA* gene. *Indels* at the gRNA7 target sites were not yet analyses by T7 assay, because both regions are homologue making it difficult to analyse *indels* for both regions separately. The occurrence of re-ligation products was tested for one possible re-ligation, the biggest fragment from the TR2 region to LUNA. No re-ligation product

could be detected. The validation of the disruption of the viral genome was then analysed by qPCR of the US8 gene. Since the possibility of re-ligation products was excluded, the loss of the US8 gene represents the loss of the viral genome. The viral genome content in the cell was decreased by around 95% in comparison to untransduced control MRC5 cells. This support the hypothesis that the viral genome is disrupted and lost. However, the knock-out of both *TRS1* and *IRS1* simultaneously would have the same effect of generating a replication-incompetent virus[509]. Therefore, theoretically only *indels* at in the *TRS1* and *IRS1* genes might be the reason of the decrease of viral genome copies inside the cells. However, this is very unlikely. The *TRS1/IRS1* genes are only targeted at one position, which is therefore prone to acquire in-frame mutations in at least one of those genes, which would lead to viral breakthrough replications. Similar inefficient inhibition of the viral replication was observed by the singleplex strategy targeting *UL122/123*. Moreover, the knock-out of *TRS1/IRS1* would only block the viral replication at delay early time points and does not interfere with the expression the MIE proteins[508,509]. However, the MIE expression in MRC5 cells containing the duplex strategy was strongly impaired. Consequently, the duplex strategy causes very likely the fragmentation and loss of the viral genome from the cells. The knock-down of an essential viral gene (multiplex) against the destruction of the entire viral genome (duplex) showed nearly the same efficiency during lytic replication. As expected, the duplex strategy reduced the amount of viral genome copies stronger than the multiplex strategy, since the multiplex strategy only inactivates the viral genome, but cannot remove it from the infected cells. Regarding the inhibition of the lytic viral replication cycle, both strategies are similarly efficient in inhibiting the IE expression. However, the virion release was stronger repressed by the duplex strategy. Further experiments are required to confirm this finding including experiments analysing mutation induced at the *TRS1/IRS1* target sides and on latently infected cells. If one of the strategies can prevent the escape of HCMV needs to be further validated in a more stable system. As discussed in chapter IV1.3.3, the transduction and selection process of the MRC5 cells is not optimal. Only very low expression levels of the Cas9 can be achieved, which is not sufficient to target all copies of the viral genome before the virus starts the replication cycle. However, the duplex strategy is less dependent on the digestion of the viral genome before the immediate early gene expression, because the loss of the viral genome would still prevent viral replication during delay early time points. Still, a higher Cas9 expression is needed once the viral genome replication is initiated to ensure that the concatemer cannot be packed into the viral particle. Furthermore, the target efficiency of the duplex strategy has to be validated during latency (see chapter V-2).

In conclusion, we have shown here that the duplex strategy efficiently degrades the viral genome during lytic replication. Based on this, we hypothesize that the duplex strategy prevents viral escape and allows the digestion of the viral genome also in latency. This provides a solid basis to further investigate this promising anti-viral CRISPR/Cas9 strategy against HCMV latent infection. The efficient destruction of the viral genome in latently infected cells will be a breakthrough approach to clear the HCMV infection from HSC used for transplantation.

# 2 Perspectives: *Prevention of HCMV reactivation from latently infected myeloid progenitor cells*

The next step is to validate the duplex CRISPR/Cas9 system in myeloid progenitor cells and to assess its efficiency against the latent viral infection. Kasumi-3 cells are an optimal model system for HCMV latent infection in HSC. They are acute myelocytic leukemia cells, which are positive for several progenitor surface markers like CD7, CD4, CD13, CD33, CD34, HLA-DR and c-Kit. In previous HCMV studies, Kasumi-3 cells have been proven permissive for HCMV, which enters into latency directly after infection[505,506]. Full latency is established and typical LAT, like *UL138* and *LUNA*, are expressed 10 days pi. Furthermore, these cells can be differentiated to monocytes/macrophages with the stimulation by TPA (Tetradecanoylphorbol-acetat) or TNF-α. This differentiation induces the reactivation of HCMV leading to IE expression and subsequent virion release. This latency model is more suitable for the proof-of-principle of the duplex strategy than primary $CD34^+$ cells, because the transduction of this cell line by LVs is very efficient. Pan *et al.* showed a transduction efficiency of 93% with a relative low MOI of 5[507].

In this HCMV latency model, we aim to infect the cells with HCMV and subsequently treat the infected cells with the LVs containing either the multiplex or duplex strategy. In this way, we will assess, if the viral genome can be cut by the Cas9 in the chromatin condensed viral genes (multiplex) or if the targeting of the viral genome is only possible in chromatin open regions (duplex). A prolonged culture of latently infected and treated cells should validate the loss of the viral genome with the duplex strategy. Furthermore, the possibility to differentiate the cells allows the analysis of the reactivation of HCMV, which should be inhibited by both strategies.

In conclusion, this model will allow us to answer, if the viral genome can be targeted during latency and if the reactivation by differentiation can be prevented by the anti-viral CRISPR/Cas9 strategies. This new concept of HCMV treatment to directly target the latent viral genome rises the opportunity to clear or inactivate the HCMV infection. In the setting of transplantation, HCMV-infected donor cells can be cleared from HCMV then impairing viral reactivation and HCMV diseases.

# VI General discussion

HCMV is a widely spread viral infection. After the primary infection, it enters into latency and persist lifelong in the host. 30 % to 100 % of the population are latently infected with HCMV[1,19]. The primary infection or the virus in latency usually do not cause severe pathologies[511]. However, the virus can reactivate in the immune deficient host, which can lead to more severe end organ diseases. Especially in HSCT, where the immune system is repressed by chemotherapy to prevent GVHD and the lymphopoiesis is not yet fully recover, the patient is at high risk of HCMV infection or reactivation. Common complications are gastrointestinal diseases or pneumonitis, which is associated with a high mortality rate. The introduction of nucleotide analogues, which inhibit the lytic viral replication, decreased the risk of HCMV reactivation significantly. The protection is very efficient in the first hundred days post transplantation, but also increases the risk of later HCMV related complications, which have often a more severe outcome[125]. Therefore, the aim of the thesis was to develop a new anti-HCMV therapy, which can prevent the reactivation of the virus and subsequently the CMV associated diseases. We used the CRISPR/Cas9 system from bacteria to cut the viral genome. Three different strategies were tested: the singleplex and the multiplex strategy target the MIE gene to induce an IE1 and IE2 knock-out leading to a replication-incompetent virus; and the duplex strategy digests the viral genome leading to the loss of the latent viral genome inside the host cells. We have shown the proof of principle that it is possible to inhibit the viral replication by targeting the viral genome with the CRISPR/Cas9 system. Only the multiplex and duplex strategy have been proven efficiently against HCMV, while the singleplex strategy still allowed viral replication.

## 1 The CRISPR/Cas9 system as an anti-viral therapy

The CRISPR/Cas9 system is a very variable tool used for genome editing and finds lately more attention as an anti-viral strategy. Several viruses have been targeted already in mainly *in vitro* studies and some *in vivo* models also showing promising results. However, so far non anti-viral CRISPR/Cas9 system has reached the state of clinical studies.

The correct choice of the target gene is important for an anti-viral CRISPR/Cas9 development. In general, there are two possibilities. Either a host factor essential for the viral replication is knocked-out or the viral genome itself is the target. Only for HIV, the targeting of the co-entry receptor have been tested with an anti-CCR5 CRISPR/Cas9 system[456]. This is possible, because HIV only infects one specific cell type ($CD4^+$ T cells) and the co-receptor for cell entry is not essential for the host. *In vitro* an *in vivo* experiments proved that $CCR5^-$ $CD4^+$ T cells are protected from the HIV infection[455,459]. For HCMV, the targeting of a host factor is not feasible. It uses various entry-receptors depending on the cell type and virus phenotype[172–174]. The knock-out of only one receptor would be inefficient to provide protection from the HCMV infection. Therefore, in this thesis we developed a CRISPR/Cas9 system targeting directly the viral genome. Typically, the CRISPR/Cas9 system is used to knock-out one specific target gene. The adaption to an anti-viral treatment depend therefore on the mutation of an essential viral gene to prevent viral replication[474]. This has been tested for HIV, HBV, HSV and the JC virus[512]. Classical targets are viral polymerases, important structure proteins or major transactivators. Importantly, motifs, which are essential for the protein function, like the YMDD motif of the reverse transcriptase of HBV[464], are the most efficient target sides. Unfortunately, the Cas9 dependency on the PAM sequence does

not always allow the direct target of this kind of motifs. Furthermore, some viruses develop various numbers of quasi-species inside the host. HIV is the typical example for this phenomenon[450]. It is a rapidly mutating virus leading to 10-20 different new quasi-species per year. In contrast, most of the other DNA viruses do not develop quasi-species, but they still can have some sequence variations. For herpesviruses, there are several different viral strains in the human population. Moreover, some viruses can be distinguished in different genotypes, like for HBV, which has up to eight genotypes varying in different viral genes. Therefore, it is essential to compare different viral strains or quasi-species to find highly conserved regions in the viral genome to ensure an anti-viral CRISPR/Cas9 system with a broad applicability to clinical isolates. Moreover, a single target site in the viral genome will lead at some point to a viral breakthrough replication, because of a silent mutation based on a nucleotide substitution or an in-frame mutation[474,482]. To increase the efficiency of a CRISPR/Cas9 system targeting essential viral genes and prevent viral escape, a multiplex approach is advantageous[485,486,513]. The simultaneous targeting of two essential genes increases the anti-viral effect significantly. Furthermore, the cleavage at two target sites mainly leads to the excision of the region in between and disrupts larger regions of the viral genome[444,473,485]. Two of our anti-HCMV CRISPR/Cas9 systems follow the idea of knocking-out an essential viral gene. The singleplex and multiplex strategy both target the *UL122/123* gene, which codes for the MIE proteins. The MIE are the first viral proteins expressed in the lytic viral replication cycle. They are the major regulators of delay early and late events of the viral repletion. The *UL122/123* gene consists of five exons, which can be alternatively spliced to two major and a couple of minor transcripts[213]. To target the two major MIE proteins (IE1 and IE2), the singleplex CRISPR/Cas9 system was designed to induce a mutation close to the start codon in order to destroy the start codon or induce an early frameshift mutation. The multiplex system is an advanced strategy based on the singleplex gRNA and two further gRNAs targeting exon 5. This strategy has been designed to prevent viral escape and increase the efficiency to be independent of frameshift mutations. As shown in hypothesis 1 (chapter IV1.3.3), the multiplex strategy is much more efficient then the singleplex strategy. It nearly abolished the virion release, while the singleplex strategy led to a breakthrough replication with comparably titres as the control cells. Therefore, we confirm the observation of other researchers that a multiplex strategy is superior over a singleplex strategy for HSV, EBV, HBV and HIV[449,466,473,474]. The major difference between most of the duplex strategies tested against singleplex strategies is that we used a multiplex strategy with three target sites in the same gene, while the other mostly target different viral genes. For EBV, a duplex strategy targeting two viral genes led to the deletion of 12 kb between the target sites[473]. In comparison, we excised the *UL122/123* viral gene of 3.5 kb. Duplex strategies against HIV usually lead to the excision of the entire viral genome (10 kb) from the host genome[444,446], if both gRNAs are highly efficient, otherwise only *indels* are acquired at the target sites[486]. The excision of the HIV genome leads to the loss of HIV from the host cells and clears the infection. To achieve similar results for non-integrative larger viruses, the viral genome needs to be targeted at multiple sites distributed throughout the entire viral genome. The multiple cuts degrade the viral genome and the fragments are removed from the infected cells. First, Wang *et al*. published a model like that[473]. EBV is a herpesvirus with a 170 kb large viral genome. During latency it produces viral oncogenes and transforms the cells to tumour cells. Targeting the viral oncogenes reduced the transformation of the cells. However, he also developed a multiplex strategy of seven gRNA, which destroyed the entire viral genome. This not only restored the transformation, but also led to the loss of the viral genome in the cell. A similar multiplex strategy of eight gRNA targeting HBV led also to the loss of the viral genome in the host cells[466]. Large extent multiplex strategies are very efficient, but their delivery is difficult and often

from low efficiency. Duplex strategies can be pack into a single viral vector or expression plasmid and are therefore more suitable for later transition to clinical applications. Our duplex strategy aims like Wang *et al*. to destroy the viral genome. It targets three positions in the viral genome and as shown in hypothesis 2 (chapter V), it is very efficient in inhibiting lytic viral replication and reduces the amount of viral genome copies inside the cells.

Another point needs to be taken into account for targeting the viral genome with the CRISPR/Cas9 system. Most of the viruses, which are of interest to be targeted with a genome editing tool, are DNA virus, which cause latent or chronic infections. Each virus family developed therefore its own mechanism to persist in the host. HIV integrates into the host genome. Therefore, targeting HIV with the CIRSPR/Cas9 system has the same features as targeting a human gene. There are only a few copies of viral genomes per cell, which means a limited number of target sites for the Cas9/gRNA complex. Consequently, a low Cas9 expression level would be sufficient. In contrast, HBV establishes chronic infections, where the viral genome persists as a circular covalently closed DNA (cccDNA). This cccDNA intermediate is very stable inside the cells. However, several groups were able to induce *indels* or destroy the cccDNA, proving the accessibility for the Cas9/gRNA complex[460,461]. Finally, herpesviruses like EBV, HSV and also HCMV persist inside the host cell as an episome. During latency, the episome is silenced and chromatin condensed[228]. Depending on the herpesvirus type, the viral genome is still accessible for the Cas9 or the chromatin condensation protects the viral genome from the cut by the Cas9[474]. EBV, a γ-herpesvirus, can be targeted by the CRISPR/Cas9 system during latency. The chromatin condensation is less condensed and several latent transcripts are produced. HSV, a α-herpesvirus, is highly specialized to neuronal cells and its genome is very high chromatin condensed in latency. It has been shown by van Diemen *et al*. that the Cas9 cannot access the HSV genome during latency. Therefore, it would be impossible to remove the HSV genome with a multiplex strategy from the infected cells. Worth to mention, the author only tested target sites of viral genes, which are silenced during latency. The effect of the Cas9 targeting latency associated proteins is so far unknown. For HCMV, a β-herpesvirus, the effect of chromatin condensation of the viral genome during latency is unknown. We analysed the chromatin condensation of the *UL122/123* gene and confirmed that it is associated with chromatin. Therefore, our duplex strategy was designed by taken the chromatin condensation into account and only targets viral genes, which are chromatin open during latency. Further evaluation of the CRISPR/Cas9 targeting of the latent HCMV genome is required.

Those anti-viral CRISPR/Cas9 strategies could be the next generation of anti-viral therapy. Moreover, it could be of advantage to use the CRISPR/Cas9 systems in combination with the actual standard treatment. Most of the standard anti-viral treatments inhibit the active viral replication. Therefore, the standard treatments decrease the amount of viral genome copies and consequently leave less target sites for the CRISPR/Cas9 systems. In this way, the anti-viral standard treatment would limit the active viral replication, while the CRISPR/Cas9 system can remove or inactivate the remaining viral genome copies. This additive effect of the two treatment systems has been tested already for HBV[464]. The standard treatment for HBV is based on nucleotide analogues. A combinatorial treatment with nucleotide analogues increased the effect of the CRISPR/Cas9 system. Nucleotide analogues are also the standard treatment for HCMV. An inhibition of the viral genome replication could enable the CRISPR/Cas9 system to target all viral genome copies. This would be very beneficial for the multiplex strategy targeting MIE. The efficiency of the multiplex strategy depends on the destruction of the *UL122/123* before the initiation of the viral replication cycle, basically in the first 3 h post infection. Because, once the MIE genes are expressed, the viral replication cycle starts and the modification of the MIE is less important. As we have shown in the

MRC5 cells, a very low expression level of the Cas9 is not sufficient to target all viral genome copies and allows viral replication. If the viral genome replication could be repressed till all viral genome copies are targeted, the HCMV genome would be silenced completely and no infectious virion could be produced. Synergistic effects of the nucleotide analogues and our CRISPR/Cas9 system needs to be further validated.

In conclusion, we designed two anti-HCMV strategies, which are very efficient to prevent viral replication and should prevent viral escape. Furthermore, the duplex strategy should be suitable to target the latent viral genome, because all target sites are chromatin open during latency. The multiplex and duplex strategy are state-of-the-art anti-viral CRISPR/Cas9 systems able to provide long-term inhibition of the viral replication.

# 2 Transmission to clinical application

The treatment model we presented in this thesis was the first step to develop an anti-HCMV therapy, which can be applied for transplant patients. Our system here is artificial, based on a stable cell line transduced with a common LV pseudotype VSV-G. For the transition of such a system to clinical application several improvements have to be considered. The delivery to the target cells needs to be optimized and off-target reduction methods need to be considered.

## 2.1 Delivery

Our primary aim was the application of our anti-HCMV CRISPR/Cas9 system in HSCT where the recipient receive a HCMV positive graft. In this case, the graft could be treated *ex vivo* and a systemic treatment would not be required.

### 2.1.1 Engineered graft for optimized CRISPR/Cas9 delivery

The graft for HSCT can be origin from different sources, i.e. bone marrow, umbilical cord blood and peripheral blood after the G-CSF mobilization of HSC. This sources can be used non-modified e.g. the injection of a mixed cell suspension containing the HCS, for the reconstitution recipient's immune system, and other myeloid cell types, like T cells, B cells, NK cells, monocytes, DC and more[514]. The delivery of the CRISPR/Cas9 system to all myeloid cells in the mixed cell suspension is right now illusionary. Actual viral vectors and transfection methods cannot target the variety of different hard-to-transduce cells. Several of those cell types could be transduced/transfected with specified vectors in a pure culture under cell type specific condition, but the treatment of this mixed cell suspension would not allow high efficient delivery to all cell types. Furthermore, very high amounts of viral vectors would be required and cell types, which are not of interest like T cells, would be treated as well. Anyway, sources of perfectly matched HLA donors are limited and the number of haploidentical transplantations from related donors (parents, siblings) is increasing[514]. Especially in this condition but also for matched unrelated donors, the injection of the graft containing high amount of T cells is correlated with an increased risk of GVHD[514,515]. Therefore, the HCS can be enriched or purified to optimize the graft. Two strategies have been tested and both are efficient to decrease the risk of GVHD. They are based on either negative selection of the HSC, where potential dangerous T cells are depleted based CD3 and CD19 staining, or positive selection of $CD34^+$ or $CD133^+$ cells[515,516]. In both cases, the graft cell suspension is stained with the concomitant antibodies with magnetic microbeads, which are then separated by the CliniMACS system of Miltenyi Biotech[515]. T-cell depletion is a method that only enriches the

HSC in the graft to around 1%, but other myeloid cells (DC, monocytes and furthers) remain in the graft[517]. In this case, that graft is still a very heterogeneous cell population, which would face similar limitations as the non-modified graft. On the other hand, HSC selection purifies the HSC to 93-97% and nearly completely removes other cell types from the graft (< 1%)[517–519]. This more homogenous population could be more easily targeted by a specialised delivery system for HSC. Unfortunately, HSC positive selection is correlated with a higher risk of graft rejection, relapse and infections because of the missing T and NK cells[519]. Several studies tested therefore the add-back of T cell as so called donor lymphocyte infusion (DLI) three to four weeks post transplantation[514,518,520]. Even though, those DLIs slightly increase the risk of GVHD, it has a beneficial effect to reduced infections and provide a graft versus tumour effect. In conclusion, to apply our anti-HCMV CRISPR/Cas9 system to HSCT, the graft needs to be positive-selected HSC for high efficient delivery and the patient could then be treated with DLIs to optimize the risk of GVHD, GVT and other infections.

### 2.1.2 Possible high efficient delivery systems to achieve prolonged gRNA/Cas9 expression

Important features for the ideal delivery are high efficiency, long-term expression of the gRNA/Cas9 and low cytotoxicity. Our actual system is based on LV transduction with an LV pseudotype VSV-G. This delivery technique provides persistent transgene expression, due to the integration into the host genome. Several clinical studies successfully used LV-modified HSC, mainly to treat rare genetic diseases[497,498]. In this case, a transgene expression of only a part of the injected cells is needed to alleviate the patient's condition. For the treatment of HCMV in HSCT, a very high transduction efficiency is needed to prevent the reactivation of the virus from the donor cells. The standard LV VSV-G has only a very low transduction efficiency for HSC. Therefore, the transduced cells need to be cultured and subsequently purified by a selection marker. A selection by FACS-sort based on a fluorescence marker or by magnetic microbeads based on a surface marker would be a possibility. So far, only the magnetic microbeads selection is approved for clinical applications. The selection based on puromycin is not optimal, because the cells need to be cultured for a much longer time frame (puromycin treatment) and are subjected to dying cells, which might release cytokines changing the phenotype or differentiation potential of the HSC. However, even the selection by fluorescence or surface marker requires a minimal culture time of two days, till the transgene is expressed. The *ex vivo* culture could interfere with the phenotype of the cells and increase the risk of contamination. A better option would be the use of an LV pseudotype, which would transduce 90 – 100% of the donor cells so that a selection step is not necessary anymore. Several LV pseudotypes could be considered here. The LV with the measles virus HF pseudotype transduces $CD34^+$ cells with an efficiency of around 80% (tested in our team in collaboration with Dr. Els Verhoyen). The difficulty with the LV-HF is its low titre in the production process, because the measles virus glycoproteins are transported purely to the cell surface for the LV particle formations[400]. This decreases the titre of the LV-HF. However, the LV-HF opened the opportunity for highly target cell/molecule specified LVs. The attachment to the cell surface and fusion of the LV particle with the plasma membrane are mediate by two proteins, the H (hemaglutinin) and F (fusion protein), respectively. Therefore, the H can be modified by mutating the natural receptor bind domain and fusing a cytokine domain or an antibody chain specific for a new surface receptor[400]. Consequently, LV with receptor-specificities for any surface receptor could be designed. So far, two LV-ΔHFs, targeting CD105 and CD133, has been designed to target specifically HSC[402,403]. The transduction efficiency is higher than the common LV VSV-G

and more importantly they mainly transduced HSC with the real progenitor phenotype for the reconstitution of the immune system. In the case of HSC selected by CD133, this could be a promising strategy to transduce the majority of the graft. However, the selection by CD34 is more common, which contain different stages of progenitor cells and the transduction by LV-ΔHF CD133 would not be sufficient. A LV-ΔHF targeting CD34 directly would be probably more optimal, but so far there is no such LV pseudotype described. Alternatively, the LV pseudotyped by the glycoproteins of the baboon retrovirus are highly efficient to transduce HSC[521]. They reach transduction efficiencies of up to 90% with a low MOI of 5. HSC maintained their differential potential and have been tested for immune reconstitution of NSG mice. To avoid the treatment and LV integration of non-infected cells, the LV could be specified for infected cells. Latently infected cells express US28, a G-protein coupled chemokine receptor, on their cell surface. A new designer LV-ΔHF targeting US28 could only transduced infected cells and spare non-infected cells. A treatment based on the US28 expression is already tested *in vitro*, based on a fusion toxin consisting of the concomitant chemokine ligand CXCL1 for US28 and a bacterial endotoxin[522]. Infected cells internalized the fusion toxin and subsequently died due to the endotoxin, while non-infected cells are not affected. The fusion of the ΔH to the CXCL1 mutant F49A[523] would theoretically create an LV highly specific to target infected cells.

Even though, LV-modified cells are already used in clinical applications, they still have the risk of transformation due the transgene integration into the host cell. There are integrase deficient LV available, but they only allow the transgene expression for a few weeks[405]. Depending on the efficiency of the anti-viral CRISPR/Cas9 system this might not be long enough to prevent viral breakthrough replication as shown for HIV[485]. Other delivery methods are also all very transient, like electroporation or transfection with expression plasmids or RNP complexes. The RNP is only present for up to 24 h and a standard expression plasmid allows a gene expression for three to four days[358]. However, if the plasmid construct contains the TR region of the HCMV, it could be maintained as long as HCMV is present inside the cells and would be replicated like the HCMV genome[234]. The region important for the genome maintenance is the TRL region rich on Sp1 binding sites, which is bound by the IE1x4 splice variant to tether the viral genome to the host genome[234]. Even though, one of our anti-HCMV CRISPR/Cas9 strategies is targeting the TR region (duplex), the gRNA target sequence is in the TRS region and is therefore not needed for the maintenance and could be excluded from the plasmid. A similar plasmid has been tested for EBV containing the oriP[473]. Therefore, the transfection or electroporation of HSC with TR-expression plasmid containing the anti-viral CRISPR/Cas9 system would be lost after three to four days in non-infected cells, but would be maintained in latently infected cells till the virus is inactivated or destroyed. In this way, the CRISPR/Cas9 system would be specifically only maintained in infected cell, which reduced the off-target effect on the total HSC population. The optimal delivery of such a plasmid would be probably by the micro fluid device with nano-silicon blades[395]. This delivery method was specifically developed for the transfection of HSC to maintain their phenotype and reduce cytotoxicity in comparison to electroporation.

Since the situation of a HCMV negative recipient and an HCMV positive donor is relatively rare, it is necessary to develop delivery methods optimized for systemic applications. Transfection and electroporation are not suitable in this case. However, even LV transduction would be very challenging to achieve high transduction efficiency in all target cells. So far, none of the above-mentioned LV systems are optimized for direct *in vivo* applications. However, the research on delivery techniques is rapidly progressing, especially since the introduction of CRISPR/Cas9, to be able to assess the full potential for this highly efficient genome editing tool.

In conclusion, the application of the CRISPR/Cas9 system is right now restricted the *ex vivo* treatment of cells and subsequent infusion into the patient. The only validated system to achieve long-term Cas9/gRNA expression, which is needed to prevent viral breakthrough replication, is the transduction with an LV. Since the transduction efficiency with the common LV VSV-G is too low, the usage of a new designer LV-ΔHF specified for HSC surface marker would be more suitable. However, so far the LV-ΔHF CD105 and CD133 do not transduce the entire HSC population and an LV-ΔHF with a broader receptor-specificity, for example CD34, does not yet exist. Therefore, the most efficient LV for HSC transduction is the baboon retrovirus pseudotype. Alternatively to the LV transduction, HSC could be transfected with an expression plasmid to avoid the uncontrolled integration of the LV transgene and carcinogenic potential. Hypothetically, an expression plasmid containing the TR region would be maintained in HCMV infected cells, as shown by Tarrant-Elorza *et al*., while it would be lost by non-infected cells.

## 2.2 Potential adverse effects and their prevention

### 2.2.1 Off-target effects on the human genome

Each gRNA/Cas9 complex has a certain off-target potential based on the homology of the spacer with non-target regions in the host genome[491]. The effect of mutations induced at these off-target sites is difficult to predict. It depends, if the off-target site is located in an intergenic region, where it could dysregulate the expression of an adjacent gene, or inside a gene, which could be subsequently disrupted. In the worst case, this could lead to the transformation of the cells. CRISPR design softwares help to choose gRNA with minimal off-target potentials. For example, the gRNAs of our multiplex strategy have between 53 and 192 potential off-target sites, while our duplex strategy is much more specific and has 4 and 76 potential off-target sites. Yet, even with highly specified gRNAs, a residual risk of off-target mutations remains. Therefore, several improvements of the Cas9/gRNA complex have been developed to avoid off-target cleavage. Two off-target optimized Cas9 variants exist, which have mutations to reduce the sequence-independent interaction with the DNA to increase the importance of the perfect gRNA/target DNA binding[373,374]. Alternatively, the gRNA can be truncated to a 17 or 18 nt spacer region to increase the specificity of the gRNA[369] and the scaffold region can be extent by 5 nt to improve the interaction with the Cas9[371]. The Cas9 variants are slightly more specific than the truncated gRNAs. However, the off-target effect is also strongly correlated with the time of Cas9 expression[358]. Since, the treatment of a viral infection requires prolonged Cas9 expression, right now only permanent Cas9 expression due to LV delivery is suitable. A lifelong Cas9 expression will definitively lead at some point to unwanted mutations in the host genome. Therefore, the Cas9 expression should be restricted to the time period needed. The before mentioned transfection of HSC with a TR-region containing Cas9/gRNA expression plasmid, would restrict the Cas9 expression time to the presence of the virus. Alternatively, an inducible Cas9 could be used. Two split Cas9 models either inducible by rapamycin treatment or by 4-hydroxytamoxifen are already tested *in vitro*[380,381]. Rapamycin is an mTOR inhibitor and is used in transplantation as an immunosuppressive drug. Some HSCT patients are up one year after transplantation (or even longer) under immunosuppressive therapy to avoid GVHD[524]. Therefore, the usage of such a treatment for the Cas9 expression regulation is not feasible and would lead to much longer Cas9 expression periods as needed to treat the viral infection and might even interfere with the management of GVHD and immune reconstitution. In contrast, 4-hydroxytamoxifen has no immunosuppressive effects. Is has been already tested *in vivo* with no adverse effect[525]. Daily oral

application as tamoxifen, *in vivo* processed to the active 4-hydroxytarmoxifen, provides a sufficient 4-hydroxytramoxifen level to activate the fusion protein, which makes it very suitable for clinical applications.

In conclusion, today the perfect design of an anti-viral CRISPR/Cas9 tool for clinical applications would be based on a delivery system specified for infected cells, either by designer LV-ΔHF specified to the viral US28 or by TR-region containing plasmids. Furthermore, precautions must be taken to avoid off-target effects. Most feasible would be a combination of the modified gRNAs to increase the gRNA/Cas9 specificity in combination with an inducible Cas9 to limit the expression time.

### 2.2.2 Cas9-induced immune reaction

The Cas9 is originally a bacterial protein[335]. For the usage in human cells it has been human codon optimized, but it is still a foreign protein. Prolonged expression *in vivo* leads to anti-Cas9 immune reactions and the loss of transduced cells[386,387]. In general, if HCS are transduced with the Cas9/gRNA and there could be a subsequent development of a Cas9-induced immune reaction, the HSC could be destroyed leading to a severe immune deficiency and anaemia. However, in HSCT the recipient is strongly immunosuppressed and might probably not be able to develop a Cas9-induced immune reaction before the reconstitution of the immune system. Still, infused T cells (DLI) could be reactive against the Cas9 and could consequently induce an autoimmune reaction against the donor HSC. Therefore, a lifelong Cas9 expression in the whole HSC population would be very risky. Possible solutions could be, either the limitation of the Cas9 expression to the time before DLI, or the delivery of the Cas9 only to HCMV-infected cells.

Inducible Cas9 transcription systems could provide the Cas9 expression for a time period by a given drug and could later be silenced by the stop of the drug administration. Typical ligand-inducible systems are steroid hormone receptor regulatory systems (ecdysone), progesterone receptor regulatory system (mifepristone) and tetracycline-dependent relatable systems (TET-ON)[526]. They are based on a ligand or drug, which can activate a transcription factor to regulate the transgene expression. To apply such a system to the CRISPR/Cas9 approach, the Cas9 would need to be under control of the inducible promoter and the ligand-inducible transcription factor (under its own promoter) would need to be included too. This lead to the delivery of three expression cassettes (Cas9, gRNAs, transcription factor). Actual viral delivery vectors are already at their maximum packing capacity with the SpCas9 and two gRNAs. An additional expression cassette might not be suitable or only in combination with smaller Cas9 orthologues. Furthermore, some of this systems are based on transcription factors from bacteria or flies and might induce likewise an immune reaction.

In contrast, the delivery of the CRISPR/Cas9 system only to HCMV-infected cells might have several advantages. HCMV directly inhibits the antigen presentation by MHC-I and MHC-II[235,238] and would therefore limit the presentation of the Cas9. Moreover, even if the infected, Cas9 positive cells would be removed by the immune system, this would still have the beneficial effect that the pool of latently infected cell would be reduced. Two possible option to specify the delivery to HCMV positive cells has been described before (chapter VI-2.1), but are not yet tested/developed.

# 3 Further applications

Originally, our anti-HCMV CRISPR/Cas9 systems were designed for the treatment of HSCT patients due to the severe pathology of HCMV reactivation in the immunosuppressed patients and the easy application to the donor cells *ex vivo*. However, HCMV is also associated with pathologies in other settings. For example, glioblastomas are positive for IE1/2, which promote the tumour progression and stem-like phenotype of the tumour stem cells[93,100,103]. In general, glioblastomas are very severe brain tumours with a short survival rate of eight to 15 month[527]. The treatment is based on surgery, radiation and chemotherapy. GCV treatment and anti-HCMV adoptive cell therapies have been tested on glioblastoma patients significantly prolonging the survival time[108,109]. Therefore, targeting the IE molecules in the tumour cells is also expected to prolong the survival. Even though, the IE molecules expressed in the tumour might be "unusual" splice variants, since they are located in the cytoplasm[528], our multiplex strategy would be ideal, because it excises the entire ORF and therefore would diminish all possible splice variants. It could be imagined that the anti-HCMV CRISPR/Cas9 system could be infused for local delivery during the brain surgery. The stereotaxic injection of the Cas9/gRNA RNP into a mouse brains has been successfully tested by using a Cas9 with several NSL signals (N- and C-terminal) to increase cell penetration[393]. Most of the neuronal cell types were transfected and only a mild immune reaction was detected. Alternatively, the local delivery via AAV serotype 1/2 is also specialised for neuronal cells. For the treatment of Huntington's disease a dual AAV approach has been tested in HD140Q-KI mouse model[435]. Such a Cas9/gRNA infusion after surgery could target the few remaining tumour cells and slow down their proliferation and tumour relapse.

Furthermore, in ICU patients HCMV reactivation has been reported and is associated with higher risks of mechanical ventilation and overall mortality[80,84]. Mainly, the virus reactivates in the lung and causes lung fibrosis and pneumonitis. There are no studies about anti-HCMV treatment on this patient group. However, theoretically, the prevention of HCMV reactivation in the lung or the protection of the lung tissue from HCMV-induced damage could improve the outcome of the patient and prevent mechanical ventilation. Possible would be a delivery of the anti-HCMV CRISPR/Cas9 system only to the lung of the patient. Different delivery methods have been already tested for an aerosol application to the lung. For example, the AAV serotype 2 and 6 are specialised to transduced lung tissue and showed 6% transduction efficiency *in vivo*[529]. Furthermore, for the treatment of cystic fibrosis, a cationic lipid delivery (GL67A) is already validated in a clinical trial II b to transfect the lung tissue with plasmid encoding the Cystic Fibrosis Transmembrane Conductance Regulator[530]. The transduction efficiency in the clinical trial was not assessed, but previous pre-clinical trials showed transgene expression in 16-53% of the bronchioles positive cells (0.3 - 38.5% positive cells/bronchiole)[531]. The so far highest transduction efficiency of 15% *in vivo* was achieved by a lentiviral vector based on the simian immunodeficiency virus pseudotyped with the glycoproteins of the Sendia virus (rSIV.F/HN) also developed for the treatment of cystic fibrosis[532,533]. A first clinical trial on this rSIV.F/HN is in preparation. If a transduction of only 15% of the lung tissue is sufficient to improve the patient's health needs to be further validated.

Further applications of the anti-HCMV CRISPR/Cas9 tools are limited based on insufficient delivery methods *in vivo*.

# VII Perspectives

In this thesis we developed three anti-viral CRISPR/Cas9 systems to be one day applied in hematopoietic stem cell transplantation to prevent the infection of a CMV seronegative recipient by CMV positive donor cells. We have shown high anti-viral efficiency for two of those strategies (multiplex and duplex) on the lytic replicating virus in a well-established cell line and in primary fibroblasts.

Short term perspectives include first of all the validation of our CRISPR/Cas9 strategies on the latent viral infection in physiological relevant cells. $CD34^+$ cells from G-CSF mobilized patients are of major interest, since they are one of the main sources for HSCT. These cells could be transduced with the LV pseudotype HF and transduced cells could be selected by FACS sort based on a fluorescence marker expression. Subsequently, target efficiency on the latent HCMV genome and possible prevention of HCMV reactivation could be validated *in vitro* while using cytokines and chloroquine treatment for reactivation[534]. Moreover, the effect of the prolonged Cas9 activity in HSC could be assessed by long-term culture and colony-forming-unit (CFU) assays. After successful *in vitro* validations, the anti-HCMV CRISPR/Cas9 strategies could be tested *in vivo* in a small animal model. NSG mice are immune deficient and allow immune reconstitution by the injection of human $CD34^+$ cells, important since HCMV only infects human cells. Two experimental designs are already tested for HCMV infection/reactivation. A protocol for Smith *et al*. is based on the reconstitution of the NSG mice by $CD34^+$ cells, improved by the implantation of an osmotic pump releasing G-CSF to mobilize the human myeloid cells, and a subsequent infection by intraperitoneal injection of infected human fibroblasts[535]. Four weeks pi, HCMV was detected by nested PCR in the bone marrow, spleen, kidney, liver and peripheral blood. This system could be used in two ways, either treating the primary fibroblast before injection to prevent lytic viral replication or to treat the $CD34^+$ cells to prevent the infection and establishment of latency in the myeloid progenitor cells. Alternatively, Hakki *et al.* successfully used $CD34^+$ cell from a HCMV positive donor and detected HCMV by qPCR in the bone marrow, liver and spleen[536]. This model is closer to the actual clinical setting.

Long-term perspectives would include the adaptation of our CRISPR/Cas9 strategies to improve specificity and transduction rates. Moreover, new delivery methods with high efficiency needs to be applied as discussed above (chapter VI2.1) based on designer LVs or TR-region containing plasmids. To test the CRISPR/Cas9 strategy for transplantation *in vivo*, the NSG mice models are imperfect. They only show the presence of the viral genome, but give no information about actual reactivation and viremia. Basically, none of the rodent CMV models allow the analysis of latency and reactivation[537]. Only, non-human primates, like rhesus monkeys, develop pathologies similar to humans[538]. They are infected mainly in their first year of life, which leads to a life-long latent infection. The RhCMV can reactivate induced by immunosuppression. The rhesus monkey model has been already used for transplantation studies[539] and could be very useful to analyse, if the virus can be inactivated by the CRISPR/Cas9 system, or at least reduce the amount of reactivating virus so far to prevent severe pathologies.

# VIII Summery in French

## UTILISATION DU SYSTEME CRISPR/CAS9 COMME STRATEGIE ANTIVIRALE CONTRE LE CYTOMEGALOVIRUS HUMAIN

Dans ce manuscrit de thèse écrit en anglais, j'expose dans un premier temps le contexte médical et scientifique qui nous a amené à développer nos études. Ainsi, je présente le cytomégalovirus humain (CMV), les pathologies qu'il induit et les traitements actuellement utilisés. Dans un deuxième temps, je présente la technologie des CRISPR/Cas9, leur origine, leur utilisation, leur vectorisation et leurs applications thérapeutiques potentielles. Enfin, je présente mes travaux de recherche et les met en perspective.

### LE CYTOMEGALOVIRUS

Le cytomégalovirus appartient à la famille des herpèsvirus (HHV-5). Sa structure comporte un génome double brin d'ADN, une capside icosaédrique et une enveloppe recouverte de glycoprotéines. Ce virus est de type non-intégratif, c'est-à-dire que lors d'une infection le génome du virus ne s'intègrera pas au génome de la cellule hôte mais restera sous forme épisomale dans le noyau de cette dernière. L'homme est le seul réservoir de ce virus dont la transmission se fait par contact étroit avec des sécrétions corporelles telles que : la salive, le sang, le sperme, les larmes, le lait maternel, les sécrétions du vagin et du col de l'utérus. Lors de l'infection, le CMV peut induire un cycle lytique aboutissant à la mort de la cellule hôte et à la production de nombreux virions pouvant propager l'infection. L'infection par le CMV peut également être latente. Le cycle réplicatif est alors réprimé et le génome viral est maintenu dans le noyau de la cellule hôte sous forme d'épisome. Une fois infecté, le sujet reste porteur du virus à l'état latent. Alors que chez des individus sains, le CMV est responsable d'une infection asymptomatique, son caractère pathogène survient chez les patients au système immunitaire affaibli (nouveau-nés, les patients immunodéprimés, transplantés ou atteints de cancer). A ce jour quatre médicaments existent pour combattre les infections à CMV. Trois d'entre eux inhibent l'ADN polymérase virale (le gancyclovir®, le foscarnet® et le cidofovir®). Le quatrième (le fomivirsen®) contient des oligonucléotides anti-sens ciblant l'ARNm codant pour la protéine Immediate–Early 2 (IE2). D'autres molécules sont actuellement en cours de développement ou testées en phase II ou III. Des stratégies pour bloquer la fixation du virus sur la cellule cible (IVIG). Le CMX001, qui correspond à du cidofovir conjugué à des lipides lui conférant une plus grande efficacité, une plus faible toxicité et la possibilité d'être pris par voie orale est testé en phase II. Le maribavir, qui inhibe spécifiquement la kinase virale UL97, ne s'est pas avéré très efficace dans le test de phase III et semble avoir des effets toxiques (notamment myélosuppression). Cependant ces traitements ne ciblent que les pools de virus réplicatifs et n'ont aucune action sur les pools de virus latents ; de plus des formes résistantes de CMV apparaissent inéluctablement. Des candidats vaccins sont également à l'étude associant une forme recombinante de la glycoprotéine B (gb) du CMV avec l'adjuvent MF59 ou bien associant deux plasmides codant pour la gB et pp65 (Transvax). Le premier de ces deux vaccins a donné des résultats encourageant chez des patients recevant une greffe de rein ou de foie, notemment dans la combinaison (D+/R-). Ces patients ont en effet une virémie qui corrèle négativement au taux d'anticorps anti-gb neutralisant (permettant ainsi de réduire la durée du traitement au ganciclovir). Le second vaccin, injecté à des patients recevant une greffe de cellules souches hématopoïétiques allogéniques, s'est montré moins efficace (pas statistiquement significatif par rapport au placébo) car malgré la diminution de la virémie il n'a

pas permis de réduire les traitements anti-viraux. De nouveau, cette stratégie vaccinale ne cible pas les pools de virus latents.

Une infection ou une réactivation du CMV survient chez la majorité des transplantés avec une sévérité très variable, allant de l'infection infra-clinique jusqu'aux atteintes viscérales pouvant toucher le greffon lui-même et mettre en jeu le pronostic vital du patient. Une primo-infection peut avoir lieu lorsque le donneur d'organe est CMV positif et le receveur CMV négatif puisque la combinaison D+/R- n'est pas une contre-indication à la transplantation. Après une greffe de cellules souches hématopoiétiques (D+/R-), 30% des receveurs vont développer une primo-infection. L'importance de la maladie à CMV est directement corrélée à la charge virale. Le CMV peut ainsi occasionner des complications particulièrement redoutables chez les receveurs de cellules souches hématopoïétiques, la plus fréquente étant la pneumonie interstitielle. La survenue d'une pneumopathie à CMV chez ces patients est de plus favorisée par le développement d'une réaction greffon-contre-hôte sévère.

La réactivation des pools de CMV latents vers un cycle lytique est dépendante de l'état de différenciation de la cellule hôte, et de l'activation du système immunitaire de l'hôte (notamment lors d'une allostimulation).

### L'ENDONUCLEASE CRISPR/CAS9

Le développement de nucléases à façon a permis d'envisager de cibler directement le génome viral pour en bloquer sa progression. Idéalement, couper l'ADN double brin du génome du CMV permettrait d'inhiber définitivement l'infection. En 2013, la première démonstration de l'utilisation du système CRISPR /Cas9 dans les cellules eucaryotes et son évidente facilité d'utilisation a ouvert de nouvelles opportunités scientifiques. En effet, contrairement aux nucléases à doigts de zinc qui étaient très onéreuses, l'utilisation du système CRISPR /Cas9 ne nécessite que l'achat d'un plasmide chez Adgène et la synthèse d'un fragment d'ARN complémentaire à la région d'ADN que l'on souhait cibler. Enfin, le développement d'outils accessibles à tous, tel que le logiciel CRISPOR, permet la sélection de gRNA de choix avec le moins de clivage non-spécifique possible. Ainsi, il semble à priori facile –et accessible- de développer et d'utiliser ces ciseaux moléculaires à façon pour créer de nouveaux modèles expérimentaux (KO, KI) et d'imaginer de nouvelles stratégies de thérapie génique ou de nouvelles thérapies anti-infectieuses. Ainsi, le système CRISPR/Cas9 a déjà été utilisé contre le HIV, l'HBV, l'HSV, le virus JC , l'EBV et le HCMV.

L'endonucléase Cas9 est dirigée par un ARN guide (gRNA) jusqu'à ce qu'une hybridation forte entre ce dernier et la séquence cible d'ADN immobilise la protéine. La coupure spécifique du double brin d'ADN ne se fait qui si un motif PAM (protospacer adjacent motif) composée des nucléotides NGG fait suite à la séquence reconnue par le gRNA. La réparation de l'ADN coupée est prise en charge par la machinerie cellulaire (Non-homologous end joining) qui est imparfaite et aboutit souvent à l'insertion ou la délétion de nucléotides.

### CHOIX DU GENE A CIBLER

La réplication et la sortie de la latence du CMV sont sous le contrôle du même gène (Immediate early gene). Lors de ma thèse, j'ai utilisé le système CRISPR/Cas9 afin de déstabiliser le génome du CMV. Dans un premier temps, nous avons voulu cibler un gène essentiel à la réplication virale et à la sortie de la latence afin bloquer le virus quel que soit son état. Nous avons donc choisi de cibler le gène *UL122/UL12*3 codant pour les molécules immediate early. En effet, lorsque les molécules IE1 et IE2 codées par le gène *UL122/UL123* ne sont pas fonctionnelles, le cycle réplicatif

s'arrête et aucun virion ne peut être produit. De plus, le Fomivirsen, premier ARN anti-sens approuvé par la FDA, cible l'ARNm codant pour la molécule IE2 et permet le contrôle efficace des rétinites à CMV.

### RESULTATS

Nous avons donc testé plusieurs gRNA ciblant : 1- l'ATG du gène *UL122/UL123* au niveau de l'exon 2 ; 2- le début et 3- la fin de l'exon 5. Nous avons comparé deux stratégies utilisant soit un, soit trois gRNAs respectivement appelées singleplex et multiplex ciblant ce même gène. La stratégie antivirale a été délivrée dans la lignée U373-MG par des lentivirus permettant une intégration du de la cassette d'expression codant pour la Cas9-GFP et pour la stratégie d'intérêt. Nous avons trié par FACS les cellules transduites et les avons ensuite infectées avec trois souches différentes de CMV (TB40, TOLEDO, VR1814). Comme attendu, la stratégie singleplex induit des insertions et délétions au site de coupure du gRNA, la stratégie multiplex quant à elle induit la délétion de 3500 paires de bases. De ce fait, la stratégie multiplex bloque efficacement l'expression du gène ciblé, la réplication virale et le relargage de nouveaux virions. La transduction de cellules primaires (MRC5) a été beaucoup plus difficile car nous n'avons pas réussi à les trier par FACS (Cas9-GFP), ni par billes magnétiques avec une expression à la surface des cellules transduites du récepteur. Nous avons donc utilisé une nouvelle construction et un agent de sélection (la Puromycine) pour sélectionner les MRC5 transduites. La résistance à l'antibiotique nous a permis de sélectionner les MRC5 transduites mais la quantité de Cas9 n'était pas suffisante pour contrôler activement le CMV, et diminuait au cours du temps.

Une autre stratégie a été développée pour cibler spécifiquement le génome à l'état latent. Cette stratégie dite « duplex » comporte deux gRNA ciblant d'une part les régions homologues TR et d'autre part le gène *LUNA*. Contrairement au gène *UL122/UL123* ; la région TR et le gène *LUNA* sont accessibles lors de la latence car la chromatine n'est que peu/pas condensée. Cette seconde stratégie conduit elle aussi à une diminution du nombre de copies de génome viral lors du cycle lytique de réplication. Il faudrait maintenant tester cette stratégie dans des cultures cellulaires infectées de façon latente (cellules souches hématopoïétiques ou lignée cellulaire telle que KASUMI-3).

### CONCLUSION ET PERSPECTIVES

Lors de mon travail de thèse, j'ai prouvé que l'utilisation de la technologie CRISPR/Cas9 permet de contrôler la réplication du virus lorsque les gRNA déstabilisent un gène essentiel ou d'éliminer une partie importante du génome viral. Les véritables enjeux à présent sont de cibler spécifiquement les cellules d'intérêt, de contrôler les niveaux d'expression de la Cas9 et de minimiser les coupures non spécifiques. Ces enjeux s'adressent à tous projet visant l'utilisation de ces nucléases à des fins thérapeutiques.

# IX List of references


1. Voigt S, Schaffrath Rosario A, Mankertz A. Cytomegalovirus Seroprevalence Among Children and Adolescents in Germany: Data From the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), 2003-2006. *Open forum Infect Dis*. 2016;3(1):ofv193. doi:10.1093/ofid/ofv193.
2. Lopo silvialopo S, Vinagre E, Palminha P, Paixão MT, Nogueira P, Freitas MG. Seroprevalence to cytomegalovirus in the Portuguese population. *Euro Surveill*. 2002;16(25). www.eurosurveillance.org:pii=19896. Accessed September 27, 2016.
3. Bate SL, Dollard SC, Cannon MJ. Cytomegalovirus seroprevalence in the United States: the national health and nutrition examination surveys, 1988-2004. *Clin Infect Dis*. 2010;50(11):1439-1447. doi:10.1086/652438.
4. Sagu MJ, Ory F De, Ramı R, Comas LG, Leo P, Sanz JC. Is there a change in cytomegalovirus seroepidemiology in Spain ? *Eur J Epidemiol*. 2004;19:85-89.
5. Adjei AA, Armah HB, Gbagbo F, Boamah I, Adu-Gyamfi C, Asare I. Seroprevalence of HHV-8, CMV, and EBV among the general population in Ghana, West Africa. *BMC Infect Dis*. 2008;8. doi:10.1186/1471-2334-8-111.
6. Rodier MH, Berthonneau J, Bourgoin A, et al. Seroprevalences of toxoplasma, malaria, rubella, cytomegalovirus, HIV and treponemal infections among pregnant women in Cotonou, Republic of Benin. *Acta Trop*. 1995;59(4):271-277. doi:10.1016/0001-706X(95)00087-U.
7. Lieberman L, Devine D V., Reesink HW, et al. Prevention of transfusion-transmitted cytomegalovirus (CMV) infection: Standards of care. *Vox Sang*. 2014;107(3):276-311. doi:10.1111/vox.12103.
8. Shigemi D, Yamaguchi S, Otsuka T, Kamoi S, Takeshita T. Seroprevalence of cytomegalovirus IgG antibodies among pregnant women in Japan from 2009-2014. *Am J Infect Control*. 2015;43(11):1218-1221. doi:10.1016/j.ajic.2015.06.026.
9. Das B, Kaur G, Basu S. Seroprevalence of cytomegalovirus antibodies among blood donors and Multitransfused recipients â€“ A study from north India. *Transfus Apher Sci*. 2014;50. doi:10.1016/j.transci.2014.02.022.
10. Griffiths PD, Mclean A, Emery VC. Encouraging prospects for immunisation against primary cytomegalovirus infection. *Vaccine*. 2001;19:1356-1362. www.elsevier.com/locate/vaccine. Accessed June 2, 2017.
11. Colugnati FA, Staras SA, Dollard SC, Cannon MJ. Incidence of cytomegalovirus infection among the general population and pregnant women in the United States. *BMC Infect Dis*. 2007;7(1):71. doi:10.1186/1471-2334-7-71.
12. Fang FQ, Fan QS, Yang ZJ, et al. Incidence of cytomegalovirus infection in Shanghai, China. *Clin Vaccine Immunol*. 2009;16(11):1700-1703. doi:10.1128/CVI.00385-08.
13. Lanzieri TM, Bialek SR, Ortega-Sanchez IR, Gambhir M. Modeling the potential impact of vaccination on the epidemiology of congenital cytomegalovirus infection. *Vaccine*. 2014;32(30):3780-3786. doi:10.1016/j.vaccine.2014.05.014.
14. Cannon MJ. Congenital cytomegalovirus (CMV) epidemiology and awareness. *J Clin Virol*. 2009;46:S6-S10. doi:10.1016/j.jcv.2009.09.002.
15. Hogea C, Dieussaert I, Effelterre T Van, Guignard A, Mols J. A dynamic transmission model with age-dependent infectiousness and reactivation for cytomegalovirus in the United

States: Potential impact of vaccination strategies on congenital infection. *Hum Vaccines Immunother*. 2015;11(7):1788-1802. doi:10.1080/21645515.2015.1016665.

16. Maingi Z, Nyamache AK, Paschale M De, et al. Seroprevalence of Cytomegalo Virus (CMV) among pregnant women in Thika, Kenya. *BMC Res Notes*. 2014;7(1):794. doi:10.1186/1756-0500-7-794.

17. Jansen MAEE, van den Heuvel D, Bouthoorn SH, et al. Determinants of Ethnic Differences in Cytomegalovirus, Epstein-Barr Virus, and Herpes Simplex Virus Type 1 Seroprevalence in Childhood. *J Pediatr*. 2016;170. doi:10.1016/j.jpeds.2015.11.014.

18. Marshall GS, Stout GG. Cytomegalovirus Seroprevalence among Women of Childbearing Age during a 10-Year Period. *Am J Perinatol*. 2005;22(7):371-376. doi:10.1055/s-2005-872590.

19. Cannon MJ, Schmid DS, Hyde TB. Review of cytomegalovirus seroprevalence and demographic characteristics associated with infection. *Rev Med Virol*. 2010;20(4):202-213. doi:10.1002/rmv.655.

20. Enders G, Daiminger A, Lindemann L, et al. Cytomegalovirus (CMV) seroprevalence in pregnant women, bone marrow donors and adolescents in Germany, 1996–2010. *Med Microbiol Immunol*. 2012;201(3):303-309. doi:10.1007/s00430-012-0232-7.

21. DOWD JB, AIELLO AE, ALLEY DE, et al. Socioeconomic disparities in the seroprevalence of cytomegalovirus infection in the US population: NHANES III. *Epidemiol Infect*. 2009;137(1):58. doi:10.1017/S0950268808000551.

22. Bates M, Brantsaeter AB. Human cytomegalovirus (CMV) in Africa: a neglected but important pathogen. *J virus Erad*. 2016;2(3):136-142. http://www.ncbi.nlm.nih.gov/pubmed/27482452. Accessed September 27, 2016.

23. Naing ZW, Scott GM, Shand A, et al. Congenital cytomegalovirus infection in pregnancy: a review of prevalence, clinical features, diagnosis and prevention. *Aust New Zeal J Obstet Gynaecol*. 2016;56(1):9-18. doi:10.1111/ajo.12408.

24. Stagno S, Pass RF, Dworsky ME, Alford CA. *Maternal Cytomegalovirus Infection and Perinatal Transmission.* Vol 25. 1982:563-576. http://www.ncbi.nlm.nih.gov/pubmed/6290121. Accessed October 10, 2016.

25. Bialas KM, Tanaka T, Tran D, et al. Maternal CD4+ T cells protect against severe congenital cytomegalovirus disease in a novel nonhuman primate model of placental cytomegalovirus transmission. *Proc Natl Acad Sci U S A*. 2015;112(44):13645-13650. doi:10.1073/pnas.1511526112.

26. Price SM, Bonilla E, Zador P, Levis DM, Kilgo CL, Cannon MJ. Educating women about congenital cytomegalovirus: assessment of health education materials through a web-based survey. *BMC Womens Health*. 2014:1-10. doi:10.1186/s12905-014-0144-3.

27. Revello MG, Tibaldi C, Masuelli G, et al. Prevention of Primary Cytomegalovirus Infection in Pregnancy ☆. *EBioMedicine*. 2015. doi:10.1016/j.ebiom.2015.08.003.

28. Stowell JD, Mask K, Amin M, et al. Cross-sectional study of cytomegalovirus shedding and immunological markers among seropositive children and their mothers. *BMC Infect Dis*. 2014;14.

29. Balcells C, Botet F, Gayete S, et al. Vertically transmitted cytomegalovirus infection in newborn preterm infants. *J Perinat Med*. 2016;44(5):485-490. doi:10.1515/jpm-2015-0325.

30. Josephson CD, Caliendo AM, Easley KA, et al. Blood Transfusion and Breast Milk Transmission of Cytomegalovirus in Very Low-Birth-Weight Infants. *JAMA Pediatr*.

2014;168(11):1054. doi:10.1001/jamapediatrics.2014.1360.

31. Wilms IR, Best AM, Adler SP, et al. Cytomegalovirus Infections among African-Americans. *BMC Infect Dis*. 2008;8(1):107. doi:10.1186/1471-2334-8-107.

32. Handsfield HH, Chandler SH, Caine V a, et al. Cytomegalovirus infection in sex partners: evidence for sexual transmission. *J Infect Dis*. 1985;151(2):344-348. doi:10.1093/infdis/151.2.344.

33. Choi YC, Hsiung GD. Cytomegalovirus infection in guinea pigs. II. Transplacental and horizontal transmission. *J Infect Dis*. 1978;138(2):197-202. http://www.ncbi.nlm.nih.gov/pubmed/210239. Accessed September 27, 2016.

34. Bowden RA, Slichter SJ, Sayers M, et al. A Comparison of Filtered Leukocyte-Reduced and Cytomegalovirus (CMV) Seronegative Blood Products for the Prevention of Transfusion-Associated CMV Infection After Marrow Transplant. *Blood*. 1995;86(9):3598-3603.

35. Puchhammer-Stockl E, Gorzer I. Human cytomegalovirus: an enormous variety of strains and their possible clinical significance in the human host. *Futur Med*. 2011;6(2):259-271. doi:10.2217/fvl.10.87.

36. Arav-Boger R. Strain Variation and Disease Severity in Congenital Cytomegalovirus Infection: In Search of a Viral Marker. *Infect Dis Clin North Am*. 2015;29(3):401-414. doi:10.1016/j.idc.2015.05.009.

37. Xanthakos SA, Schleiss MR. GLYCOPROTEIN B GENOTYPING OF CYTOMEGALOVIRUS STRAINS ISOLATED IN A PEDIATRIC POPULATION. *Pediatr Infect Dis J*. 2003;22(5):462-463.

38. Lisboa LF, Tong Y, Kumar D, et al. Analysis and clinical correlation of genetic variation in cytomegalovirus. *Transpl Infect Dis*. 2012;14(2):132-140. doi:10.1111/j.1399-3062.2011.00685.x.

39. Paradowska E, Studzińska M, Suski P, et al. Human Cytomegalovirus UL55, UL144, and US28 Genotype Distribution in Infants Infected Congenitally or Postnatally. *jouranl Med Virol*. 2015;87:1737-1748. doi:10.1002/jmv.

40. Pati SK, Pinninti S, Novak Z, et al. Genotypic diversity and mixed infection in newborn disease and hearing loss in congenital cytomegalovirus infection. *Pediatr Infect Dis J*. 2013;32(10):1050-1054. doi:10.1097/INF.0b013e31829bb0b9.

41. Nijman J, Mandemaker FS, Verboon-Maciolek MA, et al. Genotype Distribution, Viral Load and Clinical Characteristics of Infants with Postnatal or Congenital Cytomegalovirus Infection. doi:10.1371/journal.pone.0108018.

42. Ross SA, Novak Z, Pati S, et al. Mixed Infection and Strain Diversity in Congenital Cytomegalovirus Infection. *J Infect Dis*. 2011;204(7):1003. doi:10.1093/INFDIS/JIR457.

43. Ishibashi K, Tokumoto T, Tanabe K, et al. Association of the Outcome of Renal Transplantation with Antibody Response to Cytomegalovirus Strain--Specific Glycoprotein H Epitopes. *Clin Infect Dis*. 2007;45(1):60-67. doi:10.1086/518571.

44. Crough T, Khanna R. Immunobiology of human cytomegalovirus: from bench to bedside. *Clin Microbiol Rev*. 2009;22(1):76-98, Table of Contents. doi:10.1128/CMR.00034-08.

45. Kenneson A, Cannon MJ. Review meta-analysis of the epidemiology of congenital cytomegalovirus (CMV) infection. *Rev Med Virol*. 2007;17:253-276. doi:10.1002/rmv.

46. Schleiss MR. Congenital cytomegalovirus infection: molecular mechanisms mediating viral pathogenesis. *Infect Disord Drug Targets*. 2011;11(5):449-465. http://www.ncbi.nlm.nih.gov/pubmed/21827434%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3869401.

47. Lanzieri TM, Dollard SC, Bialek SR, Grosse SD. Systematic review of the birth prevalence of congenital cytomegalovirus infection in developing countries. *Int J Infect Dis*. 2014;22:44-48. doi:10.1016/j.ijid.2013.12.010.

48. Williams EJ, Embleton ND, Clark JE, Bythell M, Ward Platt MP, Berrington JE. Viral Infections: Contributions to Late Fetal Death, Stillbirth, and Infant Death. *J Pediatr*. 2013;163(2):424-428. doi:10.1016/j.jpeds.2013.02.004.

49. Stagno S, Reynolds DW, Amos CS, et al. Auditory and Visual Defects Resulting from Symptomatic and Subclinical Congenital Cytomegaloviral and Toxoplasma Infections. *Pediatrics*. 1977;59(5).

50. Hassan J, Connell J. Translational Mini-Review Series on Infectious Disease: Congenital cytomegalovirus infection: 50 years on. *Clin Exp Immunol*. 2007;149(2):205-210. doi:10.1111/j.1365-2249.2007.03454.x.

51. Boppana SB, Ross SA, Fowler KB. Congenital cytomegalovirus infection: clinical outcome. *Clin Infect Dis*. 2013;57 Suppl 4(suppl 4):S178-81. doi:10.1093/cid/cit629.

52. Pass RF, Stagno S, Myers GJ, Alford CA. Outcome of Symptomatic Congenital Cytomegalovirus Infection: Results of Long-Term Longitudinal Follow-up. *Pediatrics*. 1980;66(5).

53. Uematsu M, Haginoya K, Kikuchi A, et al. Asymptomatic congenital cytomegalovirus infection with neurological sequelae: A retrospective study using umbilical cord. *Brain Dev*. 2016;38(9):819-826. doi:10.1016/j.braindev.2016.03.006.

54. Mackenzie Dreher A, Arora N, Fowler KB, et al. Spectrum of Disease and Outcome in Children with Symptomatic Congenital Cytomegalovirus Infection. *J Pediatr*. 2014;164:855-859. doi:10.1016/j.jpeds.2013.12.007.

55. Yamamoto-Tabata T, McDonagh S, Chang H-T, Fisher S, Pereira L. Human cytomegalovirus interleukin-10 downregulates metalloproteinase activity and impairs endothelial cell migration and placental cytotrophoblast invasiveness in vitro. *J Virol*. 2004;78(6):2831-2840. doi:10.1128/JVI.78.6.2831-2840.2004.

56. Jun Y, Kim E, Jin M, et al. Human cytomegalovirus gene products US3 and US6 down-regulate trophoblast class I MHC molecules. *J Immunol*. 2000;164(2):805-811. doi:10.4049/JIMMUNOL.164.2.805.

57. Rolland M, Li X, Sellier Y, et al. PPARγ Is Activated during Congenital Cytomegalovirus Infection and Inhibits Neuronogenesis from Human Neural Stem Cells. Kalejta RF, ed. *PLOS Pathog*. 2016;12(4):e1005547. doi:10.1371/journal.ppat.1005547.

58. Marin LJ, Santos De Carvalho Cardoso E, Mara S, et al. Prevalence and clinical aspects of CMV congenital Infection in a low-income population. *Virol J*. 2016;13. doi:10.1186/s12985-016-0604-5.

59. Bale JFJ, Blackman JA, Sato Y. Outcome in children with symptomatic congenital cytomegalovirus infection. *J Child Neurol*. 1990;5(2):131-136.

60. De Kegel A, Maes L, Dhooge I, Van Hoecke H, De Leenheer E, Waelvelde H Van. Early motor development of children with a congenital cytomegalovirus infection. *Res Dev Disabil*. 2016;48:253-261. doi:10.1016/j.ridd.2015.11.014.

61. Teissier N, Delezoide A-LA-L, Mas A-EA-E, et al. Inner ear lesions in congenital cytomegalovirus infection of human fetuses. *Acta Neuropathol*. 2011;122(6):763-774. doi:10.1007/s00401-011-0895-y.

62. Ogawa H, Hashimoto H, Okano W, et al. Restricted infection of murine cytomegalovirus (MCMV) in neonatal mice with MCMV-induced sensorineural hearing loss. *J Clin Virol*.

2015;69:138-145. doi:10.1016/j.jcv.2015.06.083.

63. Gabrielli L, Bonasoni MP, Santini D, et al. Human fetal inner ear involvement in congenital cytomegalovirus infection. *Acta Neuropathol Commun*. 2013;1:1. doi:10.1186/2051-5960-1-63.

64. Ciorba A, Bovo R, Trevisi P, Bianchini C, Arboretti R, Martini A. Rehabilitation and outcome of severe profound deafness in a group of 16 infants affected by congenital cytomegalovirus infection. *Eur Arch Otorhinolaryngol*. 2009;266(10):1539-1546. doi:10.1007/s00405-009-0944-5.

65. Fortunato EA, Dell 'aquila ML, Spector DH, Cavenee WK. Specific chromosome 1 breaks induced by human cytomegalovirus. *Proc Natl Acad Sci*. 2000;97(2):853-858.

66. Schooley RT. Cytomegalovirus in the Setting of Infection with Human Immunodeficiency Virus. *Clin Infect Dis*. 1990;12(Supplement 7):S811-S819. doi:10.1093/clinids/12.Supplement_7.S811.

67. Anders H-J, Goebel FD. Neurological manifestations of cytomegalovirus infection in the acquired immunodeficiency syndrome. *Int J STD AIDS*. 1999;10(3):151-161. doi:10.1258/0956462991913817.

68. Süttmann U, Willers H, Höpken W, Gerdelmann R, Schedel I, Deicher H. Cytomegalovirus infection in HIV-1-infected individuals. *Infection*. 1988;16(2):111-114. doi:10.1007/BF01644316.

69. Robain M, Boufassa F, Hubert JB, Dussaix E, Sadeg K, Meyer L. Is cytomegalovirus infection a co-factor in HIV-1 disease progression? *Epidemiol Infect*. 2000;125(2):415-420. http://www.ncbi.nlm.nih.gov/pubmed/11117966. Accessed November 18, 2016.

70. Goplen AK, Liestøl K, Dunlop O, Bruun JN, Mæhlen J. Dementia in AIDS Patients in Oslo; The Role of HIV Encephalitis and CMV Encephalitis. *Scand J Infect Dis*. 2001;33(10):755-758. doi:10.1080/003655401317074572.

71. Salomon N, Gomez T, Perlman DC, Laya L, Eber C, Mildvan D. Clinical features and outcome of HIV-related cytomegalovirus pneumonia. *AIDS*. 1997;11(3):319-324. doi:10.1097/00002030-199703110-00009.

72. Faber DW, Wiley CA, Lynn GB, Gross JG, Freeman WR. Role of HIV and CMV in the pathogenesis of retinitis and retinal vasculopathy in AIDS patients. *Invest Ophthalmol Vis Sci*. 1992;33(8):2345-2353.

73. Jabs DA. Cytomegalovirus Retinitis and the Acquired Immunodeficiency Syndrome—Bench to Bedside: LXVII Edward Jackson Memorial Lecture. *Am J Ophthalmol*. 2011. doi:10.1016/j.ajo.2010.10.018.

74. Hutchinson S, Courtney R, Mulcahy F, O'Connor M, Mullaney S, Kilbride R. Monitoring cytomegalovirus retinitis prevalence in an HIV-seropositive cohort: the assessment of improvements observed following the introduction of highly active antiretroviral triple therapy. *Int J STD AIDS*. 1999;10(10):652-656. doi:10.1258/0956462991913213.

75. Yust I, Fox Z, Burke M, et al. Retinal and extraocular cytomegalovirus end-organ disease in HIV-infected patients in Europe : a EuroSIDA study , 1994 – 2001. *Eur J Clin Microbiol Infect Dis*. 2004:550-559. doi:10.1007/s10096-004-1160-2.

76. Agarwal A, Singh R, Sharma A, Gupta V, Dogra MR. Ocular Manifestations in Patients with Human Immunode fi ciency Virus Infection in the Pre-HAART Versus the HAART Era in the North Indian Population. *Ocul Immunol Inflamm*. 2016. doi:10.3109/09273948.2015.1133837.

77. Ford N, Shubber Z, Saranchuk P, et al. Burden of HIV-Related Cytomegalovirus Retinitis in

Resource-Limited Settings : A Systematic Review. *Clin Infect Dis*. 2013;57. doi:10.1093/cid/cit494.

78. Reed JB, Schwab IR, Gordon J, Morse LS. Regression of Cytomegalovirus Retinitis Associated With Protease-inhibitor Treatment in Patients With AIDS. *Am J Ophthalmol*. 1997;124(2):199-205.

79. Au Eong KG, Beatty S, Charles SJ. Cytomegalovirus retinitis in patients with acquired immune deficiency syndrome. *Postgrad Med J*. 1999;75(888):585-590. http://www.ncbi.nlm.nih.gov/pubmed/10621898. Accessed November 18, 2016.

80. Chiche L, Forel J-M, Roch A, et al. Active cytomegalovirus infection is common in mechanically ventilated medical intensive care unit patients. *Crit Care Med*. 2009;37(6):1850-1857. doi:10.1097/CCM.0b013e31819ffea6.

81. Mansfield S, Grießl M, Gutknecht M, Cook CH. Sepsis and cytomegalovirus: foes or conspirators? *Med Microbiol Immunol*. 2015;204(3):431-437. doi:10.1007/s00430-015-0407-0.

82. Cook CH, Yenchar JK, Kraner TO, Davies EA, Ferguson RM. Occult herpes family viruses may increase mortality in critically ill surgical patients. *Am J Surg*. 1998;176(4):357-360. http://www.ncbi.nlm.nih.gov/pubmed/9817255. Accessed September 12, 2016.

83. Papazian L, Hraiech S, Lehingue S, et al. Cytomegalovirus reactivation in ICU patients. *Intensive Care Med*. 2016;42(1):28-37. doi:10.1007/s00134-015-4066-9.

84. Limaye AP, Boeckh M. CMV in critically ill patients: pathogen or bystander? *Rev Med Virol*. 2010;20(6):372-379. doi:10.1002/rmv.664.

85. Chilet M, Aguilar G, Benet I, et al. Virological and immunological features of active cytomegalovirus infection in nonimmunosuppressed patients in a surgical and trauma intensive care unit. *J Med Virol*. 2010;82(8):1384-1391. doi:10.1002/jmv.21825.

86. Frantzeskaki FG, Karampi E-S, Kottaridi C, et al. Cytomegalovirus reactivation in a general, nonimmunosuppressed intensive care unit population: Incidence, risk factors, associations with organ dysfunction, and inflammatory biomarkers. *J Crit Care*. 2015;30(2):276-281. doi:10.1016/j.jcrc.2014.10.002.

87. Forster MR, Trgovcich J, Zimmerman P, et al. Antiviral prevention of sepsis induced cytomegalovirus reactivation in immunocompetent mice. *Antiviral Res*. 2010;85(3):496-503. doi:10.1016/j.antiviral.2009.12.004.

88. Cook CH, Zhang Y, Sedmak DD, Martin LC, Jewell S, Ferguson RM. Pulmonary cytomegalovirus reactivation causes pathology in immunocompetent mice. *Crit Care Med*. 2006;34(3):842-849. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1894751&tool=pmcentrez&rendertype=abstract. Accessed May 2, 2016.

89. Papazian L, Fraisse A, Garbe L, et al. CytomegalovirusAn Unexpected Cause of Ventilator-associated Pneumonia. *J Am Soc Anesthesiol*. 1996;84(2):280-287.

90. Chan K-S, Yang C-C, Chen C-M, et al. Cytomegalovirus colitis in intensive care unit patients: Difficulties in clinical diagnosis. *J Crit Care*. 2014;29(3):474.e1-474.e6. doi:10.1016/j.jcrc.2014.01.003.

91. Limaye AP, Kirby KA, Rubenfeld GD, et al. Cytomegalovirus Reactivation in Critically Ill Immunocompetent Patients. *JAMA*. 2008;300(4):413. doi:10.1001/jama.300.4.413.

92. Forster MR, Trgovcich J, Zimmerman P, et al. Antiviral prevention of sepsis induced cytomegalovirus reactivation in immunocompetent mice. *Antiviral Res*. 2010;85(3):496-503. doi:10.1016/j.antiviral.2009.12.004.

93. Cobbs CS, Harkins L, Samanta M, et al. Human Cytomegalovirus Infection and Expression in Human Malignant Glioma. *CANCER Res*. 2002;62:3347-3350.

94. Lehrer S, Labombardi V, Green S, Pessin-Minsley MS, Germano IM, Rosenzweig KE. No circulating cytomegalovirus in five patients with glioblastoma multiforme. *Anticancer Res*. 2011;31(3):959-960. doi:31/3/959 [pii].

95. Lucas KG, Bao L, Bruggeman R, Dunham K, Specht C. The detection of CMV pp65 and IE1 in glioblastoma multiforme. *J Neurooncol*. 2011;103(2):231-238. doi:10.1007/s11060-010-0383-6.

96. Scheurer ME, Bondy ML, Aldape KD, Albrecht T, El-Zein R. Detection of human cytomegalovirus in different histological types of gliomas. *Acta Neuropathol*. 2008;116(1):79-86. doi:10.1007/s00401-008-0359-1.

97. Ghazi A, Ashoori A, Hanley P, et al. Generation of Polyclonal CMV-specific T Cells for the Adoptive Immunotherapy of Glioblastoma. *J Immunother J Immunother*. 2012;35(2):159-168. doi:10.1097/CJI.0b013e318247642f.

98. Mitchell DA, Xie W, Schmittling R, et al. Sensitive detection of human cytomegalovirus in tumors and peripheral blood of patients diagnosed with glioblastoma. *Neuro Oncol*. 2008;10(1):10-18. doi:10.1215/15228517-2007-035.

99. Ulasov I V., Kaverina N V., Ghosh D, et al. CMV70-3P miRNA contributes to the CMV mediated glioma stemness and represents a target for glioma experimental therapy. *Oncotarget*. 2016. doi:10.18632/oncotarget.11175.

100. Soroceanu L, Matlaf L, Khan S, et al. Cytomegalovirus Immediate-Early Proteins Promote Stemness Properties in Glioblastoma. *Cancer Res*. 2015;75(15):3065-3076. doi:10.1158/0008-5472.CAN-14-3307.

101. Dziurzynski K, Wei J, Qiao W, et al. Glioma-associated cytomegalovirus mediates subversion of the monocyte lineage to a tumor propagating phenotype. *Clin Cancer Res*. 2011;17(14):4642-4649. doi:10.1158/1078-0432.CCR-11-0414.

102. Lehrer S, Green S, Ramanathan L, Rosenzweig K, Labombardi V. No consistent relationship of glioblastoma incidence and cytomegalovirus seropositivity in whites, blacks, and hispanics. *Anticancer Res*. 2012;32(3):1113-1115.

103. Cobbs CS. Cytomegalovirus and brain tumor: epidemiology, biology and therapeutic aspects. *Curr Opin Oncol*. 2013;25(6):682-688. doi:10.1097/CCO.0000000000000005.

104. Khan Z, Yaiw K-C, Wilhelmi V, et al. Human cytomegalovirus immediate early proteins promote degradation of connexin 43 and disrupt gap junction communication: implications for a role in gliomagenesis. *Carcinogenesis*. 2014;35(1):145-154. doi:10.1093/carcin/bgt292.

105. Ferguson SD, Srinivasan VM, Ghali MG, Heimberger AB. Cytomegalovirus-targeted immunotherapy and glioblastoma: hype or hope? *Immunotherapy*. 2016;8(4):413-423. doi:10.2217/imt.16.2.

106. Lee K, Jeon K, Kim J-M, et al. Downregulation of GFAP, TSP-1, and p53 in human glioblastoma cell line, U373MG, by IE1 protein from human cytomegalovirus. *Glia*. 2005;51(December 2004):1-12. doi:10.1002/glia.20179.

107. Cobbs C, Khan S, Matlaf L, et al. HCMV glycoprotein B is expressed in primary glioblastomas and enhances growth and invasiveness via PDGFR-alpha activation. *Oncotarget*. 2014;5(4):1091-1100. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4011586&tool=pmcentrez&rendertype=abstract.

108. Söderberg-Nauclér C, Rahbar A, Stragliotto G. Survival in Patients with Glioblastoma Receiving Valganciclovir. *N Engl J Med*. 2013;368(19):1767-1769. doi:10.1056/NEJMp1303698.

109. Lawler SE. Cytomegalovirus and glioblastoma; controversies and opportunities. *J Neurooncol*. 2015;123(3):465-471. doi:10.1007/s11060-015-1734-0.

110. Schuessler A, Smith C, Beagley L, et al. Autologous T-cell Therapy for Cytomegalovirus as a Consolidative Treatment for Recurrent Glioblastoma. *Cancer Res*. 2014;74(13). http://cancerres.aacrjournals.org/content/74/13/3466.long. Accessed July 27, 2017.

111. RIFKIND D, Hill, R.B. J, Schneck SA, et al. Cytomegalovirus Infection After Renal Transplantation. *Arch Intern Med*. 1965;116(4):554. doi:10.1001/archinte.1965.03870040068014.

112. Balfour HH. Cytomegalovirus: The Transplant Troll Takes Its Toll. *Clin Microbiol Newsl*. 1980;2(16).

113. Hryniewiecka E, Sołdacki D, Pączek L. Cytomegaloviral infection in solid organ transplant recipients: Preliminary report of one transplant center experience. *Transplant Proc*. 2014;46(8):2572-2575. doi:10.1016/j.transproceed.2014.09.012.

114. Humar A, Michaels M. American Society of Transplantation Recommendations for Screening, Monitoring and Reporting of Infectious Complications in Immunosuppression Trials in Recipients of Organ Transplantation. *Am J Transplant*. 2006;6(2):262-274. doi:10.1111/j.1600-6143.2005.01207.x.

115. Blyth D, Lee I, Sims KD, et al. Risk factors and clinical outcomes of cytomegalovirus disease occurring more than one year post solid organ transplantation. *Transpl Infect Dis*. 2012;14(2):149-155. doi:10.1111/j.1399-3062.2011.00705.x.

116. Linares L, Sanclemente G, Cervera C, et al. Influence of cytomegalovirus disease in outcome of solid organ transplant patients. *Transplant Proc*. 2011;43(6):2145-2148. doi:10.1016/j.transproceed.2011.05.007.

117. Rand KH, Pollard RB, Merigan TC. Increased Pulmonary Superinfections in Cardiac-Transplant Patients Undergoing Primary Cytomegalovirus Infection. *N Engl J Med*. 1983;308(2):97-100.

118. Limaye AP, Bakthavatsalam R, Kim HW, et al. Impact of cytomegalovirus in organ transplant recipients in the era of antiviral prophylaxis. *Transplantation*. 2006;81(12):1645-1652. doi:10.1097/01.tp.0000226071.12562.1a.

119. Nashan B, Gaston R, Emery V, et al. Review of Cytomegalovirus Infection Findings With Mammalian Target of Rapamycin Inhibitor-Based Immunosuppressive Therapy in De Novo Renal Transplant Recipients. *Transplant J*. 2012;93(11):1075-1085. doi:10.1097/TP.0b013e31824810e6.

120. Brennan DC, Legendre C, Patel D, et al. Cytomegalovirus incidence between everolimus versus mycophenolate in de novo renal transplants: Pooled analysis of three clinical trials. *Am J Transplant*. 2011;11(11):2453-2462. doi:10.1111/j.1600-6143.2011.03674.x.

121. Atabani SF, Smith C, Atkinson C, et al. Cytomegalovirus replication kinetics in solid organ transplant recipients managed by preemptive therapy. *Am J Transplant*. 2012;12(9):2457-2464. doi:10.1111/j.1600-6143.2012.04087.x.

122. Shanahan a, Malani PN, Kaul DR. Relapsing cytomegalovirus infection in solid organ transplant recipients. *Transpl Infect Dis*. 2009;11(6):513-518. doi:10.1111/j.1600-6143.2009.02897.x.

123. Dockrell DH, Prada J, Jones MF, et al. Seroconversion to human herpesvirus 6 following

liver transplantation is a marker of cytomegalovirus disease. *J Infect Dis*. 1997;176(5):1135-1140.

124. Grundy JE, Lui SF, Super M, et al. Symptomatic cytomegalovirus infection in seropositive kidney recipients: reinfection with donor virus rather than reactivation of recipient virus. *Lancet (London, England)*. 1988;2(8603):132-135. http://www.ncbi.nlm.nih.gov/pubmed/2899189. Accessed August 12, 2016.

125. Ciáurriz M, Zabalza A, Beloki L, et al. The immune response to cytomegalovirus in allogeneic hematopoietic stem cell transplant recipients. *Cell Mol Life Sci*. 2015. doi:10.1007/s00018-015-1986-z.

126. Boeckh M, Leisenring W, Riddell SR, et al. Late cytomegalovirus disease and mortality in recipients of allogeneic hematopoietic stem cell transplants: importance of viral load and T-cell immunity. *Blood*. 2003;101(2). doi:10.1182/blood-2002-03-0993.

127. Özdemir E, Saliba R, Champlin R, et al. Risk factors associated with late cytomegalovirus reactivation after allogeneic stem cell transplantation for hematological malignancies. *Bone Marrow Transplant*. 2007;40. doi:10.1038/sj.bmt.1705699.

128. Li A. Archives of Virology Inhibitory Effeets of Fosearnet on Herpesvirus Multiplication in Cell Culture. *Arch Virol*. 1981;209:203-209.

129. Snoeck R, Sakuma T, De Clercq E, Rosenberg I, Holy A. (S)-1-(3-hydroxy-2-phosphonylmethoxypropyl)cytosine, a potent and selective inhibitor of human cytomegalovirus replication. *Antimicrob Agents Chemother*. 1988;32(12):1839-1844. http://www.ncbi.nlm.nih.gov/pubmed/2854454. Accessed September 12, 2016.

130. Boeckh M, Nichols WG, Papanicolaou G, Rubin R, Wingard JR, Zaia J. Cytomegalovirus in hematopoietic stem cell transplant recipients: current status, known challenges, and future strategies. *Biol Blood Marrow Transplant*. 2003;9(9):543-558. doi:10.1016/S1083-8791(03)00287-8.

131. Ariza-Heredia EJ, Nesher L, Chemaly RF. Cytomegalovirus diseases after hematopoietic stem cell transplantation: a mini-review. *Cancer Lett*. 2014;342(1):1-8. doi:10.1016/j.canlet.2013.09.004.

132. van Burik JA, Lawatsch EJ, DeFor TE, Weisdorf DJ. Cytomegalovirus enteritis among hematopoietic stem cell transplant recipients. *Biol Blood Marrow Transplant*. 2001;7(12):674-679. doi:10.1053/bbmt.2001.v7.pm11787530.

133. Meyers JD, Flournoy N, Thomas ED. Risk factors for cytomegalovirus infection after human marrow transplantation. *J Infect Dis*. 1986;153(3):478-488. http://www.ncbi.nlm.nih.gov/pubmed/3005424. Accessed September 12, 2016.

134. Lilleri D, Gerna G. Strategies to control human cytomegalovirus infection in adult hematopoietic stem cell transplant recipients. *Immunotherapy*. 2016;8(9):1135-1149. doi:10.2217/imt-2015-0028.

135. Erard V, Guthrie KA, Seo S, et al. Reduced Mortality of Cytomegalovirus Pneumonia After Hematopoietic Cell Transplantation Due to Antiviral Therapy and Changes in Transplantation Practices. *Clin Infect Dis*. 2015;61(1):31-39. doi:10.1093/cid/civ215.

136. Crippa F, Corey L, Chuang EL, Sale G, Boeckh M. Virological, Clinical, and Ophthalmologic Features of Cytomegalovirus Retinitis after Hematopoietic Stem Cell Transplantation. *Clin Infect Dis*. 2001;32(2):214-219. doi:10.1086/318447.

137. Wolf DG, Lurain NS, Zuckerman T, et al. Emergence of late cytomegalovirus central nervous system disease in hematopoietic stem cell transplant recipients. 2003;101:463-465. doi:10.1182/blood-2002-07-1982.

138. Emery VC, Sabin CA, Cope A V, Gor D, Hassan-Walker AF, Griffiths PD. Application of viral-load kinetics to identify patients who develop cytomegalovirus disease after transplantation. *Lancet*. 2000;355(9220):2032-2036. doi:10.1016/S0140-6736(00)02350-3.

139. de Oliveira PGAG, Ueda MYH, Real JM, et al. Simultaneous Quantification of the 8 Human Herpesviruses in Allogeneic Hematopoietic Stem Cell Transplantation. *Transplantation*. 2015;0(0):1. doi:10.1097/TP.0000000000000986.

140. Walker CM, van Burik J-AH, De For TE, Weisdorf DJ, Weisdorf DJ, Meyers JD. Cytomegalovirus infection after allogeneic transplantation: comparison of cord blood with peripheral blood and marrow graft sources. *Biol Blood Marrow Transplant*. 2007;13(9):1106-1115. doi:10.1016/j.bbmt.2007.06.006.

141. Nishihori T, El-Asmar J, Aljurf M, Kharfan-Dabaja MA. Therapeutic strategies for cytomegalovirus in allogeneic hematopoietic cell transplantation. *Immunotherapy*. 2015;7(10):1059--1071. http://www.futuremedicine.com.gate2.inist.fr/doi/pdf/10.2217/imt.15.70. Accessed June 1, 2017.

142. Matsumura T, Narimatsu H, Kami M, et al. Cytomegalovirus Infections following Umbilical Cord Blood Transplantation Using Reduced Intensity Conditioning Regimens for Adult Patients. *Biol Blood Marrow Transplant*. 2007;13(5):577-583. doi:10.1016/j.bbmt.2006.12.454.

143. Thomson BG, Robertson KA, Gowan D, et al. Analysis of engraftment, graft-versus-host disease, and immune recovery following unrelated donor cord blood transplantation. 2000;96:2703-2711. http://www.bloodjournal.org/content/bloodjournal/96/8/2703.full.pdf. Accessed July 29, 2017.

144. Cohen T, Williams JD, Opperman TJ, Sanchez R, Lurain NS, Tortorella D. Convallatoxin-induced reduction of methionine import effectively inhibits human cytomegalovirus infection and replication. *J Virol*. September 2016:JVI.01050-16. doi:10.1128/JVI.01050-16.

145. Gabanti E, Lilleri D, Ripamonti F, et al. Reconstitution of Human Cytomegalovirus–Specific CD4+ T Cells is Critical for Control of Virus Reactivation in Hematopoietic Stem Cell Transplant Recipients but Does Not Prevent Organ Infection. *Biol Blood Marrow Transplant*. 2015;21(12):2192-2202. doi:10.1016/j.bbmt.2015.08.002.

146. Espigado I, de la Cruz-Vicente F, BenMarzouk-Hidalgo OJ, et al. Timing of CMV-specific effector memory T cells predicts viral replication and survival after allogeneic hematopoietic stem cell transplantation. *Transpl Int*. 2014;27(12):1253-1262. doi:10.1111/tri.12406.

147. Florescu DF, Qiu F, Schmidt CM, Kalil AC. A Direct and Indirect Comparison Meta-Analysis on the Efficacy of Cytomegalovirus Preventive Strategies in Solid Organ Transplant. *Clin Infect Dis*. 2014;58(6). doi:10.1093/cid/cit945.

148. Junghanss C, Boeckh M, Carter RA, et al. Incidence and outcome of cytomegalovirus infections following nonmyeloablative compared with myeloablative allogeneic stem cell transplantation, a matched control study. *Blood*. 2002;99:1978-1985.

149. Nichols WG, Corey L, Gooley T, Davis C, Boeckh M. High Risk of Death Due to Bacterial and Fungal Infection among Cytomegalovirus (CMV)–Seronegative Recipients of Stem Cell Transplants from Seropositive Donors: Evidence for Indirect Effects of Primary CMV Infection. *J Infect Dis*. 2002;185(3):273-282. doi:10.1086/338624.

150. Navarro D. The expanding role of cytomegalovirus as a human pathogen. *J Med Virol*.

December 2015. doi:10.1002/jmv.24450.

151. Pergam SA, Xie H, Sandhu R, et al. Efficiency and risk factors for CMV transmission in seronegative hematopoietic stem cell recipients. *Biol Blood Marrow Transplant*. 2012;18(9):1391-1400. doi:10.1016/j.bbmt.2012.02.008.

152. Ljungman P. The role of cytomegalovirus serostatus on outcome of hematopoietic stem cell transplantation. *Curr Opin Hematol*. 2014;21(6):466-469. doi:10.1097/MOH.0000000000000085.

153. Nichols WG, Price TH, Gooley T, Corey L, Boeckh M. Transfusion-transmitted cytomegalovirus infection after receipt of leukoreduced blood products. *Blood*. 2003;101(10). doi:10.1182/blood.

154. Marr KA, Carter RA, Boeckh M, Martin P, Corey L. Invasive aspergillosis in allogeneic stem cell transplant recipients: changes in epidemiology and risk factors. *Blood*. 2002;100(13). doi:10.1182/blood-2002-05-1496.

155. Schmidt-Hieber M, Labopin M, Beelen D, et al. CMV serostatus still has an important prognostic impact in de novo acute leukemia patients after allogeneic stem cell transplantation: a report from the Acute Leukemia Working Party of EBMT. *Blood*. 2013;122(19):3359-3364. doi:10.1182/blood-2013-05.

156. Elmaagacli AH, Steckel NK, Koldehoff M, et al. Early human cytomegalovirus replication after transplantation is associated with a decreased relapse risk: evidence for a putative virus-versus-leukemia effect in acute myeloid leukemia patients. *Blood*. 2011;118(5):1402-1412. doi:10.1182/blood-2010-08-304121.

157. Griffiths PD, Grundy JE. Molecular biology and immunology of cytomegalovirus. *Biochem J*. 1987;241(2):313-324. http://www.ncbi.nlm.nih.gov/pubmed/3036059. Accessed February 10, 2017.

158. Jenkins FJ, Roizman B. Herpes Simplex Virus 1 Recombinants with Noninverting Genomes Frozen in Different Isomeric Arrangements Are Capable of Independent Replication. *J Virol*. 1986;59(2):494-499.

159. Chee M, Rudolph SA, Plachter B, Barrell B, Jahn G. Identification of the major capsid protein gene of human cytomegalovirus. *J Virol*. 1989;63(3):1345-1353. http://www.ncbi.nlm.nih.gov/pubmed/2536837. Accessed February 10, 2017.

160. Gibson W, Baxter MK, Clopper KS. Cytomegalovirus "missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46. *J Virol*. 1996;70(11):7454-7461. http://www.ncbi.nlm.nih.gov/pubmed/8892863. Accessed February 10, 2017.

161. Gibson W, Clopper KS, Britt WJ, Baxter MK. Human cytomegalovirus (HCMV) smallest capsid protein identified as product of short open reading frame located between HCMV UL48 and UL49. *J Virol*. 1996;70(8):5680-5683. http://www.ncbi.nlm.nih.gov/pubmed/8764088. Accessed February 10, 2017.

162. Kalejta RF. Tegument proteins of human cytomegalovirus. *Microbiol Mol Biol Rev*. 2008;72(2):249-65, table of contents. doi:10.1128/MMBR.00040-07.

163. Varnum SM, Streblow DN, Monroe ME, et al. Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome. *J Virol*. 2004;78(20):10960-10966. doi:10.1128/JVI.78.20.10960-10966.2004.

164. Gretch DR, Kari B, Rasmussen L, Gehrz RC, Stinski1 MF. Identification and Characterization of Three Distinct Families of Glycoprotein Complexes in the Envelopes of Human Cytomegalovirus. *J Virol*. 1988;62(3):875-881.

165. Mach M, Kropff B, Dal Monte P, Britt W. Complex formation by human cytomegalovirus

glycoproteins M (gpUL100) and N (gpUL73). *J Virol*. 2000;74(24):11881-11892. http://www.ncbi.nlm.nih.gov/pubmed/11090188. Accessed February 10, 2017.

166. Boyle KA, Compton T. Receptor-Binding Properties of a Soluble Form of Human Cytomegalovirus Glycoprotein B. *J Virol*. 1998;72(3):1826-1833.

167. Kari B, Gehrz R. NOTES A Human Cytomegalovirus Glycoprotein Complex Designated gC-II Is a Major Heparin-Binding Component of the Envelope. *J Virol*. 1992;66(3):1761-1764.

168. Kinzler ER, Compton T. Characterization of Human Cytomegalovirus Glycoprotein-Induced Cell-Cell Fusion. *J Virol*. 2005;79(12):7827-7837. doi:10.1128/JVI.79.12.7827–7837.2005.

169. Li L, Nelson JA, Britt WJ. Glycoprotein H-related complexes of human cytomegalovirus: identification of a third protein in the gCIII complex. *J Virol*. 1997;71(4):3090-3097. http://www.ncbi.nlm.nih.gov/pubmed/9060671. Accessed February 10, 2017.

170. Irmiere A, Gibson W. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology*. 1983;130(1):118-133. doi:10.1016/0042-6822(83)90122-8.

171. Li G, Kamil JP. Viral Regulation of Cell Tropism in Human Cytomegalovirus. *J Virol*. 2015;90(2):626-629. doi:10.1128/JVI.01500-15.

172. Compton T, Nowlin DM, Cooper NR. Initiation of Human Cytomegalovirus Infection Requires Initial Interaction with Cell Surface Heparan Sulfate. *Virology*. 1993;193(2):834-841. doi:10.1006/viro.1993.1192.

173. Halary F, Amara A, Lortat-Jacob H, et al. Human cytomegalovirus binding to DC-SIGN is required for dendritic cell infection and target cell trans-infection. *Immunity*. 2002;17(5):653-664. doi:10.1016/S1074-7613(02)00447-8.

174. Soroceanu L, Akhavan A, Cobbs CS. Platelet-derived growth factor-alpha receptor activation is required for human cytomegalovirus infection. *Nature*. 2008;455(7211):391-395. doi:10.1038/nature07209.

175. Kinzler ER, Compton T. Characterization of human cytomegalovirus glycoprotein-induced cell-cell fusion. *J Virol*. 2005;79(12):7827-7837. doi:10.1128/JVI.79.12.7827-7837.2005.

176. Patrone M, Coroadinha AS, Teixeira AP, Alves PM. Palmitoylation Strengthens Cholesterol-dependent Multimerization and Fusion Activity of Human Cytomegalovirus Glycoprotein B (gB). *J Biol Chem*. 2016;291(9):4711-4722. doi:10.1074/jbc.M115.682252.

177. Ryckman BJ, Jarvis MA, Drummond DD, Nelson JA, Johnson DC. Human cytomegalovirus entry into epithelial and endothelial cells depends on genes UL128 to UL150 and occurs by endocytosis and low-pH fusion. *J Virol*. 2006;80(2):710-722. doi:10.1128/JVI.80.2.710-722.2006.

178. Haspot F, Lavault A, Sinzger C, et al. Human Cytomegalovirus Entry into Dendritic Cells Occurs via a Macropinocytosis-Like Pathway in a pH-Independent and Cholesterol-Dependent Manner. Kremer EJ, ed. *PLoS One*. 2012;7(4):e34795. doi:10.1371/journal.pone.0034795.

179. Wathen MW, Stinski MF. Temporal patterns of human cytomegalovirus transcription: mapping the viral RNAs synthesized at immediate early, early, and late times after infection. *J Virol*. 1982;41(2):462-477. http://www.ncbi.nlm.nih.gov/pubmed/6281461. Accessed February 10, 2017.

180. Anders DG, Kacica MA, Pari G, Punturieri SM. Boundaries and structure of human cytomegalovirus oriLyt, a complex origin for lytic-phase DNA replication. *J Virol*. 1992;66(6):3373-3384. http://www.ncbi.nlm.nih.gov/pubmed/1316454. Accessed February 10, 2017.

181. Xu Y, Cei SA, Rodriguez Huete A, Colletti KS, Pari GS. Human cytomegalovirus DNA replication requires transcriptional activation via an IE2- and UL84-responsive bidirectional promoter element within oriLyt. *J Virol*. 2004;78(21):11664-11677. doi:10.1128/JVI.78.21.11664-11677.2004.

182. Mettenleiter TC. Herpesvirus assembly and egress. *J Virol*. 2002;76(4):1537-1547. doi:10.1128/JVI.76.4.1537-1547.2002.

183. Mettenleiter TC, Klupp BG, Granzow H. Herpesvirus assembly: a tale of two membranes. *Curr Opin Microbiol*. 2006;9(4):423-429. doi:10.1016/j.mib.2006.06.013.

184. Isaacson MK, Compton T. Human Cytomegalovirus Glycoprotein B Is Required for Virus Entry and Cell-to-Cell Spread but Not for Virion Attachment, Assembly, or Egress. *J Virol*. 2009;83(8):3891-3903. doi:10.1128/JVI.01251-08.

185. Stinski MF, Roehr TJ, Thomsen DR, Stenberg RM, Goins WF, Stinski MF. Activation of the Major Immediate Early Gene of Human Cytomegalovirus by cis-Acting Elements in the Promoter-Regulatory Sequence and by Virus-Specific trans-Acting Components. *J Virol Proc Nati Acad Sci USA J Virol*. 1985;55(46).

186. Meier JL, Keller MJ, McCoy JJ. Requirement of multiple cis-acting elements in the human cytomegalovirus major immediate-early distal enhancer for viral gene expression and replication. *J Virol*. 2002;76(1):313-326. doi:10.1128/JVI.76.1.313-326.2002.

187. Saffert RT, Kalejta RF. Inactivating a cellular intrinsic immune defense mediated by Daxx is the mechanism through which the human cytomegalovirus pp71 protein stimulates viral immediate-early gene expression. *J Virol*. 2006;80(8):3863-3871. doi:10.1128/JVI.80.8.3863-3871.2006.

188. Nevels M, Paulus C, Shenk T. Human cytomegalovirus immediate-early 1 protein facilitates viral replication by antagonizing histone deacetylation. *Proc Natl Acad Sci U S A*. 2004;101(49):17234-17239. doi:10.1073/pnas.0407933101.

189. Tang Q, Maul GG. Mouse cytomegalovirus immediate-early protein 1 binds with host cell repressors to relieve suppressive effects on viral transcription and replication during lytic infection. *J Virol*. 2003;77(2):1357-1367. doi:10.1128/jvi.77.2.1357-1367.2003.

190. Cherrington JM, Khoury EL, Mocarskil ES. Human Cytomegalovirus ie2 Negatively Regulates x Gene Expression via a Short Target Sequence Near the Transcription Start Site. *J Virol*. 1991;65(2):887-896.

191. Macias MP, Huang L, Lashmit PE, Stinski MF. Cellular or Viral Protein Binding to a Cytomegalovirus Promoter Transcription Initiation Site: Effects on Transcription. *J Virol*. 1996;70(6):3628-3635.

192. Glass M, Busche A, Wagner K, Messerle M, Borst EM. Conditional and reversible disruption of essential herpesvirus proteins. *Nat Methods*. 2009;6(8):577-579. doi:10.1038/nmeth.1346.

193. Marchini A, Liu H, Zhu H. Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes. *J Virol*. 2001;75(4):1870-1878. doi:10.1128/JVI.75.4.1870-1878.2001.

194. Mocarski ES, Kemble GW, Lyle JM, Greaves RF. A deletion mutant in the human cytomegalovirus gene encoding IE1(491aa) is replication defective due to a failure in autoregulation. *Proc Natl Acad Sci U S A*. 1996;93(21):11321-11326. http://www.ncbi.nlm.nih.gov/pubmed/8876134. Accessed July 12, 2016.

195. Azad RF, Driver VB, Tanaka K, Crooke RM, Anderson KP. Antiviral activity of a phosphorothioate oligonucleotide complementary to RNA of the human cytomegalovirus major immediate-early region. *Antimicrob Agents Chemother*. 1993;37(9):1945-1954.

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=188097&tool=pmcentrez&rendertype=abstract. Accessed July 16, 2015.

196. White EA, Clark CL, Sanchez V, Spector DH. Small internal deletions in the human cytomegalovirus IE2 gene result in nonviable recombinant viruses with differential defects in viral gene expression. *J Virol*. 2004;78(4):1817-1830. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=369462&tool=pmcentrez&rendertype=abstract. Accessed July 16, 2015.

197. Wu J, O'Neill J, Barbosa MS. Transcription factor Sp1 mediates cell-specific trans-activation of the human cytomegalovirus DNA polymerase gene promoter by immediate-early protein IE86 in glioblastoma U373MG cells. *J Virol*. 1998;72(1):236-244. http://www.ncbi.nlm.nih.gov/pubmed/9420220. Accessed February 22, 2017.

198. Kerry JA, Priddy MA, Stenberg RM. Identification of sequence elements in the human cytomegalovirus DNA polymerase gene promoter required for activation by viral gene products. *J Virol*. 1994;68(7):4167-4176. http://www.ncbi.nlm.nih.gov/pubmed/8207792. Accessed February 16, 2017.

199. Stasiak PC, Mocarski ES. Transactivation of the Cytomegalovirus ICP36 Gene Promoter Requires the x Gene Product TRS1 in Addition to IEI and IE2. *J Virol*. 1992;66(2):1050-1058.

200. Arlt H, Lang D, Gebert S, Stamminger T. Identification of Binding Sites for the 86-Kilodalton IE2 Protein of Human Cytomegalovirus within an IE2-Responsive Viral Early Promoter. *J Virol*. 1994:4117-4125.

201. Angulo A, Ghazal P, Messerle M. The major immediate-early gene ie3 of mouse cytomegalovirus is essential for viral growth. *J Virol*. 2000;74(23):11129-11136. http://www.ncbi.nlm.nih.gov/pubmed/11070009. Accessed July 12, 2016.

202. Greaves RF, Mocarski ES. Defective growth correlates with reduced accumulation of a viral DNA replication protein after low-multiplicity infection by a human cytomegalovirus ie1 mutant. *J Virol*. 1998;72(1):366-379. http://www.ncbi.nlm.nih.gov/pubmed/9420235. Accessed July 12, 2016.

203. Gawn JM, Greaves RF. Absence of IE1 p72 protein function during low-multiplicity infection by human cytomegalovirus results in a broad block to viral delayed-early gene expression. *J Virol*. 2002;76(9):4441-4455. doi:10.1128/jvi.76.9.4441-4455.2002.

204. Nevels M, Brune W, Shenk T. SUMOylation of the human cytomegalovirus 72-kilodalton IE1 protein facilitates expression of the 86-kilodalton IE2 protein and promotes viral replication. *J Virol*. 2004;78(14):7803-7812. doi:10.1128/JVI.78.14.7803-7812.2004.

205. Zhu H, Shen Y, Shenk T. Human Cytomegalovirus IE1 and IE2 Proteins Block Apoptosis. *J Virol*. 1995;69(12):7960-7970.

206. Castillo JP, Yurochko AD, Kowalik TF. Role of human cytomegalovirus immediate-early proteins in cell growth control. *J Virol*. 2000;74(17):8028-8037. doi:10.1128/JVI.74.17.8028-8037.2000.

207. Murphy EA, Streblow DN, Nelson JA, Stinski MF. The human cytomegalovirus IE86 protein can block cell cycle progression after inducing transition into the S phase of permissive cells. *J Virol*. 2000;74(15):7108-7118. doi:10.1128/JVI.74.15.7108-7118.2000.

208. Scholz M, Doerr HW, Cinatl J. Inhibition of cytomegalovirus immediate early gene expression: a therapeutic option? *Antiviral Res*. 2001;49(3):129-145. http://www.ncbi.nlm.nih.gov/pubmed/11428240. Accessed August 24, 2016.

209. Paulus C, Krauss S, Nevels M. A human cytomegalovirus antagonist of type I IFN-dependent signal transducer and activator of transcription signaling. *Proc Natl Acad Sci U S A*.

2006;103(10):3840-3845. doi:10.1073/pnas.0600007103.

210. Taylor RT, Bresnahan WA. Human Cytomegalovirus Immediate-Early 2 Protein IE86 Blocks Virus-Induced Chemokine Expression. *J Virol*. 2006;80(2):920-928. doi:10.1128/JVI.80.2.920-928.2006.

211. Geist LJ, Monick MM, Stinski MF, Hunninghake GW. Rapid Publication Cytomegalovirus Immediate Early Genes Prevent the Inhibitory Effect of Cyclosporin A on Interleukin 2 Gene Transcription. 1992;90:2136-2140.

212. Stenberg RM, Thomsen DR, Stinski MF. Structural Analysis of the Major Immediate Early Gene of Human Cytomegalovirus. *J Virol*. 1984;49(1):190-199.

213. Stenberg RM, Witte PR, Stinski MF. Multiple Spliced and Unspliced Transcripts from Human Cytomegalovirus Immediate-Early Region 2 and Evidence for a Common Initiation Site within Immediate-Early Region 1. *J Virol*. 1985:665-675.

214. Kerry JA, Sehgal A, Barlow SW, et al. Isolation and Characterization of a Low-Abundance Splice Variant from the Human Cytomegalovirus Major Immediate-Early Gene Region. *J Virol*. 1995;69(6):3868-3872.

215. Baracchini E, Glezer E, Fish K, Stenberg RM, Nelson JA, Ghazal P. An lsoform Variant of the Cytomegalovirus Immediate-Early Auto Repressor Functions as a Transcriptional Activator. *Virology*. 1992;188:518-529.

216. Sanders RL, Del Rosario CJ, White EA, Spector DH. Internal deletions of IE2 86 and loss of the late IE2 60 and IE2 40 proteins encoded by human cytomegalovirus affect the levels of UL84 protein but not the amount of UL84 mRNA or the loading and distribution of the mRNA on polysomes. *J Virol*. 2008;82(22):11383-11397. doi:10.1128/JVI.01293-08.

217. Plachter B, Britt W, Vornhagen R, Stamminger T, Jahn G. Analysis of Proteins Encoded by IE Regions 1 and 2 of Human Cytomegalovirus Using Monoclonal Antibodies Generated against Recombinant Antigens. *Virology*. 1993;193(2):642-652. doi:10.1006/viro.1993.1172.

218. Sanchez V, Clark CL, Yen JY, Dwarakanath R, Spector DH. Viable human cytomegalovirus recombinant virus with an internal deletion of the IE2 86 gene affects late stages of viral replication. *J Virol*. 2002;76(6):2973-2989. doi:10.1128/JVI.76.6.2973-2989.2002.

219. Awasthi S, Isler J a, Alwine JC. Analysis of Splice Variants of the Immediate-Early 1 Region of Human Cytomegalovirus Analysis of Splice Variants of the Immediate-Early 1 Region of Human Cytomegalovirus. *J Virol*. 2004;78(15):8191-8200. doi:10.1128/JVI.78.15.8191.

220. Hou W, Torres L, Cruz-Cosme R, et al. Two Polypyrimidine Tracts in Intron 4 of the Major Immediate-Early (MIE) Gene Are Critical for the Gene Expression Switching from IE1 to IE2 and for Replication of Human Cytomegalovirus. *J Virol*. June 2016. doi:10.1128/JVI.00837-16.

221. Taylor-Wiedeman J, Sissons JGP, Borysiewicz~f LK, Sinclair JH. Monocytes are a major site of persistence of human cytomegalovirus in peripheral blood mononuclear cells. *J Gen Virol*. 1991;72:2059-2064. http://www.microbiologyresearch.org/docserver/fulltext/jgv/72/9/JV0720092059.pdf?expires=1488478101&id=id&accname=guest&checksum=20E939759CA681FC65B3A31062D51D2D. Accessed November 9, 2016.

222. Kondo K, Kaneshima H, Mocarski ES. Human cytomegalovirus latent infection of granulocyte-macrophage progenitors. *Proc Natl Acad Sci U S A*. 1994;91(25):11879-11883. http://www.ncbi.nlm.nih.gov/pubmed/7991550. Accessed September 12, 2016.

223. Mendelson M, Monard S, Sissons P, Sinclair J. Detection of endogenous human cytomegalovirus in CD34+ bone marrow progenitors. *J Gen Virol*. 1996;77(12):3099-3102.

doi:10.1099/0022-1317-77-12-3099.

224. Goodrum F, Jordan CT, Terhune SS, High K, Shenk T. Differential outcomes of human cytomegalovirus infection in primitive hematopoietic cell subpopulations. *Blood*. 2004;104(3):687-695. doi:10.1182/blood-2003-12-4344.

225. Maciejewski JP, St Jeor SC. Human cytomegalovirus infection of human hematopoietic progenitor cells. *Leuk Lymphoma*. 1992;33(1-2):1-13. doi:10.3109/10428199909093720.

226. Slobedman B, Mocarski ES. Quantitative analysis of latent human cytomegalovirus. *J Virol*. 1999;73(6):4806-4812. http://www.ncbi.nlm.nih.gov/pubmed/10233941. Accessed March 2, 2017.

227. Saffert RT, Penkert RR, Kalejta RF. Cellular and viral control over the initial events of human cytomegalovirus experimental latency in CD34+ cells. *J Virol*. 2010;84(11):5594-5604. doi:10.1128/JVI.00348-10.

228. Reeves MB, MacAry P a, Lehner PJ, Sissons JGP, Sinclair JH. Latency, chromatin remodeling, and reactivation of human cytomegalovirus in the dendritic cells of healthy carriers. *Proc Natl Acad Sci U S A*. 2005;102(11):4140-4145. doi:10.1073/pnas.0408994102.

229. Murphy JC, Fischle W, Verdin E, Sinclair JH. Control of cytomegalovirus lytic gene expression by histone acetylation. *EMBO J*. 2002;21(5):1112-1120. doi:10.1093/emboj/21.5.1112.

230. Saffert RT, Kalejta RF. Human cytomegalovirus gene expression is silenced by Daxx-mediated intrinsic immune defense in model latent infections established in vitro. *J Virol*. 2007;81(17):9109-9120. doi:10.1128/JVI.00827-07.

231. O'Connor CM, Vanicek J, Murphy EA. Host microRNA regulation of human cytomegalovirus immediate early protein translation promotes viral latency. *J Virol*. 2014;88(10):5524-5532. doi:10.1128/JVI.00481-14.

232. Lee SH, Caviness K, Albright ER, et al. Long and Short Isoforms of the Human Cytomegalovirus UL138 Protein Silence IE Transcription and Promote Latency. *J Virol*. 2016;90(20):9483-9494. doi:10.1128/JVI.01547-16.

233. Rossetto CCC, Tarrant-Elorza M, Pari GGS, et al. Cis and Trans Acting Factors Involved in Human Cytomegalovirus Experimental and Natural Latent Infection of CD14 (+) Monocytes and CD34 (+) Cells. Mocarski E, ed. *PLoS Pathog*. 2013;9(5):e1003366. doi:10.1371/journal.ppat.1003366.

234. Tarrant-Elorza M, Rossetto CC, Pari GS. Maintenance and Replication of the Human Cytomegalovirus Genome during Latency. *Cell Host Microbe*. 2014;16(1):43-54. doi:10.1016/j.chom.2014.06.006.

235. Wills MR, Poole E, Lau B, Krishna B, Sinclair JH. The immunology of human cytomegalovirus latency: could latent infection be cleared by novel immunotherapeutic strategies? *Cell Mol Immunol*. 2015;12(2):128-138. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4654298&tool=pmcentrez&rendertype=abstract.

236. Humby MS, O'Connor CM. Human Cytomegalovirus US28 Is Important for Latent Infection of Hematopoietic Progenitor Cells. *J Virol*. 2015;90(6):2959-2970. doi:10.1128/JVI.02507-15.

237. Jenkins C, Abendroth A, Slobedman B. A novel viral transcript with homology to human interleukin-10 is expressed during latent human cytomegalovirus infection. *J Virol*. 2004;78(3):1440-1447. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=321375&tool=pmcentrez&rendertype=abstract. Accessed February 19, 2016.

238. Cheung AKL, Gottlieb DJ, Plachter B, et al. The role of the human cytomegalovirus UL111A gene in down-regulating CD4+ T-cell recognition of latently infected cells: implications for virus elimination during latency. *Blood*. 2009;114(19):4128-4137. doi:10.1182/blood-2008-12-197111.

239. Avdic S, Cao JZ, Cheung AKL, Abendroth A, Slobedman B. Viral interleukin-10 expressed by human cytomegalovirus during the latent phase of infection modulates latently infected myeloid cell differentiation. *J Virol*. 2011;85(14):7465-7471. doi:10.1128/JVI.00088-11.

240. Poole E, Dallas SRM, Colston J, Joseph RS V, Sinclair J. Virally induced changes in cellular microRNAs maintain latency of human cytomegalovirus in CD34+ progenitors. *J Gen Virol*. 2011;92(7):1539-1549. doi:10.1099/vir.0.031377-0.

241. Reeves M, Sinclair J. Regulation of Human Cytomegalovirus Transcription in Latency: Beyond the Major Immediate-Early Promoter. *Viruses*. 2013;5:1395-1413. doi:10.3390/v5061395.

242. Poole E, Sinclair J. Sleepless latency of human cytomegalovirus. *Med Microbiol Immunol*. 2015;204(3):421-429. doi:10.1007/s00430-015-0401-6.

243. Slobedman B, Stern JL, Cunningham AL, Abendroth A, Abate DA, Mocarski ES. Impact of human cytomegalovirus latent infection on myeloid progenitor cell gene expression. *J Virol*. 2004;78(8):4054-4062. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=374258&tool=pmcentrez&rendertype=abstract. Accessed August 13, 2015.

244. Hummel M, Abecassis MM. A model for reactivation of CMV from latency. *J Clin Virol*. 2002;25:123-136. doi:10.1016/S1386-6532(02)00088-4.

245. Ibanez CE, Schrier R, Ghazal P, Wiley C, Nelson ' JA. Human Cytomegalovirus Productively Infects Primary Differentiated Macrophages. *J Virol*. 1991;65(12):6581-6588. http://jvi.asm.org/content/65/12/6581.full.pdf. Accessed March 2, 2017.

246. Söderberg-Nauclér C, Fish KN, Nelson JA. Reactivation of Latent Human Cytomegalovirus by Allogeneic Stimulation of Blood Cells from Healthy Donors. *Cell*. 1997;91(1):119-126. doi:10.1016/S0092-8674(01)80014-3.

247. Söderberg-Nauclér C, Streblow DN, Fish KN, Allan-Yorke J, Smith PP, Nelson JA. Reactivation of latent human cytomegalovirus in CD14(+) monocytes is differentiation dependent. *J Virol*. 2001;75(16):7543-7554. doi:10.1128/JVI.75.16.7543-7554.2001.

248. Söderberg-Nauclér C, Fish KN, Nelson JA. Interferon-gamma and tumor necrosis factor-alpha specifically induce formation of cytomegalovirus-permissive monocyte-derived macrophages that are refractory to the antiviral activity of these cytokines. *J Clin Invest*. 1997;100(12):3154-3163. doi:10.1172/JCI119871.

249. Prösch S, Staak K, Stein J, et al. Stimulation of the Human Cytomegalovirus IE Enhancer/Promoter in HL-60 Cells by TNFα Is Mediated via Induction of NF-κB. *Virology*. 1995;208(1):197-206. doi:10.1006/viro.1995.1143.

250. Keyes LR, Hargett D, Soland M, et al. HCMV Protein LUNA Is Required for Viral Reactivation from Latently Infected Primary CD14+ Cells. Simas JP, ed. *PLoS One*. 2012;7(12):e52827. doi:10.1371/journal.pone.0052827.

251. Biron KK. Antiviral drugs for cytomegalovirus diseases. *Antiviral Res*. 2006;71(2-3 SPEC. ISS.):154-163. doi:10.1016/j.antiviral.2006.05.002.

252. Arvin AM. *Human Herpesviruses : Biology, Therapy and Immunoprophylaxis*. Cambridge University Press; 2007.

253. Goodrich JM, Mori M, Gleaves CA, et al. Early Treatment with Ganciclovir to Prevent

Cytomegalovirus Disease after Allogeneic Bone Marrow Transplantation. *N Engl J Med*. 1991;325(23):1601-1607. doi:10.1056/NEJM199112053252303.

254. Reusser P, Einsele H, Lee J, et al. Randomized multicenter trial of foscarnet versus ganciclovir for preemptive therapy of cytomegalovirus infection after allogeneic stem cell transplantation. *Blood*. 2002;99(4). http://www.bloodjournal.org.gate2.inist.fr/content/bloodjournal/99/4/1159.full.pdf. Accessed June 2, 2017.

255. Schmidt GM, Horak DA, Niland JC, et al. A Randomized, Controlled Trial of Prophylactic Ganciclovir for Cytomegalovirus Pulmonary Infection in Recipients of Allogeneic Bone Marrow Transplants. *N Engl J Med*. 1991;324(15):1005-1011. doi:10.1056/NEJM199104113241501.

256. Boeckh M, Gooley TA, Myerson D, Cunningham T, Schoch G, Bowden RA. Cytomegalovirus pp65 Antigenemia-Guided Early Treatment With Ganciclovir Versus Ganciclovir at Engraftment After Allogeneic Marrow Transplantation: A Randomized Double-Blind Study. *Blood*. 1996;88(19):4063-4071. http://www.bloodjournal.org.gate2.inist.fr/content/bloodjournal/88/10/4063.full.pdf. Accessed June 2, 2017.

257. Ringdén O, Lönnqvist B, Paulin T, et al. Pharmacokinetics, safety and preliminary clinical experiences using foscarnet in the treatment of cytomegalovirus infections in bone marrow and renal transplant recipients. *J Antimicrob Chemother*. 1986;17(3):373-387. http://www.ncbi.nlm.nih.gov/pubmed/3009383. Accessed September 12, 2016.

258. Goodrich JM, Bowden RA, Fisher L, Keller C, Schoch G, Meyers† JD. Ganciclovir Prophylaxis To Prevent Cytomegalovirus Disease after Allogeneic Marrow Transplant. *Ann Intern Med*. 1993;118(3):173. doi:10.7326/0003-4819-118-3-199302010-00003.

259. Singh N. Antiviral drugs for cytomegalovirus in transplant recipients: advantages of preemptive therapy. *Rev Med Virol*. 2006;16(5):281-287. doi:10.1002/rmv.513.

260. Boeckh M, Ljungman P. How I treat How I treat cytomegalovirus in hematopoietic cell transplant recipients. *Blood*. 2009;113:5711-5719. doi:10.1182/blood.

261. Emery VC, Griffiths PD. Prediction of cytomegalovirus load and resistance patterns after antiviral chemotherapy. *Proc Natl Acad Sci U S A*. 2000;97(14):8039-8044. doi:10.1073/pnas.140123497.

262. de la Cámara R. CMV in Hematopoietic Stem Cell Transplantation. *Mediterr J Hematol Infect Dis*. 2016;8(1):e2016031. doi:10.4084/MJHID.2016.031.

263. Boeckh M, Nichols WG, Chemaly RF, et al. Valganciclovir for the Prevention of Complications of Late Cytomegalovirus Infection After Allogeneic Hematopoietic Cell Transplantation. *Ann Intern Med*. 2015. doi:10.7326/M13-2729.

264. Boeckh M, Ljungman P. How we treat cytomegalovirus in hematopoietic cell transplant recipients. *Blood*. 2009;113(23):5711-5719. doi:10.1182/blood-2008-10-143560.

265. Herling M, Schröder L, Awerkiew S, et al. Persistent CMV infection after allogeneic hematopoietic stem cell transplantation in a CMV-seronegative donor-to-positive recipient constellation: Development of multidrug resistance in the absence of anti-viral cellular immunity. *J Clin Virol*. 2016;74:57-60. doi:10.1016/j.jcv.2015.11.033.

266. Mousavi-Jazi M, Schloss L, Drew WL, et al. Variations in the cytomegalovirus DNA polymerase and phosphotransferase genes in relation to foscarnet and ganciclovir sensitivity. *J Clin Virol*. 2001;23(1-2):1-15. doi:10.1016/S1386-6532(01)00160-3.

267. Chou S, Marousek G, Guentzel S, et al. Evolution of mutations conferring multidrug resistance during prophylaxis and therapy for cytomegalovirus disease. *J Infect Dis*.

1997;176(3):786-789. http://www.ncbi.nlm.nih.gov/pubmed/9291334. Accessed September 12, 2016.

268. Boeckh M, Nichols WG, Papanicolaou G, Rubin R, Wingard JR, Zaia J. Cytomegalovirus in hematopoietic stem cell transplant recipients: Current status, known challenges, and future strategies. *Biol Blood Marrow Transplant*. 2003;9(1083-8791):543-558. doi:10.1016/S1083-8791(03)00287-8.

269. Bowen EF, Cherrington JM, Lamy PD, Griffiths PD, Johnson MA, Emery VC. Quantitative changes in cytomegalovirus DNAemia and genetic analysis of the UL97 and UL54 genes in AIDS patients receiving cidofovir following ganciclovir therapy. *J Med Virol*. 1999;58(4):402-407. doi:10.1002/(SICI)1096-9071(199908)58:4<402::AID-JMV13>3.0.CO;2-5.

270. Drew WL, Liu C. Repopulation of ganciclovir-resistant cytomegalovirus by wild-type virus. *Clin Transplant*. 2012;26(6):949-952. doi:10.1111/j.1399-0012.2012.01681.x.

271. FDA. Advances in CMV management: fomivirsen (Vitravene) approved. *Proj Inf Perspect*. 1998;(26):7. http://www.ncbi.nlm.nih.gov/pubmed/11366499. Accessed September 2, 2016.

272. Anderson KP, Fox MC, Brown-Driver V, Martin MJ, Azad RF. Inhibition of human cytomegalovirus immediate-early gene expression by an antisense oligonucleotide complementary to immediate-early RNA. *Antimicrob Agents Chemother*. 1996;40(9):2004-2011. http://www.ncbi.nlm.nih.gov/pubmed/8878571. Accessed June 1, 2017.

273. Biron KK, Harvey RJ, Chamberlain SC, et al. Potent and selective inhibition of human cytomegalovirus replication by 1263W94, a benzimidazole L-riboside with a unique mode of action. *Antimicrob Agents Chemother*. 2002;46(8):2365-2372. doi:10.1128/aac.46.8.2365-2372.2002.

274. Winston DJ, Young J-AH, Pullarkat V, et al. Maribavir prophylaxis for prevention of cytomegalovirus infection in allogeneic stem cell transplant recipients: a multicenter, randomized, double-blind, placebo-controlled, dose-ranging study. *Blood*. 2008;111(11). http://www.ncbi.nlm.nih.gov/pubmed/18285548. Accessed August 24, 2016.

275. Marty FM, Ljungman P, Papanicolaou GA, et al. Maribavir prophylaxis for prevention of cytomegalovirus disease in recipients of allogeneic stem-cell transplants: a phase 3, double-blind, placebo-controlled, randomised trial. *Lancet Infect Dis*. 2011;11(4):284-292. doi:10.1016/S1473-3099(11)70024-X.

276. Boeckh M, Murphy WJ, Peggs KS. Reprint of: Recent Advances in Cytomegalovirus: An Update on Pharmacologic and Cellular Therapies. *Biol Blood Marrow Transplant*. 2015;21(2):S19-S24. doi:10.1016/j.bbmt.2014.12.034.

277. Lischka P, Hewlett G, Wunberg T, et al. In vitro and in vivo activities of the novel anticytomegalovirus compound AIC246. *Antimicrob Agents Chemother*. 2010;54(3):1290-1297. doi:10.1128/AAC.01596-09.

278. Chemaly RF, Ullmann AJ, Stoelben S, et al. Letermovir for cytomegalovirus prophylaxis in hematopoietic-cell transplantation. *N Engl J Med*. 2014;370(19):1781-1789. doi:10.1056/NEJMoa1309533.

279. Buerger I, Reefschlaeger J, Bender W, et al. A novel nonnucleoside inhibitor specifically targets cytomegalovirus DNA maturation via the UL89 and UL56 gene products. *J Virol*. 2001;75(19):9077-9086. doi:10.1128/JVI.75.19.9077-9086.2001.

280. Schleiss MR, Bernstein DI, McVoy MA, et al. The non-nucleoside antiviral, BAY 38-4766, protects against cytomegalovirus (CMV) disease and mortality in immunocompromised guinea pigs. *Antiviral Res*. 2005;65(1):35-43. doi:10.1016/j.antiviral.2004.09.004.

281. Underwood MR, Harvey RJ, Stanat SC, et al. Inhibition of human cytomegalovirus DNA maturation by a benzimidazole ribonucleoside is mediated through the UL89 gene product. *J Virol*. 1998;72(1):717-725. http://www.ncbi.nlm.nih.gov/pubmed/9420278. Accessed July 11, 2017.

282. Underwood MR, Ferris RG, Selleseth DW, et al. Mechanism of action of the ribopyranoside benzimidazole GW275175X against human cytomegalovirus. *Antimicrob Agents Chemother*. 2004;48(5):1647-1651. doi:10.1128/AAC.48.5.1647-1651.2004.

283. Marty FM, Winston DJ, Rowley SD, et al. CMX001 to prevent cytomegalovirus disease in hematopoietic-cell transplantation. *N Engl J Med*. 2013;369(13):1227-1236. doi:10.1056/NEJMoa1303688.

284. Houldcroft CJ, Bryant JM, Depledge DP, et al. Detection of low frequency multi-drug resistance and novel putative maribavir resistance in immunocompromised pediatric patients with cytomegalovirus. *Front Microbiol*. 2016;7(SEP):1-11. doi:10.3389/fmicb.2016.01317.

285. Schubert A, Ehlert K, Schuler-Luettmann S, Gentner E, Mertens T, Michel D. Fast selection of maribavir resistant cytomegalovirus in a bone marrow transplant recipient. *BMC Infect Dis*. 2013;13(1):330. doi:10.1186/1471-2334-13-330.

286. Schnepf N, Corvo J, Pors MJS Le, Mazeron MC. Antiviral activity of ganciclovir and artesunate towards human cytomegalovirus in astrocytoma cells. *Antiviral Res*. 2011;89(2):186-188. doi:10.1016/j.antiviral.2010.12.002.

287. Efferth T, Marschall M, Wang X, et al. Antiviral activity of artesunate towards wild-type, recombinant, and ganciclovir-resistant human cytomegaloviruses. *J Mol Med*. 2002;80(4):233-242. doi:10.1007/s00109-001-0300-8.

288. Roy S, He R, Kapoor A, et al. Inhibition of human cytomegalovirus replication by artemisinins: effects mediated through cell cycle modulation. *Antimicrob Agents Chemother*. 2015;59(7):3870-3879. doi:10.1128/AAC.00262-15.

289. Ornaghi S, Davis JN, Gorres KL, Miller G, Paidas MJ, van den Pol AN. Mood stabilizers inhibit cytomegalovirus infection. *Virology*. 2016;499:121-135. doi:10.1016/j.virol.2016.09.012.

290. Mercorelli B, Luganini A, Nannetti G, et al. Drug Repurposing Approach Identifies Inhibitors of the Prototypic Viral Transcription Factor IE2 that Block Human Cytomegalovirus Replication. *Cell Chem Biol*. 2016;23(3):340-351. doi:10.1016/j.chembiol.2015.12.012.

291. Beelontally R, Wilkie GS, Lau B, et al. Identification of compounds with anti-human cytomegalovirus activity that inhibit production of IE2 proteins. *Antiviral Res*. 2017;138:61-67. doi:10.1016/j.antiviral.2016.12.006.

292. Strang BL. RO0504985 is an inhibitor of CMGC kinase proteins and has anti-human cytomegalovirus activity. *Antiviral Res*. 2017. doi:10.1016/j.antiviral.2017.05.004.

293. Spiess K, Jeppesen MG, Malmgaard-Clausen M, et al. Rationally designed chemokine-based toxin targeting the viral G protein-coupled receptor US28 potently inhibits cytomegalovirus infection in vivo. *Proc Natl Acad Sci*. 2015;112(27):201509392. doi:10.1073/pnas.1509392112.

294. Krishna BA, Spiess K, Poole EL, et al. Targeting the latent cytomegalovirus reservoir with an antiviral fusion toxin protein. *Nat Commun*. 2017;8:14321. doi:10.1038/ncomms14321.

295. Yang Z, Reeves M, Ye J, et al. RNase P Ribozymes Inhibit the Replication of Human Cytomegalovirus by Targeting Essential Viral Capsid Proteins. *Viruses*. 2015;7(7):3345-3360. doi:10.3390/v7072775.

296. Riddell SR, Reusser P, Greenberg PD. Cytotoxic T Cells Specific for Cytomegalovirus: A

Potential Therapy for Immunocompromised Patients. *Clin Infect Dis*. 1991;13(Supplement 11):S966-S973. doi:10.1093/clind/13.Supplement_11.S966.

297. Walter EA, Greenberg PD, Gilbert MJ, et al. Reconstitution of Cellular Immunity against Cytomegalovirus in Recipients of Allogeneic Bone Marrow by Transfer of T-Cell Clones from the Donor. *N Engl J Med*. 1995;333(16):1038-1044. doi:10.1056/NEJM199510193331603.

298. Riddell SR, Greenberg PD. Principles for Adoptne T Cell. *Annu Rev Immunol*. 1995;13:545-586.

299. Einsele H, Roosnek E, Rufer N, et al. Infusion of cytomegalovirus (CMV)–specific T cells for the treatment of CMV infection not responding to antiviral chemotherapy. *Blood*. 2002;99(11). http://www.bloodjournal.org.gate2.inist.fr/content/bloodjournal/99/11/3916.full.pdf. Accessed June 20, 2017.

300. Wills MR, Carmichael AJ, Mynard K, et al. The Human Cytotoxic T-Lymphocyte (CTL) Response to Cytomegalovirus Is Dominated by Structural Protein pp65: Frequency, Specificity, and T-Cell Receptor Usage of pp65-Specific CTL. *J Virol*. 1996;70(11):7569-7579. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC190825/pdf/707569.pdf. Accessed June 15, 2017.

301. Albiero E, Amati E, Baumeister E, Einsele H, Grigoleit GU, Rodeghiero F. Heterogeneity of Specific CD4 + and CD8 + T Cells Stimulated by CMV pp65 and IE1 Antigens. *J Immunother*. 2016;39(9). https://insights-ovid-com.gate2.inist.fr/pubmed?pmid=27631893. Accessed June 20, 2017.

302. Blyth E, Clancy L, Simms R, et al. Donor-derived CMV-specific T cells reduce the requirement for CMV-directed pharmacotherapy after allogeneic stem cell transplantation. *Blood*. 2013;121(18). doi:10.1182/blood-2012-08.

303. Gerdemann U, Katari UL, Papadopoulou A, et al. Safety and clinical efficacy of rapidly-generated trivirus-directed T cells as treatment for adenovirus, EBV, and CMV infections after allogeneic hematopoietic stem cell transplant. *Mol Ther*. 2013;21(11):2113-2121. doi:10.1038/mt.2013.151.

304. Muto T, Takeda Y, Tsukamoto S, et al. Successful treatment of cytomegalovirus enteritis after unrelated allogeneic stem cell transplantation by the infusion of *ex?vivo* -expanded CD4 + lymphocytes derived from the recipient's peripheral blood donor cells. *Transpl Infect Dis*. 2016;18(1):93-97. doi:10.1111/tid.12487.

305. Leen AM, Bollard CM, Mendizabal AM, et al. Multicenter study of banked third-party virus-speci fi c T cells to treat severe viral infections after hematopoietic stem cell transplantation Methods. *Blood*. 2013;121(26):5113-5124. doi:10.1182/blood-2013-02-486324.The.

306. Peggs KS, Thomson K, Samuel E, et al. Directly Selected Cytomegalovirus-Reactive Donor T Cells Confer Rapid and Safe Systemic Reconstitution of Virus-Specific Immunity Following Stem Cell Transplantation. *Clin Infect Dis*. 2011;52(1):49-57. doi:10.1093/cid/ciq042.

307. Jedema I, van de Meent M, Pots J, Kester MGD, van der Beek MT, Falkenburg JHF. Successful generation of primary virus-specific and anti-tumor T-cell responses from the naive donor T-cell repertoire is determined by the balance between antigen-specific precursor T cells and regulatory T cells. *Haematologica*. 2011;96(8):1204-1212. doi:10.3324/haematol.2010.039099.

308. Beloki L, Ramírez N, Olavarría E, Samuel ER, Lowdell MW. Manufacturing of highly functional and specific T cells for adoptive immunotherapy against virus from granulocyte colony-stimulating factoremobilized donors. *J Cytotherapy*. 2014;16:1390-1408. doi:10.1016/j.jcyt.2014.05.009.

309. Neuenhahn M, Albrecht J, Odendahl M, et al. Transfer of minimally manipulated CMV-specific T cells from stem cell or third-party donors to treat CMV infection after allo-HSCT. *Leukemia*. January 2017. doi:10.1038/leu.2017.16.

310. Hanley PJ, Bollard CM. Controlling cytomegalovirus: helping the immune system take the lead. *Viruses*. 2014;6(6):2242-2258. doi:10.3390/v6062242.

311. Pfirrmann V, Oelsner S, Rettinger E, et al. Cytomegalovirus-specific cytokine-induced killer cells: concurrent targeting of leukemia and cytomegalovirus. *J Cytotherapy*. 2015;17:1139-1151. doi:10.1016/j.jcyt.2015.04.011.

312. Khairallah C, Netzer S, Villacreces A, et al. γδ T cells confer protection against murine cytomegalovirus (MCMV). *PLoS Pathog*. 2015;11(3):e1004702. doi:10.1371/journal.ppat.1004702.

313. Jain NA, Lu K, Ito S, et al. The clinical and financial burden of pre-emptive management of cytomegalovirus disease after allogeneic stem cell transplantation-implications for preventative treatment approaches. *Cytotherapy*. 2014;16(7):927-933. doi:10.1016/j.jcyt.2014.02.010.

314. Schleiss MR. Cytomegalovirus vaccines under clinical development. *J virus Erad*. 2016;2(4):198-207. http://www.ncbi.nlm.nih.gov/pubmed/27781101. Accessed July 24, 2017.

315. Das B, Kaur G, Basu S. Seroprevalence of cytomegalovirus antibodies among blood donors and Multitransfused recipients – A study from north India. *Transfus Apher Sci*. 2014;50(3):438-442. doi:10.1016/j.transci.2014.02.022.

316. Elek S., Stern H. DEVELOPMENT OF A VACCINE AGAINST MENTAL RETARDATION CAUSED BY CYTOMEGALOVIRUS INFECTION IN UTERO. *Lancet*. 1974;303(7845):1-5. doi:10.1016/S0140-6736(74)92997-3.

317. Plotkin SA, Starr SE, Friedman HM, Gonczol E, Weibel RE. Protective effects of Towne cytomegalovirus vaccine against low-passage cytomegalovirus administered as a challenge. *J Infect Dis*. 1989;159(5):860-865. http://www.ncbi.nlm.nih.gov/pubmed/2540247.

318. Plotkin SA, Smiley ML, Friedman HM, et al. Towne-vaccine-induced prevention of cytomegalovirus disease after renal transplants. *Lancet (London, England)*. 1984;1(8376):528-530. http://www.ncbi.nlm.nih.gov/pubmed/6142252. Accessed August 24, 2016.

319. Plotkin SA, Starr SE, Friedman HM, et al. Effect of Towne Live Virus Vaccine on Cytomegalovirus Disease after Renal Transplant. *Ann Intern Med*. 1991;114(7):525. doi:10.7326/0003-4819-114-7-525.

320. Adler SP, Manganello A-M, Lee R, et al. A Phase 1 Study of 4 Live, Recombinant Human Cytomegalovirus Towne/Toledo Chimera Vaccines in Cytomegalovirus?Seronegative Men. *J Infect Dis*. 2016;214(9):1341-1348. doi:10.1093/infdis/jiw365.

321. Pepperl S, Münster J, Mach M, Harris JR, Plachter B. Dense bodies of human cytomegalovirus induce both humoral and cellular immune responses in the absence of viral gene expression. *J Virol*. 2000;74(13):6132-6146. doi:10.1128/JVI.74.13.6132-6146.2000.

322. Pepperl-Klindworth S, Frankenberg N, Plachter B. Development of novel vaccine strategies against human cytomegalovirus infection based on subviral particles. *J Clin Virol*. 2002;25:75-85. http://ac.els-cdn.com.gate2.inist.fr/S1386653202000999/1-s2.0-S1386653202000999-main.pdf?_tid=14eb7002-51cb-11e7-a6a2-00000aab0f26&acdnat=1497532013_cddd341e842a17be0e46d747659b5879. Accessed

June 15, 2017.

323. Kirchmeier M, Fluckiger A-C, Soare C, et al. Enveloped virus-like particle expression of human cytomegalovirus glycoprotein B antigen induces antibodies with potent and broad neutralizing activity. *Clin Vaccine Immunol*. 2014;21(2):174-180. doi:10.1128/CVI.00662-13.

324. Adler SP, Plotkin SA, Gonczol E, et al. A Canarypox Vector Expressing Cytomegalovirus (CMV) Glycoprotein B Primes for Antibody Responses to a Live Attenuated CMV Vaccine (Towne). *J Infect Dis*. 1999;180:843-846. https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/jid/180/3/10.1086/314951/2/180-3-843.pdf?Expires=1497623782&Signature=Y66nrSlGXjDa0jj5tpzoFrFQ1SKR6pepoa9jbYIpoe3OwU~-lTUCQ~6A0UvMAMP~EojhKuGqIIVgKTnazLCu75XANl45Qs4zGb9VPiBHd9Pvnxfk55EhX. Accessed June 15, 2017.

325. Berencsi K, Gyulai Z, Gö Nczö L E, et al. A Canarypox Vector–Expressing Cytomegalovirus (CMV) Phosphoprotein 65 Induces Long-Lasting Cytotoxic T Cell Responses in Human CMV-Seronegative Subjects. *J Infect Dis*. 2001;183:25-27. https://oup.silverchair-cdn.com/oup/backfile/Content_public/Journal/jid/183/8/10.1086/319680/2/183-8-1171.pdf?Expires=1497623804&Signature=OVDF-w8OyNkrFsF3k49cnK3WCaE0RRJRs0hWM5YidGTLWAZ2Yy109Wpwk17RNC-g9ujz9ht6QQsRqHmWoKpvl~3~59aV-XFeu2jcIZm~Pxd5VWmMEW9t. Accessed June 15, 2017.

326. Rosa C La, Longmate J, Martinez J, et al. MVA vaccine encoding CMV antigens safely induces durable expansion of CMV-specific T cells in healthy adults. *Blood*. 2017;129(1):114-125. doi:10.1182/blood-2016-07.

327. Pass RF, Duliegè A-M, Boppana S, et al. A Subunit Cytomegalovirus Vaccine Based on Recombinant Envelope Glycoprotein B and a New Adjuvant. *J Infect Dis*. 1999;180(4):970-975. doi:10.1086/315022.

328. Rosa C La, Wang Z, Brewer JC, et al. Preclinical development of an adjuvant-free peptide vaccine with activity against CMV pp65 in HLA transgenic mice. *Immunobiology*. 2002;100(10). doi:10.1182/blood-2002-03-0926.

329. Nakamura R, La Rosa C, Longmate J, et al. Viraemia, immunogenicity, and survival outcomes of cytomegalovirus chimeric epitope vaccine supplemented with PF03512676 (CMVPepVax) in allogeneic haemopoietic stem-cell transplantation: randomised phase 1b trial. *Lancet Haematol*. 2016;3(2):e87-98. doi:10.1016/S2352-3026(15)00246-X.

330. SCHLEISS MR, BOURNE N, JENSEN NJ, BRAVO F, BERNSTEIN DI. Immunogenicity Evaluation of DNA Vaccines That Target Guinea Pig Cytomegalovirus Proteins Glycoprotein B and UL83. *Viral Immunol*. 2000;13(2):155-167. doi:10.1089/vim.2000.13.155.

331. Kharfan-Dabaja MA, Boeckh M, Wilck MB, et al. A novel therapeutic cytomegalovirus DNA vaccine in allogeneic haemopoietic stem-cell transplantation: a randomised, double-blind, placebo-controlled, phase 2 trial. *Lancet Infect Dis*. 2012;12(4):290-299. doi:10.1016/S1473-3099(11)70344-9.

332. Schleiss MR, Stroup G, Pogorzelski K, Mcgregor A. Protection against congenital cytomegalovirus (CMV) disease, conferred by a replication-disabled, bacterial artificial chromosome (BAC)-based DNA vaccine. *Vaccine*. 2006;24:6175-6186. doi:10.1016/j.vaccine.2006.06.077.

333. Turitz Cox DB, Platt RJ, Zhang F. Therapeutic genome editing : prospects and challenges. *Nat Med*. 2015;21(2):4-6.

334. Cuculis L, Schroeder CM. A Single-Molecule View of Genome Editing Proteins: Biophysical

Mechanisms for TALEs and CRISPR/Cas9. *Annu Rev Chem Biomol Eng*. 2017;8(1):annurev-chembioeng-060816-101603. doi:10.1146/annurev-chembioeng-060816-101603.

335. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*. 2012;337(6096):816-821. doi:10.1126/science.1225829.

336. Marraffini LA, Sontheimer EJ. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat Rev Genet*. 2011;11(3):181-190. doi:10.1038/nrg2749.CRISPR.

337. Ishino Y, Shinagawa H, Makino K, Amemura M, Nakata A. Nucleotide Sequence of the iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in Escherichia coli, and Identification of the Gene Product. *J Bacteriol*. 1987;169(12):5429-5433. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC213968/pdf/jbacter00202-0107.pdf. Accessed July 31, 2017.

338. Jansen R, Embden JDA van, Gaastra W, Schouls LM. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*. 2002;43(6):1565-1575. doi:10.1046/j.1365-2958.2002.02839.x.

339. van der Oost J, Westra ER, Jackson RN, Wiedenheft B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat Rev Microbiol*. 2014;12(7):479-492. doi:10.1038/nrmicro3279.

340. Bolotin A, Quinquis B, Sorokin A, Dusko S, Correspondence E. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *microbiology*. 2005;151:2551-2561. doi:10.1099/mic.0.28048-0.

341. Haft DH, Selengut J, Mongodin EF, Nelson KE. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput Biol*. 2005;1(6):e60. doi:10.1371/journal.pcbi.0010060.

342. Horvath P, Barrangou R. CRISPR/Cas, the immune system of bacteria and archaea. *Science*. 2010;327(5962):167-170. doi:10.1126/science.1179555.

343. Barrangou R, Fremaux C, Deveau H, et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science*. 2007;315(5819):1709-1712. doi:10.1126/science.1138140.

344. Mojica FJM, Díez-Villaseñor C, García-Martínez J, Almendros C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*. 2009;155(3):733-740. doi:10.1099/mic.0.023960-0.

345. Jiang F, Doudna JA. CRISPR – Cas9 Structures and Mechanisms. *AnnuRevBiophys*. 2017;46(March):505-529. doi:DOI: 10.1146/annurev-biophys-062215-010822.

346. Cong L, Ran FA, Cox D, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science*. 2013;339(6121):819-823. doi:10.1126/science.1231143.

347. Ran FA, Hsu PD, Lin C-Y, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell*. 2013;154(6):1380-1389. doi:10.1016/j.cell.2013.08.021.

348. Kabadi a. M, Ousterout DG, Hilton IB, Gersbach C a. Multiplex CRISPR/Cas9-based genome engineering from a single lentiviral vector. *Nucleic Acids Res*. 2014;42(19):gku749-. doi:10.1093/nar/gku749.

349. Gorski SA, Vogel J, Doudna JA. RNA-based recognition and targeting: sowing the seeds of specificity. *Nat Rev Mol Cell Biol*. 2017;18(4):215-228. doi:10.1038/nrm.2016.174.

350. Chen JS, Dagdas YS, Kleinstiver BP, et al. Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*. 2017;550(7676):407-410. doi:10.1038/nature24268.

351. Rodgers K, McVey M. Error-Prone Repair of DNA Double-Strand Breaks. *J Cell Physiol*.

2016;231(1):15-24. doi:10.1002/jcp.25053.

352. Wang H, Xu X. Microhomology-mediated end joining: new players join the team. *Cell Biosci*. 2017;7(1):6. doi:10.1186/s13578-017-0136-8.

353. Seol J-H, Shim EY, Lee SE. Microhomology-mediated end joining: Good, bad and ugly. *Mutat Res Mol Mech Mutagen*. 2017. doi:10.1016/j.mrfmmm.2017.07.002.

354. Ran F, Hsu P, Wright J, Agarwala V. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc*. 2013;8(11):2281-2308. doi:10.1038/nprot.2013.143.

355. Kim JM, Kim D, Kim S, Kim J-S. Genotyping with CRISPR-Cas-derived RNA-guided endonucleases. *Nat Commun*. 2014;5:1-7. doi:10.1038/ncomms4157.

356. Brinkman EK, Chen T, Amendola M, Van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res*. 2014;42(22):1-8. doi:10.1093/nar/gku936.

357. Ding Q, Regan SN, Xia Y, Oostrom LA, Cowan CA, Musunuru K. Enhanced efficiency of human pluripotent stem cell genome editing through replacing TALENs with CRISPRs. *Cell Stem Cell*. 2013;12(4):393-394. doi:10.1016/j.stem.2013.03.006.

358. Kim S, Kim D, Cho SW, Kim J, Kim J-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Res*. 2014;24(6):1012-1019. doi:10.1101/gr.171322.113.

359. Daer R, Cutts JP, Brafman DA, Haynes KA. The impact of chromatin dynamics on Cas9-mediated genome editing in human cells. *ACS Synth Biol*. October 2016. doi:10.1021/acssynbio.5b00299.

360. He Z, Proudfoot C, Mileham AJ, McLaren DG, Whitelaw CBA, Lillico SG. Highly efficient targeted chromosome deletions using CRISPR/Cas9. *Biotechnol Bioeng*. 2015;112(5):1060-1064. doi:10.1002/bit.25490.

361. Jasin M, Haber JE. The democratization of gene editing: Insights from site-specific cleavage and double-strand break repair. *DNA Repair (Amst)*. 2016;44:6-16. doi:10.1016/j.dnarep.2016.05.001.

362. Richardson CD, Ray GJ, Dewitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. *Nat Biotechnol*. 2016;34(3). doi:10.1038/nbt.3481.

363. Koo T, Lee J, Kim J. Measuring and Reducing Off-Target Activities of Programmable Nucleases Including CRISPR-Cas9. *Mol Cells*. 2015;38(7).

364. Zischewski J, Fischer R, Bortesi L. Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. *Biotechnol Adv*. 2017;35(1):95-104. doi:10.1016/j.biotechadv.2016.12.003.

365. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. 2014;31(9):827-832. doi:10.1038/nbt.2647.DNA.

366. Lin Y, Cradick TJ, Brown MT, et al. CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Res*. 2014;42(11):7473-7485. doi:10.1093/nar/gku402.

367. Wu X, Kriz AJ, Sharp P a. Target specificity of the CRISPR-Cas9 system. *Quant Biol*. 2014;2(2):59-70. doi:10.1007/s40484-014-0030-x.

368. Singh R, Kuscu C, Quinlan A, Qi Y, Adli M. Cas9-chromatin binding information enables more accurate CRISPR off-target prediction. *Nucleic Acids Res*. 2015:1-8. doi:10.1093/nar/gkv575.

369. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 2014;32(3):279-284. doi:10.1038/nbt.2808.

370. Tycko J, Myer VE, Hsu PD. Methods for Optimizing CRISPR-Cas9 Genome Editing Specificity. *Mol Cell*. 2016;63(3):355-370. doi:10.1016/j.molcel.2016.07.004.

371. Dang Y, Jia G, Choi J, et al. Optimizing sgRNA structure to improve CRISPR-Cas9 knockout efficiency. *Genome Biol*. 2015;16:280. doi:10.1186/s13059-015-0846-3.

372. Hendel A, Bak RO, Clark JT, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat Biotechnol*. 2015;(June). doi:10.1038/nbt.3290.

373. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. *Science (80- )*. 2015;351(6268):84-88. doi:10.1126/science.aad5227.

374. Kleinstiver BP, Pattanayak V, Prew MS, et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 2016;529(7587):490-495. doi:10.1038/nature16526.

375. Kleinstiver BP, Prew MS, Tsai SQ, et al. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 2015;523(7561):481-485. doi:10.1038/nature14592.

376. Chen X, Xu F, Zhu C, et al. Dual sgRNA-directed gene knockout using CRISPR/Cas9 technology in Caenorhabditis elegans. *Sci Rep*. 2014;4:7581. doi:10.1038/srep07581.

377. Cho SW, Kim S, Kim Y, et al. Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res*. 2014;24(1):132-141. doi:10.1101/gr.162339.113.

378. Tsai SQ, Wyvekens N, Khayter C, et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat Biotechnol*. 2014;32(6):569-576. doi:10.1038/nbt.2908.

379. Wyvekens N, Topkar V V, Khayter C, Joung JK, Tsai SQ. Dimeric CRISPR RNA-Guided FokI-dCas9 Nucleases Directed by Truncated gRNAs for Highly Specific Genome Editing. *Hum Gene Ther*. 2015;26(7):425-431. doi:10.1089/hum.2015.084.

380. Davis KM, Pattanayak V, Thompson DB, Zuris JA, Liu DR. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat Chem Biol*. 2015;11(5):316-318. doi:10.1038/nchembio.1793.

381. Zetsche B, Volz SE, Zhang F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat Biotechnol*. 2015;33(2):139-142. doi:10.1038/nbt.3149.

382. Tsai SQ, Zheng Z, Nguyen NT, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat Biotechnol*. 2015;33(2):187-197. doi:10.1038/nbt.3117.

383. Tsai SQ, Nguyen NT, Malagon-Lopez J, Topkar V V, Aryee MJ, Joung JK. CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. *Nat Methods*. 2017;14(6):607-614. doi:10.1038/nmeth.4278.

384. Li L, He Z-Y, Wei X-W, Gao G-P, Wei Y-Q. Challenges in CRISPR/CAS9 Delivery: Potential Roles of Nonviral Vectors. *Hum Gene Ther*. 2015;26(7):452-462. doi:10.1089/hum.2015.069.

385. He Z-Y, Men K, Qin Z, Yang Y, Xu T, Wei Y-Q. Non-viral and viral delivery systems for CRISPR-Cas9 technology in the biomedical field. *Sci China Life Sci*. May 2017. doi:10.1007/s11427-017-9033-0.

386. Wang D, Mou H, Li S, et al. Adenovirus-Mediated Somatic Genome Editing of Pten by CRISPR

/ Cas9 in Mouse Liver in Spite of Cas9-Specific Immune Responses. 2015;26(7):432-442. doi:10.1089/hum.2015.087.

387. Chew WL, Tabebordbar M, Cheng JKW, et al. A multifunctional AAV – CRISPR – Cas9 and its host response. 2016;(September). doi:10.1038/nmeth.3993.

388. Song M. The CRISPR/Cas9 system: Their delivery, *in vivo* and *ex vivo* applications and clinical development by startups. *Biotechnol Prog*. May 2017. doi:10.1002/btpr.2484.

389. Zuris JA, Thompson DB, Shu Y, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing in vitro and in vivo. *Nat Publ Gr*. 2014;(October):1-10. doi:10.1038/nbt.3081.

390. Mout R, Ray M, Lee YW, Scaletti F, Rotello VM. In Vivo Delivery of CRISPR/Cas9 for Therapeutic Gene Editing: Progress and Challenges. *Bioconjug Chem*. 2017;28(4):880-884. doi:10.1021/acs.bioconjchem.7b00057.

391. Zhang X, Li B, Luo X, et al. Biodegradable Amino-ester Nanomaterials for Cas9 mRNA Delivery in Vitro and in Vivo. *ACS Appl Mater Interfaces*. 2017:acsami.7b08163. doi:10.1021/acsami.7b08163.

392. Rádis-Baptista G, Campelo IS, Morlighem JÉRL, Melo LM, Freitas VJF. Cell-penetrating peptides (CPPs): From delivery of nucleic acids and antigens to transduction of engineered nucleases for application in transgenesis. *J Biotechnol*. 2017;252:15-26. doi:10.1016/j.jbiotec.2017.05.002.

393. Staahl BT, Benekareddy M, Coulon-Bainier C, et al. Efficient genome editing in the mouse brain by local delivery of engineered Cas9 ribonucleoprotein complexes. *Nat Biotechnol*. 2017;35(5):431-434. doi:10.1038/nbt.3806.

394. Liang X, Potter J, Kumar S, Ravinder N, Chesnut JD. Enhanced CRISPR/Cas9-mediated precise genome editing by improved design and delivery of gRNA, Cas9 nuclease, and donor DNA. *J Biotechnol*. 2017;241:136-146. doi:10.1016/j.jbiotec.2016.11.011.

395. Ma Y, Han X, Quintana Bustamante O, et al. Highly efficient genome editing of human hematopoietic stem cells via a nano-silicon-blade delivery approach. *Integr Biol*. 2017. doi:10.1039/C7IB00060J.

396. Tabebordbar M, Zhu K, Cheng JKW, et al. In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science (80- )*. 2015;5177(DECEMBER):1-9. doi:10.1126/science.aad5177.

397. Yin H, Song C-Q, Dorkin JR, et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components in vivo. *Nat Biotechnol*. 2016;34(3):328-333. doi:10.1038/nbt.3471.

398. Gong H, Liu M, Klomp J, Merrill BJ, Rehman J, Malik AB. Method for Dual Viral Vector Mediated CRISPR-Cas9 Gene Disruption in Primary Human Endothelial Cells. *Sci Rep*. 2017;7:42127. doi:10.1038/srep42127.

399. Ran FA, Cong L, Yan WX, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015;520(7546):186-191. doi:10.1038/nature14299.

400. Buchholz CJ, Friedel T, Büning H. Surface-Engineered Viral Vectors for Selective and Cell Type-Specific Gene Delivery. *Trends Biotechnol*. 2015;33(12):777-790. doi:10.1016/j.tibtech.2015.09.008.

401. Ou W, Marino MP, Suzuki A, et al. Specific targeting of human interleukin (IL)-13 receptor α2-positive cells with lentiviral vectors displaying IL-13. *Hum Gene Ther Methods*. 2012;23(2):137-147. doi:10.1089/hgtb.2012.054.

402. Kays S-K, Kaufmann KB, Abel T, et al. CD105 is a surface marker for receptor-targeted gene

transfer into human long-term repopulating hematopoietic stem cells. *Stem Cells Dev*. 2015;24(6):714-723. doi:10.1089/scd.2014.0455.

403. Brendel C, Goebel B, Daniela A, et al. CD133-targeted gene transfer into long-term repopulating hematopoietic stem cells. *Mol Ther*. 2015;23(1):63-70. doi:10.1038/mt.2014.173.

404. Witting SR, Vallanda P, Gamble AL. Characterization of a third generation lentiviral vector pseudotyped with Nipah virus envelope proteins for endothelial cell transduction. *Gene Ther*. 2013;20(10):997-1005. doi:10.1038/gt.2013.23.

405. Ortinski PI, O 'donovan B, Dong X, Kantor B. Integrase-Deficient Lentiviral Vector as an All-in-One Platform for Highly Efficient CRISPR/Cas9-Mediated Gene Editing. 2017. doi:10.1016/j.omtm.2017.04.002.

406. Petris G, Casini A, Montagna C, et al. Hit and go CAS9 delivered through a lentiviral based self-limiting circuit. *Nat Commun*. 2017;8:15334. doi:10.1038/ncomms15334.

407. Hindriksen S, Bramer AJ, Truong MA, et al. Baculoviral delivery of CRISPR/Cas9 facilitates efficient genome editing in human cells. *PLoS One*. 2017;12(6):e0179514. doi:10.1371/journal.pone.0179514.

408. Park A, Hong P, Won ST, et al. Sendai virus, an RNA virus with no risk of genomic integration, delivers CRISPR/Cas9 for efficient gene editing. *Mol Ther Methods Clin Dev*. 2016;3:16057. doi:10.1038/mtm.2016.57.

409. Nissim L, Perli SD, Fridkin A, Perez-Pinera P, Lu TK. Multiplexed and Programmable Regulation of Gene Networks with an Integrated RNA and CRISPR/Cas Toolkit in Human Cells. *Mol Cell*. 2014;54(4):698-710. doi:10.1016/j.molcel.2014.04.022.

410. Mefferd AL, Kornepati AVR, Bogerd HP, Kennedy EM, Cullen BR. Expression of CRISPR/Cas single guide RNAs using small tRNA promoters. *RNA*. 2015;21(9):1683-1689. doi:10.1261/rna.051631.115.

411. Yoshioka S, Fujii W, Ogawa T, Sugiura K, Naito K. Development of a mono-promoter-driven CRISPR/Cas9 system in mammalian cells. *Sci Rep*. 2015;5(December):18341. doi:10.1038/srep18341.

412. Xu T, Li Y, Van Nostrand JD, He Z, Zhou J. Cas9-based tools for targeted genome editing and transcriptional control. *Appl Environ Microbiol*. 2014;80(5):1544-1552. doi:10.1128/AEM.03786-13.

413. Unniyampurath U, Pilankatta R, Krishnan MN. RNA Interference in the Age of CRISPR: Will CRISPR Interfere with RNAi? *Int J Mol Sci*. 2016;17(3):291. doi:10.3390/ijms17030291.

414. Thakore PI, Ippolito AMD, Song L, et al. Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. *Nat Methods*. 2015;(october):1-9. doi:10.1038/nmeth.3630.

415. Lei Y, Zhang X, Su J, et al. Targeted DNA methylation in vivo using an engineered dCas9-MQ1 fusion protein. *Nat Commun*. 2017;8:16026. doi:10.1038/ncomms16026.

416. Xiong T, Meister GE, Workman RE, et al. Targeted DNA methylation in human cells using engineered dCas9-methyltransferases. *Sci Rep*. 2017;7(1):6732. doi:10.1038/s41598-017-06757-0.

417. Balboa D, Weltner J, Eurola S, Trokovic R, Wartiovaara K, Otonkoski T. Conditionally Stabilized dCas9 Activator for Controlling Gene Expression in Human Cell Reprogramming and Differentiation. *Stem cell reports*. 2015;5(3):448-459. doi:10.1016/j.stemcr.2015.08.001.

418. Mali P, Aach J, Stranges PB, et al. CAS9 transcriptional activators for target specificity

screening and paired nickases for cooperative genome engineering. *Nat Biotechnol*. 2013;31(9):833-838. doi:10.1038/nbt.2675.

419. Wang H-X, Li M, Lee CM, et al. CRISPR/Cas9-Based Genome Editing for Disease Modeling and Therapy: Challenges and Opportunities for Nonviral Delivery. *Chem Rev*. 2017:acs.chemrev.6b00799. doi:10.1021/acs.chemrev.6b00799.

420. Chen B, Gilbert L a, Cimini B a, et al. Dynamic Imaging of Genomic Loci in Living Human Cells by an Optimized CRISPR / Cas System. *Cell*. 2013;155(7):1479-1491. doi:10.1016/j.cell.2013.12.001.

421. Anton T, Bultmann S, Leonhardt H, Markaki Y. Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. *Nucleus*. 2014;5(2):163-172. doi:10.4161/nucl.28488.

422. Deng W, Shi X, Tjian R, Lionnet T, Singer RH. CASFISH: CRISPR/Cas9-mediated in situ labeling of genomic loci in fixed cells. *Proc Natl Acad Sci U S A*. 2015;112(38):11870-11875. doi:10.1073/pnas.1515692112.

423. Ran FA, Cong L, Yan WX, et al. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 2015. doi:10.1038/nature14299.

424. Friedland AE, Baral R, Singhal P, et al. Characterization of Staphylococcus aureus Cas9: a smaller Cas9 for all-in-one adeno-associated virus delivery and paired nickase applications. *Genome Biol*. 2015;16(1):257. doi:10.1186/s13059-015-0817-8.

425. Kleinstiver BP, Prew MS, Tsai SQ, et al. Broadening the targeting range of Staphylococcus aureus CRISPR-Cas9 by modifying PAM recognition. *Nat Biotechnol*. 2015;33(12):1293-1298. doi:10.1038/nbt.3404.

426. Xu K, Ren C, Liu Z, et al. Efficient genome engineering in eukaryotes using Cas9 from Streptococcus thermophilus. *Cell Mol Life Sci*. 2015;72(2):383-399. doi:10.1007/s00018-014-1679-z.

427. Müller M, Lee CM, Gasiunas G, et al. Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol Ther*. 2016;24(3):636-644. doi:10.1038/mt.2015.218.

428. Hou Z, Zhang Y, Propson NE, et al. Efficient genome engineering in human pluripotent stem cells using Cas9 from Neisseria meningitidis. *Proc Natl Acad Sci U S A*. 2013;110(39):15644-15649. doi:10.1073/pnas.1313587110.

429. Cebrian-Serrano A, Davies B. CRISPR-Cas orthologues and variants: optimizing the repertoire, specificity and delivery of genome engineering tools. *Mamm Genome*. 2017;0(0):1-15. doi:10.1007/s00335-017-9697-4.

430. Kim E, Koo T, Park SW, et al. In vivo genome editing with a small Cas9 orthologue derived from Campylobacter jejuni. *Nat Commun*. 2017;8:1-12. doi:10.1038/ncomms14500.

431. Zetsche B, Gootenberg JS, Abudayyeh OO, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell*. 2015;163(3):759-771. doi:10.1016/j.cell.2015.09.038.

432. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*. 2016:1-19. doi:10.1038/nature17945.

433. Zhong G, Wang H, Li Y, Tran MH, Farzan M. Cpf1 proteins excise CRISPR RNAs from mRNA transcripts in mammalian cells. *Nat Chem Biol*. 2017;(June):1-5. doi:10.1038/nchembio.2410.

434. Men K, Duan X, He Z, Yang Y, Yao S, Wei Y. CRISPR / Cas9-mediated correction of human genetic disease. *Sci China Life Sci*. 2017;9. doi:10.1007/s11427-017-9032-4.

435. Yang S, Chang R, Yang H, et al. CRISPR/Cas9-mediated gene editing ameliorates neurotoxicity in mouse model of Huntington's disease. *Proc Natl Acad Sci U S A*. 2017;104(43):17204-17209. doi:10.1172/JCI92087.

436. Cyranoski D. CRISPR gene-editing tested in a person for the first time. *Nature*. November 2016. doi:10.1038/nature.2016.20988.

437. Xu L, Zhao L, Gao Y, Xu J, Han R. Empower multiplex cell and tissue-specific CRISPR-mediated gene manipulation with self-cleaving ribozymes and tRNA. *Nucleic Acids Res*. 2017;45(5). doi:10.1093/nar/gkw1048.

438. Hu X. CRISPR/Cas9 System and Its Applications in Human Hematopoietic Cells. *Blood Cells, Mol Dis*. 2016. doi:10.1016/j.bcmd.2016.09.003.

439. Yue J, Gou X, Li Y, Wicksteed B, Wu X. Engineered Epidermal Progenitor Cells Can Correct Diet-Induced Obesity and Diabetes. *Cell Stem Cell*. 2017;21(2):256-263.e4. doi:10.1016/j.stem.2017.06.016.

440. Li HL, Fujimoto N, Sasakawa N, et al. Precise correction of the dystrophin gene in duchenne muscular dystrophy patient induced pluripotent stem cells by TALEN and CRISPR-Cas9. *Stem Cell Reports*. 2015;4(1):143-154. doi:10.1016/j.stemcr.2014.10.013.

441. Huang Z, Tomitaka A, Raymond A, Nair M. Current application of CRISPR/Cas9 gene-editing technique to eradication of HIV/AIDS. *Gene Ther*. 2017. doi:10.1038/gt.2017.35.

442. Soppe JA, Lebbink RJ. Antiviral Goes Viral: Harnessing CRISPR/Cas9 to Combat Viruses in Humans. *Trends Microbiol*. 2017. doi:10.1016/j.tim.2017.04.005.

443. Hütter G, Bodor J, Ledger S, et al. CCR5 Targeted Cell Therapy for HIV and Prevention of Viral Escape. *Viruses*. 2015;7(8):4186-4203. doi:10.3390/v7082816.

444. Ebina H, Misawa N, Kanemura Y, Koyanagi Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. *Sci Rep*. 2013;3:2510. doi:10.1038/srep02510.

445. Khalili K, Gordon J, Kaminski R, Cosentino L, Hu W. Genome editing strategies: potential tools for eradicating HIV-1/AIDS. *J Neurovirol*. 2015:310-321. doi:10.1007/s13365-014-0308-9.

446. Hu W, Kaminski R, Yang F, et al. RNA-directed gene editing specifically eradicates latent and prevents new HIV-1 infection. *Proc Natl Acad Sci*. 2014;111(31):11461-11466. doi:10.1073/pnas.1405186111.

447. Zhu W, Lei R, Le Duff Y, et al. The CRISPR/Cas9 system inactivates latent HIV-1 proviral DNA. *Retrovirology*. 2015;12:22. doi:10.1186/s12977-015-0150-z.

448. Kaminski R, Bella R, Yin C, et al. Excision of HIV-1 DNA by gene editing: a proof-of-concept in vivo study. *Gene Ther*. 2016;(April):1-6. doi:10.1038/gt.2016.41.

449. Yin C, Zhang T, Qu X, et al. In Vivo Excision of HIV-1 Provirus by saCas9 and Multiplex Single-Guide RNAs in Animal Models. *Mol Ther*. 2017;25(5). doi:10.1016/j.ymthe.2017.03.012.

450. Dampier W, Nonnemacher MR, Sullivan NT, Jacobson JM, Wigdahl B. HIV Excision Utilizing CRISPR/Cas9 Technology: Attacking the Proviral Quasispecies in Reservoirs to Achieve a Cure. *MOJ Immunol*. 2014;1(4). doi:10.15406/moji.2014.01.00022.

451. Zhang Y, Yin C, Zhang T, et al. CRISPR/gRNA-directed synergistic activation mediator (SAM) induces specific, persistent and robust reactivation of the HIV-1 latent reservoirs. *Sci Rep*. 2015;5:16277. doi:10.1038/srep16277.

452. Saayman SM, Lazar DC, Scott TA, et al. Potent and targeted activation of latent HIV-1 using the CRISPR/dCas9 activator complex. *Mol Ther*. 2016;24(August):488-498. doi:10.1038/mt.2015.202.

453. Limsirichai P, Gaj T, Schaffer D V. CRISPR-mediated activation of latent HIV-1 expression. *Mol Ther*. 2015;(November). doi:10.1038/mt.2015.213.

454. Liao H-K, Gu Y, Diaz A, et al. Use of the CRISPR/Cas9 system as an intracellular defense against HIV-1 infection in human cells. *Nat Commun*. 2015;6:6413. doi:10.1038/ncomms7413.

455. Li C, Guan X, Du T, et al. Inhibition of HIV-1 infection of primary CD4+ T cells by gene editing of CCR5 using adenovirus-delivered CRISPR/Cas9. *J Gen Virol*. April 2015. doi:10.1099/vir.0.000139.

456. Wang W, Ye C, Liu J, Zhang D, Kimata JT, Zhou P. CCR5 gene disruption via lentiviral vectors expressing Cas9 and single guided RNA renders cells resistant to HIV-1 infection. *PLoS One*. 2014;9(12):e115987. doi:10.1371/journal.pone.0115987.

457. Kang H, Minder P, Park MA, Mesquitta W-T, Torbett BE, Slukvin II. CCR5 Disruption in Induced Pluripotent Stem Cells Using CRISPR/Cas9 Provides Selective Resistance of Immune Cells to CCR5-tropic HIV-1 Virus. *Mol Ther Acids*. 2015;4(12):e268. doi:10.1038/mtna.2015.42.

458. Mandal PK, Ferreira MR, Rossi DJ, et al. Efficient Ablation of Genes in Human Hematopoietic Stem and Effector Cells using CRISPR/Cas9. *Cell Stem Cell*. 2014;15(5):643-652. doi:10.1016/j.stem.2014.10.004.

459. Xu L, Yang H, Gao Y, et al. CRISPR/Cas9-Mediated CCR5 Ablation in Human Hematopoietic Stem/Progenitor Cells Confers HIV-1 Resistance In Vivo. *Mol Ther*. 2017;25(8). doi:10.1016/j.ymthe.2017.04.027.

460. Wang J, Xu Z-W, Liu S, et al. Dual gRNAs guided CRISPR/Cas9 system inhibits hepatitis B virus replication. *World J Gastroenterol*. 2015;21(32):9554. doi:10.3748/wjg.v21.i32.9554.

461. Karimova M, Beschorner N, Dammermann W, et al. CRISPR/Cas9 nickase-mediated disruption of hepatitis B virus open reading frame S and X. *Sci Rep*. 2015;5:13734. doi:10.1038/srep13734.

462. Sakuma T, Masaki K, Abe-Chayama H, Mochida K, Yamamoto T, Chayama K. Highly multiplexed CRISPR-Cas9-nuclease and Cas9-nickase vectors for inactivation of hepatitis B virus. *Genes to Cells*. 2016;21(11):1253-1262. doi:10.1111/gtc.12437.

463. Kurihara T, Fukuhara T, Ono C, et al. Suppression of HBV replication by the expression of nickase- and nuclease dead-Cas9. *Sci Rep*. 2017;7(1):6122. doi:10.1038/s41598-017-05905-w.

464. Kennedy EM, Bassit LC, Mueller H, et al. Suppression of hepatitis B virus DNA accumulation in chronically infected cells using a bacterial CRISPR/Cas RNA-guided DNA endonuclease. *Virology*. 2015;476:196-205. doi:10.1016/j.virol.2014.12.001.

465. Li H, Sheng C, Wang S, et al. Removal of Integrated Hepatitis B Virus DNA Using CRISPR-Cas9. *Front Cell Infect Microbiol*. 2017;7:91. doi:10.3389/fcimb.2017.00091.

466. Liu X, Hao R, Chen S, Guo D, Chen Y. Inhibition of Hepatitis B Virus by CRISPR/Cas9 System via Targeting the Conserved Regions of Viral Genome. *J Gen Virol*. April 2015. doi:10.1099/vir.0.000159.

467. Ramanan V, Shlomai A, Cox DBT, et al. CRISPR/Cas9 cleavage of viral DNA efficiently suppresses hepatitis B virus. *Sci Rep*. 2015;5(April):10833. doi:10.1038/srep10833.

468. Dong C, Qu L, Wang H, Wei L, Dong Y, Xiong S. Targeting hepatitis B virus cccDNA by CRISPR / Cas9 nuclease efficiently inhibits viral replication. 2015;(April):1-8.

469. Zhen S, Hua L, Liu Y-H, et al. Harnessing the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR-associated Cas9 system to disrupt the hepatitis B

virus. *Gene Ther*. 2015;(July 2014):1-9. doi:10.1038/gt.2015.2.

470. Lin S-R, Yang H-C, Kuo Y-T, et al. The CRISPR/Cas9 System Facilitates Clearance of the Intrahepatic HBV Templates In Vivo. *Mol Ther Nucleic Acids*. 2014;3:e186. doi:10.1038/mtna.2014.38.

471. Scott T, Moyo B, Nicholson S, et al. ssAAVs containing cassettes encoding SaCas9 and guides targeting hepatitis B virus inactivate replication of the virus in cultured cells. *Sci Rep*. 2017;7(1):7401. doi:10.1038/s41598-017-07642-6.

472. Schiffer JT, Swan DA, Stone D, Jerome KR. Predictors of hepatitis B cure using gene therapy to deliver DNA cleavage enzymes: a mathematical modeling approach. *PLoS Comput Biol*. 2013;9(7):e1003131. doi:10.1371/journal.pcbi.1003131.

473. Wang J, Quake SR. RNA-guided endonuclease provides a therapeutic strategy to cure latent herpesviridae infection. *Proc Natl Acad Sci*. 2014;111(36):13157-13162. doi:10.1073/pnas.1410785111.

474. van Diemen FR, Kruse EM, Hooykaas MJG, et al. CRISPR/Cas9-Mediated Genome Editing of Herpesviruses Limits Productive and Latent Infections. *PLoS Pathog*. 2016;12(6):e1005701. doi:10.1371/journal.ppat.1005701.

475. Hu Z, Yu L, Zhu D, et al. Disruption of HPV16-E7 by CRISPR/Cas system induces apoptosis and growth inhibition in HPV16 positive human cervical cancer cells. *Biomed Res Int*. 2014;2014:612823. doi:10.1155/2014/612823.

476. Kennedy EM, Kornepati AVR, Goldstein M, et al. Inactivation of the human papillomavirus E6 or E7 gene in cervical carcinoma cells by using a bacterial CRISPR/Cas RNA-guided endonuclease. *J Virol*. 2014;88(20):11965-11972. doi:10.1128/JVI.01879-14.

477. Yu L, Wang X, Zhu D, et al. Disruption of human papillomavirus 16 E6 gene by clustered regularly interspaced short palindromic repeat/Cas system in human cervical cancer cells. *Onco Targets Ther*. 2015;8:37-44. doi:10.2147/OTT.S64092.

478. Zhen S, Hua L, Takahashi Y, Narita S, Liu YH, Li Y. In vitro and in vivo growth suppression of human papillomavirus 16-positive cervical cancer cells by CRISPR/Cas9. *Biochem Biophys Res Commun*. 2014;450(4):1422-1426. doi:10.1016/j.bbrc.2014.07.014.

479. Zhen S, Lu J-J, Wang L-J, et al. In Vitro and In Vivo Synergistic Therapeutic Effect of Cisplatin with Human Papillomavirus16 E6/E7 CRISPR/Cas9 on Cervical Cancer Cell Line. *Transl Oncol*. 2016;9:498-504. doi:10.1016/j.tranon.2016.10.002.

480. O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA. Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature*. 2014;516(7530):263+. doi:10.1038/nature13769.

481. Price A a., Sampson TR, Ratner HK, Grakoui A, Weiss DS. Cas9-mediated targeting of viral RNA in eukaryotic cells. *Proc Natl Acad Sci*. 2015;9:201422340. doi:10.1073/pnas.1422340112.

482. Wang G, Zhao N, Berkhout B, Das AT. CRISPR-Cas9 can inhibit HIV-1 replication but NHEJ repair facilitates virus escape. *Mol Ther*. 2016;24(3):522-526. doi:10.1038/mt.2016.24.

483. Wang Z, Pan Q, Gendron P, et al. CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and Accelerate Viral Escape. *Cell Rep*. 2016. doi:10.1016/j.celrep.2016.03.042.

484. Yoder KE, Bundschuh R. Host Double Strand Break Repair Generates HIV-1 Strains Resistant to CRISPR/Cas9. *Sci Rep*. 2016;6:29530. doi:10.1038/srep29530.

485. Wang G, Zhao N, Berkhout B, Das AT. *A Combinatorial CRISPR-Cas9 Attack on HIV-1 DNA Extinguishes All Infectious Provirus in Infected T Cell Cultures*. Vol 17. 2016. doi:10.1016/j.celrep.2016.11.057.

486. Lebbink RJ, de Jong DCM, Wolters F, et al. A combinational CRISPR/Cas9 gene-editing approach can halt HIV replication and prevent viral escape. *Sci Rep*. 2017;7:41968. doi:10.1038/srep41968.

487. Yuen K-S, Chan C-P, Wong N-HM, et al. CRISPR/Cas9-mediated genome editing of Epstein-Barr virus in human cells. *J Gen Virol*. 2015;96(Pt 3):626-636. doi:10.1099/jgv.0.000012.

488. Angulo A, Kerry D, Huang H, et al. Identification of a boundary domain adjacent to the potent human cytomegalovirus enhancer that represses transcription of the divergent UL127 promoter. *J Virol*. 2000;74(6):2826-2839. doi:10.1128/JVI.74.6.2826-2839.2000.

489. Bellec J, Bacchetta M, Losa D, Anegon I, Chanson M, Nguyen T. *CFTR* Inactivation by Lentiviral Vector-mediated RNA Interference and CRISPR-Cas9 Genome Editing in Human Airway Epithelial Cells. *Curr Gene Ther*. 2015;15(5):447-459. doi:10.2174/1566523215666150812115939.

490. Canver MC, Bauer DE, Dass A, et al. Characterization of Genomic Deletion Efficiency Mediated by Clustered Regularly Interspaced Palindromic Repeats (CRISPR)/Cas9 Nuclease System in Mammalian Cells * . *J Biol Chem*. 2014;289(31):21312-21324. doi:10.1074/jbc.M114.564625.

491. Hsu PD, Scott DA, Weinstein JA, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol*. 2013;31(9):827-832. doi:10.1038/nbt.2647.

492. Chenouard V, Brusselle L, Heslan J-M, et al. A Rapid and Cost-Effective Method for Genotyping Genome-Edited Animals: A Heteroduplex Mobility Assay Using Microfluidic Capillary Electrophoresis. *J Genet Genomics*. 2016;43(5):341-348. doi:10.1016/j.jgg.2016.04.005.

493. Genini E, Gerna G, Labò N, et al. In vitro selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells. *J Gen Virol*. 2001;82(6):1429-1438. doi:10.1099/0022-1317-82-6-1429.

494. Glass M, Busche A, Wagner K, Messerle M, Borst EM. Conditional and reversible disruption of essential herpesvirus proteins. *Nat Methods*. 2009;6(8):577-579. doi:10.1038/nmeth.1346.

495. Wroblewska Z, Wellish MC, Wolinsky JS, Gilden D. Comparison of human cytomegalovirus growth in MRC-5 human fibroblasts, brain, and choroid plexus cells in vitro. *J Med Virol*. 1981;8(4):245-256. http://www.ncbi.nlm.nih.gov/pubmed/6278069. Accessed July 23, 2017.

496. Stamminger T, Puchtler E, Fleckenstein B. Discordant expression of the immediate-early 1 and 2 gene regions of human cytomegalovirus at early times after infection involves posttranscriptional processing events. *J Virol*. 1991;65(5):2273-2282. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=240576&tool=pmcentrez&rendertype=abstract.

497. Cartier N, Hacein-Bey-Abina S, Bartholomae CC, et al. Hematopoietic stem cell gene therapy with a lentiviral vector in X-linked adrenoleukodystrophy. *Science*. 2009;326(5954):818-823. doi:10.1126/science.1171242.

498. Aiuti A, Biasco L, Scaramuzza S, et al. Lentiviral hematopoietic stem cell gene therapy in patients with Wiskott-Aldrich syndrome. *Science*. 2013;341(6148):1233151. doi:10.1126/science.1233151.

499. Biron KK, Harvey RJ, Chamberlain SC, et al. Potent and Selective Inhibition of Human Cytomegalovirus Replication by 1263W94, a Benzimidazole L-Riboside with a Unique Mode of Action. *Antimicrob Agents Chemother*. 2002;46(8):2365-2372.

doi:10.1128/AAC.46.8.2365-2372.2002.

500. Hamilton ST, Milbradt J, Marschall M, Rawlinson WD. Human cytomegalovirus replication is strictly inhibited by siRNAs targeting UL54, UL97 or UL122/123 gene transcripts. *PLoS One*. 2014;9(6):e97231. doi:10.1371/journal.pone.0097231.

501. Gustems M, Borst E, Benedict CA, et al. Regulation of the transcription and replication cycle of human cytomegalovirus is insensitive to genetic elimination of the cognate NF-kappaB binding sites in the enhancer. *J Virol*. 2006;80(19):9899-9904. doi:10.1128/JVI.00640-06.

502. Slobedman B, Mocarski ES. Quantitative Analysis of Latent Human Cytomegalovirus. 1999;73(6):4806-4812.

503. Humby MS, O'Connor CM. HCMV US28 is important for latent infection of hematopoietic progenitor cells. *J Virol*. 2015;(December):JVI.02507-15. doi:10.1128/JVI.02507-15.

504. Rossetto CCC, Tarrant-Elorza M, Pari GGS, et al. Cis and Trans Acting Factors Involved in Human Cytomegalovirus Experimental and Natural Latent Infection of CD14 (+) Monocytes and CD34 (+) Cells. Mocarski E, ed. *PLoS Pathog*. 2013;9(5):e1003366. doi:10.1371/journal.ppat.1003366.

505. Albright ER, Kalejta RF. Myeloblastic cell lines mimic some but not all aspects of human cytomegalovirus experimental latency defined in primary CD34+ cell populations. *J Virol*. 2013;87(17):9802-9812. doi:10.1128/JVI.01436-13.

506. O'Connor CM, Murphy EA. A myeloid progenitor cell line capable of supporting human cytomegalovirus latency and reactivation, resulting in infectious progeny. *J Virol*. 2012;86(18):9854-9865. doi:10.1128/JVI.01278-12.

507. Pan C, Zhu D, Wang Y, et al. Human Cytomegalovirus miR-UL148D Facilitates Latent Viral Infection by Targeting Host Cell Immediate Early Response Gene 5. Flemington EK, ed. *PLOS Pathog*. 2016;12(11):e1006007. doi:10.1371/journal.ppat.1006007.

508. Ziehr B, Vincent HA, Moorman NJ. Human Cytomegalovirus pTRS1 and pIRS1 Antagonize Protein Kinase R To Facilitate Virus Replication. *J Virol*. 2016;90(8):3839-3848. doi:10.1128/JVI.02714-15.

509. Marshall EE, Bierle CJ, Brune W, Geballe AP. Essential role for either TRS1 or IRS1 in human cytomegalovirus replication. *J Virol*. 2009;83(9):4112-4120. doi:10.1128/JVI.02489-08.

510. Song Y, Yuan L, Wang Y, et al. Efficient dual sgRNA-directed large gene deletion in rabbit with CRISPR/Cas9 system. *Cell Mol Life Sci*. 2016;73(15):2959-2968. doi:10.1007/s00018-016-2143-z.

511. Boeckh M, Adam P. Geballe. Cytomegalovirus: pathogen, paradigm, and puzzel. *J Clin Invest*. 2011;121(5). doi:10.1172/JCI45449.The.

512. Wahid B, Usman S, Ali A, et al. Therapeutic Strategies of Clustered Regularly Interspaced Palindromic Repeats-Cas Systems for Different Viral Infections. *Viral Immunol*. 2017;30(8):1-8. doi:10.1089/vim.2017.0055.

513. van Diemen FR, Lebbink RJ. CRISPR/Cas9, a powerful tool to target human herpesviruses. *Cell Microbiol*. 2017;19(2):e12694. doi:10.1111/cmi.12694.

514. Li HW, Sykes M. Emerging concepts in haematopoietic cell transplantation. *Nat Rev Immunol*. 2012;12(6):403-416. doi:10.1038/nri3226.

515. Nayak D, Roth TL, Mcgavern DB. How do I perform hematopoietic progenitor cell selection? 2014;56(2):367-402. doi:10.1146/annurev-immunol-032713-120240.Microglia.

516. Huenecke S, Bremm M, Cappel C, et al. Optimization of individualized graft composition: CD3/CD19 depletion combined with CD34 selection for haploidentical transplantation. *Transfusion*. 2016;56(9):2336-2345. doi:10.1111/trf.13694.

517. Lang P, Schumm M, Greil J, et al. A comparison between three graft manipulation methods for haploidentical stem cell transplantation in pediatric patients: preliminary results of a pilot study. *Klin Padiatr*. 2005;217(6):334-338. doi:10.1055/s-2005-872529.

518. Kobbe G, Fenk R, Neumann F, et al. Transplantation of allogeneic CD34+-selected cells followed by early T-cell add-backs: Favorable results in acute and chronic myeloid leukemia. *Cytotherapy*. 2004;6(6):533-542. doi:10.1080/14653240410005375.

519. Bornhäuser M, Platzbecker U, Theuser C, Hölig K, Ehninger G. CD34+-enriched peripheral blood progenitor cells from unrelated donors for allografting of adult patients: High risk of graft failure, infection and relapse despite donor lymphocyte add-back. *Br J Haematol*. 2002;118(4):1095-1103. doi:10.1046/j.1365-2141.2002.03731.x.

520. Ferrá C, Rodríguez-Luaces M, Gallardo D, et al. Individually adjusted prophylactic donor lymphocyte infusions after CD34-selected allogeneic peripheral blood stem cell transplantation. *Bone Marrow Transplant*. 2001;28(10):963-968. doi:10.1038/sj.bmt.1703277.

521. Girard-Gagnepain A, Amirache F, Costa C, et al. Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-stimulated and resting HSCs. *Blood*. 2014;124(8):1221-1231. doi:10.1182/blood-2014-02-558163.

522. Spiess K, Jeppesen MG, Malmgaard-Clausen M, et al. Rationally designed chemokine-based toxin targeting the viral G protein-coupled receptor US28 potently inhibits cytomegalovirus infection in vivo. *Proc Natl Acad Sci*. 2015;112(27):201509392. doi:10.1073/pnas.1509392112.

523. Spiess K, Jeppesen MG, Malmgaard-Clausen M, et al. Rationally designed chemokine-based toxin targeting the viral G protein-coupled receptor US28 potently inhibits cytomegalovirus infection in vivo. *Proc Natl Acad Sci U S A*. 2015;112(27):8427-8432. doi:10.1073/pnas.1509392112.

524. Chakrbarti A, Jaiswal SR, Chakrabarti S. Sirolimus as Long-term Graft-versus-host-disease Prophylaxis in Haploidentical Hematopoietic Stem Cell Transplant Recipients for Non-malignant Disorders is Associated with High Incidence of Acneiform Lesions. *Indian J Dermatol*. 2015;60(6):588-591. doi:10.4103/0019-5154.169131.

525. Wilson CH, Gamper I, Perfetto A, Auw J, Littlewood TD, Evan GI. The kinetics of ER fusion protein activation in vivo. *Oncogene*. 2014;33(40):4877-4880. doi:10.1038/onc.2014.78.

526. Goverdhana S, Puntel M, Xiong W, et al. Regulatable gene expression systems for gene therapy applications: progress and future challenges. *Mol Ther*. 2005;12(2):189-211. doi:10.1016/j.ymthe.2005.03.022.

527. Anjum K, Shagufta BI, Abbas SQ, et al. Current status and future therapeutic perspectives of glioblastoma multiforme (GBM) therapy: A review. *Biomed Pharmacother*. 2017;92:681-689. doi:10.1016/j.biopha.2017.05.125.

528. Lucas KG, Lei @bullet, @bullet B, Bruggeman R, Dunham K, Specht C. The detection of CMV pp65 and IE1 in glioblastoma multiforme. doi:10.1007/s11060-010-0383-6.

529. Halbert CL, Rutledge EA, Allen JM, Russell DW, Miller AD. Repeat transduction in the mouse lung by using adeno-associated virus vectors with different serotypes. *J Virol*. 2000;74(3):1524-1532. doi:10.1128/JVI.74.3.1524-1532.2000.

530. Alton EWFW, Armstrong DK, Ashby D, et al. Repeated nebulisation of non-viral CFTR gene therapy in patients with cystic fibrosis: a randomised, double-blind, placebo-controlled, phase 2b trial. *Lancet Respir Med*. 2015;3(9):684-691. doi:10.1016/S2213-2600(15)00245-3.

531. McLachlan G, Davidson H, Holder E, et al. Pre-clinical evaluation of three non-viral gene

transfer agents for cystic fibrosis after aerosol delivery to the ovine lung. *Gene Ther*. 2011;18(10):996-1005. doi:10.1038/gt.2011.55.

532. Mitomo K, Griesenbach U, Inoue M, et al. Toward gene therapy for cystic fibrosis using a lentivirus pseudotyped with Sendai virus envelopes. *Mol Ther*. 2010;18(6):1173-1182. doi:10.1038/mt.2010.13.

533. Alton EWFW, Beekman JM, Boyd AC, et al. Preparation for a first-in-man lentivirus trial in patients with cystic fibrosis. *Thorax*. 2017;72(2):137-147. doi:10.1136/thoraxjnl-2016-208406.

534. Rauwel B, Jang SM, Cassano M, Kapopoulou A, Barde I, Trono D. Release of human cytomegalovirus from latency by a KAP1 / TRIM28 phosphorylation switch. *Elife*. 2015:1-20. doi:10.7554/eLife.06068.

535. Smith MS, Goldman DC, Bailey AS, et al. Granulocyte-colony stimulating factor reactivates human cytomegalovirus in a latently infected humanized mouse model. *Cell Host Microbe*. 2010;8(3):284-291. doi:10.1016/j.chom.2010.08.001.

536. Hakki M, Goldman DC, Streblow DN, et al. HCMV infection of humanized mice after transplantation of G-CSF-mobilized peripheral blood stem cells from HCMV-seropositive donors. *Biol Blood Marrow Transplant*. 2014;20(1):132-135. doi:10.1016/j.bbmt.2013.10.019.

537. Kern ER. Pivotal role of animal models in the development of new therapies for cytomegalovirus infections. *Antiviral Res*. 2006;71(2-3):164-171. doi:10.1016/j.antiviral.2006.05.018.

538. Itell HL, Kaur A, Deere JD, Barry PA, Permar SR. Rhesus monkeys for a nonhuman primate model of cytomegalovirus infections. *Curr Opin Virol*. 2017;25:126-133. doi:10.1016/j.coviro.2017.08.005.

539. Lo DJ, Anderson DJ, Weaver TA, et al. Belatacept and sirolimus prolong nonhuman primate renal allograft survival without a requirement for memory T cell depletion. *Am J Transplant*. 2013;13(2):320-328. doi:10.1111/j.1600-6143.2012.04342.x.

540. Arvin A, Campadelli-Fiume G, Mocarski E, et al. *Human Herpesviruses*. Cambridge University Press; 2007. http://www.ncbi.nlm.nih.gov/pubmed/21348071. Accessed February 15, 2017.

541. Modrow S, Falke D, Truyen U, Schätzl H. *Molekulare Virologie*. Heidelberg: Spektrum Akademischer Verlag; 2010. doi:10.1007/978-3-8274-2241-5.

542. Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. *Nat Methods*. 2013;10(10):957-963. doi:10.1038/nmeth.2649.

543. Shao Y, Guan Y, Wang L, et al. CRISPR/Cas-mediated genome editing in the rat via direct injection of one-cell embryos. *Nat Protoc*. 2014;9(10):2493-2512. doi:10.1038/nprot.2014.171.

# X Appendix

## 1 Publication

Multiplex CRISPR/Cas9 system impairs HCMV replication by excising an essential viral gene

Janina Gergen[1,2], Flora Coulon[1,2], Alison Creneguy[1,2], Nathan Elain-Duret[1,2], Alejandra Gutierrez[3], Olaf Pinkenburg[4], Els Verhoeyen [3,5], Ignacio Anegon[1,2], Tuan Huy Nguyen[1,2], Franck Albert Halary[1,2*] and Fabienne Haspot [1,2*]

[1] Centre de Recherche en Transplantation et Immunologie UMR 1064, INSERM, Université de Nantes, Nantes, France
[2] Institut de Transplantation Urologie Néphrologie (ITUN), CHU Nantes, Nantes, France
[3] CIRI – International Center for Infectiology Research, Team EVIR, Inserm, U1111, Université Claude Bernard Lyon 1, CNRS, UMR5308, Ecole Normale Supérieure de Lyon, Université Lyon, F-69007, Lyon, France.
[4] Institut für medizinische Mikrobiologie, und Krankenhaushygiene, Universität Marburg, Hans-Meerwein-Strasse 2, 35043 Marburg, Germany
[5] Inserm, U1065, Centre Méditerranéen de Médecine Moléculaire (C3M), équipe "contrôle métabolique des morts cellulaires", Nice, 06204, France
*co-senior authors

Correspondence should be addressed to F.H. (fabienne.haspot@univ-nantes.fr)

Corresponding author address:
Dr Fabienne Haspot
30 Boulevard Jean Monnet
44093 Nantes Cedex 01
France.
Phone 0033(0)240087410
Fax 0033(0)240087411
fabienne.haspot@univ-nantes.fr

Short Title: Multiplexed CRISPR/Cas9 against HCMV

Abstract

Anti-HCMV treatments used in immunosuppressed patients, reduce viral replication but resistant viral strains can emerge. Moreover, latently infected cells are not targeted by the drugs. We designed two anti-viral CRISPR/Cas9 strategies to target the *UL122/123* gene, the key regulator of lytic replication and reactivation from latency. The *singleplex* strategy contains one gRNA to target the start codon, the *multiplex* strategy contains three gRNA to excise the complete *UL122/123* gene. Primary fibroblast and U373-MG cells were transduced with a lentiviral vectors coding for the Cas9 and one or three gRNAs. Both strategies showed mutations in the target gene and concomitant reduction of IE protein expression in primary fibroblasts. Further detailed analysis in U373-MG cells showed, that the *singleplex* strategy induced 50% *indels* in the viral genome leading to a reduction of the IE protein expression. The *multiplex* strategy excised the IE gene in 90% of all viral genomes and thus led to the inhibition of the IE protein expression. Consequently, viral genome replication and late protein expression was reduced by 90%. Finally, the production of new viral particle was nearly abrogated. In conclusion, a *multiplex* anti-*UL122/123* CRISPR/Cas9 system can target the viral genome efficiently enough to significantly prevent viral replication.

Introduction:

The human Cytomegalovirus (HCMV) primary infection or reactivation can cause severe pathologies in non-immunocompetent individuals(1),(2). In hematopoietic stem cell transplantation (HSCT), HCMV active replication is the major source of transplant-related morbidity and mortality. Up to one third of the patients with HCMV reactivation develop a CMV disease with possible end organ diseases(3),(4). The currently available treatments(5), which target the viral DNA polymerase, are based on nucleotide analogues (Ganciclovir(6) and Cidofovir(7)) and on a non-competitive inhibitor (Foscarnet(8)). The occurrence of Ganciclovir or Foscarnet-resistant viral strains(9),(10) urge the development of innovative strategies. Moreover, the current treatments just target the lytic replicating virus but have no impact on the latent viral pool, thus preventing any complete clearance of the virus.

The CRISPR/Cas9 system is an easy, fast and highly potent genome-editing tool. Originally found as an adaptive immune system in bacteria and archaea against phages and plasmids(11), it is now adapted for the use in eukaryotic cells to a suitable two-component system consisting of a Cas9 endonuclease and a chimeric guide RNA (gRNA)(12). The CRISPR/Cas9 system has also been proposed to be used as an anti-viral strategy to fight latent or chronic viral infections(13–18).

In this study, we hypothesized that disrupting the *UL122/123* gene with a CRISPR/Cas9 system based on one or three gRNAs will prevent viral replication. The *UL122/123* gene encodes several immediate early molecules (IE)(19), which are the first and most essential proteins responsible for the initiation of the viral replication cycle(20),(21). Indeed, the mutation or shutdown of the *UL122/123* gene leads to a non-replicative virus(22,23). We induced site-specific cleavage in the *UL122/123* gene by using one gRNA, while we deleted 3300 bp in the *UL122/123* gene from the viral genome by using a *multiplex* strategy with three gRNAs. This led to a strong reduction of IE expression and consequently to the inhibition of late viral protein

expression. Overall, the production of new virions was reduced by up to 98% by the *multiplex* anti-IE CRISPR/Cas9 system. This innovative approach could be used to clear HCMV from infected hematopoietic stem cells before their transplantation to a seronegative recipient.

Results:

**Anti-IE CRISPR/Cas9 system reduces IE expression in primary fibroblasts**

In order to prevent HCMV replication, two anti-HCMV CRISPR/Cas9 strategies based on one (singleplex) or three gRNAs (multiplex) were developed to knockout the *UL122/123* gene encoding the major immediate early proteins (Figure 1). Each gRNA position has been chosen on conserved region (Supplementary Tables 1, 2 and 3). Those IE proteins are the most essential key regulators of the viral replication(20,21).

MRC5 primary fibroblast cells were transduced with the LV type 1 (Supplementary Figure 1) containing either one of the anti-HCMV CRISPR/Cas9 system or an unspecific gRNA/Cas9 as an internal control. Cas9 positive cells were selected via puromycin treatment prior to be infected with Toledo (MOI 0.1) to assess the effect of the anti-IE strategies on the viral replication. Two days pi, the viral genome was analyzed for *indels* by T7-endonuclease assay (Figure 2a). The efficiency of the *singleplex* cutting was calculated as described elswhere(24–26) (Table 1). Twenty-nine percent of the viral genome had *indels* at the target site. To analyze the efficiency of the *multiplex* strategy, a PCR spanning exons 2 to 5 was performed, which gave a 3862 bp amplicon. Interestingly, while this WT amplicon was strongly detected in control MRC5s and in MRC5s expressing the unsp. gRNA; *multiplex* MRC5s showed a weaker WT band. Moreover, a weak amplification of 500 bp was also observed (Figure 2b), which probably represents a deletion of 3300 bp in the *UL122/123* gene between the target sites of gRNA1 and gRNA3. When quantifying the weaker band (500 bp) in comparison with the wild type band (3862 bp), around 5% of the viral genome copies showed this deletion in the *UL122/123* gene. In PCR, the amplification of small products are favored in comparison to longer fragments, so that these percentages might not represent the exact quantity of mutations in the viral genome extracted from HCMV-infected MRC5 cells. However, the presence of the small PCR fragments still proofs that a part of the viral genome has a larger deletion in the

*UL122/123* gene.

The effect of those mutations in the *UL122/123* gene on the IE expression was then analyzed by western blot. The *singleplex* and *multiplex* strategies enabled a strong reduction of the IE1 expression and nearly abrogated the IE2 expression (Figure 2c). The unexpected strong decrease of IE protein expression by the *multiplex* strategy is probably due to a combination of large deletions (5%) as analyzed by PCR and *indels* at each target site itself. Furthermore, the cut by the gRNA/Cas9 and the subsequent repair take at least 5 hours(27), which delay the IE expression, even if the gene is correctly repaired.

The analysis of the Cas9 expression by western blot showed, that the *singleplex* and *multiplex* cells contained only a very low amount of Cas9 as compared to unsp. gRNA strategy. Thus, we questioned the stability of the Cas9 expression in transduced MRC5s. As shown in Figure 2(d) and (e), the Cas9 expression level was reduced during consecutive subculture steps. After three passages, the Cas9 expression was decreased by 60% (*singleplex* and *multiplex*) to 87% (unsp. gRNA). This Cas9 expression over time seemed not sufficient to prevent late events of the viral replication.

Here we show that anti-HCMV CRISPR/Cas9 strategies disturb the viral genome at the target site, which results in a strong decrease of IE protein expression.

**Stable expression of the anti-IE CRISPR/Cas9 system induces mutations in the *UL122/123* gene in HCMV-infected U373-MG cells**

Since the Cas9 expression was not stable in MRC5 cells, we decided to perform a more detailed analysis of the efficiency of the *singleplex* and *multiplex* strategies in a HCMV-permissive astrocytoma cell line (U373-MG). These cells support a full lytic replication cycle of HCMV. We designed new LVs expressing a gRNA cassette and a Cas9 fused to the GFP(28) (Supplementary Figure 1). Those three LVs type 2 were used to transduce U373-MG cells at

MOIs varying from five to ten. Subsequently, these transduced U373-MG cells were FACS-sorted based on their Cas9-GFP$^{high}$ expression. All three cell lines had a similar mean of fluorescence (MFI = 1985-2203) for Cas9-GFP expression, which represents a similar expression level of the CRISPR/Cas9 system in the different cell lines (Supplementary Figure 2).

The three transduced U373-MG cell lines were infected with three different low passage HCMV strains: TB40GFP, Toledo and VR1814(29). The HCMV viral genome was extracted eight days pi to analyze mutations induced by the gRNA/Cas9. The efficiency of both, the *singleplex* and *multiplex* strategies, was assessed as described before with the MRC5. The *singleplex* strategy yielded 30–50% *indels* (Figure 3a and b and Table 2) thus slightly higher than in the MRC5. To confirm that the mutations were induced by the *singleplex* strategy, the target region of gRNA1 was analyzed by Sanger sequencing. Small *indels* were detected around the cleavage site of gRNA1 (Figure 3e). For the *multiplex* strategy, a major amplification of 500 bp (Figure 3c and d) representing the deletion of the target region between gRNA1 and gRNA3, and a smear above representing smaller deletions, were detected. The quantification of this band and the smear above in comparison to the WT band revealed that up to 95% of the viral genome was affected by the *multiplex* strategy (Table 2). Importantly, all three viral strains tested were similarly efficiently targeted showing the universal usage of our anti-*IE* gRNAs (Table 2). Overall, the *multiplex* strategy was more efficient than the *singleplex* strategy and showed significant higher yields of mutations in the viral genome.

**Dramatic decrease of IE protein expression in HCMV-infected U373-MG cells expressing gRNA/Cas9**

We analyzed, if the induction of mutations in the IE gene led to a concomitant reduction of IE expression in the different U373-MG cell lines, two and eight days pi, with three different viral

strains. The unsp. gRNA cell line was equally permissive to HCMV infection compared to untransduced control cells for the three HCMV strains tested suggesting that there was no effect of the Cas9/unsp. gRNA on the viral infection (Figure 4a and b). HCMV-infected *singleplex* U373-MG cells showed a reduction of IE positive cells of up to 50 % with TB40GFP or Toledo. The *multiplex* strategy was significantly more efficient than the *singleplex* and reduced the amount of IE positive cells by 75-85% (Figure 4a and b). The endotheliotropic HCMV strain VR1814 could only be used at a low MOI (0.1). In this conditions, the singleplex strategy was already significant to reduce the IE expression by up to 75 % (Figure 4a and b). The decrease of IE positive cells by the *multiplex* strategy reached up to 95% for VR1814. In comparision to the strains, TB40GFP and Toledo, at a low MOI, the effect on IE reduction was also stronger than with a MOI of 1 (Supplementary Figure 3). Subsequent analyses for TB40GFP and Toledo were done with an MOI of 1 to strongly challenge the anti-viral CRISPR/Cas9 system. Comparing the effect of both strategies between day two and day eight pi, it appears that the decrease of IE expression was stable over time when cells were infected with TB40-GFP. The IE expression decreased significantly over time in cells harboring the *multiplex* strategy infected with Toledo (MOI 1).

A western blot analysis was performed to analyze both major IE splice variants (IE1 and IE2). The expression of both IE variants was impaired by the anti-*UL122/123* CRISPR/Cas9 system with a higher effect on IE2 than on IE1. Importantly, the IE2 expression was undetectable with the *multiplex* strategy for Toledo and VR1814 (Figure 3c), which could point out a possible knock-out of IE2. We also analyzed the expression of the Cas9 by western blot and could confirm a stable and comparable expression level of the Cas9 in all three U373-MG cell lines. Mutations induced by the *singleplex* and *multiplex* strategies led to a significant and stable decrease of the number of IE positive cells over time and to almost undetectable levels of the IE2 protein.

**The *multiplex* strategy is superior over the *singleplex* strategy to inhibit the viral genome replication and late protein expression**

IE proteins are transactivators and induce the production of delayed early proteins, essentials for the genome replication, and for the production of structure proteins, needed for the assembly of new virions(22,23). Thus, the disruption of IE expression abrogates the progression of the viral replication cycle. We analyzed eight days pi, the effect of the anti-*IE* CRISPR/Cas9 strategies on the genome replication using a qPCR on the *US8* gene. While the *singleplex* strategy was only effective for VR1814 at a low MOI with a decrease of 80% of viral genome copies, we detected 70 to 90 % less viral genome in U373-MG containing the *multiplex* strategy compared to the untransduced cells for all viral strains (MOI of 1 or 0.1) (Figure 5a).

Furthermore, we analyzed the expression of the viral envelope glycoprotein B (gB), an IE-dependent late viral antigen, by intracellular FACS eight days pi. The untransduced control U373-MG cells infected with Toledo harbored around 6.5% of gB positive cells (Figure 5b). The use of gRNA1 alone only slightly decreased the percentage of gB positive cells, while the *multiplex* strategy nearly abrogated the gB expression (Figure 5b and c). The expression level of gB for TB40GFP (MOI 1) and VR1814 (MOI 0.1) was not high enough to be detected by FACS analysis.

Overall the progression of the viral replication cycle was dramatically impaired by the *multiplex* strategy, as shown by a strong reduction of the genome replication and by the decreased expression of the late envelope glycoprotein B.

**The anti-HCMV *multiplex* strategy strongly impairs virion release from U373-MG cells**

The *multiplex* anti-*IE* CRISPR/Cas9 system efficiently decreased gB expression and genome replication. To assessed the production of infectious viral particles, we established a trans-

infection plaque assay based on direct cell-to-cell transfer of the virus, because U373-MG cells only poorly secrete HCMV particles in the extracellular space(30). Control U373-MG cells and the unsp. gRNA cell line, infected with TB40GFP reached a trans-infection plaque-titer of around 2500 plaques/10^5 cells (Figure 6a). *Singleplex* U373-MG cells released 32 % less infectious virions (1600 plaques/10^5 cells). Importantly, targeting the *UL122/123* gene with the *multiplex* strategy decreased virion release by 80 % in average (Figure 6a) (436 plaques/10^5 cells). In comparison, Toledo was produced in much higher amounts by the control and unsp. gRNA cells (7287 plaques/10^5 cells) (Figure 6b). While the *singleplex* U373-MG cells released 67 % less infectious virions (4925 plaques/10^5 cells), *multiplex* U373-MG cells showed a remarkably 98 % inhibition of virion release (156 plaques/10^5 cells, Figure 6b).

In conclusion, the *multiplex* anti-HCMV CRISPR/Cas9 system strongly inhibits the production of infectious viral particles and efficiently prevents viral spreading *in vitro.*

**Discussion**

HCMV is a widely spread infection in the human population and can cause severe end organ diseases in immunosuppressed patients such as solid-organ or HSC transplanted patients. Treatments exist, but are only efficient on replicative virus and have no effect on the latent virus pool. Here, we proposed an antiviral strategy expected to be applicable for both lytic and latent viral infections. Using CRISPR/Cas9, we excised several exons from the essential *UL122/123* gene and further blocked IE-dependent steps of the viral replication life cycle.

We designed two CRISPR/Cas9 strategies to knock-out the *UL122/123* viral gene based on one or three gRNAs. The *UL122/UL123* gene encodes the Immediate Early molecules, IE1 and IE2, which are the first molecules expressed during the replication cycle and are essentials for the end of latency. Whereas IE2 is known to be essential for the viral replication(20,21),(23)

and is expressed at first in the lytic replication cycle(31); IE1 is more responsible for transcriptional activation of immediate early and delay early promoters by the inhibition of HDACs (32),(33) and is only essential for infections with a very low MOI(34–36). Furthermore, the IE molecules are necessary for the initiation of replication from latency and a splice variant of IE1 is essential for the viral genome maintenance during latency(37). The destruction of the *UL122/123* gene would therefore not only be efficient to inhibit lytic replicating virus, but would also prevent reactivation form latency and persistence in the host cell. Since IE molecules also influence the host cells in terms of cell cycle regulation and cytokine release(38,39), the inhibition of the expression of those molecule would therefore protect the cells from those side effects. Moreover, Formivirsen, an approved anti-CMV-retinitis drug, is based on an antisense oligonucleotide targeting the *UL122/123* gene and is efficiently blocking local HCMV replication(40). Consequently, the *UL122/123* gene is a suitable target for an anti-HCMV CRISPR/Cas9 system. Targeting the HCMV genome by the CRISPR/Cas9 system has been already investigated by Van Diemen *et al.*(18). Several gRNAs targeting delay early genes, which are involved in the viral genome replication, were tested as *singleplex* approach and achieved short time inhibition of viral replication. Here, we improved the HCMV-targeting by choosing an earlier target gene (immediate early) and three gRNAs on the same gene to block viral protein expression and to prevent the further steps of replication cycle.

In this study, we challenged cells, pretreated with the *singleplex* or the *multiplex* strategy, with three different HCMV viral strains. When cells were infected at a low MOI (0.1), the singleplex strategy was efficient at decreasing the expression of IE molecules. Interestingly, the analysis of the splice variant expression showed a possible knock-out of IE2 and a very strong decrease of IE1. Both variant are essential at this MOI, which explains the subsequent inhibition of the replication cycle. As expected, the reduction of IE molecules by the *singleplex* strategy for the viral strains Toledo and TB40GFP at a higher MOI (1) is not sufficient to

prevent viral replication. The expression of a low amount of IE2 under this conditions, was probably sufficient to start the replication cycle and could be boosted by the positive autoregulation of IE1 as already proposed by others(35). Van Diemen and colleagues also used a singleplex strategy anti-HCMV to target delay early genes with different efficacies at impairing the viral replication, even when a very low MOI (0.05) was used(18). Furthermore, simultaneously targeting of the viral genome with several gRNAs completely abolishes the viral cycle as shown so far for HSV-1, HIV and HBV(18,41,42). In line with this, we confirmed that targeting HCMV with a multiplex strategy was more efficient than a single gRNA. It abrogated IE expression at low and high MOI and led to a subsequent blockage of the viral replication cycle. Importantly, our *multiplex* strategy was effective on the three viral strains tested thus opening perspectives for its use for clinical application.

The use of RNAs-guided endonuclease offers advantages over the actual standard treatment for HCMV infections, which are Ganciclovir and Foscarnet. They block the productive infection of HCMV by targeting the viral polymerase UL54(43). This improves the health of patients facing HCMV diseases significantly, but side effects such as nephrotoxicity and myelosuppression are essential problems for the patient. Developing CRISPR/Cas9 strategies targeting the viral genome with low/no homology to the human genome should be less toxic and have no proven myelosuppressive effects.(15) Moreover, nowadays several high-fidelity Cas9s(44,45) have been shown to significantly reduce off-targets. As previously mentioned, controlling IE proteins expression was already used in clinic for CMV retinitis in HIV-1 patients before the development of highly active anti-retroviral therapy. The limit of such a strategy involving oligonucleotides is that the effect is only transient and usually does not completely inhibit the protein expression. As shown by Hamilton and colleagues(46), the knock-down of HCMV by siRNA targeting *UL122/123* mRNAs reduces viral replication and virion release. However, the application of siRNA is very transient and would not prevent

HCMV replication over a longer time course. In contrast, the mutations or deletions induced by the CRISPR/Cas9 system are permanent and can provide a long-term protection, if all viral genome copies are efficiently targeted.

Drug-resistances to Ganciclovir and Foscarnet are due to mutations in the kinase UL97 or polymerase UL54 genes(43),(47). Escape mutations against the antiviral-CRISPR/Cas9 *singleplex* system has been shown before on HIV(48,49), and on HCMV(18). For example, the proposed anti-HCMV CRISPR/Cas9 systems by Van Diemen *et al.* using one gRNA targeting essential viral genes involved in viral genome replication gave rise to viral escape mutations. Those viral genomes harbored in-frame mutations, after being targeted by the anti-viral gRNA/Cas9.The probability of an escape mutation would be significantly lower with a *multiplex* strategy, because the cut within several targets leads nearly always to the deletion of parts or the complete targeted region and not just to small *indels*. Moreover, two studies on HIV have also shown that the *duplex* strategy can prevent escape mutations and viral breakthrough replication(41,50). They have proven that the combination of several gRNAs diminishes the probability of in-frame mutations and that a longer exposure to Cas9/gRNAs increases the frequency of bigger deletions in the viral genome. Furthermore, a more extensive *multiplex* strategy was successfully used against EBV, whose genome of 170 kbp could be destroyed entirely in Raji cells by the use of seven gRNAs simultaneously(13). Our multiplex strategy is expected to prevent the viral escape as described for the other viruses. It induced mainly large deletions (80-95%) after an exposure of only eight days. Moreover, to reach this goal, high Cas9 expression is needed to target all copies of the viral genome before the expression of the IE molecules, which occurs as early as three hours after HMCV infection (31,51). It has been shown by Richardson *et al*, that the Cas9 stays for around 5.5 h attached to DNA after cleavage and therefore is not available to cut further target sites(27). During lytic replication, the viral genome copy number increases rapidly

exponential and it might not be possible for the Cas9/gRNA complex to target all copies. During natural latency in mononuclear cells from G-CSF mobilized blood or bone marrow no more than 13 viral genome copies are present per cell (52). As a consequence, low Cas9/gRNA expression is expected to target all viral genome copies in a manageable exposure time.

In conclusion, we gave the proof-of-concept that targeting the *UL122/123* gene of HCMV genome with a *multiplex* strategy is efficient to affect the viral genome and to inhibit the virion release up to 98%. In this study, we showed that even a *singleplex* strategy is efficient at inhibiting viral replication, if a low MOI is used. The *multiplex* strategy is superior over a single gRNA at low and high MOIs. Thus, these results pave the way to the development of a promising new therapeutic strategy that could be applicable to treat a hematopoietic stem cell suspension. Challenges for such a pre-emptive CRISPR/Cas9 therapy involves an optimized *ex vivo* delivery system, the selection of the targeted cells and the use of high-fidelity Cas9(44,45).

Methods:

*gRNA design:*

The three gRNA were designed to target the *UL122/123* gene close to the start codon and at the beginning and end of exon 5 in order to target all IE splice variants (Figure 1). Therefore, the sequence of the UL122/123 gene exon 2 and 5 was entered into the CRISPOR software (http://crispor.tefor.net) and the human genome was used as a reference for the calculation of off target sites. 14 different sequences of HCMV genomes available of NCBI were aligned for the potential target region to find conserved regions. gRNAs were selected based on a high selectivity score and low off target potential assessed by the software and in conserved regions of the viral genome. gRNA1 5'GGACTCCATCGTGTCAAGGACGG3'; gRNA2 5'GT CCTGGATGGCTGCCTCGATGG3'; gRNA3 5'GGTGCTACTGGAATCGATACCGG3'. For the unspecific gRNA we used a different DNA virus as input sequence and chose again a gRNA with low off target potential. Unspecific (unsp.) gRNA 5'GAATTTCACCCTGACAAAGGGGG3';

*Cells and Virus:*

MRC5 (primary fibroblasts, RD-Biotech S.A.S, France) and U373-MG (astrocytoma cell-line) cells were cultured in DMEM (Gibco life technology, USA); 10% (v/v) FCS (eurobio, France), 100 U/mL Penicillin/streptomycin (Gibco life technology, USA) 2 mM L-Glutamine (Sigma, USA), 10 mM HEPES (Gibco life technology, USA), so called complete medium later on. MRC5 were transduced with concentrated lentiviral vector (LV) type 1 containing the Cas9, a puromycin resistance and one or three gRNAs. MRC5 were transduced with supernatant of LV diluted at 1:100 in complete medium supplemented with 4 ng/µL polybrene and spinoculated at 1000g for 90 min at 33 °C. The inoculum was replaced after 8 h with fresh complete medium. Three days post transduction, MRC5 were treated with 2 µg/mL puromycin for two days

(Merck KGaA, Darmstadt, Germany) and subsequently cultured in complete medium. After two weeks, cells were a second time selected with puromycin (0,5 µg/mL) every two days for a period of 9 days and further maintained in complete medium. U373-MG were transduced with a LV type 2 containing Cas9-GFP(28) and one or three gRNAs with MOIs ranging between five to ten. After a three-week expanding phase, p24 negative transduced cells were FACS sorted based on their GFP level.

Viral stocks of TB40GFP, Toledo and VR1814 (kindly provided by Dr Giada Frascaroli and Pr Christian Sinzger, Ulm, Germany) were produced on MRC5 infected with a MOI of 0.01 and incubated in a low-FCS (2%) medium. Viral supernatants were harvested between seven to eight days post-infection (pi) and either used directly for infection or purified and concentrated by ultracentrifugation (24000 rpm, for 2.5 h at 4°C) on a 20% sucrose cushion (Qbiogene, USA). Viruses were tittered on MRC5 by a FACS analysis of IE positive cells 2 days pi.

*Infection of transduced and control cells*

MRC5 were plated in a 6-well plate (Falcon, Corning Incorporation, USA) at a density of 1.5x10^5 cells/mL and U373-MG were plated in a 12-well plate (Falcon, Corning Incorporation, USA) at a density of 3x10^5 cells/ml, one day prior to the HCMV-infection. Subsequently, cells were incubated for 2 h with the inoculum (MOI 1-0.1) and further cultured in fresh complete medium as defined earlier.

*PCR and T7-Assay*

Cellular and viral DNA were isolated with the NucleoSpin TriPrep Kit (Macherey-Nagel, Düren, Germany). The target regions of the *UL122/123* were PCR-amplified by either a small PCR (660 bp amplicon, Primers F: GTTCTCGTTGCAATCCTCGGTCAC; R: CGTGGCGGTAGGGTATGTGT) spanning the start codon for the IE or a bigger PCR

amplifying the whole *UL122/123* region (3862 bp amplicon; Primers F: ACATGAGGGGGAGAAGGACA; R: CGTGGCGGTAGGGTATGTGT). For the T7-assay, the small PCR was purified by the NucleoSpin column (Macherey Nagel, Düren, Germany). 200 ng of purified PCR product were denatured for five minutes at 95 °C and slowly re-annealed in three steps consisting of 15 s at 95 °C, 15 s at 85 °C and 30 s at 25 °C followed by a 30 min digestion at 37 °C with T7 endonuclease (New England Biolab Inc., UK). The reaction was stopped by adding 2 µL 0.25 M EDTA and samples were analyzed by capillary electrophoresis on a Caliper LabChip GX device (PerkinElmer). The concentration and purity of each band was measured in comparison to an internal marker, which allowed us to quantify our digested and wild type (WT) bands. Percentage of *indels* were calculated based on the formula form Hsu et al(53).

*TA cloning and Sequencing*

PCR products of the infected U373-MG cells four days pi were inserted into an empty ampicillin/kanamycin vector via TA cloning and transformed into competent bacteria with the StrataClone PCR cloning KIT (Agilent Technologie Devision, USA) following the manufacturer's instructions. After overnight incubation at 37 °C, positive clones were chosen by blue/white selection and send to sequencing to MWG (Eurofins Genomics GmbHm Ebersfeld, Germany).

*Virion release analysis by trans-infection plaque assay*

MRC5 were plated in a 24-well plate (Falcon, Corning Incorporation, USA) at a density of 2*10^5 cells/mL to be used the day after for trans-infection plaque assay. Eight days after HCMV-infection, U373-MG cells were harvested, counted, serially diluted (from 10^5 to 1 cell per well) and seeded over the MRC5 monolayer in duplicate. After overnight incubation, liquid

media was replaced by 0.8% agarose (Sigma, USA) in MEM (Gibco life technology, USA) (10% FCS, 100 U/mL Penicillin/Streptomycin, 2 mM L-Glutamine, 10 mM HEPES). After 7-14 days, plaques were observed by phase-contrast microscopy and counted.

*See supplement section of Material and Methods*

Acknowledgments

The authors would like to thank Laurent Tesson for technical discussions, Lucas Brusselle for technical support and Sophie Conchon for critical review of this paper. We thank Gisèle Froment, Didier Nègre and Caroline Costa from the lentivectors production facility / SFR BioSciences Gerland - Lyon Sud (UMS3444/US8)

This work was supported by the Labex IGO project (n° ANR-11-LABX-0016-01) funded by the «Investissements d'Avenir» French Government program, managed by the French National Research Agency (ANR); a grant from the Fondation Centaure ; a research grant from the European Society of clinical Microbiology and Infectious Diseases (ESCMID), a Junior basic grant from the European Society of Organ transplantation (ESOT). This work was realized in the context of the IHU-Cesti project, which received French government financial support managed by the Agence Nationale de la Recherche via the ''Investment Into The Future'' program ANR-10-IBHU-005. The IHU- Cesti project is also supported by Nantes Metropole and the Pays de la Loire Region.

Authorship Contributions

JG designed, performed, analyzed and interpreted experiments and wrote the paper. FC and designed, performed, analyzed experiments. NED performed, analyzed experiments. OP performed experiments. AC, EV and JG produced lentiviral vectors. IA and THN participated in experimental design and data discussion and contributed to the paper. FAH and FH designed experiments, interpreted results, wrote the paper and oversaw all aspects of the work. All authors reviewed and approved the manuscript.

Disclosure of Conflicts of Interest

Conflict-of-interest disclosure: the authors declare no conflict of interest disclosure.

**References**


1. Crough T, Khanna R. Immunobiology of human cytomegalovirus: from bench to bedside. Clin Microbiol Rev [Internet]. 2009 Jan [cited 2015 Mar 20];22(1):76–98, Table of Contents. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2620639&tool=pmcentrez&rendertype=abstract
2. Boeckh M, Adam P. Geballe. Cytomegalovirus: pathogen, paradigm, and puzzel. J Clin Invest. 2011;121(5).
3. Meyers JD, Flournoy N, Thomas ED. Risk factors for cytomegalovirus infection after human marrow transplantation. J Infect Dis [Internet]. 1986 Mar [cited 2016 Sep 12];153(3):478–88. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3005424
4. Lilleri D, Gerna G. Strategies to control human cytomegalovirus infection in adult hematopoietic stem cell transplant recipients. Immunotherapy [Internet]. 2016 Sep [cited 2016 Sep 12];8(9):1135–49. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27485084
5. Biron KK. Antiviral drugs for cytomegalovirus diseases. Antiviral Res. 2006;71(2–3 SPEC. ISS.):154–63.
6. Schmidt GM, Horak DA, Niland JC, Duncan SR, Forman SJ, Zaia JA. A randomized, controlled trial of prophylactic ganciclovir for cytomegalovirus pulmonary infection in recipients of allogeneic bone marrow transplants; The City of Hope-Stanford-Syntex CMV Study Group. N Engl J Med [Internet]. 1991 Apr 11 [cited 2016 Sep 12];324(15):1005–11. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1848679

7. Snoeck R, Sakuma T, De Clercq E, Rosenberg I, Holy A. (S)-1-(3-hydroxy-2-phosphonylmethoxypropyl)cytosine, a potent and selective inhibitor of human cytomegalovirus replication. Antimicrob Agents Chemother [Internet]. 1988 Dec [cited 2016 Sep 12];32(12):1839–44. Available from: http://www.ncbi.nlm.nih.gov/pubmed/2854454

8. Ringdén O, Lönnqvist B, Paulin T, Ahlmén J, Klintmalm G, Wahren B, et al. Pharmacokinetics, safety and preliminary clinical experiences using foscarnet in the treatment of cytomegalovirus infections in bone marrow and renal transplant recipients. J Antimicrob Chemother [Internet]. 1986 Mar [cited 2016 Sep 12];17(3):373–87. Available from: http://www.ncbi.nlm.nih.gov/pubmed/3009383

9. Chou S, Marousek G, Guentzel S, Follansbee SE, Poscher ME, Lalezari JP, et al. Evolution of mutations conferring multidrug resistance during prophylaxis and therapy for cytomegalovirus disease. J Infect Dis [Internet]. 1997 Sep [cited 2016 Sep 12];176(3):786–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9291334

10. Sullivan V, Coen DM. Isolation of foscarnet-resistant human cytomegalovirus patterns of resistance and sensitivity to other antiviral drugs. J Infect Dis [Internet]. 1991 Oct [cited 2016 Sep 12];164(4):781–4. Available from: http://www.ncbi.nlm.nih.gov/pubmed/1654362

11. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science [Internet]. 2007 Mar 23 [cited 2014 Jul 10];315(5819):1709–12. Available from: http://www.sciencemag.org.gate2.inist.fr/content/315/5819/1709.long

12. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A

programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science [Internet]. 2012 Aug 17 [cited 2016 Aug 12];337(6096):816–21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22745249

13. Wang J, Quake SR. RNA-guided endonuclease provides a therapeutic strategy to cure latent herpesviridae infection. Proc Natl Acad Sci [Internet]. 2014 Sep 9 [cited 2015 Aug 20];111(36):13157–62. Available from: http://www.pnas.org/cgi/doi/10.1073/pnas.1410785111

14. Lin S-R, Yang H-C, Kuo Y-T, Liu C-J, Yang T-Y, Sung K-C, et al. The CRISPR/Cas9 System Facilitates Clearance of the Intrahepatic HBV Templates In Vivo. Mol Ther Nucleic Acids [Internet]. 2014 Jan [cited 2015 Aug 19];3:e186. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4221598&tool=pmcentrez&rendertype=abstract

15. Mandal PK, Ferreira MR, Rossi DJ, Cowan CA, Ferreira LMRR, Collins R, et al. Efficient Ablation of Genes in Human Hematopoietic Stem and Effector Cells using CRISPR/Cas9. Cell Stem Cell [Internet]. 2014 Nov [cited 2014 Nov 7];15(5):643–52. Available from: http://www.sciencedirect.com/science/article/pii/S193459091400455X

16. Ebina H, Misawa N, Kanemura Y, Koyanagi Y. Harnessing the CRISPR/Cas9 system to disrupt latent HIV-1 provirus. Sci Rep [Internet]. 2013 Jan [cited 2014 Jul 14];3:2510. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3752613&tool=pmcentrez&rendertype=abstract

17. Stone D, Niyonzima N, Jerome KR. Genome editing and the next generation of antiviral therapy. Hum Genet [Internet]. Springer Berlin Heidelberg; 2016 Sep 8 [cited 2016 Sep 14];135(9):1071–82. Available from:

http://link.springer.com/10.1007/s00439-016-1686-2

18. van Diemen FR, Kruse EM, Hooykaas MJG, Bruggeling CE, Schürch AC, van Ham PM, et al. CRISPR/Cas9-Mediated Genome Editing of Herpesviruses Limits Productive and Latent Infections. PLoS Pathog [Internet]. 2016 Jun [cited 2016 Jul 5];12(6):e1005701. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27362483

19. Awasthi S, Isler J a, Alwine JC. Analysis of Splice Variants of the Immediate-Early 1 Region of Human Cytomegalovirus Analysis of Splice Variants of the Immediate-Early 1 Region of Human Cytomegalovirus. J Virol. 2004;78(15):8191–200.

20. Marchini A, Liu H, Zhu H. Human cytomegalovirus with IE-2 (UL122) deleted fails to express early lytic genes. J Virol [Internet]. 2001 Feb [cited 2016 Jul 12];75(4):1870–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11160686

21. Angulo A, Ghazal P, Messerle M. The major immediate-early gene ie3 of mouse cytomegalovirus is essential for viral growth. J Virol [Internet]. 2000 Dec [cited 2016 Jul 12];74(23):11129–36. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11070009

22. Glass M, Busche A, Wagner K, Messerle M, Borst EM. Conditional and reversible disruption of essential herpesvirus proteins. Nat Methods [Internet]. 2009 Aug [cited 2015 Jul 16];6(8):577–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/19578384

23. White EA, Clark CL, Sanchez V, Spector DH. Small internal deletions in the human cytomegalovirus IE2 gene result in nonviable recombinant viruses with differential defects in viral gene expression. J Virol [Internet]. 2004 Feb [cited 2015 Jul 16];78(4):1817–30. Available from:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=369462&tool=pmcentrez&rendertype=abstract

24. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol [Internet]. 2013 Sep [cited 2014 Jul 10];31(9):827–32. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3969858&tool=pmcentrez&rendertype=abstract

25. Ran F, Hsu P, Wright J, Agarwala V. Genome engineering using the CRISPR-Cas9 system. Nat Protoc [Internet]. 2013;8(11):2281–308. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24157548%5Cnhttp://www.nature.com/nprot/journal/v8/n11/abs/nprot.2013.143.html

26. Chenouard V, Brusselle L, Heslan J-M, Remy S, Ménoret S, Usal C, et al. A Rapid and Cost-Effective Method for Genotyping Genome-Edited Animals: A Heteroduplex Mobility Assay Using Microfluidic Capillary Electrophoresis. J Genet Genomics [Internet]. 2016 May 20 [cited 2017 Jul 17];43(5):341–8. Available from: http://linkinghub.elsevier.com/retrieve/pii/S1673852716300297

27. Richardson CD, Ray GJ, Dewitt MA, Curie GL, Corn JE. Enhancing homology-directed genome editing by catalytically active and inactive CRISPR-Cas9 using asymmetric donor DNA. Nat Biotechnol [Internet]. 2016 [cited 2017 Jun 23];34(3). Available from: http://www.nature.com.gate2.inist.fr/nbt/journal/v34/n3/pdf/nbt.3481.pdf

28. Bellec J, Bacchetta M, Losa D, Anegon I, Chanson M, Nguyen T. *CFTR* Inactivation by Lentiviral Vector-mediated RNA Interference and CRISPR-Cas9 Genome Editing in Human Airway Epithelial Cells. Curr Gene Ther [Internet]. 2015;15(5):447–59. Available from:

http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1566-5232&volume=15&issue=5&spage=447

29. Genini E, Gerna G, Labò N, Percivalle E, Baldanti F, Grazia Revello M, et al. In vitro selection of human cytomegalovirus variants unable to transfer virus and virus products from infected cells to polymorphonuclear leukocytes and to grow in endothelial cells. J Gen Virol [Internet]. Microbiology Society; 2001 Jun 1 [cited 2016 Aug 3];82(6):1429–38. Available from: http://jgv.microbiologyresearch.org/content/journal/jgv/10.1099/0022-1317-82-6-1429

30. Wroblewska Z, Wellish MC, Wolinsky JS, Gilden D. Comparison of human cytomegalovirus growth in MRC-5 human fibroblasts, brain, and choroid plexus cells in vitro. J Med Virol [Internet]. 1981 [cited 2017 Jul 23];8(4):245–56. Available from: http://www.ncbi.nlm.nih.gov/pubmed/6278069

31. Stamminger T, Puchtler E, Fleckenstein B. Discordant expression of the immediate-early 1 and 2 gene regions of human cytomegalovirus at early times after infection involves posttranscriptional processing events. J Virol [Internet]. 1991;65(5):2273–82. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=240576&tool=pmcentrez&rendertype=abstract

32. Nevels M, Paulus C, Shenk T. Human cytomegalovirus immediate-early 1 protein facilitates viral replication by antagonizing histone deacetylation. Proc Natl Acad Sci U S A [Internet]. 2004 Dec 7 [cited 2016 Sep 15];101(49):17234–9. Available from: http://www.ncbi.nlm.nih.gov/pubmed/15572445

33. Tang Q, Maul GG. Mouse cytomegalovirus immediate-early protein 1 binds with host cell repressors to relieve suppressive effects on viral transcription and

replication during lytic infection. J Virol [Internet]. American Society for Microbiology (ASM); 2003 Jan [cited 2016 Sep 15];77(2):1357–67. Available from: http://www.ncbi.nlm.nih.gov/pubmed/12502852

34. Gawn JM, Greaves RF. Absence of IE1 p72 protein function during low-multiplicity infection by human cytomegalovirus results in a broad block to viral delayed-early gene expression. J Virol [Internet]. American Society for Microbiology (ASM); 2002 May [cited 2016 Jul 12];76(9):4441–55. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11932411

35. Mocarski ES, Kemble GW, Lyle JM, Greaves RF. A deletion mutant in the human cytomegalovirus gene encoding IE1(491aa) is replication defective due to a failure in autoregulation. Proc Natl Acad Sci U S A [Internet]. National Academy of Sciences; 1996 Oct 15 [cited 2016 Jul 12];93(21):11321–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/8876134

36. Glass M, Busche A, Wagner K, Messerle M, Borst EM. Conditional and reversible disruption of essential herpesvirus proteins. Nat Methods. 2009;6(8):577–9.

37. Tarrant-Elorza M, Rossetto CC, Pari GS. Maintenance and Replication of the Human Cytomegalovirus Genome during Latency. Cell Host Microbe. 2014;16(1):43–54.

38. Taylor RT, Bresnahan WA. Human Cytomegalovirus Immediate-Early 2 Protein IE86 Blocks Virus-Induced Chemokine Expression. J Virol [Internet]. American Society for Microbiology; 2006 Jan [cited 2017 Feb 14];80(2):920–8. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16378994

39. Castillo JP, Yurochko AD, Kowalik TF. Role of human cytomegalovirus immediate-early proteins in cell growth control. J Virol [Internet]. American Society for Microbiology; 2000 Sep [cited 2017 Feb 20];74(17):8028–37. Available from:

http://www.ncbi.nlm.nih.gov/pubmed/10933712

40. FDA. Advances in CMV management: fomivirsen (Vitravene) approved. Proj Inf Perspect [Internet]. 1998 Dec [cited 2016 Sep 2];(26):7. Available from: http://www.ncbi.nlm.nih.gov/pubmed/11366499

41. Wang G, Zhao N, Berkhout B, Das AT. A Combinatorial CRISPR-Cas9 Attack on HIV-1 DNA Extinguishes All Infectious Provirus in Infected T Cell Cultures. Cell Reports. 2016.

42. Liu X, Hao R, Chen S, Guo D, Chen Y. Inhibition of Hepatitis B Virus by CRISPR/Cas9 System via Targeting the Conserved Regions of Viral Genome. J Gen Virol [Internet]. 2015 Apr 22 [cited 2015 Aug 20]; Available from: http://www.ncbi.nlm.nih.gov/pubmed/25904148

43. Biron KK, Harvey RJ, Chamberlain SC, Good SS, Smith AA, Davis MG, et al. Potent and Selective Inhibition of Human Cytomegalovirus Replication by 1263W94, a Benzimidazole L-Riboside with a Unique Mode of Action. Antimicrob Agents Chemother [Internet]. American Society for Microbiology; 2002 Aug 1 [cited 2016 Aug 24];46(8):2365–72. Available from: http://aac.asm.org/cgi/doi/10.1128/AAC.46.8.2365-2372.2002

44. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, Zheng Z, et al. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. Nature [Internet]. Nature Publishing Group; 2016;529(7587):490–5. Available from: http://www.nature.com/doifinder/10.1038/nature16526

45. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. Rationally engineered Cas9 nucleases with improved specificity. Science (80- ) [Internet]. 2015 Dec 1 [cited 2015 Dec 2];351(6268):84–8. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4714946&tool=pmc

entrez&rendertype=abstract

46. Hamilton ST, Milbradt J, Marschall M, Rawlinson WD. Human cytomegalovirus replication is strictly inhibited by siRNAs targeting UL54, UL97 or UL122/123 gene transcripts. PLoS One [Internet]. 2014 [cited 2016 Sep 15];9(6):e97231. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24887060

47. Mousavi-Jazi M, Schloss L, Drew WL, Linde A, Miner RC, Harmenberg J, et al. Variations in the cytomegalovirus DNA polymerase and phosphotransferase genes in relation to foscarnet and ganciclovir sensitivity. J Clin Virol. 2001;23(1–2):1–15.

48. Wang Z, Pan Q, Gendron P, Zhu W, Guo F, Cen S, et al. CRISPR/Cas9-Derived Mutations Both Inhibit HIV-1 Replication and Accelerate Viral Escape. Cell Rep [Internet]. Elsevier; 2016 [cited 2016 Apr 8]; Available from: http://www.cell.com/article/S2211124716302984/fulltext

49. Wang G, Zhao N, Berkhout B, Das AT. CRISPR-Cas9 can inhibit HIV-1 replication but NHEJ repair facilitates virus escape. Mol Ther [Internet]. 2016 Jan 22 [cited 2016 Jan 30];24(3):522–6. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4786927&tool=pmcentrez&rendertype=abstract

50. Lebbink RJ, de Jong DCM, Wolters F, Kruse EM, van Ham PM, Wiertz EJHJ, et al. A combinational CRISPR/Cas9 gene-editing approach can halt HIV replication and prevent viral escape. Sci Rep [Internet]. Nature Publishing Group; 2017 Feb 8 [cited 2017 Feb 23];7:41968. Available from: http://www.nature.com/articles/srep41968

51. Gustems M, Borst E, Benedict CA, Pérez C, Messerle M, Ghazal P, et al. Regulation of the transcription and replication cycle of human cytomegalovirus is insensitive to genetic elimination of the cognate NF-kappaB binding sites in the enhancer. J

Virol [Internet]. American Society for Microbiology (ASM); 2006 Oct [cited 2017 Jul 23];80(19):9899–904. Available from: http://www.ncbi.nlm.nih.gov/pubmed/16973595

52. Slobedman B, Mocarski ES. Quantitative Analysis of Latent Human Cytomegalovirus. 1999;73(6):4806–12.

53. Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. 2014;31(9):827–32.

Table 1: **Relative quantification of CRISPR-induced mutations in *UL122/123* gene in MRC5 cells**

| | ***Mean in % ±SD*** | | | |
|---|---|---|---|---|
| ***HCMV strain*** | *control* | *unsp. gRNA* | *singleplex* [a] | *multiplex* [b] |
| *Toledo* | 0 ±0 | 0 ±0 | 29.02 ±2.31 | 5.36 ±0.86 |

Mean percentages of *indels* are presented. [a] The percentage of mutation for the *singleplex* strategy is analyzed based on the T7 assay and quantification of the PCR products and cleavage products by the Caliper microfluidic bioanalyzer. [b] Bigger delections induced by the *multiplex* strategy is analyzed by PCR and quantified by the Caliper microfluidic bioanalyzer (n=2 or 3 independent experiments for each transduced MRC5s).

Table 2: **Relative quantification of CRISPR-induced mutations in *UL122/123* gene in U373-MG cells**

| | ***Mean in % ±SD*** | | | |
|---|---|---|---|---|
| ***HCMV strain*** | *control* | *unsp. gRNA* | *singleplex* [a] | *multiplex* [b] |
| *TB40-GFP* | 0 | 0 | 50.63 ±9.25 | 95.18 ±5.47 |
| *Toledo* | 0.28 ±0.69 | 0.09 ±0.22 | 31.18 ±5.18 | 92.14 ±4.69 |
| *VR1814* | 0.86 ±1.62 | 1.83 ±3.09 | 46.46 ±11.78 | 80.00±8.58 |

Mean percentages of *indels* are presented. [a] The efficiency of the *singleplex* strategy to induce mutations is analyzed based on the T7 assay and quantified by the Caliper microfluidic bioanalyzer. [b] Detection of deletion induced by the *multiplex* strategy is analyzed by PCR and quantified by the Caliper microfluidic bioanalyzer (n=3 independent experiments per virus strain).

Figure 1: **Design of the HCMV-targeting gRNAs**: Scheme of the targeted *UL122/123* gene and its major splice variants for IE1 and IE2 with the position of the three designed gRNAs anti-HCMV and their corresponding sequences (Scissors: gRNA/Cas9 complex).

Figure 2: **Anti-HCMV CRISPR/Cas9 system induces mutations resulting in the decrease of IE protein expression in primary fibroblasts:** MRC5 were transduced with one of the three LV type 1 and selected by puromycin treatment (2µg/mL) for 2 days. Control (untransduced) and puromycin-resistant MRC5 were subcultured prior to be infected with Toledo (MOI 0.1). Two days pi, proteins and DNA were extracted from the infected cells via TriPrep Kit. a) Viral DNA extracts were PCR-amplified at the target region. Amplicons were subsequently subjected to the T7 endonuclease to detect *indels* induced by the *singleplex* strategy. b) PCR amplicons of the whole IE gene were analyzed to detect bigger deletions induced by the *multiplex* strategy. The arrows highlight the *indels* (*singleplex*) and bigger deletions (*multiplex*) induced by the anti-HCMV CRISPR/Cas9 strategies. (one out of three independent experiments is shown). c) Western Blot analysis of the IE and Cas9 expression 2 days pi (one representative western blot out of 3 independent experiments is shown). d) At each passage, proteins were extracted with the TriPrep Kit and the Cas9 expression was assessed by Western Blot. e) Relative quantification of Cas9 expression based on the Western Blot (d) normalized by housekeeping protein expression of actin. pt: post transduction.

Figure 3: **Mutations in the *UL122/123* gene induced by the CRISPR/Cas9 anti-HCMV in U373-MG cell line:** Control and transduced U373-MG cells were infected with HCMV (Toledo, MOI 1) and cultured for eight days. Viral DNA was extracted and PCR

amplified. a) T7-assay was performed on the PCR amplicon of Exon 2 to detect *indels* induced by the *singleplex* strategy. b) Electrogramm of the T7 assay by the Caliper LabChip analysis for the in the unsp. gRNA and *singleplex* strategies. c) Large deletions induced by the *multiplex* strategy were highlighted by analyzing the whole *UL122/123* gene amplicon. d) Electrogramm of the PCR by the Caliper LabChip analysis identify a major amplicon of 500 bp and a smear above for the *multiplex* strategy. Arrows highlight the *indels* (*singleplex*) and bigger deletions (*multiplex*) induced by the anti-HCMV CRISPR/Cas9 strategies. One representative experiment out of three is shown for Toledo, similar data were found with TB40GFP and VR1814 (n=3 independent experiments per virus strain) LM, lower marker; UL, upper marker. e) Sequence analysis of the mutations induced by the *singleplex* strategy in the viral genome four days pi. Black: protospacer + PAM; bold: start codon; grey: insertions; grey-white: substitution.

Figure 4: **Decrease of IE expression by the HCMV targeting CRISPR/Cas9 systems**: Control and transduced U373-MG were infected with HCMV and harvested at two or eight days pi. a) Representative FACS histograms of intranuclear IE expression eight days pi are shown for all U373-MG cell lines and three different viral strains. The grey histogram represents uninfected U373-MG cells. b) IE expression on the different U373-MG cell lines normalized to HCMV-infected control U373-MG cells (dash line) (n=4 to 5 independent experiments). One-way ANOVA, multiple comparison tests, were performed to compare the results within the different cell lines and are presented in the table under each graph. Mann Whitney tests were performed to analyze each cell lines over time (day 2 pi *vs* day 8 pi). The only statistical difference is noted in the graph. c) Western Blot analysis of protein extraction obtained using TriPrep kit eight

days pi (one representative western blot out of 3 independent is shown for each virus strain as well as for the uninfected control U373-MG cells).

Figure 5: **Reduced progression of the viral replication cycle by the *multiplex* strategy**: Transduced and control U373-MG cells infected with HCMV were harvested eight days pi a) Relative viral genome quantification normalized to HCMV-infected control U373-MG cells (dash line) (n=3 independent experiments, +/- SD) One-way ANOVA, multiple comparison test, was performed and significant differences in comparison to the control are mentioned. b) Cells were FACS stained for total gB expression. Representative dot plots of total gB expression after infection with Toledo. c) gB expression normalized to HCMV-infected control U373-MG cells (dash line) (triangles: unsp. gRNA; dots: *singleplex*; diamonds: *multiplex*) (n=4 independent experiments). One-way ANOVA, multiple comparison test, was performed.

Figure 6: **Inhibition of virion release by the CRISPR/Cas9 anti-HCMV**: Trans-infection plaque assay of infected U373-MG cells over MRC5 incubated in solid media. Plaque formation was observed seven to 14 days post trans-infection with TB40GFP-infected U373-MG cell lines (a) or Toledo-infected U373-MG cell lines (b). Each symbol represents medians of duplicates obtained in independent experiments. One-way ANOVA, multiple comparison test, was performed and only significant differences are mentioned in the figure.

Figure 1

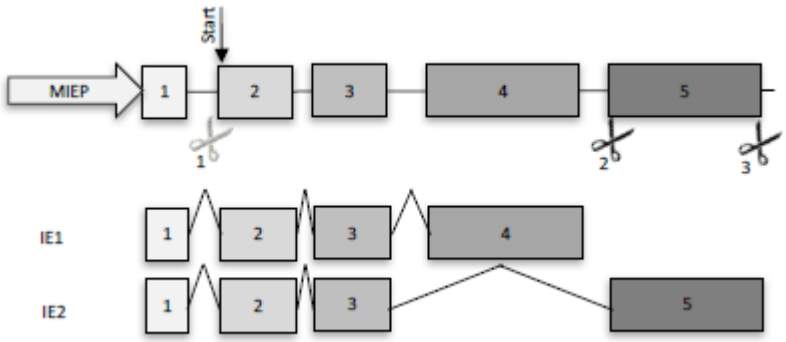


Figure 2

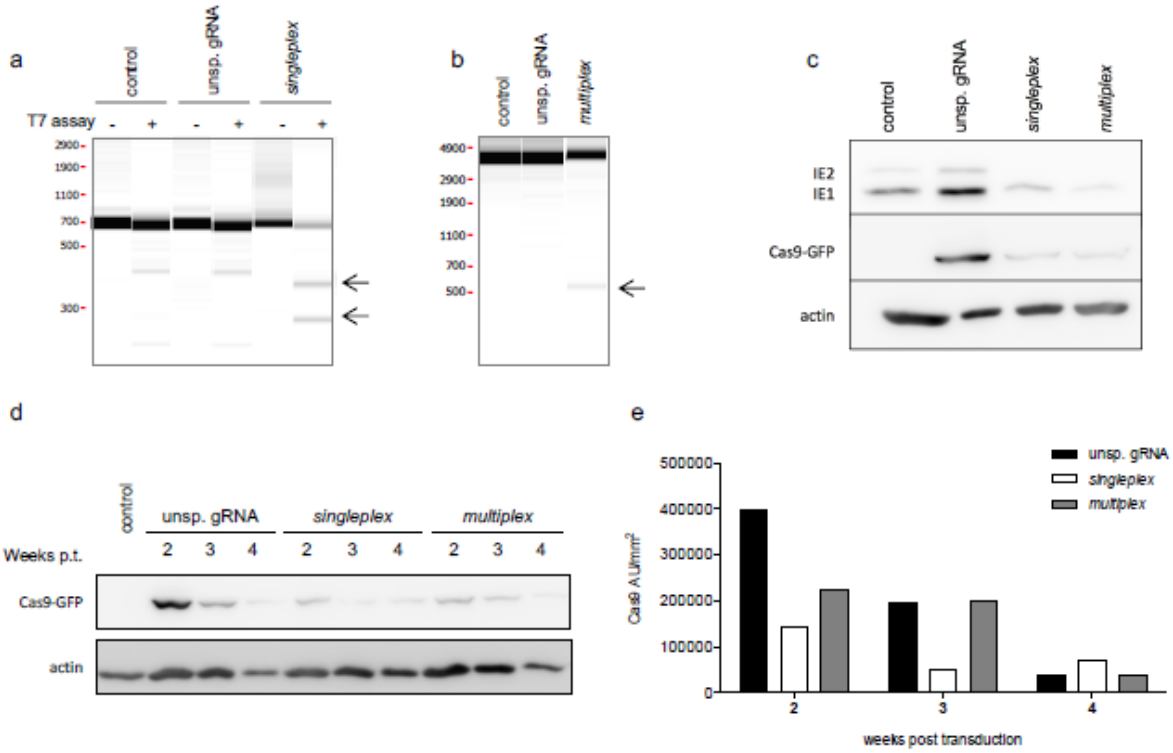
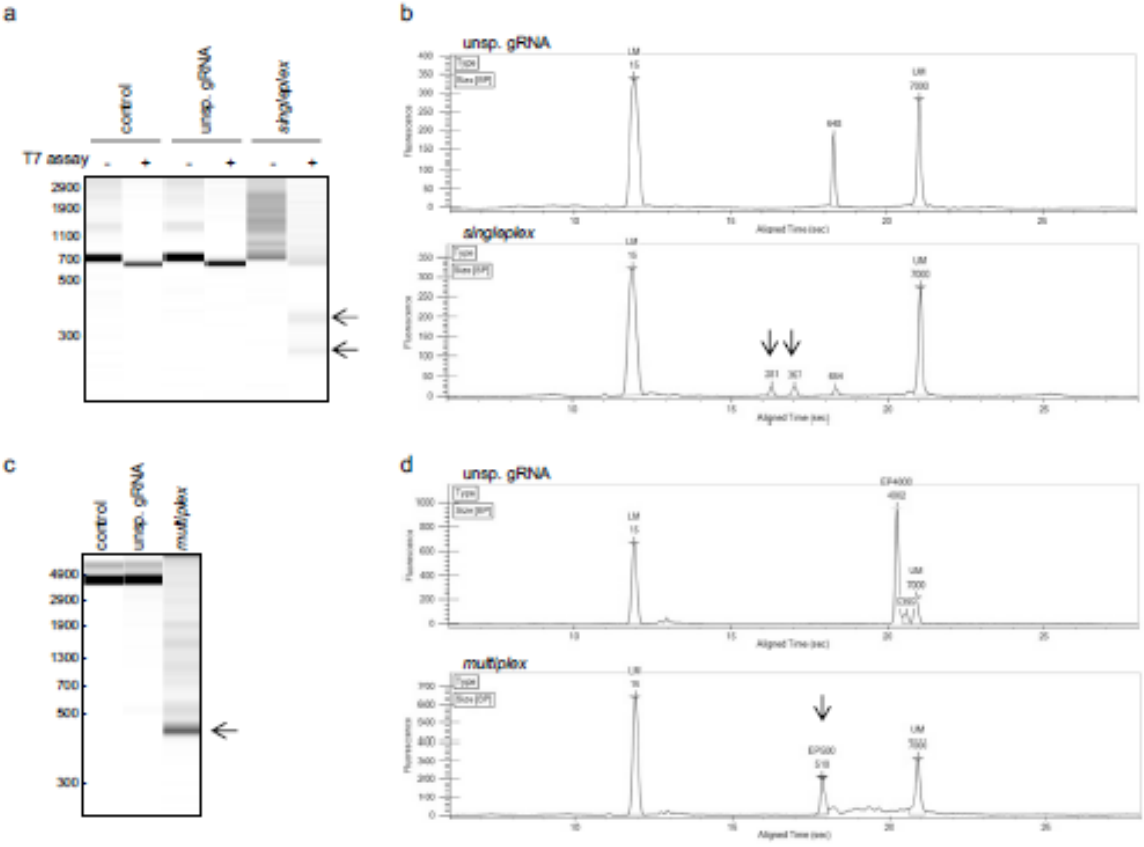

a
control
unsp. gRNA
singleplex
T7 assay
b
control
unsp. gRNA
multiplex
c
control
unsp. gRNA
singleplex
multiplex
IE2
IE1
Cas9-GFP
actin
d
control
unsp. gRNA
singleplex
multiplex
Weeks p.t.
2 3 4 2 3 4 2 3 4
Cas9-GFP
actin
e
unsp. gRNA
singleplex
multiplex
Cas9 A.U/mm²
weeks post transduction

Figure 3




e

| | |
|---|---|
| WT sequence | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAA |
| Sample 1 | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAA |
| Sample 2 | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTGCAGAA |
| Sample 3 | CTTGGCAGAGGACTCCATCGTGTCAAGGACGGTGACTCAGAA |
| Sample 4 | CTTGGCAGAGGACTCCATCGTGTCAA---CGGTGACTGCAGAA |
| Sample 5 | CTTGGCAGAGGACTCCATCGTGTCA-GGACGGTGACTGCAGAA |
| Sample 6 | CTTGGCAGAGGACT----------GTGACTGCAGAA |
| Sample 7 | CTTGGCAGAGG----------GGACGGTGACTGCAGAA |
| Sample 8 | CTTGGCAGAGGACTCCATCGTGTCA-GGACGGTGACTGCAGAA |
| Sample 9 | CTTGGCAGAGGACTCCATCGTGTCAA--ACGGTGACTGCAGAA |
| Sample 10 | CTTGGCAGAGGACTCCATCGTGTCAAAGGACGGTGACTGCAGAA |
| Sample 11 | CTTGGCAGAGGACTCCATCGTGTCAAAGGACGGTGACTGCAGAA |
| Sample 12 | CTTGGCAGAGGACTCCATCGTGTCAAAGGACGGTGACTGCAGAA |
| Sample 13 | CTTGGCAGAGGACTCCATCGTGTCAAAAGGACGGTGACTGCAGAA |

Figure 4

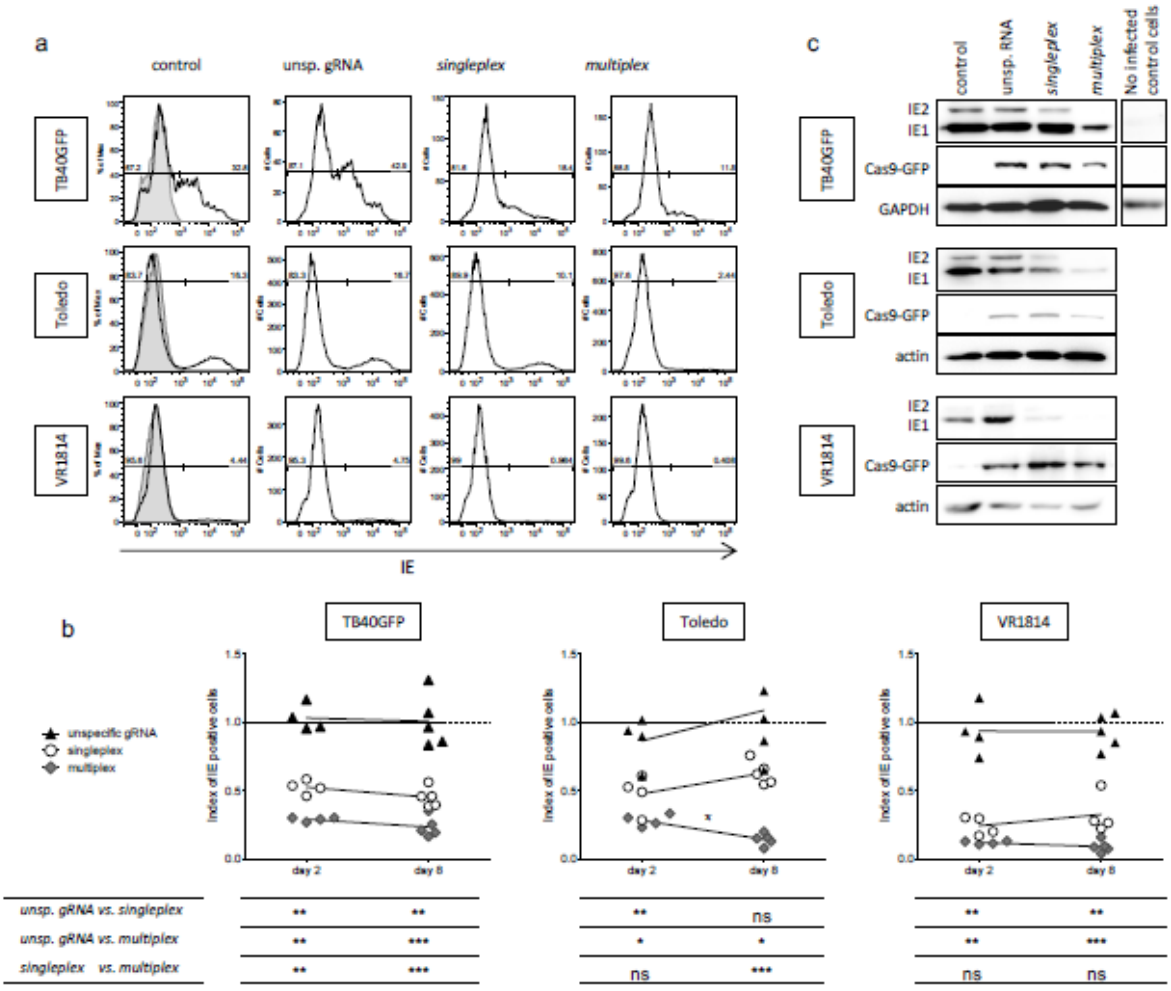

a
control
unsp. gRNA
singleplex
multiplex
TB40GFP
Toledo
VR1814
IE
c
control
unsp. RNA
singleplex
multiplex
No infected control cells
IE2
IE1
Cas9-GFP
GAPDH
actin
b
TB40GFP
Toledo
VR1814
unspecific gRNA
singleplex
multiplex
Index of IE positive cells
day 2
day 8
unsp. gRNA vs. singleplex
unsp. gRNA vs. multiplex
singleplex vs. multiplex
** ** ** ns ** **
** *** * * ** ***
** *** ns *** ns ns

Figure 5

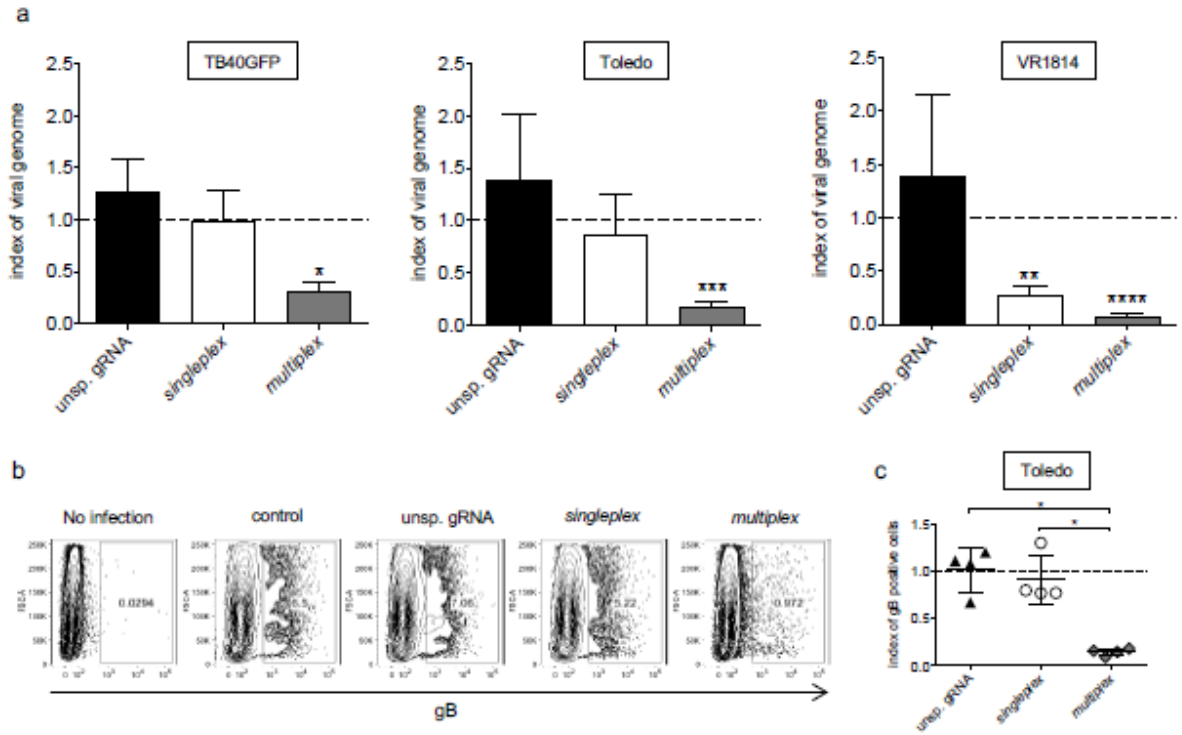
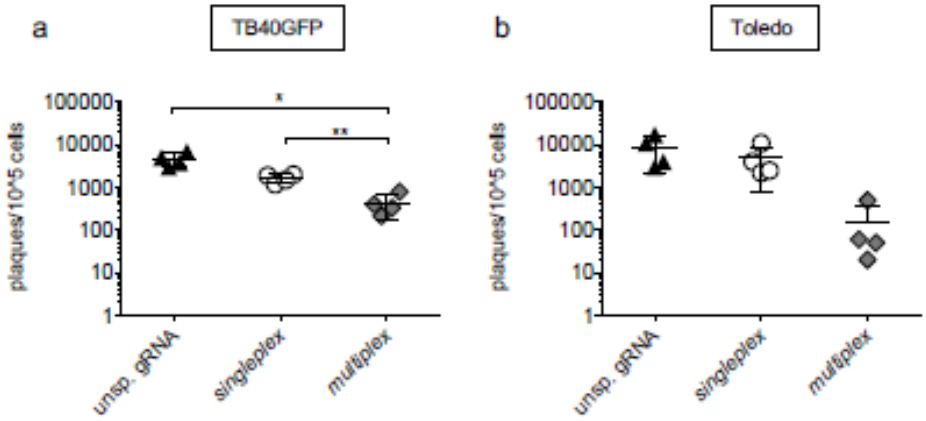
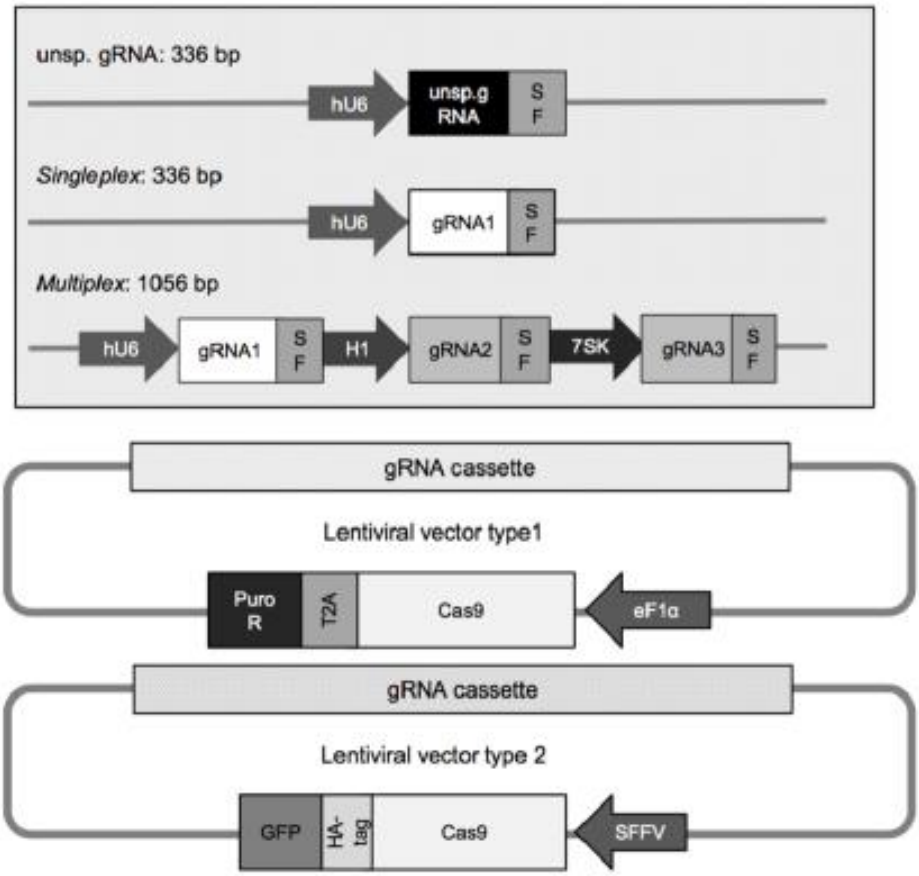

a
TB40GFP
Toledo
VR1814
index of viral genome
2.5
2.0
1.5
1.0
0.5
0.0
*
***
**
****
unsp. gRNA
singleplex
multiplex
b
No infection
control
unsp. gRNA
singleplex
multiplex
0.0294
6.5
7.05
5.22
0.972
gB
c
Toledo
index of gB positive cells
1.5
1.0
0.5
0.0
*
*
unsp. gRNA
singleplex
multiplex

Figure 6

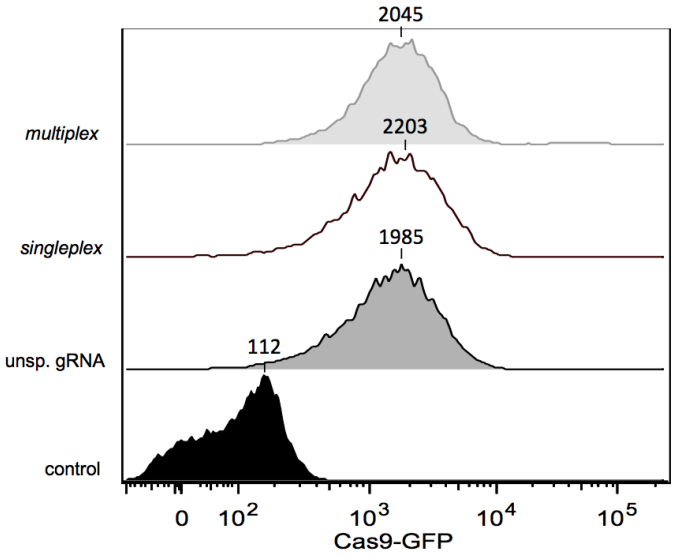

a
TB40GFP
b
Toledo
plaques/10^5 cells
100000
10000
1000
100
10
1
*
**
unsp. gRNA
singleplex
multiplex

**Supplementary Figure 1**



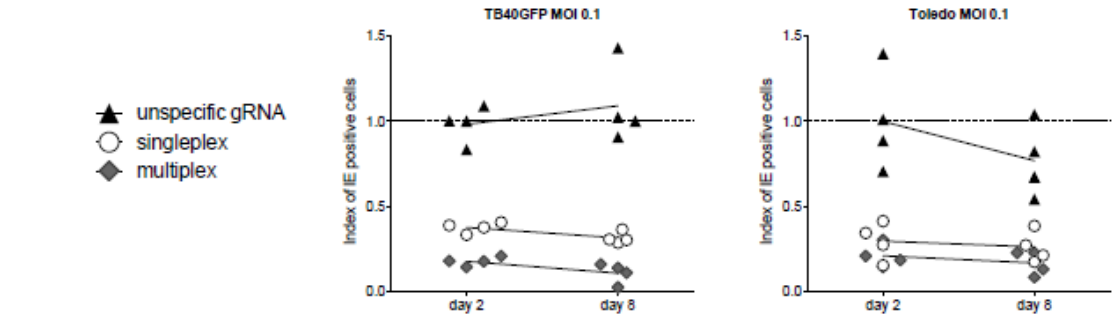

**Supplementary Figure 1: Constructs of HCMV-targeting CRISPR/Cas9 system**: Scheme of the CRISPR/Cas9 constructs including the different gRNA cassettes: unsp. gRNA, *singleplex* and *multiplex* and the two lentiviral vectors: type1 with Cas9-T2A-puromycin resistance and type 2 with Cas9-GFP fusion protein. Arrows represent the different promoters. SF: scaffold

**Supplementary Figure 2**




**Supplementary Figure 2: Generation of U373-MG cell lines expressing the CRISPR/cas9 strategies.** U373-MG cells transduced with one of the three lentiviral vectors type 2 were FACS-serted based on the Cas9-GFP$^{high}$ expression. Post-sort analysis of each U373-MG cell line is presented. Mean of fluorescence is indicated above each histogram.

**Supplementary Figure 3**




| | TB40GFP MOI 0.1 | | Toledo MOI 0.1 | |
|---|---|---|---|---|
| | day 2 | day 8 | day 2 | day 8 |
| *unsp. gRNA vs. singleplex* | ** | * | * | ns |
| *unsp. gRNA vs. multiplex* | ** | * | * | * |
| *singleplex vs. multiplex* | * | ** | ns | ns |

**Supplementary Figure 3: Decrease of IE expression by the anti-HCMV CRISPR/Cas9 systems at low MOI**: Control and transduced U373-MG were infected with HCMV at a MOI of 0.1 and harvested at two or eight days pi. The different U373-MG cell lines were stained for intranuclear IE expression and analyzed by FACS. The fractions of IE positive cells were normalized to HCMV-infected control U373-MG cells (dash line) (n=3 independent experiments. One-way ANOVA, multiple comparison tests, were performed to compare the results within the different cell lines and are presented in the table under each graph. Mann Whitney tests were performed to analyze each cell lines over time (day 2 pi *vs* day 8 pi). Only statistical difference are noted in the graph.

**Supplementary Table 1: Alignment of the target region of IE exon 2 for gRNA 1**

| | |
|---|---|
| **TB40/E - KF297339.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| **3157 - GQ221974.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAA-G- |
| **3301 - GQ466044.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| AD169 - **FJ527563.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| AF1 - **GU179291.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| HAN13 - **GQ221973.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| HAN38 - **GQ396662.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| JHC - **HQ380895.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAA-G- |
| JP - **GQ221975.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAGGA |
| Toledo - **GU937742.2** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| Towne - **FJ616285.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| U8 - **GU179288.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGATTATCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAA-G- |
| U11 - **GU179290.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGACGGTGACTGCAGAAAAGA |
| VR1814 - **GU179289.1** | GCACCTTGGAGGAAGGGCCCTCGTCAGGGTTGTCAGGGTCCATCTTTCTCTTGGCAGAGGACTC**CAT**CGTGTCAAGGA*CGGT*GACTGCAGAAAAGA |

**Supplementary Table 2: Alignment of the target region in EI exon 5 for gRNA2**

| | |
|---|---|
| **TB40/E - KF297339.1** | GTACAAGAGTCCATGTCTCT--TTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| **3157 - GQ221974.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| **3301 - GQ466044.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| AD169 - **FJ527563.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| AF1 - **GU179291.1** | GTACAAGAGTCCATGTCTCT--TTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| HAN13 - **GQ221973.1** | GTACAAGAGTCCATGTCTCT--TTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| HAN38 - **GQ396662.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTTCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| JHC - **HQ380895.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| JP - **GQ221975.1** | GTACAAGAGTCCATGTCTCT--TTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| Toledo - **GU937742.2** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTTAGGGTGTC |
| Towne - **FJ616285.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| U8 - **GU179288.1** | GTACAAGAGTGCATGTCTCT--TTCCAGTTTTTCACTTACTGAGATTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTATC |
| U11 - **GU179290.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |
| VR1814 - **GU179289.1** | GTACAAGAGTCCATGTCTCTCTTTCCAGTTTTTCACTTACTGAGACTTGTTCCTCAGGTCCTGGATGGCTGCCTCGATGGCCAGGCTCAGGGTGTC |

**Supplementary Table 3: Alignment of the target region in IE exon 5 for gRNA3**

| | |
|---|---|
| **TB40/E - KF297339.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| **3157 - GQ221974.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| **3301 - GQ466044.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGTGGCCGGTGCTACTGGAATCGATACCGGCA |
| AD169 - **FJ527563.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| AF1 - **GU179291.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| HAN13 - **GQ221973.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| HAN38 - **GQ396662.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| JHC - **HQ380895.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| JP - **GQ221975.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| Toledo - **GU937742.2** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| Towne - **FJ616285.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| U8 - **GU179288.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| U11 - **GU179290.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |
| VR1814 - **GU179289.1** | GACGCTGGTGGGGGTCGGCTTGTTAAGAGGGGCGCTGCTAACGCTGCAAGAGTGGGTTGTCAGCGTGGGGCCGGTGCTACTGGAATCGATACCGGCA |

**Supplemental Material and Methods**

*Cloning of the lentiviral vectors*

The different gRNAs used in this study were designed with the web tools from *Zhang Lab, MIT 2015* (http://crispr.mit.edu) and the TEFOR/CNRS network (http://tefor.net/crispor/crispor.cgi). Unspecific (unsp.) gRNA 5'GAATTTCACC CTGACAAAGGGGG3'; gRNA1 5'GGACTCCATCGTGTCAAGGACGG3'; gRNA2 5'GT CCTGGATGGCTGCCTCGATGG3'; gRNA3 5'GGTGCTACTGGAATCGATACCGG3' The gRNA1 and the unsp. gRNA were synthetized as oligonucleotids and cloned into the pX330 (Addgene). Further the gRNA1 or the unsp. gRNA and their promoter were PCR-amplified with the following primers F: 5'ATATGAATTCTT TTGCTCACATGTGAGGGC3', R: 5'ATATGAATTCCGCGCTAAAAACGGACTAGC3'. The amplicon was then cloned in the LV type 2 expressing the *S. pyrogenes* Cas9-GFP[489] by EcoRI. The expression cassette encoding the three gRNAs including their own promoters (U6, H1 and 7SK)[348] for the *multiplex* strategy was synthetized (GeneScript, Piscataway, USA) and cloned by EcoRI in the LV type 2.

The LV Cas9 Puro (Addgene plasmid # 52961) contains the Cas9-P2A-Puromycin as well as a U6 promotor for gRNA expression. The gRNA1 and the unsp. gRNA were synthetized as oligonucleotids and cloned into the LV type 1 expressing the *S. pyrogenes* Cas9 Puro following the Zhang's Lab protocol. The *mutliplex* strategy was PCR amplified form the LV Cas9-GFP by the following primers: F: TATAttaattaaacgcgtGAGGGCCTATTTCC, R: TATAgaattccgtacgaAAAAAAGCACCGA and cloned into the LV Cas9 Puro via PacI and EcoRI.

*qPCR*

Cellular and viral DNA were isolated with the NucleoSpin TriPrep Kit (Macherey-Nagel, Düren, Germany) and the viral genome was quantified by a *US8* qPCR with the following primers F: GGCACCAAATGCAGAGTGAG, R: AAGCCGTATTCCGTTTGCG, and probe: TGGTCCAAGTCCGTGGGCACC (FAM-BHQ-1, Eurofins Genomics GmbH, Ebersberg, Germany). The absolute quantification was performed based on a plasmid standard (10^6 to 10 copies/well) with the TaqMan™ Universal PCR Master Mix (Thermo Fisher Scientific, Waltham, MA USA) at 45 cycles of 94°C 20 s, 57°C 20s and 72°C 20s. Because the percentage of infected cells varied from one experiment to another, the genome copy numbers were first normalized to the amount of total cellular DNA and then expressed as an index of the amount of viral genomes in HCMV-infected untransduced U373-MG cells.

*FACS*

Cells were stained with Live/Dead Fixable Dead Cell Stain Kit (Invitrogen- Thermo Fisher Scientific, USA), fixed in 3.2 % PFA for 10 min on ice and then permeabilized with PBS / 3 % BSA / 0.2 % Triton for 30 min on ice. Intracellular IE was detected by either an anti-HCMV mAb (clone MAB810R; Millipore, Germany) or an anti-IE/E CMV antibody (Argene Biomérieux, France). HCMV glycoprotein B (gB) was detected intracellularly by a mouse anti-CMV gB antibody (1-M-12, Santa Cruz, USA). An anti-mouse IgG antibody conjugated to Alexa® 647 (BD Biosciences, USA) was used as a secondary antibody in all staining experiments presented in this study.

*Western Blot*

MRC5 (two days pi) or U373-MG (eight days pi) were harvested and the proteins were extracted by the NucleoSpin TriPrep Kit (Macherey-Nagel, Germany) according to

manufacturer's protocol. 15 µL of the samples were separated by SDS-PAGE and transferred via semi-dry western blot on a nitrocellulose membrane (GE Healthcare life science, UK) for the U373-MG lysate or via liquid transfer on a PVDF membrane (Millipore) for the MRC5 lysates. The membrane was blocked for 2 h in 5 % milk in TBST. IE were detected by the mouse anti-CMV antibody (MAB810R, Millipore, Germany) and a donkey anti-mouse HRP antibody (Jackson Immuno Research Labs, USA). After being washed, the membrane was incubated 5 min with the SuperSignal™ West Pico Chemiluminescent Substrate (Thermo Fisher Scientific, USA) and the signal revealed by the Luminescent Image analyzer LAS-4000 (FujiFilm, Japan). Following a short wash in TBST, antibodies were removed from the membrane by Restore Western Blot Stripping Buffer (Thermo Fisher Scientific, USA) for 30 min at room temperature and were used for subsequent detections. GAPDH or actin detection was used as a housekeeping protein with mouse anti-GAPDH (6C5, Santa Cruz, USA) or the mouse anti-actin (C4, Santa Cruz, USA) antibodies and a secondary antibody donkey anti-mouse HRP. The Cas9-GFP from the U373-MG lysate was detected with a rabbit anti-GFP antibody (Invitrogen- Thermo Fisher Scientific, USA) and revealed by a goat anti-rabbit HRP secondary antibody (Jackson Immuno Research Labs, USA). The Cas9 form the MRC5 lysate was detected with the anti-Cas9-A647 (clone 7A9-3A3 Alexa 488, Cell signaling, The Nederland) and revealed with a secondary antibody donkey anti-mouse HRP. For the quantification of the proteins on the membrane, the pictures were analyzed by the software GIMP 2. The background signal of the membrane was subtracted and the signal intensity of each band was calculated in an arbitrary unit/$mm^2$. The Cas9 expression was normalized to the actin expression.

# 2 Oral Presentations

| | |
|---|---|
| 05/2017 Nantes | "Advances in transgenic animal models and techniques" Fifth meeting of the Transgenic Rat ImmunoPhenomic platform of Nantes |
| | **Selected oral presentation (and poster)** "*A multiplex CRIPSR/Cas9 system which inhibits the human Cytomegalovirus replication*" |
| 10/2016 Hamburg | "6th European Congress of Virology (ECV)" by the European Society for Virology (ESV) |
| | **Best poster presentations:** "*Inhibition of Human Cytomegalovirus infection by the CRIPSR/Cas9 system*" |
| 05/2016 Washington | "ASGCT 19th Annual Meeting" by the American Society of Gene & Cell Therapy |
| | **Selected oral presentation:** "*The CRISPR/Cas9 system as an anti-viral treatment to prevent primary infection by HCMV positive hematopoietic stem cells*" |
| 12/2015 Nantes: | Annual Meeting of the Doctoral School « Biologie-Santé » Nantes/Angers |
| | **Selected oral presentation**: "*A new antiviral strategy against HCMV using genome editing tools*" |
| 07/2015 Nantes: | "Transgenic animals and genetic engineering techniques" of The Transgenic Rat Immunophenomic Platform facility of SFR François Bonamy, Biogenouest and IBiSA |
| | **Selected oral presentation**: "*Genome editing: a new antiviral strategy against HCMV*" |

# 3 Poster presentations

| | |
|---|---|
| 10/2016 Hamburg | "6th European Congress of Virology (ECV)" by European Society for Virology (ESV) |
| | **Best poster:** "*Inhibition of Human Cytomegalovirus infection by the CRIPSR/Cas9 system*" |
| 06/2016 Boston | "FOCIS 2016" of the Federation by Clinical Immunology Societies (FOCIS) |
| | "*The CRISPR/Cas9 system as an anti-viral treatment to prevent primary infection by Human Cytomegalovirus positive hematopoietic stem cells*" |

# 4 Awards

| | |
|---|---|
| 10/2016 Hamburg | **Poster Award** ECV 2016 from Life Science Nord |
| 10/2016 Nantes | **Travel Grant** from the of the Doctoral School « Biologie-Santé » Université de Nantes for the ECV 2016 in Hamburg |
| 05/2016 Nantes | **Travel grant** from Labex IGO for the "ASGCT 19th Annual Meeting" in Washington DC |
| 12/2015 Nantes | **First prize** for oral presentations at the Annual Meeting of the Doctoral School « Biologie-Santé » Nantes/Angers |

UNIVERSITE BRETAGNE LOIRE

UNIVERSITÉ DE NANTES

# Thèse de Doctorat

Janina GERGEN

## The CRISPR/Cas9 system as an anti-viral strategy against the human Cytomegalovirus

## Utilisation du système CRISPR/Cas9 comme stratégie antiviral contre le Cytomégalovirus humain

**Abstract**

The human cytomegalovirus (HCMV) primary infection is usually asymptomatic but leads to latent infection of blood progenitor cells. Immunocompromised patients are at high risks of HCMV reactivation, which is associated with severe end organ diseases and increased mortality in transplant patients. Standard anti-viral treatments based on nucleotide analogues decreased the occurrence of HCMV reactivation and diseases, but induce side effects and drug-resistant viral strains. In this thesis, we introduced new anti-viral approaches based on the CRISPR/Cas9 gene editing tool. Two strategies are designed to target the *UL122/123* gene of HCMV encoding the immediate early proteins, essential for lytic viral replication and reactivation from latency. We validated that the disruption of the *UL122/123* gene by the CRISPR/Cas9 system to abrogate viral replication. The *multiplex* CRISPR/Cas9 system (three gRNA) was much more efficient than the *singleplex* approach targeting the same gene. Target gene expression, concomitant genome replication and virion release were significantly impaired by the *multiplex* strategy. A further anti-HCMV CRISPR/Cas9 system was developed specifically to target the HCMV genome during latency. Two gRNAs target the viral genome at three target sites: LUNA, essential for reactivation, and the two homolog TR regions. We verified this duplex strategy on the lytic replicating virus and detected mutations at the target site as well as the reduction of viral genome copy number. In conclusion, the anti-HCMV strategies based on two or three gRNAs efficiently blocked viral replication. This provides the basis for the development of an anti-HCMV CRISPR/Cas9 therapy.

**Human cytomegalovirus, CRISPR/Cas9, Immediate Early, Multiplex, Anti-viral therapy, Latency, LUNA, TR region**

---

**Résumé**

L'infection primaire par le cytomégalovirus (CMV) humain est asymptomatique. Le sujet infecté reste cependant porteur du virus à l'état latent. Le CMV ne se réactive que sporadiquement chez l'individu immunocompétent. Chez les patients immunodéprimés, la réactivation du CMV peut induire des maladies à CMV touchant des organes vitaux et peut mettre en jeu le pronostic vital du patient. Les traitements standards sont efficaces mais leurs effets secondaires et l'apparition de souches virales résistantes relancent l'engouement pour le développement de nouvelles thérapies. Lors de ma thèse, j'ai utilisé le système CRISPR/Cas9 afin de déstabiliser le génome du CMV. Nous avons choisi de cibler le gène *UL122/UL123* codant pour les molécules immediate early essentielles au cycle lytique réplicatif et à la sortie de la latence. Nous avons comparé deux stratégies utilisant soit un soit trois gRNAs, respectivement appelées singleplex et multiplex ciblant ce même gène. Alors que le singleplex induit des insertions et délétions au site de coupure du gRNA, la stratégie multiplex induit la délétion de 3500 paires de bases du gène ciblé. De ce fait, la stratégie multiplex bloque efficacement l'expression du gène ciblé, la réplication virale et le relargage de nouveaux virions. Une autre stratégie a été développée pour cibler spécifiquement le génome à l'état latent. Les deux régions homologues TR et le gène LUNA sont ciblés par deux gRNAs. Cette seconde stratégie résulte elle aussi en une diminution du nombre de copies de génome viral lors du cycle lytique de réplication. Il est désormais possible d'envisager de nouvelles solutions thérapeutiques anti-HCMV avec la stratégie CRISPR/Cas9.

**Cytomégalovirus humain, CRISPR/Cas9, Molécules Immediate early, Multiplex, Stratégies thérapeutiques antiviral, Latence, LUNA, Région TR**

---