Impact of cultivation conditions on *N*-glycosylation of influenza A virus hemagglutinin, on quasispecies composition, and on immunogenicity of virus preparations

Dissertation

Zur Erlangung des akademischen Grades

Doktoringenieurin

(Dr.-Ing.)

von Dipl.-Ing. Jana V. Rödig

geb. am: 24.01.1980 in Berlin

genehmigt durch die Fakultät für Verfahrens- und Systemtechnik der Otto-von-Guericke-Universität Magdeburg

Promotionskommission:

Prof. Dr.-Ing. Andreas Seidel-Morgenstern (Vorsitz)

Prof. Dr.-Ing. Udo Reichl (Gutachter)

Prof. Michael Butler (Gutachter)

Dr.-Ing. Holger Lübben (Gutachter)

eingereicht am: 02.12.2013

Promotionskolloquium am: 24.04.2014

Zusammenfassung

Noch heute zählt die Impfung zur wichtigsten Influenzaprävention. Influenzaimpfstoffe werden in embryonierten Hühnereiern oder in tierischer Zellkultur hergestellt. Klare Vorteile der zellkulturbasierten Produktion sind die Unabhängigkeit von Eilieferanten, das Ausschließen möglicher anaphylaktischer Reaktionen auf Eiproteine sowie die Möglichkeit, die Produktion schneller an den aktuellen Bedarf anzupassen. In allen Prozessen stellt das Glykoprotein Hämagglutinin (HA) das Hauptantigen dar. HA ist in der viralen Hülle hoch abundant und löst aufgrund seiner hohen Immunogenität schützende Immunantworten aus. Viele Eigenschaften eines Glykoproteins, wie z. B. Immunogenität, Antigenität, Rezeptorbindungsspezifitäten und Stabilität, können jedoch von dessen Glykosylierung entscheidend beeinflusst werden. Ein grundsätzliches Verständnis der Einflussgrößen verschiedener zellkulturbasierter Kultivierungsbedingungen auf die *N*-Glykosylierung ist daher essentiell zur Verbesserung des Prozessdesigns z. B. zur Auswahl der Zelle.

Der Einfluss der Kultivierungsbedingungen auf das HA *N*-Glykosylierungsmuster des Influenzavirus A (IVA) wird in dieser Studie mittels kapillarer Gelelektrophorese mit Laser-induzierter Fluoreszenzdetektion- (CGE-LIF-) basierter Glykoanalytik untersucht. In den resultierenden, gut reproduzierbaren Fingerabdrücken repräsentiert jeder Peak mindestens eine bestimmte *N*-Glykanstruktur. Auf genomischer Ebene erlaubt die Methode der Pyrosequenzierung darüber hinaus die Charakterisierung der viralen Quasispecies. Im Rahmen von *in vitro* und *in vivo* Mausstudien wird abschließend in einem transgenen HA-spezifischem T cell Rezeptor (TCR-HA) Modell sowie im BALB/c *wildtyp* (*wt*) Modell der Einfluß verschiedener Glykovarianten auf die Immunogenität verschiedener Viruspräperationen adressiert.

Diese Studie zeigt, dass vor allem die Wahl des Produktionssystems (MDCK, Vero, AGE1.CR.pIX, Cap, MDCK.SUS1, MDCK.SUS2, MDCK.SUS3 Zellinien und embryonierte Hühnereier) und des Virusstammes (IVA PR/8/34, H1N1;

California/07/2009-Reassortant, H1N1pandemic; IVA Uruguay/716/2007-Reassortant, H3N2, IVA Victoria/210/2009-Reassortant, H3N2) das *N*-Glykosylierungsmuster des HA entscheidend beeinflussen und sowohl das Vorkommen als auch die relativen Häufigkeiten verschiedener *N*-Glykanstrukturen bestimmen. Bemerkenswert ist, dass eine Adaptation von adhärent wachsenden Madin Darby canine kidney (MDCK) Produktionszellen an serumfreies Zellwachstum nur relative Häufigkeiten HA-assoziierter *N*-Glykane beeinflusst, während eine Adaptation an Suspensionswachstum in serumfreiem Medium das *N*-Glykosylierungsmuster grundlegend verändert. Im Allgemeinen ist zur Prozessoptimierung, oft eine Adaptation des Saatviruses an die Produktionszelle notwendig, um optimale Virusausbeuten zu erzielen. Diese Arbeit zeigt, dass der Adaptationsstatus des Viruses lediglich die relative Häufigkeit HA-assoziierter Glykane beeinflusst. Die Steigerung der Ausbeute ist auf erworbene Mutationen während der Adaptation vor allem im HA, aber auch im viralen Nukleoprotein (NP), in der Neuraminidase (NA) und im nicht-strukturellen Protein 1 (NS1) zurückzuführen. Darüber hinaus wird gezeigt, dass der Lieferant/Ursprung des zur Virusproduktion verwendeten Saatviruses, die Viruspassage, die Kultivierungsgefäße und –medien, sowie der Erntezeitpunkt nur einen geringen Einfluss auf relative Häufigkeiten der HA-assoziierten *N*-Glykane ausüben. Die Relevanz solcher Glykoanalyse für die Qualitätssicherung in der Impfstoffproduktion wird in einem transgenen TCR-HA Mausmodell, sowie in einem BALB/c *wt* Modell hervorgehoben. Ein Vergleich von MDCK- und Vero-spezifisch glykosylierten Viruspräperationen des IVA PR/8/34 (H1N1) zeigt einen signifikanten Einfluss der *N*-Glykosylierung auf die Immunogenität *in vitro* und *in vivo*. Außerdem deuten die Daten auf eine ausgeprägtere Unterstützung der humoralen Immunantwort durch die MDCK Zell-spezifische Glykosylierung, sowie auf eine stärkere Förderung der zellulären Immunantwort durch die Vero Zell-spezifische Glykosylierung hin.

Abstract

Vaccination and hygiene measures still represent the best strategies to prevent influenza virus infection. Manufactures produce influenza vaccines in different host systems, i.e. either in fertilized chicken eggs or in different mammalian cell lines. Advantages of cell culture-based virus production include independence from egg supply, prevention of anaphylactic reactions caused by egg proteins as well as the ability to rapidly scale-up and -down to better match vaccine demand. In all processes, the viral glycoprotein hemagglutinin (HA) is purified as the major vaccine antigen. HA is highly abundant in the envelope of influenza viruses, and able to induce strong and protective immune responses. Quality characteristics of glycoproteins, such as immunogenicity, antigenicity, protein stability and receptor-binding specificity can strongly depend on the glycan composition with respect to *N*-glycan structures attached as well as their relative abundances. A fundamental understanding of the impact of cultivation conditions is necessary to support process design, e.g. regarding the choice of host system, in cell culture-derived influenza vaccine production.

In this study the impact of cultivation conditions on the HA *N*-glycosylation pattern of influenza A virus (IVA) is investigated by capillary gel electrophoresis with laser-induced fluorescence detection- (CGE-LIF-) based glycoanalysis. As a result, well reproducible *N*-glycan fingerprints are obtained, in which one peak corresponds to at least one distinct *N*-glycan structure. Furthermore, for characterization of the viral quasispecies, next-generation pyrosequencing is applied. Finally, the importance of *N*-glycosylation on immunogenicity of virus preparations is addressed *in vitro* as well as *in vivo* using transgenic HA-specific T cell receptor (TCR-HA) and *wild type* (*wt*) BALB/c mouse models.

The results demonstrate that peak presence as well as peak abundance mainly depend on the host system (MDCK, Vero, AGE1.CR.pIX, Cap, MDCK.SUS1, MDCK.SUS2, MDCK.SUS3 cell lines and embryonated hens' eggs) and the virus strain (IVA PR/8/34, H1N1; California/07/2009-reassortant, H1N1pandemic; IVA Uruguay/716/2007-reassortant, H3N2, IVA Victoria/210/2009-reassortant, H3N2)

chosen. Interestingly, the adaptation of adherently growing Madin Darby canine kidney (MDCK) cells to serum-free cell growth only slightly affects relative abundances of HA-associated *N*-glycan structures. In contrast, the adaptation to suspension growth in serum-free medium alters HA *N*-glycan fingerprints drastically with respect to relative abundances as well as *N*-glycan structure presence. In particular, the total number of different *N*-glycan structures is reduced, and the *N*-glycans show a tendency towards smaller structures. In general, for process optimization, the adaptation of virus seed is often necessary for sufficient virus yields. This work demonstrates that the adaptation status of the virus hardly affects the HA *N*-glycosylation fingerprint, only showing changes in relative *N*-glycan structure abundances. However, after virus adaptation, acquired mutations, in particular within the HA, allowed increased virus replication and hence improved final virus titers. Additional mutations are detected within the viral nucleoprotein (NP), the neuraminidase (NA) and the non-structural protein 1 (NS1). Changes in process conditions, including virus passages, virus suppliers/origin, virus production media, virus production vessels as well as time points of harvest only affect relative abundances of HA-associated *N*-glycans. Finally, the relevance of glycoanalysis for quality control and assessment for vaccine production is highlighted in transgenic TCR-HA as well as *wt* BALB/c mouse models. A comparison of MDCK and Vero cell-specific glycosylated virus preparations demonstrates that *N*-glycosylation has a marked impact on immunogenicity *in vitro* as well as *in vivo*. Furthermore, results suggest that MDCK cell-specific glycosylation more promotes the humoral immune response whereas Vero cell-specific *N*-glycosylation seems to more promote the cellular immune response.

Abbreviations, indices and symbols

2-AA	2-amino-antranillic acid
AA	amino acid
Ab	antibody
2-AB	2-aminobenzamide
ADCC	antibody dependent cell-mediated cytotoxicity
Ag	antigen
AGE1.CR.pIX	immortalized and modified designer cell line from ProBioGen AG, Berlin, Germany; originating from Muscovy Duck
ANTS	8-aminonaphthalene-1,3,6-trisulfonic
APC	antigen presenting cell
APTS	8-aminopyrene-1,3,6-trisulfonic acid
Asn	asparagine
BALB/c mice	albino, laboratory-bred strain of the house mouse (*mus musculus*)
BCR	B cell receptor
bp	base pairs
°C	degree Celsius
Cap	immortalized designer cell line from Cevec Pharmaceuticals GmbH, Cologne, Germany; originating from primary human amniocytes
CD	cluster of differentiation
cDNA	coding DNA
CE	capillary electrophoresis
CGE	capillary gel electrophoresis
CLR	C-type lectin receptor
CPSF	cleavage and polyadenylation specificity factor
cRNA	complementary RNA
CTL	cytotoxic T lymphocytes
Da	dalton
ΔRPH	difference of relative peak height

IΔRPHI	absolute value of ΔRPH
DC	dendritic cell
DC-SIGN	DC-specific intercellular adhesion molecule-3-grabbing non-integrin
DNA	deoxyribonucleic acid
dNTP	deoxinucleoside triphosphate
ddNTP	dideoxinucleoside triphosphate
DMSO	dimethylsulfoxid
DO	dissolved oxygen
Dol	dolichol
dpi	days post immunization
dsRNA	double-stranded RNA
e.g.	for example (*exempli gratia*)
elF4G1	eukaryotic initiation factor 4 G1
ELISA	enzyme-linked immunosorbent assay
ELISPOT	enzyme-linked immunosorbent spot assay
EMA	European medicines agency
emPCR	emulsion PCR
EndoH	endonuclease H
ER	endoplasmatic reticulum
et al.	and others (*et alii*)
FACE	fluorescence-assisted carbohydrate electrophoresis
FDA	U.S. food and drug administration
FLI	Friedrich-Loeffler-Institut, Greifswald – Insel Riems, Germany
g	gravity constant (earth: g = 9.81 m/s^2)
GDP	guanine diphosphate
Glc	glucose
GlcNAc	*N*-acetylglucosamine
GMP	good manufacturing practice
h	hour
HA	hemagglutinin
HAI	hemagglutination inhibition

HAU	HA units
HCD	high confidence difference
HILIC	hydrophilic interaction chromatography
HPAEC	high performance (or high pH) anion exchange chromatography
hpi	hours post infection
HPLC	high-performance liquid chromatography
hps	hours post stimulation
H_2O_{MQ}	ultrapure water (Millipore)
i.e.	that is (*id est*)
IFN	interferon
IL	interleukin
i.p.	intraperitoneal
i.v.	intravenous
IVA	influenza A virus
IVA-California	IVA reassortant California/07/2009 (H1N1pandemic)
IVA-PR8	IVA Puerto Rico/8/34 (H1N1)
IVA-Uruguay	IVA reassortant Uruguay/716/2007 (H3N2) x IVA-PR8/34 (H1N1)
IVA-Victoria	IVA reassortant Victoria/210/2009 (H3N2) x IVA-PR8/34 (H1N1)
kDa	kilodalton
L	liter
LAIV	live attenuated influenza virus
LIF	laser-induced fluorescence
LOD	limit of detection
M1	matrix protein 1
M2	matrix protein 2 (also termed proton channel protein M2)
mAb	monoclonal antibody
Man	mannose
MDCK	Madin Darby canine kidney
MHC	major histocompatibility complex
min	minute

MGL	macrophage galactose-type lectin
MMR	macrophage mannose receptor
µL	microliter
mL	milliliter
moi	multiplicity of infection
MPI-KG	Max Planck Institute for Colloids and Interfaces, Potsdam – Golm, Germany
mRNA	messenger RNA
MS	mass spectrometry
MTU	migration time units
MTU'	normalized migration time units (equivalent of bp)
MW	molecular weight
NA	neuraminidase
$NaBH_3CN$	sodium cyanoborohydride
NEP	nuclear export protein (also termed NS 2)
NIBSC	National Institute for Biological Standards and Control
NIBSC-strain	IVA-PR8 purchased from NIBSC
NP	nucleoprotein
NS1	non-structural protein 1
NS2	non-structural protein 2 (aslo termed NEP)
OST	oligosaccharyltransferase
P	phosphate
PA	polymerase acidic protein
PAD	pulsed amperometric detection
PAGE	polyacrylamide gel electrophoresis
PAS	periodic acid -Schiff
PB1	polymerase basic protein 1 (also termed RNA-directed RNA polymerase catalytic subunit)
PB1-F2	protein PB1-F2
PB2	polymerase basic protein 2
PCR	polymerase chain reaction

PGC	porous graphitized carbon
PNGaseF	peptide: *N*-glycosidase F
PP	pyrophosphate
RFU	relative fluorescence units
RIG-I	retinoic inducible gene I
RKI	Robert Koch Institute
RKI-strain	IVA-PR8 obtained from RKI
RNA	ribonucleic acid
RNP	ribonucleoprotein complex
RP	reverse phase
RPH	relative peak height
rpm	rounds per minute
RT	room temperature
RT-PCR	reverse transcriptase PCR
s	second
SEC	size exclusion chromatography
SEM	standard error of mean
Ser	serine
SD	standard deviation
SDS	sodium dodecyl sulfate
SOP	standard operating procedure
SP-D	surfactant protein D
sstDNA	single stranded template DNA
$TCID_{50}$	tissue culture infectious dose 50 %
TCR	T cell receptor
Th	T helper cell
Thr	threonine
t_{mig}	migration time
TPH	total peak height
Treg	regulatory T cell
TRIM25	tripartite motif-containing protein 25

UDP	uridine diphosphate
vRNA	viral RNA
WHO	World Health Organization
wt	wild type

Acknowledgements

All people, who have more or less contributed to this work, are cordially aknowledged. In particular:

I thank Prof. Dr.-Ing. U. Reichl for the possibility to work under excellent conditions during my Ph.D. studies. Furthermore, I very much appreciated the opportunity for development as well as the steady willingness for constructive discussions.

I thank Prof. Dr. M. Butler and Dr.-Ing. H. Lübben for reviewing this thesis.

I thank Dr. rer. nat. E. Rapp and Dr. rer. nat. habil. Y. Genzel for excellent supervision and/or support during my work as well as for the fruitful discussions and suggestions for scientific publications.

I thank Dr. rer. nat. D. Höper, Dr. rer. nat. B. Lepenies, Dr.-Ing. B. Hundt and J. Hütter for the very successful, exciting, data-rich and pleasant cooperations.

I thank Prof. Dr. rer. nat. Th. Schüler for fruitful discussions as well as suggestions and therefore for his contribution to this work.

I thank all technical assistances, especially I. Behrendt, S. König, C. Best, F. Hasewinkel and C. Ziemann for the excellent team work and support.

I thank my supervised students H. Kaffka and M. Kampke for their motivation and dedication and therefore for their contribution to this work.

I thank all my colleagues, especially M. Rüger, R. Janke, B. Heynisch and S. Freund for the pleasant working atmosphere and all scientific and non-scientific discussions.

Last but not least, special thank goes to my whole family, who have supported me all the way through.

Index of contents

1 Introduction and motivation of work

Many biopharmaceuticals such as monoclonal antibodies (mAbs), growth factors, protein hormones, therapeutic enzymes and coagulation factors are glycoproteins. Since it is known that glycosylation impacts essential characteristics such as stability, activity as well as immunogenicity of these proteins, regulatory agencies like the U.S. Food And Drug Administration (FDA) and European medicines agency (EMA) demand *N*-glycosylation profiles for quality control and drug release. Interestingly, however, such regulations are still absent for vaccines, where glycoproteins are often the main antigens. In contrast to other biopharmaceuticals, there is only little known so far about the impact of variations in cultivation condition on the *N*-glycosylation of vaccine components. Accordingly, little is known about the impact of differential glycosylation on vaccine efficacy and safety (e.g. [1-3]).

So far, our understanding of conditions affecting protein glycosylation, limits our ability to control final product glycosylation. In general, direct cultivation factors such as the production system, the medium composition, the pH as well as indirect factors such as the availability of sugar-nucleotides within the cell, the residence time of the *N*-glycan in the Golgi, the host cell's glycosylation machinery and the three-dimensional protein structure have been described to markedly impact glycosylation site occupancy and/or types of glycan structures attached.

The aim of this work was to investigate the impact of various process conditions on the hemagglutinin (HA) *N*-glycosylation pattern of influenza virus A (IVA) and determine possible consequences on characteristics of virus preparations for vaccine production. Therefore, high-performance capillary gel electrophoresis with laser-induced fluorescence detection- (CGE-LIF)-based glycoanalysis is applied for the comparison of HA-associated glycan pools. Different upstream processing steps for IVA production are investigated: **seed virus** (with respect to passages, strains, suppliers and adaptation status), **host cell** (with respect to cell

line and adaptation status), **cultivation scale** and **vessel** (with respect to T-flask, roller bottle and stirred tank reactor, STR, cultivations in varying volumes) and **virus production medium** (with respect to composition and trypsin activity). The main focus is laid on MDCK cell-derived IVA PR/8/34 (H1N1, in the following termed as IVA-PR8) preparations. However, for multiple cultivation conditions such as virus adaptation status, production vessels, etc. other production systems (Vero, AGE1.CR.pIX, MDCK.SUS2 cell lines; embryonated hens' eggs) or virus strains (IVA reassortant California/07/2009-like, H1N1pandemic; IVA reassortant Uruguay/716/2007-like, H3N2; IVA reassortant Victoria/210/2009-like, H3N2) complement and confirm the trend of MDCK cell-derived IVA-PR8 data.

For seed virus adaptation, factors are addressed, leading to higher virus titers in shorter time frames. Therefore, **CGE-LIF-based glycoanalysis** characterizes HA *N*-glycosylation and next-generation **pyrosequencing** confirms stable potential *N*-glycosylation sites during virus adaptation and allows the characterization of quasispecies composition. Pyrosequencing is performed in cooperation with Dr. D. Höper from the Friedrich-Loeffler-Institut (FLI, Greifswald - Insel Riems, Germany). Altogether, these data contribute to a better understanding of the common requirement for virus seed adaptation to production cell lines.

Finally, the **impact of *N*-glycosylation on immunogenicity** is addressed, which is of particular relevance for influenza vaccine potency and efficiency. This is done in cooperation with Dr. B. Lepenies and J. Hütter from the Max Planck Institute for Colloids and Interfaces (MPI-KG, Potsdam-Golm, Germany). Therefore, differentially glycosylated virus preparations are characterized in a variety of *in vitro* and *in vivo* **immunogenicity assays** (e.g. whole spleen cell assays, adaptive T cell transfer, etc.) using a transgenic mouse model expressing a $HA_{111-119}$-specific T-cell receptor (TCR-HA) presented by class II major histocompatibility complex (MHCII) molecules or a BALB/c *wt* model, respectively.

Finally, potential 'rescue mutations', allowing increased virus yields after virus adaptation, are identified. Moreover, no-, low- and high- impact process

conditions that contribute to HA *N*-glycosylation are identified. Such information is of high value since it not only allows estimating consequences of production process modifications but also allows evaluating process failures. Immunogenicity studies allow for identification of possible consequences of altered *N*-glycosylation with respect to T cell activation, proliferation and induction of HA-specific antibody (Ab) levels. Finally, the importance of *N*-glycosylation assessment for influenza virus preparations is affirmed and favorable characteristics of IVA-PR8 glycovariants are identified.

Many aspects of this thesis have been published as first/shared-first [1, 4-7] or co-author [8, 9]. Within this work quotations of these publications will not be indicated specifically. Whenever parts of phrases, phrases, parts of paragraphs or paragraphs are used, the reference will only be given after the phrase or paragraph by the reference number [reference]. Throughout this work, text from first/shared-first or co-author articles/book chapters is generally only quoted, if it was primarily written by myself for the publication. A detailed overview of publications is given in section 13, including a statement on authorship.
Since J. Hütter (shared-first co-author of [1]) may use parts of the published work for her dissertation, too, it will additionally be indicated in the beginning of a section, if text, analogous text content, structure, figures, figure legends or parts of figure legends were taken from the paper published together with J. Hütter, D. Höper, P.H. Seeberger, E. Rapp and B. Lepenies, January 2013 in *J. Immunol*. [1].

2 Theory and background

2.1 Influenza virus

Influenza virus belongs to the family of *Orthomyxoviridae*. The virions are pleomorphic, i.e. they vary in size and shape between spherical and filamentous appearances of 80 nm to 120 nm in diameter [10]. Influenza is an enveloped, negative-sense (complementary to mRNA) RNA virus with a segmented genome, coding for up to 14 viral proteins. The genome of all influenza viruses encodes for the following seven proteins: the polymerase acidic protein (**PA**), the polymerase basic protein 1 (**PB1**) and the polymerase basic protein 2 (**PB2**), nucleoprotein (**NP**), matrix protein **M1**, and the non-structural proteins **NS1** and **NS2**. PA, PB1 and PB2 represent the three subunits of the RNA-dependent RNA polymerase. The polymerase complex together with NP is associated with the viral RNA, forming a ribonucleotideprotein complex (RNP) for each segment (figure 1). The antigenic specificity of NP determines the type (A, B or C) of the influenza virus [11]. The matrix protein M1 is a structural protein. NS1 was described to inhibit mRNA transport from the nucleus and to act as an interferon (IFN) antagonist inhibiting antiviral host responses. In contrast, NS2 carries out functions during nuclear export (hence NS2 is also referred to as nuclear export protein, NEP). Besides these seven proteins, the eight segments of the genome of influenza A and B viruses (figure 1) additionally encode two glycoproteins the hemagglutinin (**HA**) and the neuramidase (**NA**), which are integrated into the viral lipid envelope. HA mediates host cell binding and following membrane fusion, whereas NA finally liberates virus progeny from the host cell. Influenza A viruses also encode for the ion channel protein **M2**. It is located within the viral envelope and is described to be activated by low pH of the endosomes, allowing protons enter the virions' core. In particular, it is highly specific for H^+ ions [12]. Moreover, non-essential proteins derived from alternative splicing or reading frames have been described, recently. These include: **PB1-F2**, **N40**, **PA-X** and **M42** [13-15].

Influenza A viruses are further categorized into different subtypes based on the antigenic specificity of their surface antigens HA and NA. Strain designations of

influenza viruses contain the type, the host of origin (only if non-human), geographical origin, strain number, year of isolation and for influenza A viruses the antigenic description of the HA and NA proteins (i.e the subtype) in parenthesis, e.g. influenza A virus/duck/USSR/695/1976 (H2N3) [11]. Up to date, 17 different HA and ten different NA subtypes are described, although not all possible combinations of both proteins occur [16].

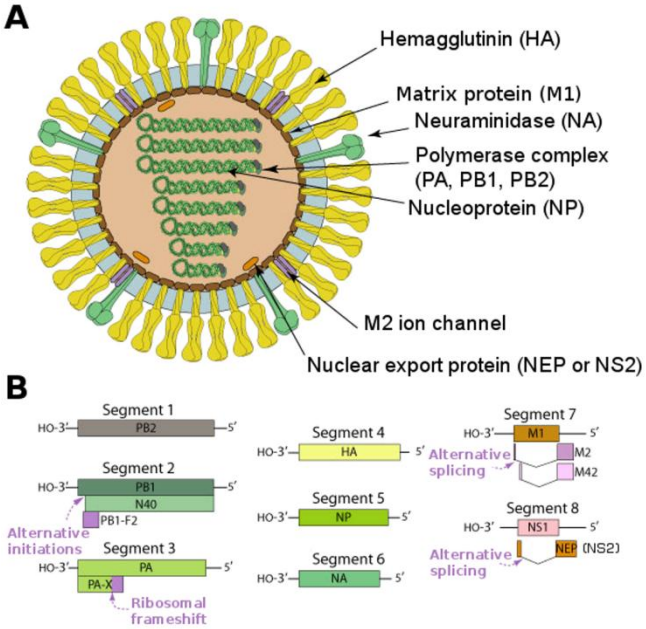


figure 1: Influenza A virus.
(A) The genome consists of 8 segments, coding for up to 14 proteins. (B) Alternative reading frames (segments 2, due to alternative initiations; segment 3, due to ribosomal frameshifts; attached boxes) and splice variants (segments 7 and 8; free boxes) allow one sequence segment to code for different proteins. Modified and reprinted with permission [17].

2.2 Influenza – a threatening human pathogen

So far, only influenza virus B and two IVA subtypes, i.e. H1N1 and H3N2, have been described to generally circulate in humans. An infection may cause severe illness, potentially leading to death. Once in a while pandemic outbreaks claim plenty of victims. Beside the pandemic threads, seasonal epidemics periodically demand profound economic losses, numerous hospitalizations and deaths each year. Up to date, hygiene and vaccination represent the best measures to prevent infection and resulting possible health complications. The periodical reoccurrence of pandemic and epidemic influenza outbreaks and hence the need

for seasonal vaccine reformulation is mainly attributed to the virus' ability to rapidly adapt to new environments.

2.3 Virus adaptation

On the one hand, new influenza variants can originate from the virus' ability to newly reassort (genetic shift) [4, 18]. On the other hand, the high error rate of the viral polymerase raises constantly new virus variants [19], which only differ in single or few amino acid positions, resulting in variations of the virus genome. Further, natural selection leads to the adaptation of a given virus as an evolutionary response to 'new-host-pressure' [18]. The frequency of a virus variant in a population largely depends on its ability to survive and reproduce – i.e. its fitness [20]. However, if coupled to high fitness genotypes, low fitness virus variants can be maintained at higher levels than expected [4, 21].

Virus adaptation is one of the most important processes in virus evolution, and a crucial factor to be taken into account for seasonal and pandemic vaccine production. Escape from immune pressure, balancing host cell receptor binding avidity of infecting virus with the release of progeny virus as well as adjustment to altered endosomal pH-values or to different, specific sialic acid containing host cell receptors have been described as driving forces for adaptation processes in virus evolution [22-25]. On the one hand, adaptation allows the virus to cross species boarders, evade immune or therapeutic pressures and optimize its replication in a given host system [26]. On the other hand, it challenges manufacturers to adapt emerging strains to existing egg-based or cell-culture-based system processes to obtain maximum yields for formulation of potent vaccines [4, 22, 27].

2.4 Quasispecies

In general, due to the viral polymerase's error-prone nature, influenza replication and hence also adaptation processes, result in the co-existence of related virus subpopulations on the genomic and hence often on the proteomic level. Such a population of related, though differing virus variants is referred to as a quasispecies [28-30]. The consensus sequence of such a quasispecies

represents the most frequent information for each position. Often, no single virus of a vius population carries the consensus sequence, because no mutation is present at a sufficient high level.

During most infections, i.e. during cell culture-based virus production as well as during illness, more than one virus particle is involved. Furthermore, each infected cell produces plenty of slightly differeing progeny viruses. Altogether, this suggests that hardly one seed virus stock matches the other. Moreover, it implies that a virus strain obtained from one supplier most likely not matches with the same virus strain obtained from another supplier. Hence, varying virus production yields during vaccine production processes and differing courses of disease during illness may result from such differing viral quasispecies compositions. Only recently, new deep sequencing methods, e.g. next generation pyrosequencing, allow the detection of different variants within such a quasispecies.

2.5 Glycovariants

The variation on the genomic level is further increased by the complexity of protein *N*-glycosylation of the two viral surface proteins hemagglutinin (HA) and neuramidase (NA). Glycoproteins can be considered as a collection of different glycoforms or glycosylation variants [4, 31]. They vary in glycosylation site occupancy (macroheterogeneity) and in structure and composition of sugar residues (microheterogeneity) attached to the protein backbone [4].

2.6 Influenza virus replication

The glycoprotein HA plays a key role in virus replication and therefore is often affected during adaptation processes, resulting in altered amino acid sequences and/or altered HA *N*-glycosylation. However, other viral proteins also contribute to virus replication and may also be affected during virus adaptation. In order to help interpreting virus genome sequencing data presented in the result section, a short overview of the IVA replication cycle will be given in the following.

The replication cycle of influenza A viruses comprises virus **adsorption**, **entry**, **uncoating**, mRNA synthesis (**transcription**) and **replication of viral RNA**,

synthesis of viral proteins (**translation**), virus **assembly, budding** and final **release** of virus progeny [32, 33].

2.6.1 Adsorption, entry and uncoating

In the first step the virus binds to the host cell, which is mediated by interactions of the receptor binding domain on the distal tip of the viral HA molecule with sialic acid residues of host cell receptors (figure 2). Here, different HA subtypes (H1-17) have different specificities for sialic acid bond to galactose by either α2,3 or α2,6 linkage. A change of this specificity, e.g. due to substituted amino acid residues within the receptor binding pocket, can allow to cross species borders. In cells of birds intestines for instance α2,3 linkages are predominant, whereas in cells of human trachea mainly α2,6 linkages are expressed. In contrast, cells in pigs trachea contain both α2,3 and α2,6 sialic acid linkages. This co-expression of α2,3 and α2,6 linkages makes them susceptible for avian and human influenza strains and turns them into a mixing vessel facilitating genetic/antigenic drift (via

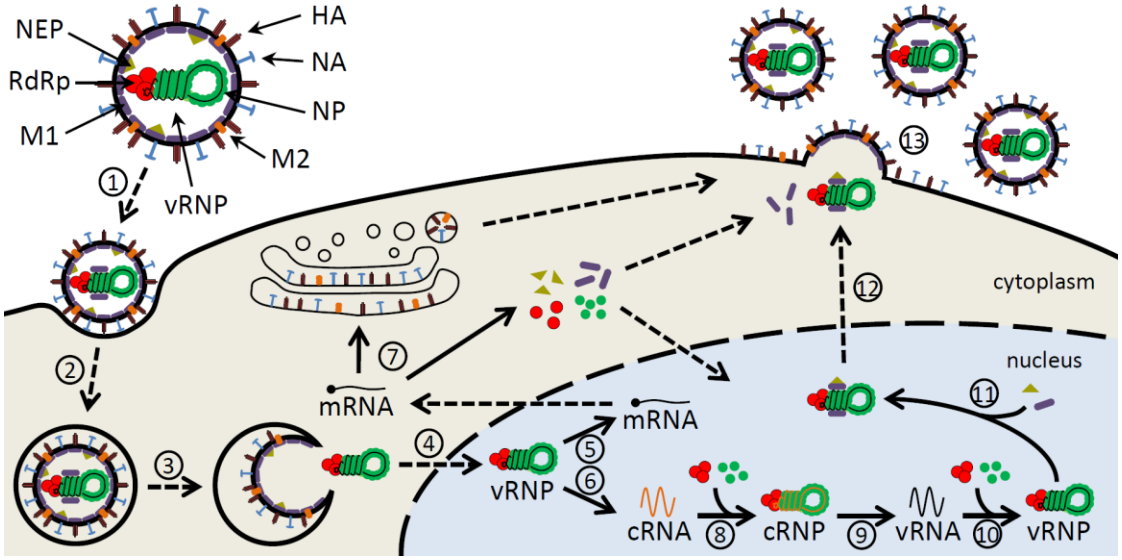


figure 2: Schematic influenza A virus replication cycle.
The model was simplified by displaying only one vRNP within the virus particle, instead of eight and by omitting nonstructural proteins. Transport processes are indicated by dashed arrows, whereas synthesis and protein binding are indicated by solid arrows. The virus replication cycle comprises (1) virus adsorption, (2) entry by endocytosis, (3) uncoating, (4) nuclear import, (5) transcription and (6) replication of viral RNA (cRNA synthesis), (7) translation of viral proteins, (8) encapsidation of cRNA with newly synthesized NP and polymerase proteins, (9) replication of viral RNA (vRNA synthesis), (10) encapsidation of vRNA with newly synthesized NP and polymerase proteins, (11) M1 and NEP binding mediate (12) nuclear export, (13) virus assembly, budding and final release [33]. Modified and reprinted with permission. Copyright 2012, American Society for Microbiology.

mutation) and shift (reassortment). Furthermore, tissue tropism is determined by a cleavage site within the HA molecule. The HA molecule, as it is synthesized, is referred to as HA_0 (molecular weight, MW, approximately 70 kDa). HA_0 must be cleaved e.g. by host proteases into a HA_1 (MW approximately 50 kDa) and a HA_2 (MW approximately 20 kDa) subunit to activate virus infectivity [34, 35]. In natively folded, cleaved HA molecules the HA_1 and HA_2 subunits are linked with a single disulfide bond and are considered to be in a metastable state. Adsorbed influenza virus is internalized within vesicles by receptor-mediated endocytosis. These vesicles begin to fuse with endosomes [12]. On the one hand increasingly acidic pH values activate the M2 ion channel [36] and permit the H^+ ion flow into the virion, destabilizing protein-protein interactions and RNP - M1 interactions [12, 37, 38]. On the other hand the low-pH induces a conformational change of the cleaved, metastable HA molecule - especially the HA_2 subunit refolds – leading to fusion of viral and cellular membrane [12] and thus allowing the viral RNA to enter the host cell's cytoplasm.

2.6.2 Transcription and replication of the virus genome

Nuclear localization signals within viral proteins trigger specific interactions with nuclear transport complexes (so called importins) that interact with nuclear pore complexes of the host cell, thus allowing the transport of viral RNPs into the cellular nucleus ([39], figure 2). The synthesis of viral mRNA is dependent on the cellular RNA polymerase II. The PB2 subunit of the heterotrimeric viral polymerase recognizes and binds the 5' cap of cellular mRNA transcripts of the polymerase II [40-42]. In a next step, the endonuclease active part of the PA subunit cleaves off the capped RNA fragments from the transcripts. This procedure is also known as cap-snatching [42-44]. The fragments of 10 to 13 nucleotides serve as primers for the viral polymerase and are required for the initiation of viral mRNA synthesis [12]. PB1 binds the vRNA, which serves as template. A conserved domain within the PB1 subunit catalyzes the elongation of the mRNA [42, 45, 46] until a stretch of 4 to 7 uridine residues is reached, where

transcription is terminated and polyadenylation occurs. Consequently, viral and host cell mRNAs are structurally indistinguishable.

In contrast to viral mRNA synthesis, the replication of the full-length RNA variants, the template cRNA and the vRNA, do not require any primers and are not terminated at the poly A site (reviewed in [12]).

2.6.3 Translation

PB1, PB2, PA, NP, M1 and NS2 proteins are essential for the nuclear export of the vRNA (figure 2). These proteins are translated from the viral mRNA in the cytoplasm and are, as described before, actively transported (with the exception of NS2, which is supposed to be small enough for diffusion through pores) into the cellular nucleus. Here, vRNA, PA, PB1, PB2 and NP assemble to form new RNPs. M1 is proposed to bind the RNP complex and NS2 in turn is proposed to bind M1. NS2 also interacts with exportin1 (also known as chromosome region maintenance protein 1, CRM1), which in turn binds a small GTPase called Ran, which again needs to be associated with GTP before the whole RNP-M1-NS2-CRM1-RanGTP-complex can leave the nucleus (reviewed in [39]). Additionally, also NS1 carries nuclear localization signals and is transported into the host's nucleus, where it does not interact with the RNP-export complex but antagonizes the cellular immune response [39, 47].

The viral membrane proteins HA, NA and M2 are synthesized on membrane-bound ribosomes and are flipped upon a signal sequence across the membrane into the endoplasmatic reticulum (ER, figure 2). The glycoproteins HA and NA are furthermore modified by *N*-linked glycosylation in the ER and the Golgi whereas HA and M2 are additionally palmitoylated in the Golgi [12].

2.6.4 Assembly, budding and release

After completion of processing during the transport from the ER through the *cis*-, *mid*- to *trans*-Golgi, HA, NA and M2 are transported by separate transport vesicles to the plasma membrane domains ([48], figure 2). In polarized epithelial cells, such as MDCK or primary bronchial and lung epithelial cells, influenza viruses assemble at the apical surface of the cells [49], in so-called lipid rafts

(cholesterol- and sphingolipid-enriched regions within the plasma membrane). HA and NA both possess their own signal(s) for apical transport and lipid raft association [49]. Also M2, M1 and vRNPs assemble at the budding site, though there is still some discourse about the exact mechanisms (reviewed in [49, 50]). Also, the exact processes involved in budding as well as final scission of the new progeny viruses from the cellular membrane are still discussed (reviewed in [49, 50]). Finally NA prevents virus progeny to aggregate to itself or the cell surface by removing sialic acid residues thus allowing the spread of virus progeny to other cells [12, 51].

2.6.5 Role of non-structural proteins

So far, it is unknown how the multiple functions of NS1 contribute to the IVA phenotype (e.g. reviewed in [52]). One function of NS1 is cleavage and polyadenylation specificity factor- (CPSF-) binding, hence suppressing cellular gene expression. A second function suppresses the export of cellular mRNAs into the cytoplasm, impairing cellular protein synthesis and thereby suppressing the host's IFN response. A third function is dsRNA- and tripartite motif-containing protein25- (TRIM25-) binding, which prevent retinoic acid inducible gene I (RIG-I) mediated IFN-induction. The forth function is the activation of the phosphatinositol 3-kinase/Akt pathway suppressing apoptosis in infected cells. The interaction of NS1 with eukaryotic initiation factor 4 G1 (eIF4G1) represents a fifth function, stimulating the translation of viral transcripts.

2.7 Anti-influenza drugs and influenza vaccines

All food and drug administration (FDA) approved pharmaceutical substances to treat influenza infections interfere with the virus replication cycle: Zanamivir (Relenza) and oseltamivir phosphate (Tamiflu) are NA inhibitors, whereas amantadine (Symmetrel) and rimantadine (Flumadine) block the M2 ion channels [53]. However, due to the virus' ability to rapidly adapt to environmental pressures, developing drug resistances may cause therapy failing. Therefore, prevention of infection in the first place is highly recommended. In this regard, hygiene and vaccination represent the best measures. Though, also for

vaccination, the virus' ability to rapidly adapt to changing environments makes seasonal reformulations of influenza vaccines necessary. In general, the annual production circles start with the definition of virus strains recommended for the next season's vaccine formulation by the World Health Organization (WHO). Often the recommendation matches the circulating strains, however some risks remain that wrong virus strains are selected. For virus seed preparation multiple virus variant selection steps are required: In order to minimize the risk of contamination with other human pathogens, clinical specimens of the strains recommended by the WHO are usually blind-passaged in embryonated chicken eggs by WHO Collaborating Centers. In general, human isolates replicate poorly in eggs. The manufacturers then usually select variants that replicate well to be reassorted to high-yield laboratory viruses to generate virus seeds used in production. Due to possible antigenic drift during each of these steps antigen identity testing and sequence analyses are required [4, 54, 55].

Currently, most commercially available influenza vaccines are inactivated vaccines produced in egg/cell culture. These are classified into whole virus, split (by detergent disrupted virus particles) or subunit (purified HA and NA) vaccines. Seasonal vaccine production requires approximately six month from the definition of WHO to commercialization [54]. Principally, seasonal vaccine formulations are trivalent, i.e. they comprise two IVA strains and one influenza virus B strain. In contrast, the production of a pandemic vaccine, which is usually monovalent, only comprising the pandemic strain, takes about five months [56].

In some places such as the United States, alternatively to inactivated vaccines cold-adapted, live attenuated influenza whole virus (LAIV) vaccines are available [57, 58]. These are usually administered as nasal spray.

In addition, other vaccine platforms using recombinant proteins/peptides, DNA, virus-like particles-, virosomes or vector-based vaccines have recently been described [59]. However, most liscensed vaccines, so far, are not from those other platforms.

2.8 Cell culture-based inactivated influenza vaccines

So far, most influenza vaccines are still produced in embryonated hen's eggs, though an increasing number of cell culture-based processes are being established (see section 2.8.1). Advantages of such cell culture-based processes include independence from egg supply, skills and equipment is transferable to production of other cell culture-derived vaccines such as rabies, enormous reduction of infectious solid waste (approx. 80 % of mass of eggs), possible rapid process adjustment to better match supply with vaccine demand e.g. during pandemics and finally cell culture-derived vaccines bear no risk of anaphylactic reactions caused by egg proteins [6, 60].

2.8.1 Host cells used in production

So far, few cell culture-based processes using MDCK cells have been licensed (e.g. Novartis Behring for Optaflu® and Celtura). MDCK cells have been comprehensively characterized since their establishment in 1958 by Madin and Darby [6]: absence of contaminating viruses, resistance to prion infections and ideal properties for influenza virus replication have been demonstrated [61]. Suspension as well as adherently growing cell clones are being used [6, 62]. As an alternative, other cell lines such as Vero (Baxter International Inc. for Celvapan®, PreFluCel®, Vepacel®; [61, 63-66]), AGE1.CR [67], AGE1.CR.pIX [7, 67], PER.C6 [61, 68, 69], EBx [61], Eb14® [62], PBS-1 [70] or SJPL cells [61, 62, 71] have been described for influenza virus production [6]. Here, the importance of host choice shall be emphazised, since the selected production cell line not only affects virus propagation dynamics and hence harvest time points and virus yields but most likely also the *N*-glycosylation pattern of HA [6, 72, 73]. However, for virus seed preparation in all egg- and cell culture-based processes, multiple virus adaptation steps are usually required to achieve optimal yields.

2.8.2 Preparation of virus working seeds

In general, human IVA isolates replicate poorly in eggs. Nevertheless, in order to minimize the risk of contamination with adventitious agents, clinical specimens of the strains recommended by the WHO for the next season's vaccine formulation are usually blind-passaged in embryonated chicken eggs by WHO Collaborating Centers. Usually, the manufacturers select virus variants that replicate well in the particular final production host system, e.g. egg, MDCK cells, Vero cells, etc. and reassort these to even higher yielding laboratory viruses to generate virus seeds used in production (figure 3, [4, 54, 55]).

Alternatively, propagation of human influenza virus isolates in mammalian cells would circumvent the passaging in embryonated hen's eggs and thus prevent the selection of mutations in the HA, causing altered antigenic properties [74, 75]. One option for cell culture-derived virus production is the use of comprehensively characterized MDCK cells. However, here the reduction of contaminating pathogens due to broad species border may be smaller. The use of avian cell lines such as AGE1.CR, AGE1.CR.pIX [76] and EB66 [77] for isolation and production would probably combine the reduction of adventitious agents with the advantages of cell culture technologies, eliminating the need for additional adaptation steps [6, 7].

2.8.3 Vaccine production process

While virus seeds for production are generated, production cells are expanded to desired quantities (figure 3): in lab and pilot scale sufficient virus is produced under good manufacturing practice (GMP) conditions for phase I and II clinical trials to demonstrate the product's immunogenicity and safety; in industrial scale for final commercial vaccine production [60]. After cell expansion (cell growth phase), cells are infected with virus seed (virus production phase). 24 hours post infection (hpi) to 96 hpi the virus is harvested. During following downstream processing the harvest is inactivated (e.g. by formaldehyde or β-propiolactone), virus is concentrated, purified (e.g. using diafiltration or centrifugation) and contaminating host cell DNA may be degraded, e.g. by benzonase treatment or

removed by ion-exchange chromatography. For split and subunit vaccines the virus is again diluted. Added detergents such as Triton X-100, sodium lauryl sulphate or Tween 80 basically extract proteins from the viral envelope. Purification may be performed by e.g. sucrose gradient or diafiltration and the dilution in the formulation buffer result in final vaccines. Sometimes, adjuvants, stabilizers and/or preservatives are added [60, 78]. The final inactivated, cell-culture-derived vaccine for human use must not contain more than 10 ng DNA per dose [79]. Furthermore, each dose should contain at least 15 µg HA per strain and the HA - total protein content ratio should be within the limits approved by national regulatory authorities [79]. However, total protein content including HA should not exceed 100 µg per strain, i.e. 300 µg per dose of a trivalent vaccine. For subunit vaccines the total protein content is even limited to a maximum of 40 µg per strain [79]. Additionally to HA, vaccines usually contain the lower abundant viral glycoprotein NA, which represents the second important antigen in influenza vaccines.

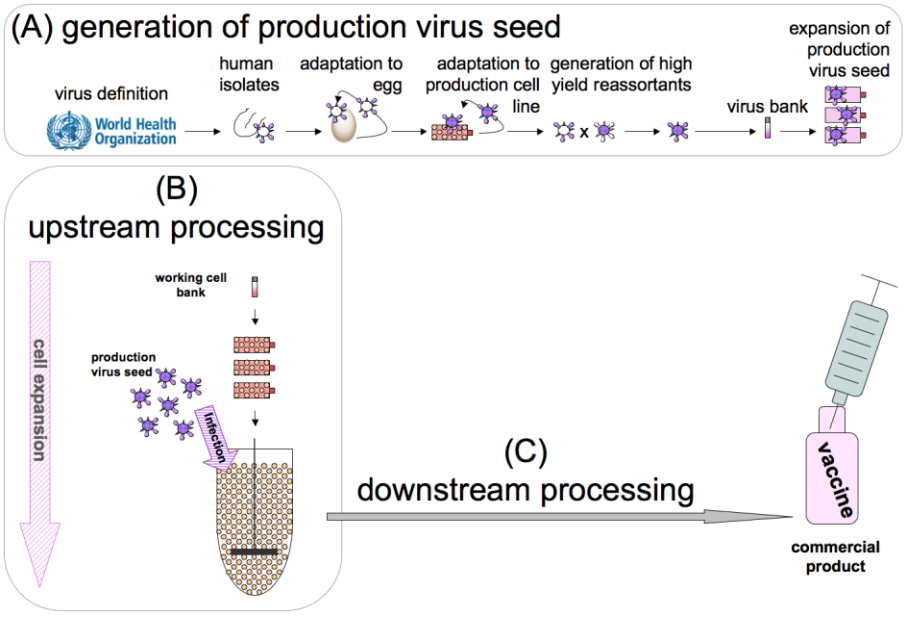


figure 3: Cell culture-based influenza vaccine production process.
(A) The generation of the production virus seed: The WHO defines the strains for the next year's influenza vaccine. Human isolates of these recommended strains are blind-passaged in embryonated chicken eggs by WHO Collaborating Centers and are distributed to vaccine manufacturers. Here, well replicating variants are selected for the specific production system. These are reassorted with a high yield laboratory strain such as IVA PR/8/34 (H1N1) to high yield production strains, carrying the recommended HA and NA antigens. (B) Upstream processing: cell expansion from the cell bank to lab or industrial scale for final vaccine production. (C) Downstream processing and final fill and finish make the commercial product.

Altogether, the ready to administer vaccine has passed through a multitude of production steps. Hence, various up- and downstream process conditions such as production cells, production scales, media, production temperatures, harvest time points, etc. may vary between different established processes (reviewed in [62]). However, all processes have the vaccine's major antigen in common - the HA, which is able to induce strong and protective immune responses. Here, questions arise concerning the impact of cultivation conditions such as virus adaptation, media, production scales, etc. on the *N*-glycosylation of HA and its impact on virus immunogenicity and hence on vaccine quality and savety. Glycoanalysis of NA is beyond the scope of this work and furthermore, due to lower abundance in the virus particle, bigger sample volumes would be required.

2.9 N-glycoproteins and their synthesis

Many antigenic proteins used for vaccination are *N*-glycoproteins. These glycoproteins such as HA from IVA (figure 4) often play key roles in virus replication, e.g. by enabling the attachment to and the infection of host cells. Glycoproteins can be considered as a collection of different glycoforms or glycosylation variants [31], varying in microheterogeneity as well macroheterogeneity [4]. With respect to the HA of IVA, depending on various factors such as protein conformation and host, various *N*-glycan structures of the high mannose, the hybrid and the complex type have been detected attached to the HA protein [73, 80, 81]. In general, differences in *N*-glycosylation may impact proteins' characteristics such as Ab dependent cell-mediated cytotoxicity (ADCC, [82]), specific activity [83], antigenicity [83-85], binding avidity [86], specificity [87], immunogenicity and virulence [88]. For IVA in particular, *N*-glycosylation of HA was described to impact on protein folding, receptor binding activity, -avidity and –specificity, evasion of host immunity, protein cleavability as well as the recognition by the host's innate immunity e.g. via calcium-dependent (C-type) lectins (summarized in [88]).

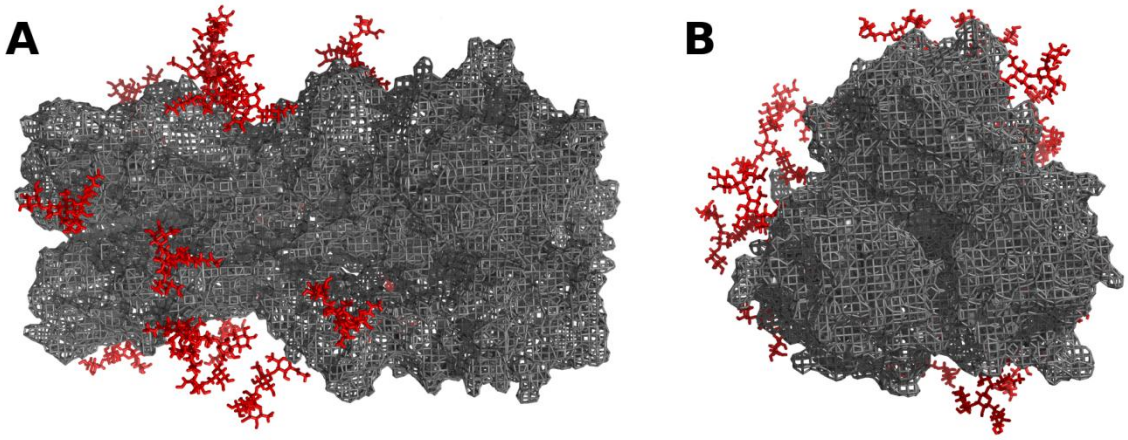


figure 4: Three-dimensional, spherical structure of trimeric, *N*-glycosylated IVA-PR8 HA. Attached *N*-glycans are couloured in red. (A) Side and (B) top view. The PDB entry 1RU7 and Pymol (v0.99, DeLano Scientific LLC, California, USA) software were used for structure display. *N*-glycan modeling was performed with GlyProt [89].

Viruses use the host cells for protein synthesis as well as the host cells' glycosylation machinery for modifying their glycoproteins. The process of protein *N*-glycosylation takes place in different cell compartments: Briefly, in the cytoplasm of eukaryotes monosaccharyl-transferases stepwise attach seven monosaccharides (a to g, figure 5) from nucleotide sugar donors (UDP-acetylglucosamines, UDP-GlcNAc; GDP-Mannose, GDP-Man) to an ER membrane-bound lipid carrier (dolichol-pyrophosphate): two GlcNAc followed by five Man residues forming two branches (reviewed in [90]). Subsequently, the sugar moiety is flipped into the lumen of the ER by a still controversially discussed mechanism (reviewed in [91]). In the ER, additional four Man residues (h to k, figure 5) and three Glc residues (l to n, figure 5) are attached by different glycosyltransferases from dolichylphosphate-linked monosaccharides (Dol-P-Man, Dol-P-Glc), finally forming the tri-antennary tetradecasaccharide (Glc3Man9GlcNAc2) core *N*-glycan structure [92]. The last glucose residue (n, figure 5) is required for recognition by the oligosaccharyltransferase (OST) complex [93], which transfers the core *N*-glycan structure co-translationally from the membrane-bound dolichol-pyrophosphate to the side chain nitrogen of the Asn of the *N*-glycosylation motif Asn-X-Ser/Thr of a growing polypeptide chain. In this motif, X represents any amino acid but proline. Proline is suggested to

impede the formation of a required loop, bringing the hydroxyl groups of Ser/Thr into closer contact with Asn. This close contact increases the nucleophilicity of Asn [94, 95], allowing the covalent attachment by a *N*-glycosidic bond of the sugar core *N*-glycan structure to the protein backbone. Gavel *et al.* estimated, that 90 % of such potential *N*-glycosylation sites are actually glycosylated [96], which leads to differing glycosylation site occupancy – so-called macroheterogeneity of glycoproteins. In mammalian cells two OST complexes are expressed, varying in substrate selectivity [97]: one uses complete whereas the second may also use incomplete oligosaccharide core structures and has a higher maximal reaction velocity. As soon as the oligosaccharide is attached to the protein-backbone a glucosidase (type I) detaches the last Glc residue (n, figure 5), and a further glucosidase (type II) removes the second Glc residue (m, figure 5). The monoglucosylated core structure is bound by calnexin and/or calreticulin, which support proper protein folding. The removal of the remaining Glc residue (l, figure 5) by glucosidase II allows properly folded proteins to leave the ER and enter the Golgi complex. Incorrectly folded proteins are either modified with a new Glc residue allowing the binding to calnexin and/or calreticulin again or ER-associated degradation is initiated (reviewed in [98]). Mannose residues are trimmed off (f, g, i, k, figure 5) by different mannosidases and a GlcNAc residue is added to mannose residue d (figure 5). After removal of mannose residues (h, j, figure 5) further sugars such as GlcNAc, galactose, sialic acid and fucose residues are added in the Golgi complex by a variety of different glycosyltransferases [99]. The Microheterogeneity of glycoproteins describes the resulting variance of attached glycans with respect to their sugar residue composition. Depending on the extent and the types of modification within the Golgi complex, final *N*-glycans are divided into three classes: high mannose, hybrid and complex *N*-glycan structures. But even complex IVA-derived glycoproteins, e.g. HA molecules, lack sialic acid residues, which is probably attributed to the neuramidase activity [73, 100, 101].

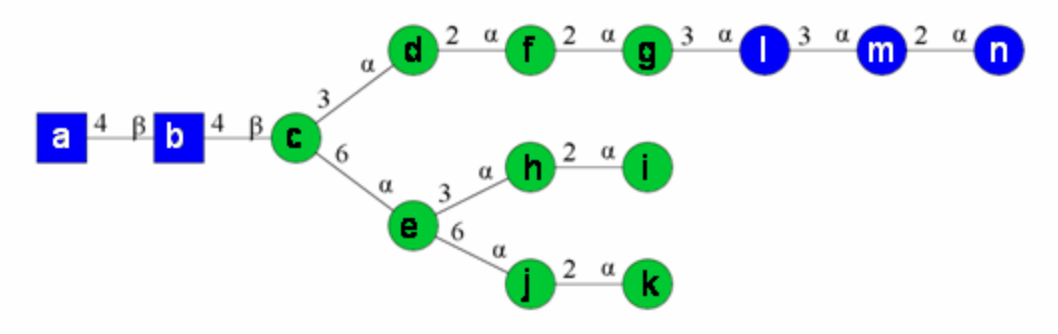


figure 5: Scheme of *N*-linked precursor oligosaccharide.
In the cytoplasm, monosaccharyltransferases stepwise attach monosaccharides (a-g; ■,*N*-acetylglucosamines, GlcNAc; ●, mannose, Man) to ER membrane-bound dolichol-pyrophosphate. Subsequently the sugar moiety is flipped into the lumen of the ER. Different glycosyltransferases stepwise add additional four Man and three glucose (●, Glc) residues (h-n). The oligosaccharyltransferase (OST) complex transfers this precursor *N*-glycan structure to the nitrogen side chain of an asparagine of the *N*-glycosylation motif. Glucosidase I removes the last (n) and glucosidase II removes the second Glc residue (m). The first Glc residue (l) now allows binding to calnexin and/or calreticulin, supporting proper protein folding. Glucosidase II then also removes the first Glc residue (l), allowing the properly folded protein to leave the ER.

At first sight, this highly conserved biosynthetic pathway appears rather energetically unfavorable. Why should a core oligosaccharide be build up and straight after its translocation to a polypeptide chain be trimmed down again, just to re-synthesize it once again with different sugars? The answer is that the different stages of *N*-glycosylation serve important functions, such as proper protein folding and quality control in the ER, intracellular transport and targeting in the ER as well as in the Golgi complex and finally various different functions specific for the mature glycoprotein [90].

What does all this mean for the HA of the IVA-PR8 used throughout this study? Each HA monomer of the RKI-strain carries seven potential *N*-glycosylation sites, five within the HA_1 (AA positions 27, 28, 40, 286, 304) and two within the HA_2 chain (AA positions 498, 557). However, an analysis of the HA AA sequence by NetNGlyc 1.0 [102] predicts no *N*-glycosylation on residue 27 due to a low *N*-glycosylation potential of 0.4 [103]. Since naturally HA occurs as a trimeric structure, this makes 18 potentially *N*-glycosylated sites for each HA trimer (figure 6). Taking micro- and macroheterogeneity into account makes numerous different isoforms of one and the same HA protein possible.

The *N*-glycans attached are likely to determine protein characteristics of HA: e.g. it was decribed that HA with terminal mannose induces lower hemagglutination

inhibition (HAI) than HA with complex structures or single GlcNAc residues [2, 3]. Furthermore, Wang et al. showed that HA carrying single GlcNAc residues induced Abs with higher binding affinity and neutralization activity than fully glycosylated HA [104]. Interestingly and somewhat contradictory, Lin *et al.* reported that high mannose glycan structures lead to higher levels of HA-specific Ab titers due to different antigen presentation efficiencies [2]. However, not only the type of attached glycans but also the numbers and/or positions of potential HA *N*-glycosylation sites may significantly impact immunogenicity as well as antigenicity. Sun *et al.* demonstrated that the introduction of HA *N*-glycosylation sites attenuated highly virulent viruses, whereas the removal of sites resulted in increased virulence of lower virulent strains [88]. Furthermore, other factors such as the presence of a HA polybasic cleavage site may contribute to the virus' pathogenicity.[105].

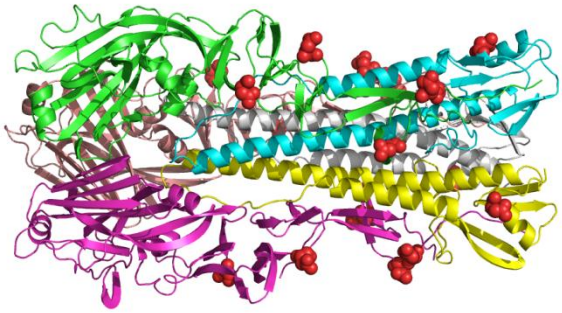


figure 6: Three-dimensional cartoon structure of trimeric IVA-PR8 HA.
HA_1 chains are coloured in green, pink and brown, whereas HA_2 chains are coloured in yellow, cyan and grey. Potentially *N*-glycosylated asparagine residues are highlighted as red spheres. The PDB entry 1RU7 and Pymol (v0.99, DeLano Scientific LLC, California, USA) software were used for structure display.

2.10 Analytics

2.10.1 N-glycan analysis

N-glycosylation modifications are as complex and diverse as the methods available for their characterization. Therefore, a complete overview and detailed description of methods goes far beyond the scope of this work. However, a sketch of available methods for glycoanalysis is given in the following:

So far, no single method is able to provide all structural and site-specific information for the complete characterization of a glycoprotein, i.e. sugar residue

sequence assigned to a specific *N*-glycosylation site. All glycosylated molecules can be detected by the periodic acid-Schiff (PAS) reaction [106], by which sugars are oxidized and subsequently react with the Schiff reagent to a pink color. Alternatively, glycan structures whether attached to or released from proteins can be detected by specific lectins or antibodies (Ab). Due to the availability of a broad range of different specific lectins and Ab some structural information can be gained by evaluating glycan's binding characteristics [107]. In microarray approaches this principle of lectin or Ab-based *N*-glycan detection was transferred to high-throughput applications [108]. Furthermore, different glycoforms of a protein, varying in molecular weight and/or isoelectric point, can be visualized by one- or two-dimensional gel electrophoresis.

The *N*-glycans may be released from the protein backbone either chemically by ozonolysis (e.g. hydrazinolysis) or enzymatically by specific enzymes, e.g. the peptide: *N*-glycosidase F (PNGaseF, [109], cleaves Asn-GlcNAc linkage of almost all *N*-glycans) or endoglycosidase H (EndoH, cleaves GlcNAc-GlcNAc linkage in core of high mannose and some hybrid *N*-glycans). In contrast to all other available enzymes, PNGaseF removes almost all types of *N*-glycans and the cleavage between Asn-GlcNAc retains the complete glycan structure. Other enzymes are favoured for structure characterization whereas PNGaseF predominantly serves unspecific *N*-glycan removal. However, factors such as accessibility of the *N*-glycan and the glycan structure itself (e.g. α1-3 core fucosylation) may inhibit PNGaseF digestion.

Unlabeled mono- and oligosaccharides may be detected by pulsed amperometric detection (PAD, reviewed in [110]) or mass spectrometry (reviewed in [111]). Otherwise, glycans can be radio-labeled or fluorescence-labeled for detection [109]. The most common approach is fluorescence-labeling by reductive amination commonly using sodium cyanoborohydride ($NaBH_3CN$, [112-114]) or more recently the non-toxic reducing agent 2-picoline borane [115]. Depending on the label requirements diverse labels are available such as 2-amino-antranillic acid (2-AA), 2-aminobenzamide (2-AB), 8-aminonaphthalene-1,3,6-trisulfonic acid (ANTS) and 8-aminopyrene-1,3,6-trisulfonic acid (APTS, reviewed in [111]).

After labeling, the *N*-glycans may be separated in an electric field for instance by gel electrophoresis (fluorescence-assisted carbohydrate electrophoresis, FACE, [116], kit commercially available from www.prozyme.de), capillary electrophoresis (CE, [117], charge-based separation) or capillary gel electrophoresis (CGE, [112, 113], size-based and charge-based separation, separation of α- and β-isomers possible). An alternative is the chromatographic separation. Here, multiple methods have been described including capillary electrochromatography ([118, 119], charged-based separation with chromatographic interaction for increased selectivity), gas-liquid chromatography (determination of monosaccharide linkage and position in complex *N*-glycans, [120]), or most commonly by adsorption using high-performance liquid chromatography (HPLC, [121]). Most widely applied is the HPLC in hydrophilic interaction chromatography mode (HILIC, reviewed in [122]). For this, a database was established by Rudd *et al.*, allowing 2-AB labeled *N*-glycan structure assignment [123]. Other modes that have been applied for *N*-glycan characterization include anion-exchange HPLC [124], high-performance (or high pH) anion exchange chromatography (HPAEC), porous graphitized carbon (PGC) HPLC and reversed phase (RP) HPLC (reviewed in [111, 125]). Furthermore, mixed modes such as HILIC in combination with anion exchange separations have been published [126].

HPLC-based techniques may be combined with mass spectrometry (MS) detection for on-line or off-line *N*-glycan structure elucidation. However, the great advantage of the CGE-LIF-based separation technique is its suitability for high-throughput applications using multiplex DNA sequencers (e.g. ABI PRISM 3100-Avant Genetic Analyzer from Applied Biosystems, Foster City, California, USA). However, the identification of the detected *N*-glycan structures is elaborate if no detailed database for structure identification is available and requires sequential exoglycosidase digestions as well as sample reanalysis. Since such a detailed and comprehensive database for *N*-glycan structure identification is currently being established in our group, glycoprofiling within this study was performed by CGE-LIF-based glycoanalysis.

2.10.2 CGE-LIF-based glycoanalysis

As already addressed in the previous section, CGE-LIF-based glycoanalysis is used in this work to investigate the impact of different process conditions on the HA *N*-glycosylation. The following terms will be used frequently within this study: The ***N*-glycosylation pattern** of a **glycoprotein** includes information regarding the structures and abundances of attached sugars and if available information regarding the location of modification. As described earlier, glycoproteins are collections of different glycoforms/glycovariants, which vary in attached oligosaccharide structures (**microheterogeneity**) as well as in potential glycosylation site (Asn – X – Ser/Thr) occupancy (**macroheterogeneity**). Due to this heterogeneity and due to limitations of the available glycoanalysis methods, the *N*-glycosylation pattern of a glycoprotein can only be approximated by profiling the *N*-glycosylation of a **glycoprotein pool** (figure 7A). Specific enzymes such as PNGaseF allow release of the ***N*-glycan pool** of such a glycoprotein pool (set of different glycoforms of a protein, figure 7B). CGE-LIF-based glycoprofiling of a labeled *N*-glycan pool results in **electropherograms**, where **relative fluorescence units** (**RFU**) are plotted over **datapoints**. These *N*-glycosylation electropherograms show high batch-to-batch as well as day-to-day variations regarding the x-axis (figure 7C) due to minimal changes in buffer concentration, temperature, etc.. However, normalization of raw data to an internal DNA basepair (bp) standard results in ***N*-glycosylation fingerprints** (figure 7D), also referred to as ***N*-glycan fingerprints**, of a glycoprotein pool (patented strategy, [127, 128]). In such fingerprints RFU are plotted over **normalized migration time units** (**MTU'**, figure 7D), which are equivalent to **bp**. In electropherograms as well as in fingerprints one peak corresponds to at least one distinct *N*-glycan structure. Also relative quantification of detected *N*-glycan peaks is possible. Therefore, the y-axis is normalized by calculating the **total peak height** (**TPH**, summed peak height of all specific peaks) and setting it to 100 %. Consequently the **relative peak height** (**RPH**) in % of each single peak corresponds to its **relative *N*-glycan structure abundance** (figure 7E). By

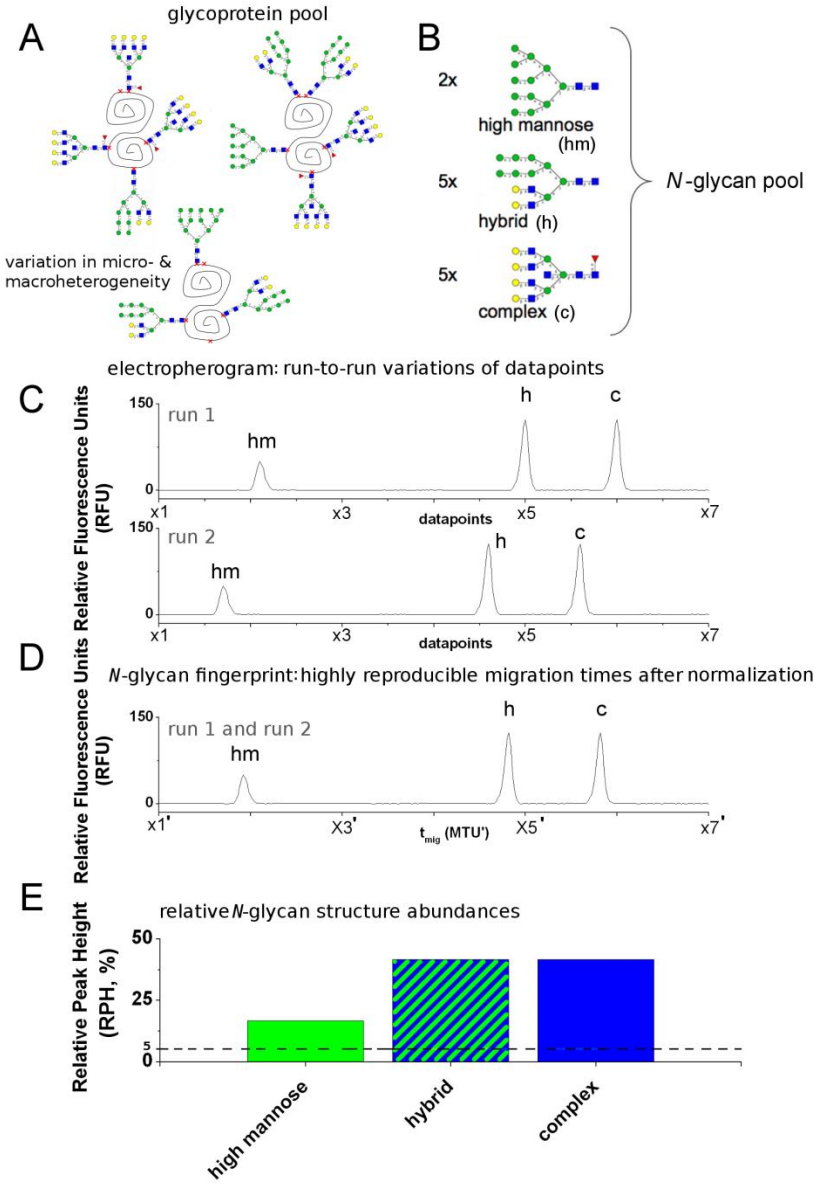


figure 7: Relevant terms and data processing steps for CGE-LIF-based *N*-glycoanalysis.
(A) Glycoproteins can be considered as a collection of different glycoforms varying in attached sugars (microheterogeneity) as well as glycosylation site occupancy (macroheterogeneity). The *N*-glycosylation pattern of a glycoprotein can only be approximated. (B) Depending on their composition, *N*-glycans are classified into high mannose (hm), hybrid (h) and complex (c) structures. (C) CGE-LIF-based glyco-profiling results in electropherograms, where relative fluorescence units (RFU) are plotted over datapoints. Datapoints vary dramatically from run to run. (D) Data processing and normalization to an internal standard result in reproducible *N*-glycosylation/*N*-glycan fingerprints of a given glycoprotein pool, in which RFU are plotted over migration times (t_{mig}) in normalized migration time units (MTU'). In electropherograms as well as in fingerprints one peak corresponds to at least one distinct *N*-glycan structure. Furthermore, large glycan structures have high migration times unless they are sialated (not in case of IVA HA). Every sialiation introduces charges reducing the migration time. (E) Relative quantification of detected *N*-glycan structures. Total peak height (TPH, sum height of all specific peaks) is set to 100 %. Consequently, relative peak heights (RPH) in % of the different peaks correspond to relative *N*-glycan structure abundances. Peaks < 5 % represent low abundant, whereas peaks > 5 % represent high abundant *N*-glycan structures.

definition, peaks with RPH < 5 % represent low abundant and peaks with RPH > 5 % represent high abundant *N*-glycan structures.

2.10.3 Next-generation pyrosequencing for characterizing viral quasispecies compositions

Beside the characterization of attached *N*-glycans to the major vaccine antigen HA, the genomic sequences of virus preparations should be determined during virus adaptation to exclude changing potential *N*-glycosylation sites and furthermore, monitor genetic variability/heterogeneity of the virus population [4]. Several methods are available for sequencing:

The oldest method for DNA-sequencing is based on **chain termination** and was developed by Federick Sanger, why it is also referred to as **Sanger sequencing**. Basically, a template DNA is amplified in a polymerase chain reaction (PCR), but using additionally to the deoxynucleoside triphosphates (dNTPs) low concentrations of fluorescently labeled dideoxynucleoside triphosphates (ddNTPs) with a specific label for each of the four bases. Consequently, by random ddNTP incorporation chains of all different lengths are amplified. These DNA chains are separated by size, e.g. using a capillary DNA sequencer. The sequence of the specific label signals corresponds to the base sequence of the template DNA. Principally, Sanger sequences the consensus of a population, in which low abundant variants disappear.

In contrast, the more recently developed techniques of **next generation sequencing** sequences variants, from which the consensus is formed. This allows characterizing a population consisting of differing genotypes, e.g. the composition of a viral quasispecies. All next generation sequencing techniques use the principle of **sequencing by synthesis**. One is called **Illumina sequencing** and uses fluorescently labeled, reversible terminator-bound dNTPs, minimizing incorporation bias. During each chain elongation cycle all four terminator-bound dNTPs are present. As a dNTP is added to the complementary growing DNA-strand and the fluorescently labeled terminator is cleaved off, the base-specific fluorescent signal is detected and a next fluorescently labeled,

reversible terminator-bound dNTP can be incorporated, and so on. Read length of approximately 200 bp can be achieved. Another technique is called **next-generation pyrosequencing** using the Roche 454. Here, one dNTP is offered for chain elongation after the other. Whenever a dNTP is incorporated into the growing DNA chain, pyrophosphate and hydrogen are released. The released pyrophosphate is enzymatically converted to ATP, which is used in the Luciferin/Luciferase reaction, producing oxyluceferin and light. The latter can be detected and signal intensity is proportional to the number of incorporated bases. So briefly, whenever a base cycle results in a light signal, this base was incorporated. For a detailed scheme of the next-generation pyrosequencing, see section 12.10 in the supplementary. A clear advantage of next-generation pyrosequencing is that it combines a relatively long read length of approximately 400 bp with a sequence depth (number of sequences obtained for one position) of up to hundreds of reads per base position. Other techniques such as **Ion Torrent Semiconductor Sequencing** have been described, too. As for next-generation pyrosequencing, one dNTP is offered after the other for complementary chain elongation. Here, the released hydrogen ion after chain elongation is measured by an ion-sensitive field-effect transistor (ISFET) sensor. As for next-generation pyrosequencing a proportionally higher signal (here electrical) indicates the presence of homopolymer repeats. Read-lengths of up to 200 bp have been reported (reviewed in [129, 130]).

Within this work next-generation pyrosequencing was applied to unravel quasispecies composition for the whole coding IVA-PR8 genomes, predominantly due to the long read lengths. In this work the limit of detection (LOD) of different virus variants is most probably determined by the RT-PCR, which is performed prior sequencing to transcribe viral RNA into cDNA for sequencing.

2.11 Immunogenicity and adaptive immune responses

The combination of different methods such as CGE-LIF-based glycoanalysis and pyrosequencing allows a comprehensive characterization of different virus

preparations regarding HA *N*-glycosylation and quasispecies composition. Then, questions rise concerning the impact of virus variations on the immunogenicity and hence on the efficacy of different virus preparations for vaccine production.

So what is immunogenicity or better what makes an antigen (Ag) immunogenic? To define an Ag such as HA immunogenic, different conditions must be met (figure 8): first the Ag needs to be taken up by antigen presenting cells (APC), second the Ag needs to be processed within the APC, third processed Ag fragments need to be presented by the APC, fourth the presented Ag fragment needs to be recognized by a specific T cell expressing a specific T cell receptor (TCR) and fifth the specific T cell needs to be activated by the APC to proliferate and differentiate [131]. In the following, a brief overview of the immune system, in particular of adaptive immunity, will be given, outlining key factors and intersections, relevant for the understanding and interpretation of data obtained within this study [1].

The immune system is divided into two, often closely interacting parts: the innate immune system (no memory function; e.g. dendritic cells, DC, $CD11c^+$; macrophages, $CD19^+$) and the adaptive immune system (with memory function; B and T cells, [132-134]). The adaptive immune system is further divided into a cellular arm (T cell-mediated, figure 8A) and a humoral arm (B cell-mediated, figure 8B).

2.11.1 Cellular immunity

Whenever an Ag, e.g. a viral vaccine after intramuscular application, enters the organism, it is taken up by Ag presenting cells (APC) in the peripheral tissues. APC are represented either by different cells of the innate immune system such as macrophages and DC or by B cells, which belong to the humoral arm of the adaptive immune system. Once the Ag is taken up, the APC migrates to secondary lymphoid tissues (e.g. spleen, lymph nodes). Meanwhile, the Ag is processed into small peptide fragments, which are loaded onto major histocompatibility complexes (MHC) of the type I or II. These peptide-MHCI/II-complexes are transported to the cell surface, where they are presented (figure

8A, [135]). In general, with respect to DCs exogenous Ag (e.g. inactivated vaccines) is usually presented by MHCII, whereas endogenous Ag is usually presented by MHCI (e.g. infected cells). However, the ability of DCs in particular to cross-present exogenous Ag results in MHCI-presentation of exogenous Ag peptides. Whenever a naive T cell, expressing a specific TCR for the presented peptide fragment presented, encounters the Ag presenting DC the TCR binds to the peptide-MHC-complex. In general, T helper cells ($CD4^+$) bind to MHCII-peptide-complexes, whereas cytotoxic T cells ($CD8^+$) bind to MHCI-peptide-complexes. Additional binding of different cell surface markers triggers intracellular signaling series, leading to IL-2 secretion, expression of IL-2 receptors (IL-2R), T cell activation with CD69 expression (very early T cell activation marker, [136]), T cell proliferation and T cell differentiation into effector cells ([134, 135, 137, 138], figure 8). However, usually TCR signals are often insufficient for complete T cell activation and costimulation by other interactions such as CD28/B7.1 (B7.2) are required for full T cell activation and survival [139]. The autocrine interaction of IL-2 secretion and IL-2R expression plays an essential role for the stimulation of Ag-specific T cells to proliferate, differentiate into effector cells and to survive. Different effector cell subtypes have been characterized, differing in their secretion of cytokines, their expression of cell marker molecules and their function (figure 8A): On the one hand, there are $CD8^+$ cytotoxic T lymphocytes (CTL), which kill infected cells and which can activate/induce cytotoxic functions in macrophages and granulocytes [135]. On the other hand there are $CD4^+$ T helper cells. These are divided into different cell types: low abundant cell populations such as T helper cells (Th) 9, Th17 and follicular Th (Th_{FH}, [137]) and rather high abundant populations such as Th1, Th2 and regulatory T cells (Tregs). However, few exceptions ($CD4^-$ Th cells) were described for special Treg subtypes [137, 138]. Nevertheless, upon activation each Th population may be discriminated by specific markers: Th1 cells express T-bet and secrete multiple cytokines e.g. IL-2 and IFNγ. Altogether, the Th1 response activates/induces macrophages, CTL and the production of B cell IgG Ab [137]. Th1 responses are generally associated with strong CTL responses,

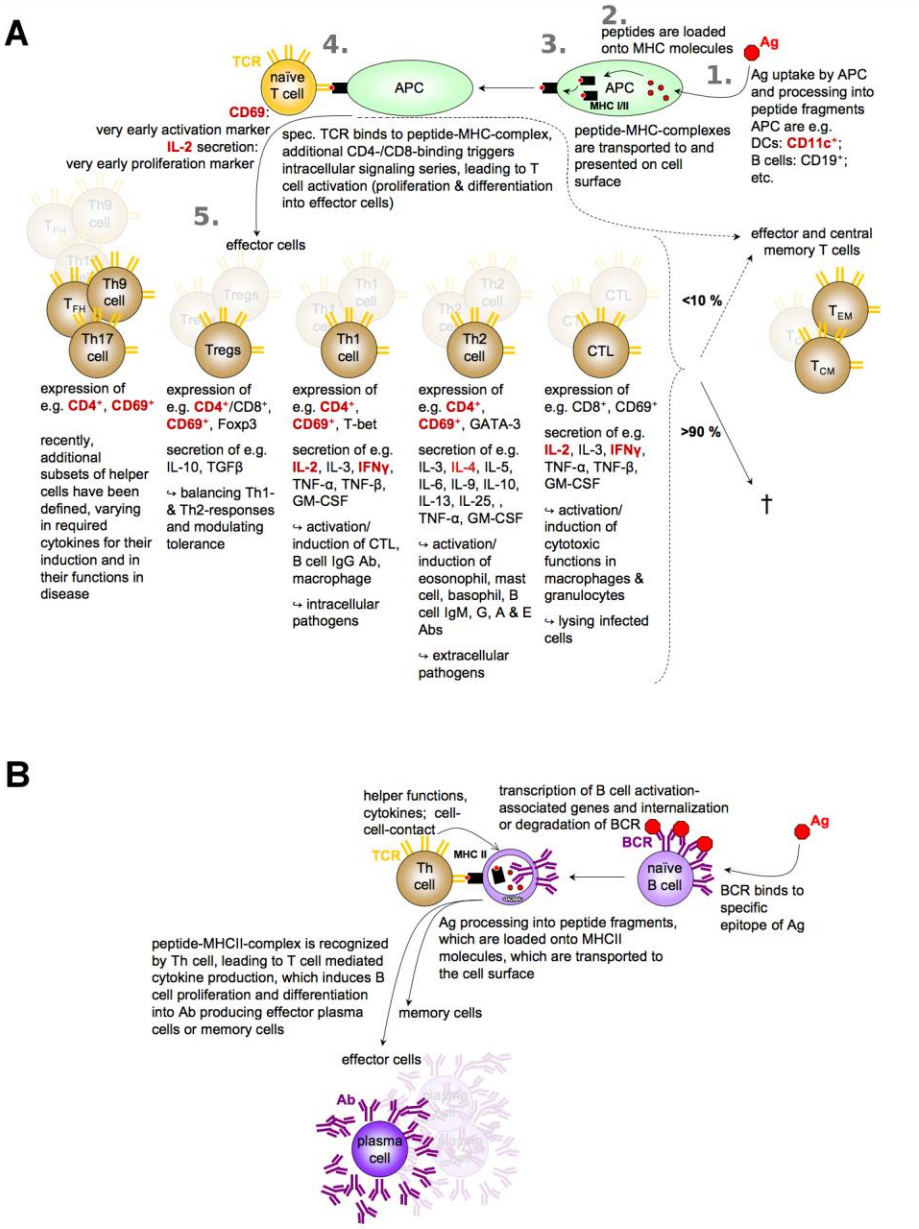


figure 8: Adaptive immune responses.
Cytokines and cell surface marker proteins quantified in this study are highlighted in red. **(A)** T cell-mediated cellular arm of the adaptive immune response. **1.** An antigen (Ag) is taken up by an Ag presenting cell (APC), **2.** processed and **3.** presented on the APC cell surface. **4.** When a naive, T cell binds with its specific T cell receptor (TCR) to the presented Ag the T cell is **5.** activated, proliferates and differentiates into different effector cell types such as cytotoxic T lymphocytes (CTL, CD8+) or different T helper (Th, mostly CD4+) cells, e.g. Th1, Th2, Th9, Th17. So far, it is not known, whether memory cells (T_{EM}: effector memory cells, T_{CM}: central memory cells) derive from effector cells or whether they differentiate directly from naive T cells (·····►). **(B)** B cell-mediated humoral arm of the adaptive immune response. Once a naive B cell binds with its specific B cell receptor (BCR) to the Ag, it is taken up, processed and presented on the B cell surface. When a naive, Ag-primed Th cell binds to the Ag-presenting B cell it provides helper functions, secretes cytokines and allows cell-cell-contact by CD40-CD154-interactions. Altogether this allows B cell proliferation and differentiation into antibody (Ab) producing plasma cells (effector cells) or memory cells.

predominantly fighting intracellular pathogens [135]. In contrast, Th2 responses are associated with Ab-mediated immune responses, predominantly fighting extracellular pathogens [135, 140]. Activated Th2 cells express GATA-3 and secrete cytokines such as IL 4, IL 5, IL 6, IL 9, IL 10, IL 13 and IL 25 [137, 141-144]. The Th2 response activates/induces eosinophils, mast cells, basophils and the production of B cell IgM, G, A and E Ab [137]. However, most important for successful vaccination is the development of memory cells, since most effector cells will die once the Ag/pathogen has been cleared. So far, it is not fully understood, whether memory cells develop from naive T cells in parallel to the effector cells (divergent) or whether they develop from the effector cells (linear, [145]). Though, up to date only little research focuses on T cell-mediated vaccine function [146].

In this study, CD69 expression of CD4+ T-cells and IL-2, IL-4 and IFNγ secretion was measured to characterize T cell responses in vitro. The use of TCR-HA transgenic mice (i.e. homozygote mice only express a TCR which is specific for the HA110-120 peptide) allowed an easy quantification of HA110-120-specific T cells.

2.11.2 Humoral immunity

So far, most vaccine approaches concentrate on the elucidation of neutralizing Ab: Whenever a naive B cell encounters its specific Ag, it binds with its B cell receptor (BCR) to the specific epitope (figure 8B). A signaling series leads to the transcription of B cell activation-associated genes and the internalization or degradation of the BCR. The Ag is processed into small peptide fragments, which are loaded onto MHCII molecules and which are transported to the cell surface, where they are presented. Whenever a fragment-specific, antigen-primed T helper cell is encountered, it binds with its specific TCR to the specific BCR. Binding induces signaling series leading to complete T cell activation. The activated T cell carries out helper functions, such as cytokine secreting and cell-cell-interaction, finally leading to B cell activation, proliferation and differentiation

into effector cells – the Ab-producing plasma cells - and into memory cells (figure 8B, [135, 140, 141]).

In this study, B cell responses induced *in vivo* by differentially glycosylated IVA preparations in wild type (*wt*) BALB/c mice, were characterized by HA-specific Ab titers as well as by the ability of induced Ab to inhibit hemagglutination, i.e. the ability to block IVA binding to red blood cells.

3 Materials and methods

3.1 Cell lines and cultivation conditions

An overview of cell lines used during this study is given in table 1 along with details including cell growth media and cultivation vessels. Briefly, unless stated otherwise, adherently growing MDCK (#84121903, ECACC, Salisbury, UK) and adherently growing Vero (#88020401, ECACC) cells were cultivated under serum-containing conditions in Glasgow Minimum Essential Medium (GMEM, #22100-093, Invitrogen/Gibco, Darmstadt, Germany), supplemented with 5.5 g/L glucose (#X997.3, Roth, Karlsruhe, Germany), 2 g/L peptone (#MC33, IDG, Lancashire, UK), 10 % fetal calf serum (FCS, #10270-106, Invitrogen/Gibco) and 4 mg/mL $NaHCO_3$ (#6885.3, Roth) until confluence in T75-flasks (50 mL, 37 °C, 5 % CO_2), T175-flasks (150 mL, 37 °C, 5 % CO_2), roller bottles with closed caps (250 mL, 37 °C), 1.2 L-stirred tank reactors (STR, cellferm-pro® DasGip, Jülich, Germany) with a working volume of 800 mL and parameters set to pO_2 40 %, pH 7.2, 55 rpm, 37 °C as well as 5 L-STR (Biostat C5, B. Braun Biotech International GmbH, Melsungen, Germany) with a working volume of 5 L and parameters set to pO_2 40 %, pH 7.2, 50 rpm, 37 °C (SOP see section 12.9.5). In stirred systems the adherently growing cells were cultivated on microcarriers (2 g/L, CytodexTM1, #17-0448-03, GE Healthcare, Uppsala, Sweden). For adaptation to serum-free growth MDCK cells were cultivated in Episerf medium (#10732-022, Invitrogen/Gibco) until at least three consecutive passages of cells were successfully propagated using 100 % Episerf medium in total. Suspension growth adapted MDCK cells, namely MDCK.SUS1, MDCK.SUS2 and MDCK.SUS3 [142] were cultivated in chemically defined, protein- and peptide-free SMIF8 medium or in SMIF8pgd (both available from Gibco/Invitrogen by contact through Prof. Dr. K. Scharfenberg, FH Emden/Leer, Germany) in shaker flasks (250 mL with baffles with a working volume of 100 mL, 185 rpm, 37 °C, 5 % CO_2). The new human-derived, suspension cell line CAP was cultivated in PEM medium (#12661-013, Invitrogen/Gibco) supplemented with 4 mM glutamine (#G3126, Sigma-Aldrich, Steinheim, Germany) and 4 mM pyruvate

(#P8574, Sigma-Aldrich) in 1 L-STR (Biostat B-Plus, Sartorius AG, Göttingen, Germany) with a working volume of 1 L and parameters set to pO_2 40 %, pH 7.1-7.2, 120 rpm, 37 °C. Duck-derived AGE1.CR.pIX suspension cells (ProBioGen AG, Berlin, Germany, [76]) were grown in 250 mL shaker flasks with baffles in a total volume of 50 mL in a chemically defined medium (CD-U2, [143]), supplemented with 2 mM alanine-glutamine (#G8541, Sigma-Aldrich) and 10 ng/mL insulin like growth factor (#91590C, Sigma-Aldrich). Flasks were shaken at 185 rpm, 37 °C and 5 % CO_2 [6, 7].

Standard operating procedures (SOPs) for thawing (section 12.9.1) of cells and passaging (section 12.9.2) of adherent MDCK cells including the SOP for trypsin preparation (1.25 g trypsin Invitrogen/Gibco, #27250018 for 2500 mL 1x trypsin solution, section 12.9.7), required for cell passaging are compiled in the supplementary section.

Generally, before infection adherent cells were washed three times with phosphate buffered saline (PBS, SOP see section 12.9.3) for complete removal of FCS. Cell growth medium of suspension cells was either not changed before infection or partly/fully replaced by fresh medium (see table 1 for details).

table 1: Overview of culture conditions and media used for influenza virus A (IVA) production.

IVA PR/8/34 (H1N1) from the Robert Koch Institute (RKI, RKI-strain) or the National Institute for Biological Standards and Control (NIBSC, NIBSC-strain) as well as IVA-California/07//2009-like (H3N2, IVA-California), IVA Uruguay//716/2007-like (H3N2, IVA-Uruguay) and IVA Victoria/210/2009-like (H1N1, IVA-Victoria) reassortants were produced in cell culture or embryonated hens' eggs.

Comparison		Host Cell	Seed Virus	Cell Growth Medium	Virus Production Medium	Trypsin Activity	Cultivation Vessel	Virus Harvest (hpi)	Purification Method for Labelled *N*-Glycans	Medium Exchange Prior Infection	Moi or Volume for Volume-Based Infection
harvest time points	harvest time point (RKI in Vero)	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	48	SEC	yes	0.1 or 1 mL
		Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	SEC	yes	0.1 or 1 mL
		Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	120	SEC	yes	0.1 or 1 mL
		Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	360	SEC	yes	0.1 or 1 mL
	harvest time point (RKI in MDCK)	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL
		MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	SEC	yes	1 mL
	harvest time point (Uruguay in MDCK)	MDCK	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	5L-STR	24	SEC	yes	0.05
		MDCK	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	5L-STR	72	SEC	yes	0.05
host cells	Vero	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	HILIC	yes	0.08
	MDCK	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	HILIC	yes	1
	Cap	Cap	RKI-strain	PEM°	PEM° + trypsin	3x 10^{-5} units/cell	1L-STR	72	SEC	no	0.025
	AGE1.CR.pIX	AGE1.CR.pIX	RKI-strain	CD-U2[#]	CD-U2[#]+ trypsin	1x 10^{-6} units/cell	250mL shaker w baffles	24	SEC	no	1 mL
	egg[%]	egg	RKI-strain	egg	egg	0 U/mL	egg	96	HILIC	%	nd
host cell adaptation	serum <-> serum-free	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	SEC	yes	1 mL
		MDCK	RKI-strain	Episerf"	Episerf" + trypsin	5 U/mL	roller bottle	96	SEC	no	1 mL
	serum <-> serum-free suspension growth I	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	HILIC	yes	1
		MDCK.SUS1	RKI-strain	SMIF8-Pgd[§]	SMIF8-Pgd[§] + trypsin	1x 10^{-5} units/cell	250mL shaker w baffles	72	HILIC	yes	188
		MDCK.SUS2	RKI-strain	SMIF8-Pgd[§]	SMIF8-Pgd[§] + trypsin	1x 10^{-5} units/cell	250mL shaker w baffles	72	HILIC	yes	197.3
		MDCK.SUS3	RKI-strain	SMIF8-Pgd[§]	SMIF8-Pgd[§] + trypsin	1x 10^{-5} units/cell	250mL shaker w baffles	72	HILIC	yes	192.6
	serum <-> serum-free suspension growth II	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.2 mL
		MDCK.SUS1	NIBSC-strain	SMIF8-Pgd[§]	SMIF8-Pgd[§] + trypsin	12.5 U/mL	250mL shaker w baffles	48	HILIC	partly (90 %)	1.41
		MDCK.SUS2	NIBSC-strain	SMIF8-Pgd[§]	SMIF8-Pgd[§] + trypsin	12.5 U/mL	250mL shaker w baffles	48	HILIC	partly (90 %)	1.17
		MDCK.SUS3	NIBSC-strain	SMIF8-Pgd[§]	SMIF8-Pgd[§] + trypsin	12.5 U/mL	250mL shaker w baffles	48	HILIC	partly (90 %)	1.75
viruses		MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.2 mL
		MDCK	IVA-California	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	48	HILIC	yes	0.2 mL
		MDCK	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.05
		MDCK	IVA-Victoria	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	48	HILIC	yes	0.2 mL
virus passages	passage 1	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	0.333333
	passage 2	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 3	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 4	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 5	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 6	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 7	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 8	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 9	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
	passage 10	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	24	SEC	yes	1 mL of previous passage
virus suppliers		MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	2.33 or 0.2 mL
		MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.028 or 0.2 mL

* GMEM-medium was supplemented with 5.5 g/L glucose, 2 g/L peptone and 4 mg/mL $NaHCO_3$.

° PEM-medium was supplemented with 4 mM glutamine and 4 mM pyruvate.

" Episerf was supplemented with 20 mM glucose, 2 mM glutamine and 2 mM pyruvate.

[#] CD-U2 [143] was supplemented with 2 mM alanine-glutamine, 10 ng/mL insulin like growth factor.

[$, §] SMIF8 and SMIF8-Pdg are available from Gibco/Invitrogen by contact through Prof. Dr. K. Scharfenberg (FH Emden/Leer, Germany).

[%] virus was produced in cooperation with IDT Biologika GmbH, Dessau-Rosslau, Germany

SEC: size exclusion chromatography, HILIC: hydrophilic interaction chromatography

table 1 continued

Comparison			Host Cell	Seed Virus	Cell Growth Medium	Virus Production Medium	Trypsin Activity	Cultivation Vessel	Virus Harvest (hpi)	Purification Method for Labelled *N*-Glycans	Medium Exchange Prior Infection	Moi or Volume for Volume-Based Infection
virus adaptation	virus adaptation (RKI, Vero)	**passage 1**	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	24	SEC	yes	2
		passage 2	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	360	SEC	yes	1 mL of previous passage
		passage 3	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	120	SEC	yes	0.4 mL of previous passage
		passage 4	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	96	SEC	yes	0.5 mL of previous passage
		passage 5	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	72	SEC	yes	0.5 mL of previous passage
		passage 6	Vero	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	48	SEC	yes	0.5 mL of previous passage
		passage 7	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	96	SEC	yes	0.5 mL of previous passage
		passage 8	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	72	SEC	yes	0.5 mL of previous passage
		passage 9	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	96	SEC	yes	0.5 mL of previous passage
		passage 10	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	72	SEC	yes	0.5 mL of previous passage
		passage 11	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	96	SEC	yes	0.5 mL of previous passage
	virus adaptation (NIBSC, Vero)	**passage 1**	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	24	SEC	yes	0.0294
		passage 2	Vero	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	360	SEC	yes	1 mL of previous passage
		passage 3	Vero	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	120	SEC	yes	0.4 mL of previous passage
		passage 4	Vero	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.25 mL of previous passage
		passage 5	Vero	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.25 mL of previous passage
		passage 6	Vero	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	48	SEC	yes	0.25 mL of previous passage
		passage 7	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.25 mL of previous passage
		passage 8	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.25 mL of previous passage
		passage 9	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.25 mL of previous passage
		passage 10	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.25 mL of previous passage
		passage 11	MDCK	NIBSC-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.25 mL of previous passage
	virus adaptation (Uruguay, Vero)	**passage 1**	MDCK	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	24	SEC	yes	0.2 mL virus seed
		passage 2	Vero	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	1 mL of previous passage
		passage 3	Vero	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	264	SEC	yes	0.4 mL of previous passage
		passage 4	Vero	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	120	SEC	yes	0.25 mL of previous passage
		passage 5	Vero	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.25 mL of previous passage
		passage 6	Vero	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	SEC	yes	0.25 mL of previous passage
	virus adaptation (RKI, AGE1.CR.pIX)	**seed**	MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	24	SEC	yes	1 mL
		passage 1	AGE1.CR.pIX	RKI-strain	CD-U2$^{\#}$	CD-U2$^{\#}$+ trypsin	1x 10^{-6} units/cell	250mL shaker w baffles	24	SEC	no	0.001
		passage 2	AGE1.CR.pIX	RKI-strain	CD-U2$^{\#}$	CD-U2$^{\#}$+ trypsin	1x 10^{-6} units/cell	250mL shaker w baffles	24	SEC	no	1 mL of previous passage
		passage 3	AGE1.CR.pIX	RKI-strain	CD-U2$^{\#}$	CD-U2$^{\#}$+ trypsin	1x 10^{-6} units/cell	250mL shaker w baffles	24	SEC	no	1 mL of previous passage
		passage 4	AGE1.CR.pIX	RKI-strain	CD-U2$^{\#}$	CD-U2$^{\#}$+ trypsin	1x 10^{-6} units/cell	250mL shaker w baffles	24	SEC	no	1 mL of previous passage
	virus adaptation (California, MDCK)	**passage 1**	egg	IVA-California	egg	egg	nd	nd	nd	HILIC	%	nd
		passage 2	MDCK	IVA-California	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	HILIC	yes	0.15 mL from ampulle
		passage 3	MDCK	IVA-California	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	HILIC	yes	0.2 mL of previous passage
		passage 4	MDCK	IVA-California	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	48	HILIC	yes	0.2 mL of previous passage
	virus adaptation (Victoria, MDCK)	**passage 1**	egg	IVA-Victoria	egg	egg	nd	nd	nd	HILIC	%	nd
		passage 2	MDCK	IVA-Victoria	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	HILIC	yes	0.15 mL from ampulle
		passage 3	MDCK	IVA-Victoria	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	72	HILIC	yes	0.2 mL of previous passage
		passage 4	MDCK	IVA-Victoria	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	48	HILIC	yes	0.2 mL of previous passage
	virus adaptation (RKI, egg)	**passage 1**[%]	egg	RKI-strain	egg	egg	0 U/mL	egg	96	HILIC	%	pool of various mois
		passage 2[%]	egg	RKI-strain	egg	egg	0 U/mL	egg	96	HILIC	%	pool of various mois
		passage 3[%]	egg	RKI-strain	egg	egg	0 U/mL	egg	96	HILIC	%	pool of various mois
cultivation vessels			MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	2.33 or 0.2 mL
			MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	96	SEC	yes	2 or 0.4 mL
			MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	roller bottle	96	SEC	yes	1 mL
			MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	1 L-STR	96	SEC	yes	0.025
			MDCK	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	24	SEC	yes	0.25 or 0.25 mL
			MDCK	IVA-Uruguay	GMEM* + FCS	GMEM* + trypsin	5 U/mL	5 L-STR	24	SEC	yes	0.05
virus production media	**media compositions**		MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T175	72	SEC	yes	2.8
			MDCK	RKI-strain	GMEM* + FCS	Episerf* + trypsin	5 U/mL	T175	72	SEC	yes	2.8
			MDCK	RKI-strain	GMEM* + FCS	SMIF8[$] + trypsin	5 U/mL	T175	72	SEC	yes	2.8
	trypsin concentration		MDCK	RKI-strain	GMEM* + FCS	GMEM*	0 U/mL	T75	120	SEC	yes	0.025
			MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL	T75	96	SEC	yes	0.025
			MDCK	RKI-strain	GMEM* + FCS	GMEM* + trypsin	5 U/mL every 24 h	T75	96	SEC	yes	0.025

3.2 Viruses and infection conditions

Depending on the production cell line, either MDCK cell-, Vero cell- or AGE1.CR.pIX cell-adapted influenza A virus PR/8/34 (H1N1) - in the following referred to as **IVA-PR8** - was used for infections. In the scope of this study IVA-PR8 from two different suppliers was used: one was purchased from the National Institute for Biological Standards and Control (NIBSC, #06/114, South Mimms, UK), which is also referred to as **NIBSC-strain**. The other was provided by the Robert Koch Institute (RKI, Berlin, Germany, Amp. 3138) and is referred to as **RKI-strain**. Moreover, the reassortant influenza A virus California/07/2009 (H1N1pandemic, #09/176, NIBSC), **IVA-California,** the reassortant influenza A virus Uruguay/716/2007 (H3N2) x IVA-PR8/34 (H1N1, #07/360, NIBSC), **IVA-Uruguay** and the reassortant influenza A virus Victoria/210/2009 (H3N2) x IVA-PR8/34 (H1N1, #12/112, NIBSC), **IVA-Victoria** was used as indicated (table 1).
Details concerning virus production are summarized in table 1. The SOP for virus propagation in culture flasks is provided in the supplementary (section 12.9.6). Briefly, unless indicated otherwise, virus was produced in non-serum containing cell growth medium supplemented with a final activity of 5 U/mL trypsin (Invitrogen/Gibco, #27250-18, Darmstadt, Germany). Trypsin was prepared in PBS according to the activity given by the supplier (for SOP see section 12.9.8). Generally, volume-based infections using 1 mL virus seed for roller bottles and shaker flasks, 0.4 mL/0.5 mL for T175- and 0.2 mL/0.25 mL for T-75-flasks were applied (for details see table 1). For microcarrier-based cultivations a multiplicity of infection (moi) of usually 0.025 was applied. CAP suspension cells were also infected with a moi of 0.025. For infection of AGE1.CR.pIX cells $1x\ 10^{-6}$ units trypsin (Invitrogen/Gibco) were added per cell and a moi of 0.001 (passage 1) was used. During all virus adaptation experiments and all virus passaging experiments subsequent passages were infected using an appropriate volume (depending on the vessel) of the supernatant of the previous passage (for details see table 1).

3.3 Virus quantification by hemagglutination- (HA-) assay

Influenza virus from cell culture supernatant was titrated by hemagglutination based on the method from Kalbfuss *et al.* [144]. Samples were serially diluted $1:2^{0.5}$ in PBS resulting in a final volume of 100 µL using round-bottomed 96-well plates. 1.9 – 2.1 x 10^6 red blood chicken cells in a total volume of 100 µL in PBS (SOP see section 12.9.11) were added and incubated for 3 h to 4 h at room temperature. Light extinction was measured at 700 nm and plotted over the logarithm of the inversed dilutions. The end point was defined as the point of inflection. The point of inflection represents the last dilution (d) showing complete hemagglutination and corresponds to 1 HAU (HAU = HA Units = $-\log_{10}$ d/100 µL). A sample's HA-titer is calculated by considering the dilution factor of the end point. Consequently, if e.g. the end point is reached at a dilution of 1/64 ($1/10^{1.8}$), this corresponds to an HA-titer of 1.8 HAU for the undiluted sample. The detailed SOP can be found in the supplementary (section 12.9.12, [4]).

3.4 β-propiolactone inactivation

For immunogenicity studies and prior to native in-solution virus deglycosylation MDCK- and Vero cell-derived virus containing culture broths were harvested 96 hpi and clarified by centrifugation (100 g for 20 min, 4000 g for 35 min and 10.000 g for 45 min, Avanti J-20XP, Beckman Coulter, Krefeld, Germany). Clarified supernatants were stored at -80 °C and were chemically inactivated by β-propiolactone (β-PL, #33672.01, Serva Electrophoresis, Heidelberg, Germany) as described by Kalbfuss *et al.* [86, 145]. Briefly, pH of culture broth was stabilized by the addition of 25 mM (final) HEPES-buffer (pH 7.5; stock), and β-PL prediluted in HEPES-buffer was added to the clarified culture broth (final β-PL concentration 3 mM). The broth was transferred into a new vessel in order to ensure proper mixing and was incubated at 37 °C for 24 h. Inactivation was confirmed when two consecutively infected T75-flasks of confluent MDCK cells (1 mL inactivated broth + 50 mL medium) exhibited HA-titers according to the dilution used during infection, *i.e.* when latest, the second passage reached a titer of 0 HAU. The detailed SOP for β-PL inactivation can be found in the

supplementary (section 12.9.13). All steps were carried out under sterile conditions.

3.5 Protein quantification by bicinchoninic acid assay

A sample's protein content was determined by bicinchoninic acid assay (BCA assay, #23227, Thermo Fisher Scientific Inc., Rockford, Illinois, USA, [146]) according to instructions provided by the manufacturer for micro-plate assays.

3.6 N-glycosylation pattern analysis

3.6.1 Workflow

HA *N*-glycosylation pattern analysis was performed as described earlier using a CGE-LIF-based approach ([4, 6, 7, 73, 112, 147, 148], figure 9). According to Schwarzer *et al.* CGE-LIF-based glycoanalysis has a limit of detection (LOD) of at least 5 fmol/L and a linear dynamic range from 2 pmol/L up to 2 nmol/L as determined by dilution series using the fluorescent marker molecule APTS [112]. Furthermore, they determined a SD of 0.3 % to 1.8 % for the reproducibility of RPH for bovine IgG-derived *N*-glycan fingerprints [112]. With respect to migration times, long-term reproducibility was described to be below 0.08 min [112], which corresponds to roughly 13.9 datapoints (about 2.9 datapoints per second for four capillary 3100 ABI PRISM Genetic Analyzer).

During the time of this work another student improved the purification step of labeled *N*-glycans from excess label and impurity-salts towards high-throughput applications within his student's project (figure 9, step 6). The optimized protocol for sample preparation substituted the SEC by HILIC (figure 9, step 6). Therefore, in this work both methods (SEC, HILIC) were used as indicated in table 1. However, according to [149] the applied purification method - SEC or HILIC – does not impact *N*-glycan structure presence or absence. It may merely affect relative *N*-glycan structure abundances.

After harvest of the virus-containing supernatant, the culture broth was cleared by centrifugation (usually 100 g for 20 min, 4000 g for 35 min, 10.000 g for 45 min). Virus isolation was performed at 31000 rpm (g force at average

diameter 5.24 mm: about 56300 g; Type 70Ti rotor, Beckman Coulter) in OptiSeal tubes (32.4 mL, #361625, Beckman Coulter) for 90 min at 4 °C (figure 9, step 1). Finally, the virus containing pellet was resuspended in 100 mM Tris (pH 7) and virus preparations were stored at -80 °C [8].

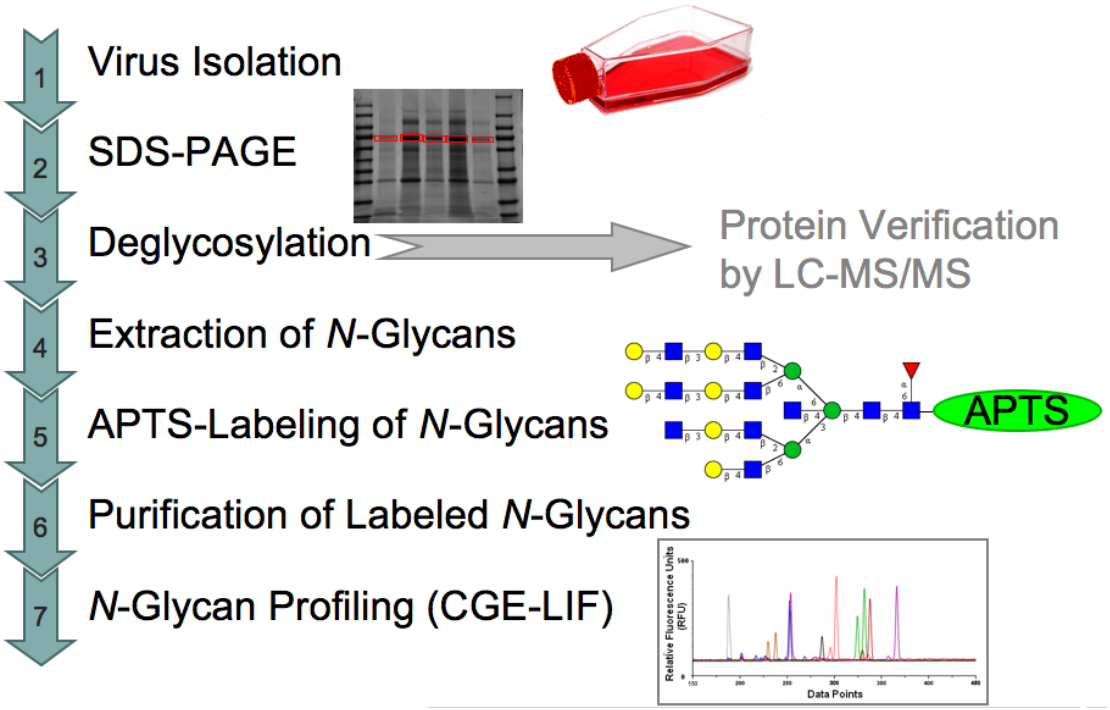


figure 9: Workflow of CGE-LIF-based *N*-glycosylation pattern analysis for IVA-derived HA. After virus harvest and isolation by centrifugation (1), viral proteins are separated in a SDS-PAGE (2). HA at approximately 70 kDa is cut out and *N*-glycans are enzymatically cut off the protein backbone (3). The protein may as well be analyzed by e.g. LC-MS/MS in order to verify protein identity. *N*-glycans are extracted (4) and labeled with APTS (5). Excess label and impurity-salts are removed by SEC or HILIC (6). Finally, the labeled *N*-glycans are separated by CGE and detected by LIF, resulting in capillary electropherograms (7), which are further processed into highly reproducible *N*-glycan fingerprints (workflow according to [112]).

Approximately 10 µg - 25 µg total virus protein (e.g. BCA, section 3.5) was mixed in a ratio 3:1 with 4x non-reducing sample buffer (125 mM Tris, 4 % SDS, 20 % glycerole, 8 M urea, bromophenol blue) and was applied to a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, #161-1158, Bio-Rad Laboratories, Inc., Munich, Germany). Subsequently, viral proteins were separated according to their size at 15 mA per gel for approximately 95 min (figure 9, step 2). Subsequently, the gels were transferred for 1 h into a fixation solution (40 % ethanol, 7.5 % acetic acid), washed in ultra pure water (H_2O_{MQ}) for re-hydration, and stained in a colloidal coomassie solution (0.5 g coomassie

brilliant blue, 10 mL H_2O_{MQ}, 50 g ammonium sulfate, 6 mL phosphoric acid (85 %), filled up to 490 mL with H_2O_{MQ}, 125 mL methanol) over night at room temperature. Afterwards, the gels were washed multiple times in H_2O_{MQ} until protein bands were clearly visible. All incubations were performed on a see-saw-shaker. The stained SDS-PAGEs were stored sealed at 4 °C [8].

In a next step, HA-bands were cut out of the gel, divided into pieces and transferred into 1.5 mL tubes. During multiple incubation steps the gel pieces were washed, reduced and alkylated using different ammonium bicarbonate-, acetic acid-, methanol-, acetonitrile-, dithiothreitol- and iodacetamide- solutions [112, 147]. Finally, for in-gel-deglycosylation, the gel pieces were incubated with 6 µL PNGaseF (#7367, Sigma-Aldrich, Taufkirchen, Germany) and 54 µL 50 mM ammonium bicarbonate over night at 37 °C (figure 9, step 3). For *N*-glycan extraction (figure 9, step 4), the supernatant was transferred into a new tube and remaining *N*-glycans were extracted by three rounds of sonication (30 min., 10 °C, level 6, Ultrasonic cleaner, Model USC6000, VWR), each in 200 µL H_2O_{MQ}, pooling all extracts in the new tube and then desiccated [8].

Then, the extracted *N*-glycans were labeled (figure 9, step 5) with 8-aminopyrene-1,3,6-trisulfonic acid (APTS) by reductive amination (5 µL 20 mM APTS in 15 % acetic acid, 5 µL 1M $NaBH_3CN$ in 15 % acetic acid or the non-toxic alternative: 2 µL H_2O_{MQ}, 2 µL 20 mM APTS in 3,6 M citric acid, 2 µL 0.2 M picoline borane in dimethylsulfoxid, DMSO) overnight in the dark at 37 °C whilst shaking. Excess label was removed by SEC for $NaBH_3CN$-based labeling or by HILIC for picoline borane-based labeling (figure 9, step 6). Briefly, for SEC a 2 mL toyopearl® (#19808, Tosoh Bioscience, Stuttgart, Germany) column was packed and labeled samples were applied. During following step centrifugation H_2O_{MQ} was continuously added. Finally, labeled *N*-glycans eluted before excess-label and impurity-salts [112]. A detailed SOP for glycoanalysis is given in the supplementary ($NaBH_3CN$-based: section 12.9.14; picoline borane-based: section 12.9.15). For HILIC, Bio-Gel P10 (#150-4144, Bio-Rad) was equilibrated by multiple washing steps using different acetonitril/H_2O_{MQ} solutions. Sample (diluted in 80:20, acetonitril: H_2O_{MQ}) addition and incubation (5 min., shaking)

was followed by several washing steps with acetonitril/H_2O_{MQ} (80 %/20 %) solutions with or without 100 mM triethylamine (TEA) supplementation. Finally, the labeled *N*-glycans were eluted using H_2O_{MQ} (SOP in section 12.9.16, [8]).

For glycosylation pattern profiling (figure 9, step 7) labeled *N*-glycans were separated by CGE (see table 2 for settings) and monitored by LIF detection using an ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystems, Foster City, California, USA). The resulting electropherograms were normalized to an internal bp-standard with respect to migration times (x-axis). The resulting *N*-glycan fingerprints correspond to the actual *N*-glycosylation pattern of the analyzed glycoproteins and are plotted in relative fluorescence units (RFU) over the migration time (t_{mig}) in normalized migration time units (MTU'), which are equivalent to base pairs (bp; MTU' = bp). In original electropherograms as well as in *N*-glycan fingerprints one peak corresponds to at least one distinct *N*-glycan structure [6].

In general, data processing was performed as published previously [1, 4-7, 73, 112, 150]. For all other data an additional baseline correction was performed as described by Rödig *et al*. [6]. Therefore, baseline correction was performed by asymmetric least squares smoothing according to Eilers et al. [151], using the parameters p = 0.01 for asymmetry and $\lambda=10^9$ for smoothness. Blank samples were approximately normally distributed and the difference between mean and median was always below 0.05 for all signal intensities. Therefore, the noise threshold of the signal was set to the signal intensity where the difference between mean and median was < 0.05 for all data points below the threshold [6].

As published before [4, 7, 73, 112], the range of t_{mig} = 300 MTU' - 450 MTU' in HA *N*-glycan fingerprints was defined for MDCK cell-specific peaks. Within this range, all *N*-glycan peaks were generally taken for MDCK cell-derived IVA-PR8 comparisons, which reached in at least one of the comparable patterns (table 3) a peak height of more than 10x the baseline noise threshold (for peak annotation see figures of MDCK cell-derived *N*-glycan fingerprints, e.g. figure 16). Between different IVA-PR8 comparisons, the total number of annotated peaks may vary slightly (e.g. 15, 16) due to low abundant peaks droping below the 10x the noise

threshold. For all further comparisons such as different host cells as well as adherence versus suspension growth for MDCK cells, all peaks between 150 MTU' and 450 MTU' exhibiting in at least one of the *N*-glycan fingerprints peak heights of more than 10x the noise threshold were included. However, during virus adaptation to AGE1.CR.pIX cells, all peaks were annotated and compared, which exhibited at least 20 % of the RFU of the highst peak [7]. Whatever comparison stategy was applied (usually as published previously) the conclusions are comparable [6].

For quantitative comparison, additional normalization of the LIF-signal (y-axis) may be performed. Therefore, the set of major *N*-glycan peaks was defined and the total peak height (TPH, summed height of all major peaks) was calculated and set to 100 %. Accordingly, *N*-glycan abundances can be described by percentage of relative peak height (RPH, in %), allowing direct qualitative and quantitative comparisons of different HA *N*-glycosylation patterns. In the following, peaks were classified as either low- (RPH < 5 %) or high- abundant (RPH ≥ 5 %). In general, peaks were numbered as published previously [8].

table 2: Settings for ABI PRISM 3100- and 3130-Avant Genetic Analyzer for CGE-LIF-based glycoanalysis

Oven Wait Time	1200	Voltage_Numbers_Of_Steps	20
Poly_Fill_Vol	66	Voltage_Step_Interval	30
Cap_Fill_Vol	46	DC_Data_Delay	50
Leak_Threshold	25	Run_Voltage	15.0
Max_Current	300	DC_Run_Time	6000 or 7200 or 7800
Current_Tolerance	100	First_Read_Out_Time	160
Run_Current	100	Second_Read_Out_Time	160
Voltage_Tolerance	0.6	DC_RS_Plate_Type	A 12 8
PreRun_Voltage	15.0	DC_RS_CSData	Cap1 A1
DC_PreRun_Time	600	DC_Scale_Divisor	8.0
Injection_Voltage	15.0	DC_Down_Sample	2
DC_Injection_Time	5	DC_Laser_Power_Setting	15.0

3.6.2 Stability of HA N-glycan fingerprint over different harvest time points

Preliminary experiments showed that virus with a HA-value lower 1.8 may fail later glycoanalysis (data not shown). Therefore, if possible only samples with a minimum HA-titer of 1.8 were analyzed. With differences in replication dynamics, for instance during adaptation to a new cell line, i.e. slow replication dynamics

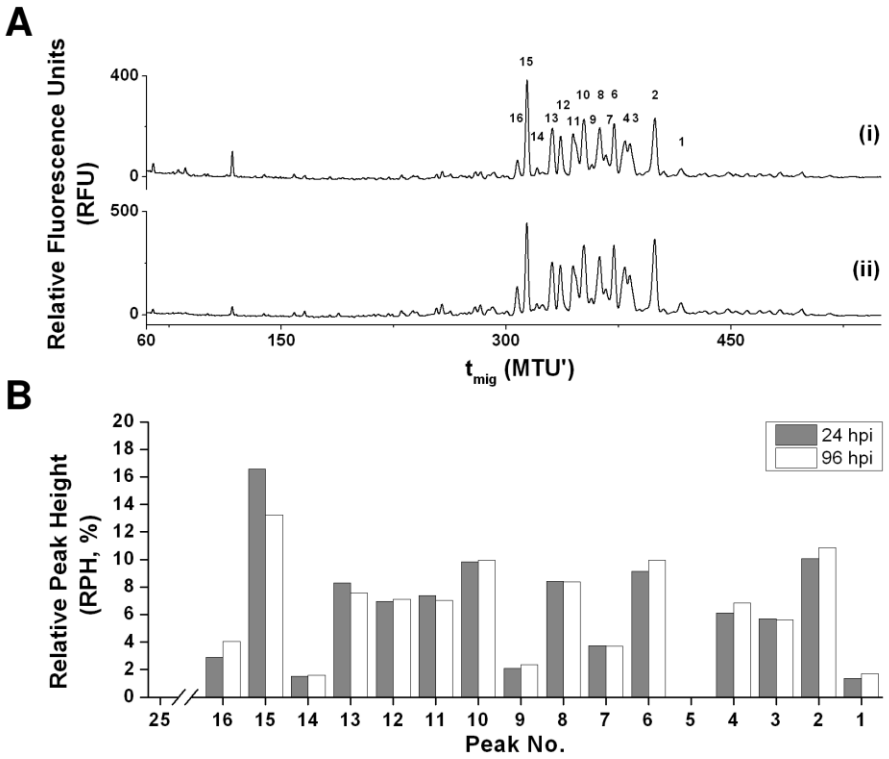


figure 10: Impact of harvest time point on the HA *N*-glycosylation pattern of MDCK cell-derived RKI-strain.
(A) Shifted overlay of *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). An overall stability of the HA *N*-glycosylation pattern is demonstrated for 24 hpi (i) and 96 hpi (ii). HA *N*-glycan fingerprints from both harvest time points exhibited the same 15 numbered main peaks (peak no.: 1-4, 6-16) with migration times between 300 MTU' and 420 MTU'. (B) Relative peak heights (RPH) of the 15 dominating peaks (no.: 1-4, 6-16; numbering according to virus adaptation from MDCK to Vero cells and back, section 4.6.2) are represented by grey (24 hpi) or white (96 hpi) columns. Modified and reprinted with permission [4].

during initial adaptation steps and fast progress of infection towards the end, different time points for sampling were required to achieve a minimum of 1.8 HAU [4]. Hence, the impact of harvest time point on HA *N*-glycosylation is essential for the set-up and later interpretation of experiments investigating the impact of virus adaptation (result section). Therefore, as part of materials and

methods the stability of HA *N*-glycosylation patterns over different harvest time points was investigated for MDCK cell-adapted IVA-PR8 (sampling 24 hpi and 96 hpi) and IVA-Uruguay (sampling 24 hpi and 72 hpi) in MDCK cells and for Vero cell-adapted IVA-PR8 (sampling 48 hpi, 96 hpi, 120 hpi and 360 hpi) in Vero cells.

For all sampling time points of IVA-PR8 during MDCK cell-based virus replication, the HA *N*-glycan fingerprint was dominated by the same 15 main peaks (no. 1 - 4, 6 - 16; figure 10A), exhibiting maximum differences in RPH ≤ 3.3 % (figure 10B, $|\Delta RPH_{max}|$ for peak15, [4]).

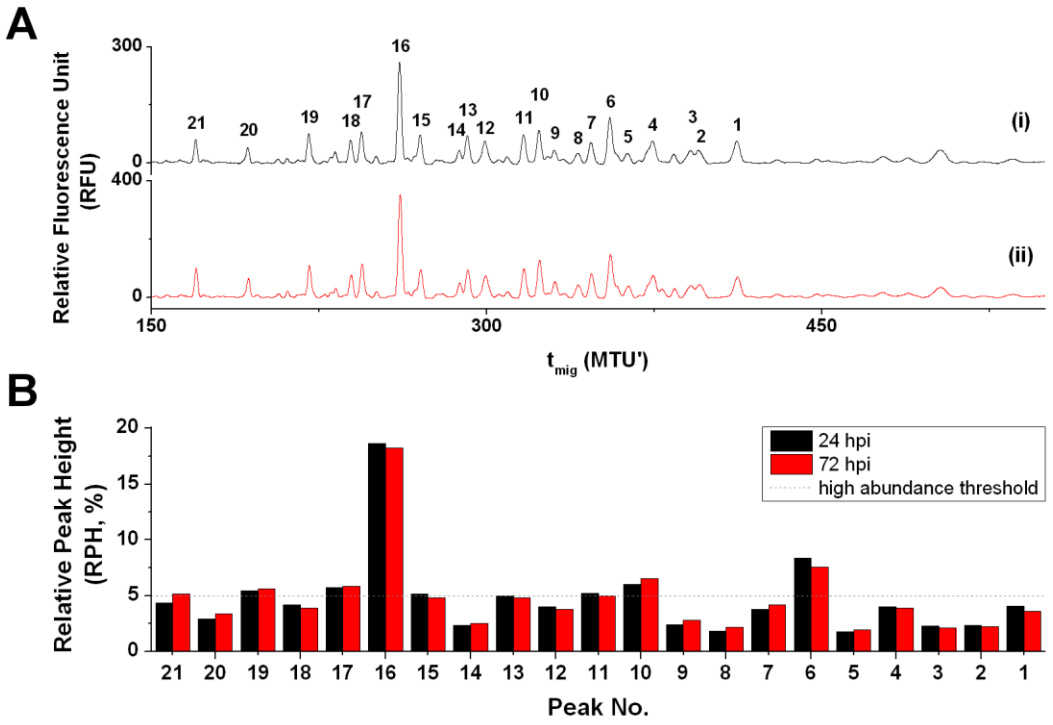


figure 11: Impact of harvest time point on the HA *N*-glycosylation pattern of MDCK cell-derived IVA-Uruguay.
(A) Shifted overlay of *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). Overall stability of HA *N*-glycan fingerprints is demonstrated within the range from 24 hpi (i) to 72 hpi (ii). Both harvest time points exhibited the same 21 numbered main peaks with migration times between 150 MTU' and 420 MTU' (also see [5]). (B) Relative peak heights (RPH) of the 21 dominating peaks are represented by black (24 hpi) or red (72 hpi) columns.

Similar results were obtained for the impact of harvest time point on the HA *N*-glycosylation pattern of MDCK cell-derived IVA-Uruguay (figure 11): Virus harvested at either 24 hpi or 72 hpi exhibited 21 MDCK cell-specific peaks above

the 10x noise threshold between 160 MTU' and 400 MTU'. Overall, differences in relative structure abundance of all peaks above the 10x noise threshold were rather small with a maximal difference of ≤ 0.8 % RPH (figure 11B, $|\Delta RPH_{max}|$ for peak 6, table 3). This strongly suggests that HA of virus particles released in the supernatant is stable over the time window relevant for influenza virus production [5, 152, 153].

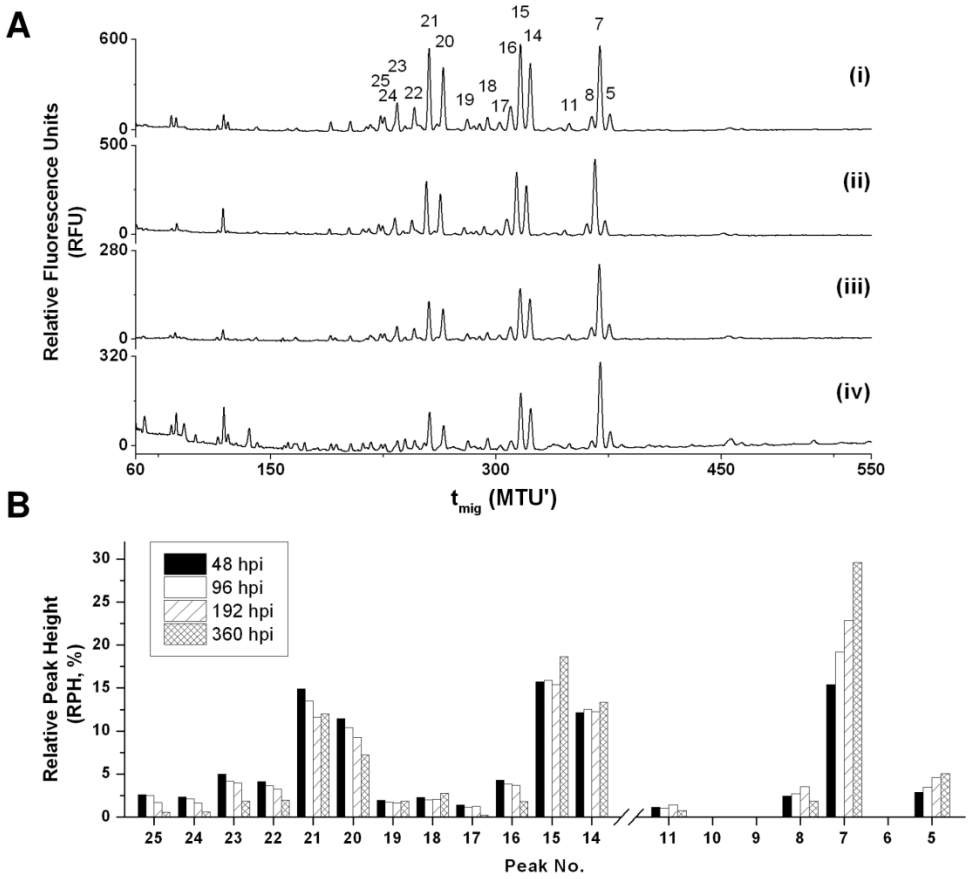


figure 12: Impact of harvest time point on the HA *N*-glycosylation pattern of Vero cell-derived RKI-strain.
(A) Shifted overlay of *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). An overall stability of the HA *N*-glycosylation pattern is demonstrated from 48 hpi (i) over 96 hpi (ii) and 192 hpi (iii) to 360 hpi (iv). The harvest time point has an impact on the RPH of the 16 numbered main peaks (peak no.: 5, 7, 8, 11, 15-25; numbering according to virus adaptation from MDCK to Vero cells and back, section 4.6.2)), exhibiting normalized migration times between 220 MTU' and 380 MTU'. (B) For each harvest time point, the relative peak heights (RPH) of all 16 dominating peaks (no.: 5, 7, 8, 11, 15-25) are represented by a column. Modified and reprinted with permission [4].

For relative quantification of attached *N*-glycans to Vero cell-derived RKI-strain 16 Vero cell-specific peaks (no. 5, 7, 8, 11, 15 - 25, [4] figure 12A) were defined. Regarding the complete time span (sampling points 48, 96, 192, 360 hpi),

differences of RPH were ≤14.2 % (figure 12B, $|\Delta RPH_{max}|$ for peak 7). Considering only the time span until 96 hpi, comparable maximal differences in RPH to the time series of the RKI-strain in MDCK cells were observed ($|\Delta RPH_{max}| \leq 3.8$ %, table 3). In particular, the RPH of peaks 5 and 7 steadily increased (almost twofold for peak 7) and the RPH of peaks 16, 20, 22 to 25 steadily decreased over time. Taking into account the comparatively long sampling period of 360 h, the higher maximal differences in RPH may be caused by differential *N*-glycan degradation or synthesis. Another possible reason may be slightly changing growth conditions in roller bottles for increasing cultivation times with decreasing pH values. This may also explain why the time series for IVA-Uruguay, which was performed in a pH-controlled 5 L-STR exhibited the lowest differences in RPH. However, even despite those changes of RPH, the HA *N*-glycosylation pattern itself was stable over 360 h [4].

Overall, these results demonstrate a highly reproducible HA *N*-glycosylation pattern for MDCK as well as Vero cell-derived HA from the RKI-strain and IVA-Uruguay with respect to number of present major *N*-glycan structures. However, the bigger the investigated time frame was, the bigger were maximal differences of relative *N*-glycan structure abundances ($|\Delta RPH_{max}|$ values) of structures (peaks) present in the different glycan pools.

3.6.3 Distribution of attached N-glycan structures over the HA_1 and HA_2 subunits in the HA_0-molecule

A part of this work focused on the impact of different process conditions on the *N*-glycosylation pattern of the IVA HA_0-molecule. The HA_0-molecule comprises two subunits, the HA_1- and the HA_2- subunits, which are connected by disulfide bonds. Altogether the HA_0 of the RKI-strain comprises seven potential *N*-glycosylation sites. Five of these are located within the HA_1-molecule and two are located within the HA_2-molecule. For later data interpretation it is of interest whether the different MDCK cell-specific *N*-glycans are distributed equally over all potential *N*-glycosylation sites or not. An equal distribution would lead to matching *N*-glycan fingerprints of HA_0, HA_1 and HA_2. Therefore, *N*-glycan

fingerprints of HA_0, HA_1 and HA_2 were characterized within the material and method section by applying reducing (HA_1, HA_2) and non-reducing conditions during SDS-PAGE sample processing. Resulting *N*-glycan fingerprints

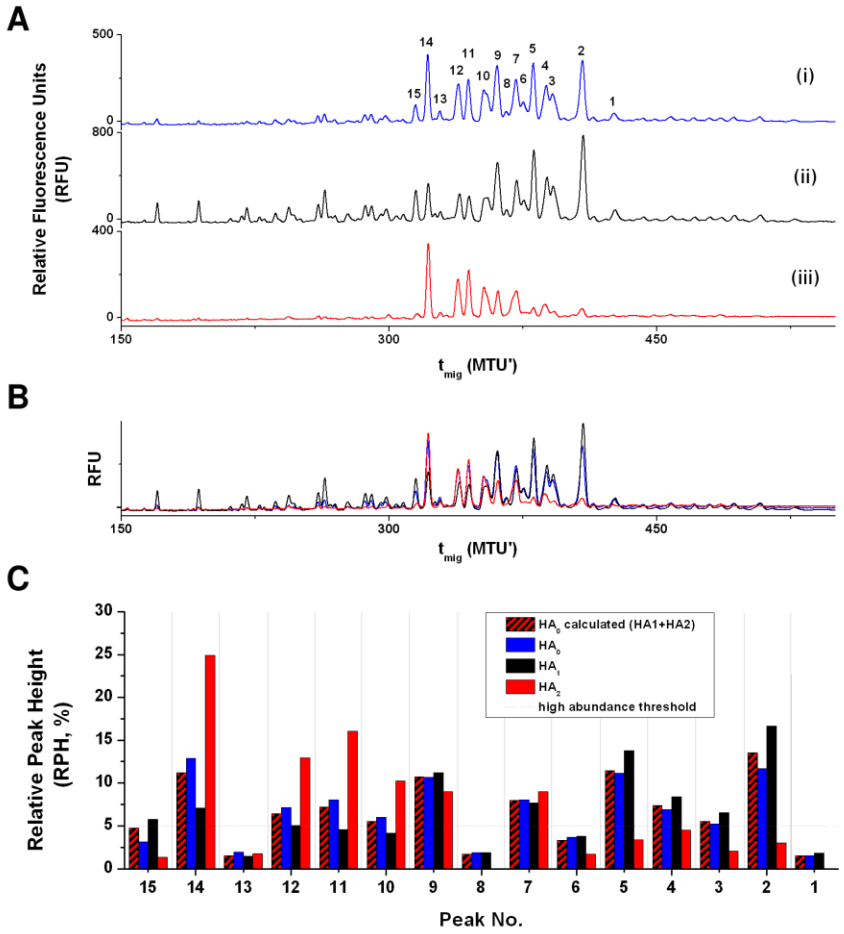


figure 13: Heterogeneity of HA *N*-glycosylation of MDCK cell-derived RKI-strain.
Shifted overlays (A) and direct overlays (B) of *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). *N*-glycosylation fingerprints from HA_0- (i), HA_1- (ii) and HA_2- (iii) molecules. MDCK cell-specific peaks between 300 MTU' and 420 MTU' are annotated (no. 1 to 15). (C) Relative quantification of MDCK cell-specific peaks of a calculated HA_0 (summed-up HA_1- and HA_2-fingerprint), the analyzed HA_0-, HA_1- and HA_2-molecule. HA_1 shows a tendency towards larger *N*-glycan structures whereas HA_2 shows a tendency towards smaller structures. HA_0 calculated (sum of HA_1 and HA_2 *N*-glycan fingerprints) is comparable to the analyzed HA_0 with respect to relative *N*-glycan structure abundances exhibiting maximum differences of 1.9 % (peak 2).

demonstrated that all MDCK cell-specific peaks (no. 1 to 15, figure 13A, B) detected for HA_0 were also present on HA_1, slightly favoring the larger structures with higher migration times indicating that not every *N*-glycan structure is equally distributed over all potential *N*-glycosylation sites. This is consistent with the

fingerprint of HA_2 exhibiting a tendency towards smaller structures. Furthermore, the low abundant peaks 1 and 8 were not detected for HA_2. The maximum difference of relative peak abundance for HA_0, HA_1 and HA_2 was 17.8 % for peak 14. Summation of HA_1 and HA_2 fingerprints (HA_0 calculated, figure 13C) resulted in comparable relative peak abundances as observed for the analyzed HA_0, showing maximum differences between HA_0 and HA_0calculated of 1.9 % RPH (peak 2). These findings demonstrate that the HA_0 *N*-glycosylation pattern, which is investigated throughout this study, results from combining HA_1 and HA_2 *N*-glycosylation patterns.

3.7 Native influenza virus deglycosylation

The following section 3.7 contains text, analogous text content and structure taken from a paper published together with J. Hütter, D. Höper, P.H. Seeberger, E. Rapp and B. Lepenies, January 2013 in *J. Immunol. [1].* As described before (section 1, last two paragraphs), scentences and/or paragraphs containing quotations are not indicated specifically. The reference will only be given after the phrase or paragraph by the number of the quoted reference.

For immunogenicity studies natively deglycosylated virus preparations, i.e. natively folded virus proteins in a non-glcosylated state, were required in order to investigate the impact of HA *N*-glycosylation on immunogenicity. Therefore, an aliquot of the MDCK and Vero cell-adapted RKI-strain was natively deglycosylated in-solution. All buffers and enzymes used for deglycosylation were purchased from Sigma-Aldrich (Steinheim, Germany) unless otherwise stated. After ultracentrifugation, the virus pellet was resuspended in 160 µL virus infection medium and 6.7 µL protease inhibitor (40x, #11777700, Roche, Mannheim, Germany), 50 µL reaction buffer (#R9150), 10 µL endoglycosidase F2 (#E0639), 10 µL endoglycosidase F3 (#E2264) and 10 µL α-galactosidase (#G8507) were added and the mixture was shaken at 450 rpm and 37 °C for 24 h in the dark. Then, 10 µL reaction buffer (#R9025) and 10 µL endoglycosidase F1 (#E9762) were added and shaking was continued at 450 rpm at 37 °C for 24 h in the dark. In the following, 10 µL reaction buffer (#R0266), 10 µL α-mannosidase

(#M7257), 10 µL α-neuraminidase (#N8271), 10 µL β-*N*-acetylglucosaminidase (#A6805), 20 U β-galactosidase (#G5160), 2 µL α-galactosidase and 2 µL endoglycosidase F3 were added and the mixture was again shaken at 450 rpm and 37 °C for 24 h in the dark. As before, the virus was isolated by ultracentrifugation at 31000 rpm and 4 °C for 90 min. The pelleted virus was resuspended in 100 mM Tris, pH 7 and stored at -80 °C. A detailed protocol is attached in the supplementary (see section 12.9.17, [1]).

3.8 Pyrosequencing and sequence evaluation

Within this study pyrosequencing was performed in cooperation with Dr. D. Höper (FLI, Greifswald - Insel Riems, Germany). The Genome Sequencer FLX instrument (Roche, Mannheim, Germany) was used for virus genome sequencing of IVA-PR8 during three independent virus adaptation series to different host cells (principle of next-generation pyrosequencing workflow is summarized in section 12.10, http://www.genomicdisorders.nl/downloads/GSFLX_Poster1.pdf to http://..._Poster4.pdf): one adaptation was published by Genzel *et al.* [72], demonstrating that at least three virus passages in the new Vero cell host are required to achieve high maximum virus titers and faster increase to maximum HA-titers (figure 14A). The final passage 4 of this adaptation was stored (-80 °C) as in-house Vero-adapted virus seed and was used for immunogenicity studies (see sections 3.10, 4.9) performed in cooperation with Dr. B. Lepenies and J. Hütter (MPI-KG, Potsdam-Golm, Germany). Two further adaptation series – one using the RKI-strain and one using the NIBSC-strain - comprised forward-adaptation over five consecutive virus passages from MDCK to Vero cells as well as back-adaptation to MDCK cells (five passages; figure 14B; see also section 4.6).

Pyrosequencing of the Vero-adapted seed virus [72] for immunogenicity studies was performed as follows: The DNA was prepared as described by Höper *et al.* [154] and resulting DNA was fragmented according to manufacturer's instruction. DNA fragments were converted to a GS FLX Titanium library using SPRIworks Fragment Library System II (Beckman Coulter, Krefeld, Germany), a SPRIworks

Fragment Library Kit II (Beckman Coulter) and a GS FLX Titanium Rapid Library MID Adaptor (Roche). The KAPA Library Quant Roche 454 Titanium Universal Kitsystem (Kapa Biosystems, Cape Town, South Africa) and the Bio-Rad CFX96 Real-Time PCR System (Bio-Rad, München, Germany) was used for library quantification. Finally emPCR was performed with 0.08 copies per bead. For final sequencing of the amplified library Titanium chemistry was used. Raw data analysis was performed using the GS FLX software suite (v. 2.5.3; Roche). For sequence assembly, primer sequences were cut off the raw data as described previously [154]. For quasispecies analysis, the GS FLX reference mapper software (version 2.5.3; Roche) and primer-trimmed, raw sequencing reads were used.

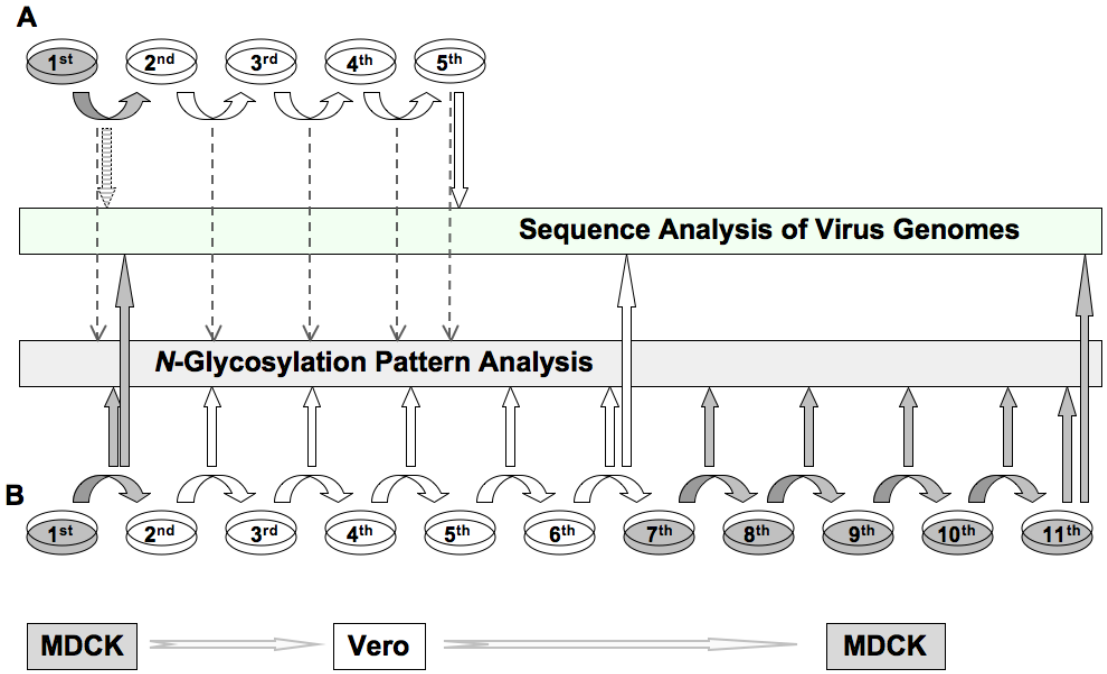


figure 14: Scheme of IVA-PR8 adaptation and sampling for glycoanalysis and next-generation pyrosequencing.
(A) Adaptation of the RKI-strain performed and published by Genzel *et al.* [72]. The MDCK cell-derived seed virus (1st passage) is matching the seed virus from B and is therefore indirectly sequenced (striped arrow), too. Virus from the 5th passage, which serves as in-house Vero cell-adapted virus seed was sequenced within this study (white arrow). HA *N*-glycan fingerprints were published by Genzel *et al.* and are not part of this work (dashed arrows). (B) The MDCK cell-derived RKI-strain as well as the MDCK cell-derived NIBSC-strain were adapted to Vero cells and back (published in [4]). Sampling for HA *N*-glycosylation pattern analysis and for next-generation pyrosequencing is indicated (short and long arrows, respectively).

Pyrosequencing of samples from the forward and backward adaptation (section 4.6) was performed as published before [4]: Briefly, DNA was prepared according

to Höper *et al.* [155], with slight modifications as specified by Leifer and colleagues [156]. Reverse-transcriptase- (RT-) PCR was performed as published by Höper *et al.* [154]. The Transcriptor High Fidelity cDNA Synthesis Kit (Roche, Mannheim, Germany) was used for reverse transcription of the viral RNA genome segments. cDNA was amplified using iProof High-Fidelity Master Mix (Bio-Rad Laboratories GmbH, München, Germany). Pyrosequencing libraries were prepared as described by Wiley *et al.* [157]. Subsequently, DNA was bound to library capture beads. Release of the unbound DNA strands resulted in single-stranded template DNA (sstDNA) library, which was used to estimate DNA quality as well as the amount required for following duplicate emulsion PCRs (emPCR). Bead-bound DNA amplification by emPCRs was performed using the GS emPCR kit I (Roche, Mannheim, Germany), applying two copies per bead. Beads were recovered, enriched and finally sequenced using the Genome Sequencer FLX instrument and the GS LR70 sequencing kit (Roche, Mannheim, Germany). After sequencing, primer sequences were trimmed off the raw read data and GS FLX sequence assembly software newbler (version 2.3; Roche, Mannheim, Germany) was used for sorting and assembling reads into contigs. Such contigs are sets of overlapping reads – finally each set displaying one virus genome segment. Quasispecies analysis was performed by aligning primer-trimmed, raw sequencing reads to the reference sequence using the GS FLX reference mapper software (version 2.3; Roche, Mannheim, Germany) applying the default parameters [4].

High confidence differences (HCD) were defined by a combination of flow signal, quality score and difference type information [4]: Required conditions for HCD included the existence of more than two non-duplicate reads showing the difference and the existence of both forward and reverse reads showing the difference, unless there are more than six reads of quality scores over 20 or 30, if the difference is at least a 5mer (GS FLX Software Manual 2.17.1.14). In order to remove sequences which have resulted from the same microreactor (water drop in water-oil emulsion) the software groups reads with matching start sequences to one read.

HCDs of the forward and backward adaptation series for segment 4 were taken from these grouped results as published in 2011 [4]. A disadvantage of this grouping is the accumulation of matching reads within the mean read length, when prior to pyrosequencing genomic material is transcribed and amplified by PCR. Therefore, all other HCDs have been extracted from ungrouped data. The trends of grouped HCDs for segment 4 are comparable with the ungrouped results, but may differ in their percentage-values. Only one subpopulation (10 %) carrying a silent base substitution (A189G = G63G) was not detected by using the grouped reads for the RKI-strain in passage 11.

The limit of detection (LOD) for this method is most probably determined by the sensitivity of prior PCRs. Höper *et al.* demonstrated a minimum sensitivity for fragments from segment 1 and segment 3 of roughly 5.6 x 10^3 copies/µL [155].

All sequences of segment 4 were uploaded to the GISAID EpiFlu database. The accession numbers are provided in the supplementary (table S 2). All further original sequence data can be requested from Dr. D. Höper (FLI, Greifswald – Insel Riems, Germany). The consensus DNA- and translated AA-sequences are provided in the supplementary (section 12.7 and section 12.8).

In the following, amino acid substitutions during the adaptation processes will be annotated by the one letter amino acid code for the original residue (in passage 1), followed by the position number of the residue, followed by the one letter code of the substituting amino acid in the later passage. Insertions and deletions are indicated by a minus sign (-) at the first or last position, respectively [4].

3.9 Sequence alignment, cDNA translation and prediction of N-glycosylation sites

Sequence alignments were performed using the Universal Protein Resource (UniProt, http://www.uniprot.org/align/, [158]). cDNA sequences were translated using ExPASy (Expert Protein Analysis System) the Bioinformatics Resource Portal from the Swiss Institute of Bioinformatics (SIB) at http://web.expasy.org/translate/ [159].

Potential *N*-glycosylation sites were predicted using the NetNGlyc 1.0 Server (http://www.cbs.dtu.dk/services/NetNGlyc/, [160]). Here, all *N*-glycosylation sequons (Asn-Xaa-Ser/Thr) are defined as potential *N*-glycosylation sites.

3.10 Immunogenicity studies using TCR-HA-transgenic mice

TCR-HA transgenic mice and BALB/c *wt* mice were used for immunogenicity studies performed at the group for glycoimmunology at the MPI-KG (Potsdam-Golm, Germany). *In vitro* studies included TCR-HA transgenic spleen cell assays whereas for monitoring T cell proliferation TCR-HA transgenic T cells were adoptively transferred to BALB/c *wt* mice. HA-specific Ab induction was quantified in BALB/c *wt* mice after prime-boost immunization with different glycovariants. For more details see [1] or contact Dr. B. Lepenies (MPI-KG, Potsdam-Golm, Germany).

4 Results

The first part of this work investigates the impact of various process conditions on the HA *N*-glycosylation pattern of IVA-PR8 by CGE-LIF-based *N*-glycan fingerprinting. Throughout all tested process conditions the IVA-PR8 was replicated in MDCK cells [4, 6]. Moreover, most results were confirmed and supplemented by CGE-LIF-based glycoanalysis data from other viruses and/or host cells [5, 7]. An overview of all conditions tested and applied process set-ups is given in table 1. Tested conditions include varying host cells, host cell adaptation stati, virus strains, virus passages, virus suppliers, virus adaptation stati, cultivation vessels and varying virus production media. Results during these studies suggested that other factors than HA *N*-glycosylation essentially contribute to improved virus titers after virus adaptation. Therefore, in a second part, the impact of virus adaptation was investigated in more detail using next-generation pyrosequencing in cooperation with Dr. D. Höper (FLI, Greifsald – Insel Riems, Germany, [4]). Finally, in the last part, the impact of differentially glycosylated virus preparations on immunogenicity was investigated in mice models (*in vitro* as well as *in vivo*; in cooperation with Dr. B. Lepenies and J. Hütter, MPI-KG, Potsdam-Golm, Germany; [1]).

4.1 Host cell

Up to date manufacturers produce influenza vaccines in different host systems, e.g. in embryonated chicken eggs or in mammalian cell lines (e.g. MDCK, Vero). Firstly, this raises questions concerning differential HA *N*-glycosylation caused by the selected host system. Secondly, host cell-specific differences in HA *N*-glycosylation would lead to the question wether they impact vaccine quality (e.g. immunogenic properties) or not.

To firstly investigate the impact of host cell on HA *N*-glycosylation, the RKI-strain was produced in MDCK, Vero, Cap, AGE1.CR.pIX cells and in embryonated chicken’s eggs (in cooperation with Dr. B. Hundt, IDT Biologika GmbH, Dessau-Rosslau, Germany). CGE-LIF-based glycoanalysis of the HA was performed and HA *N*-glycan fingerprints were compared. The comparison demonstrated strong

host cell-specificity of HA *N*-glycan fingerprints: all fingerprints, i. e. Vero; MDCK, Cap and AGE1.CR.pIX cell-derived as well as embryonated chicken egg-derived (produced in cooperation with Dr. B. Hundt; IDT Biologika GmbH, Dessau-Rosslau, Germany), differed significantly with respect to *N*-glycan structure presence and hence to relative *N*-glycan structure abundance (figure 12). Interestingly, MDCK cell-derived HA showed a tendency towards larger *N*-glycan structures in contrast to all other hosts. Small *N*-glycan structures below 300 MTU' were especially favored, when virus was produced in embryonated chickens eggs.

These results confirm previous studies of our group reporting host cell-specificity of HA *N*-glycosylation for MDCK and Vero cell-derived IVA-PR8 [73]. Furthermore, these results extend the host cell-specific nature of HA *N*-glycosylation for embryonated chicken eggs-, AGE1.CR.pIX and Cap cell-derived IVA-PR8.

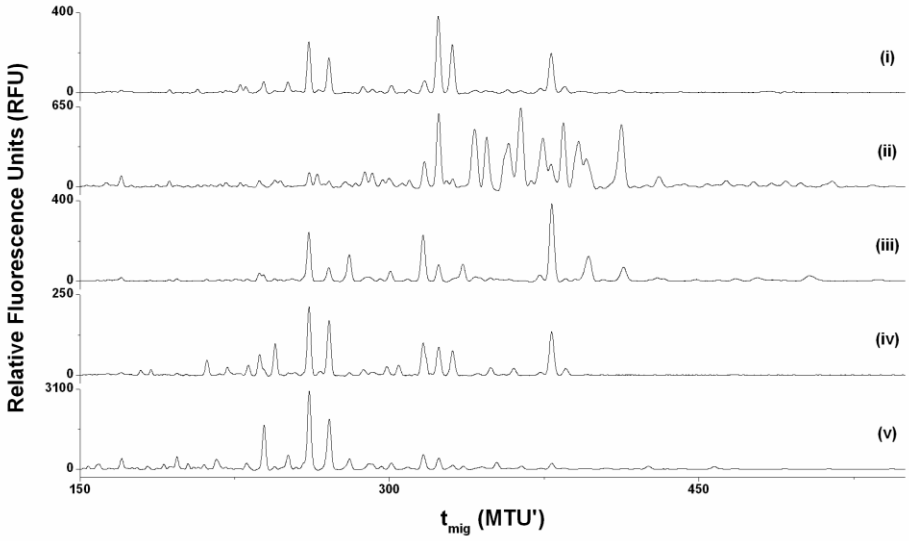


figure 15: Impact of host cells on the HA *N*-glycosylation pattern.
Shifted overlay of *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). IVA-PR8 (RKI-strain) was produced in adherent Vero cells (i) of African green monkey origin or in adherent MDCK cells (ii) of canine origin. Furthermore, suspension cells such as human-derived Cap cells (iii) or duck-derived AGE1.CR.pIX cells (iv) were used for virus production. Finally, cell culture-derived virus was compared with virus produced in embryonated hen's eggs (v, produced in cooperation with Dr. B. Hundt, IDT Biologika GmbH, Dessau-Rosslau, Germany).

4.2 Host cell adaptation

4.2.1 Cell growth in serum-containing and serum-free medium

For the production of biologicals the use of animal-derived products is problematic due to risks involved with adventitious agents. Therefore, most commercial processes avoid the need for animal-derived medium compounds; i.e. generally serum, by host cell adaptation to serum-free growth. Here, HA *N*-glycosylation of the RKI-strain produced in non-adapted, serum-requiring MDCK cells was compared with virus produced in MDCK cells adapted to serum-free growth in Episerf medium. In all *N*-glycan fingerprints, 16 MDCK cell-specific peaks were detected between 300 MTU' and 450 MTU' (figure 16A), which have been described before [4, 7, 73]. Relative quantification of these peaks revealed a subset of 10 high abundant *N*-glycan peaks for non-adapted cells (2 - 5, 7, 9, 11 – 13 and 15, figure 15B). For serum-free growth adapted cells a slightly different subset of 10 high abundant *N*-glycan peaks (2, 5, 7, 9 – 15) was found (figure 16B). The differences of RPH were in the range between 0.3 % and 8.5 % ($|\Delta RPH_{max}|$ for peak 14, figure 15B, table 3, [6]).

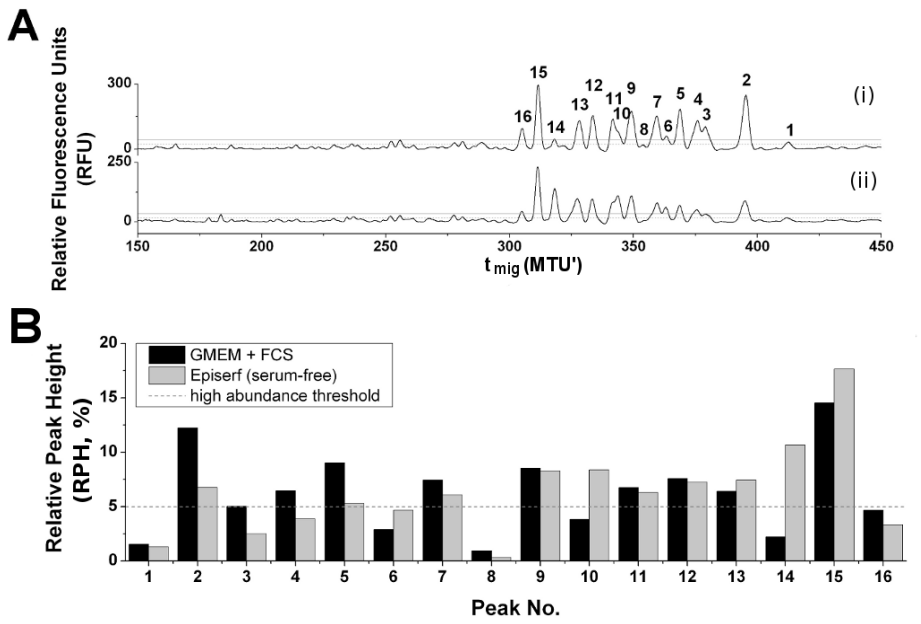


figure 16: Impact of host cell adaptation to serum-free growth on the HA *N*-glycosylation pattern of the RKI-strain.

(A) Shifted overlay of HA *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). MDCK cell-specific peaks between 300 MTU' and 450 MTU' exceeding the 10x baseline noise threshold in at least one of the direct comparable fingerprints (table 3, +) are annotated. The non-adapted, serum requiring MDCK cell line (i) was adapted to growth in serum-free Episerf medium (ii). (B) Relative peak abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined low abundant if RPH < 5 % (---). Modified and reprinted with permission [6].

table 3: Differences of relative peak height (IΔRPHI) of all annotated peaks of all performed fingerprint comparisons (numbered according to sections in the manuscript).

In the following, an example is shown for the calculation of IΔRPHI (section 3.6.2; harvest time point, RKI in MDCK; peak no. 1): the RPH of peak 1 is 1.4 % at 24 hpi (figure 10, B, grey column), whereas at 96 hpi the RPH of peak 1 is 1.7 % (figure 10B, white column). Hence, IΔRPHI = I1.4 % - 1.7 %I = 0.3 %. If more than two peaks are compared (e.g. section 3.6.2; harvest time point, RKI in Vero; peak 5), IΔRPHI is calculated from the highest (figure 11, grid column, 360 hpi) and the lowest abundant peak (figure 11, black column, 48 hpi). Maximal difference ($I\Delta RPH_{max}I$) of each experiment is highlighted in bold.

section	3.6.2				4.2.1	4.2.2		
peak	harvest time point RKI in MDCK*	harvest time point URU in MDCK#	harvest time point RKI in Vero*	harvest time point RKI in Vero until 96 hpi*	serum↔ serum-free+	serum↔serum-free suspension (MDCKadh,SUS1,2,3)'	serum-free suspension (SUS1,2,3)'	serum-free suspension (SUS1,2)'
no.	|ΔRPH| (%)	|ΔRPH| (%)	|ΔRPH| (%)	|ΔRPH| (%)	|ΔRPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)
1	0.3	0.5			0.2	1.5	0.0	0.0
2	0.8	0.1			5.5	7.3	7.3	0.0
3	0.0	0.2			2.5	1.8	0.9	0.8
4	0.7	0.2			2.6	8.1	2.7	0.2
5		0.2	2.2	0.6	3.7	4.0	1.8	0.7
6	0.8	**0.8**			1.8	6.6	1.8	0.8
7	0.0	0.4	**14.2**	**3.8**	1.3	9.2	0.0	0.0
8	0.0	0.3	1.7	0.3	0.6	9.7	9.7	0.0
9	0.3	0.4			0.3	3.3	2.3	0.3
10	0.1	0.5			4.5	9.5	9.5	1.0
11	0.3	0.2	0.6	0.1	0.4	11.4	3.6	0.0
12	0.1	0.2			0.3	2.7	1.5	0.1
13	0.7	0.2			1.0	7.2	2.9	0.2
14	0.1	0.2	1.2	0.4	**8.4**	4.3	4.3	1.1
15	**3.3**	0.3	3.3	0.2	3.1	9.2	9.2	1.1
16	1.2	0.4	2.5	0.4	1.3	2.1	2.1	0.6
17		0.1	1.1	0.3		14.5	10.2	**1.9**
18		0.3	0.7	0.3		**18.1**	**18.1**	0.2
19		0.1	0.3	0.2		3.7	2.3	1.6
20		0.5	4.2	1.0		1.9	0.6	0.3
21		0.8	3.3	1.4		1.6	1.6	0.3
22			2.1	0.5		2.3	0.0	0.0
23			3.1	0.8		6.8	6.8	0.9
24			1.7	0.2		1.9	0.0	0.0
25			2.0	0.1		14.2	3.7	1.7
26						2.0	1.2	1.0
27								
28								
29								
30								

*, #, +, ‘, °, ^, § matching tags indicate identical peak annotation. Hence a comparison across different experiments is possible.

table 3 continued

section	4.3			4.4	4.5	4.6.1 (RKI-strain)		4.6.1 (NIBSC-strain)		4.6.2 (RKI-strain)	
peak	all strains°	H1N1 strains°	H3N2 strains°	diff. virus passages (no. 1-10)*	RKI↔ NIBSC+	MDCK-derived passages (no. 1, 7-11)*	Vero-derived passages (no. 2-6)*	MDCK-derived passages (no. 1, 7-11)*	Vero-derived passages (no. 2-6)*	all passages (no. 1-5)^	AGE1.CR.pIX-derived passages (no. 2-5)^
no.	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)
1	2.5	1.5	1.0	0.8	0.2	0.5	0.0	0.6	0.0	12.3	0.0
2	2.4	0.9	2.1	2.8	**5.0**	4.0	0.0	3.7	0.0	3.5	0.0
3	1.5	0.7	0.7	1.4	0.5	1.4	0.0	0.8	0.0	8.6	0.0
4	8.3	0.7	0.6	1.6	3.4	2.9	0.0	2.0	0.0	3.1	2.3
5	2.6	1.9	1.1		2.9	0.0	4.2	0.0	3.9	12.9	0.0
6	1.1	0.2	0.1	1.3	0.4	5.6	0.0	8.5	0.0	8.8	2.8
7	9.4	3.5	0.2	0.9	1.6	1.6	**21.0**	4.4	**14.8**	6.2	0.0
8	4.4	2.2	4.4	1.0	1.8	0.7	3.6	0.8	1.1	10.1	0.0
9	9.4	3.7	0.6	1.1	0.8	1.2	0.0	1.3	0.0	2.9	2.2
10	7.8	7.4	0.1	1.8	0.1	2.4	0.0	7.0	0.0	4.1	0.0
11	2.4	0.7	2.3	1.4	0.2	1.6	2.6	2.5	0.7	10.5	0.0
12	15.3	**12.4**	3.8	1.5	0.4	3.7	0.0	3.4	0.0	3.8	0.0
13	3.5	3.5	0.5	1.9	0.7	3.5	0.0	3.9	0.0	4.9	1.9
14	1.4	1.4	1.0	1.1	1.4	2.9	7.1	4.8	4.0	2.2	1.3
15	3.6	1.0	3.6	**3.5**	1.5	**9.5**	5.9	**9.9**	3.7	2.0	0.8
16	3.5	1.8	3.5	1.5	0.4	3.5	4.1	3.0	2.6	3.0	1.2
17	3.6	0.5	3.1			0.0	2.0	0.0	0.7	2.3	2.3
18	3.9	3.2	0.1			0.0	4.1	0.0	1.0	1.8	1.0
19	3.3	0.9	2.5			0.0	0.6	2.4	2.4	17.0	4.6
20	3.2	0.4	2.5			0.0	5.6	0.9	4.1	3.2	0.0
21	1.2	0.5	1.2			0.0	8.2	5.9	3.2	**26.7**	**13.1**
22	**25.2**	0.7	**9.8**			0.0	1.0	0.0	1.4	9.3	2.0
23	5.2	5.2	1.5			0.0	5.2	0.0	1.3	5.3	1.4
24	3.7	0.0	0.1			0.0	1.5	0.0	1.5	3.9	2.0
25	2.6	2.0	0.5			0.0	4.9	0.0	2.0	5.7	2.5
26	7.5	2.3	2.5								
27	2.0	2.0	0.4								
28	4.5	2.5	1.6								
29	6.2	0.9	2.8								
30	2.5	2.5	0.6								

*, #, +, ‘, °, ^, § matching tags indicate identical peak annotation. Hence a comparison across different experiments is possible.
RKI: RKI-strain
NIBSC: NIBSC-strain
SUS1, 2, 3. MDCK.SUS1, MDCK.SUS2, MDCK.SUS3
IVA-Uru: IVA-Uruguay

table 3 continued

section	4.6.3 (RKI-strain)	4.6.4 (IVA-Uru)		4.6.5 (IVA-California)	4.6.5 (IVA- Viktoria)	4.7 (RKI-starin)	4.7 (IVA-Uru)	4.8.1	4.8.2
peak	egg-derived passages (no. 1-3)	all passages (seed, no. 1-5)[§]	Vero-derived passages (no. 1-5)[§]	MDCK-derived passages (no. 1-3)	MDCK-derived passages (no. 1-3)	scales & vessels[+]	scales & vessels[#]	virus production media composition[+]	trypsin activity of virus production media[+]
no.	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)	|Δ RPH| (%)
1	2.3	6.9	4.0	1.3	0.5	2.5	0.7	1.5	1.4
2	2.5	2.2	1.8	0.6	0.7	**7.3**	1.6	3.5	1.1
3	0.9	3.4	2.0	1.0	0.2	5.3	0.1	1.7	0.9
4	0.8	2.2	0.9	2.5	1.1	2.5	0.6	2.7	1.9
5	1.2	4.0	2.1	1.5	0.8	3.7	0.8	1.7	4.6
6	0.7	1.0	1.0	1.1	1.5	0.4	**6.6**	1.6	0.7
7	3.2	2.1	0.7	**3.1**	1.1	1.2	0.9	0.7	0.6
8	1.2	3.5	2.6	1.0	1.3	0.2	0.9	0.3	0.8
9	2.0	3.9	2.2	2.9	1.2	5.5	0.8	3.6	5.2
10	0.5	**14.9**	**5.1**	0.6	1.1	0.9	3.7	1.3	1.6
11	**8.3**	5.5	2.8	3.0	0.6	2.5	2.0	1.3	2.6
12	0.9	6.2	0.0	1.5	0.2	5.4	0.4	**5.7**	4.0
13	5.5	2.0	0.0	0.9	0.2	2.1	3.2	0.2	6.2
14	2.3	2.6	1.3	0.2	0.5	0.7	3.6	2.6	**14.0**
15	1.9	2.8	0.9	0.6	0.4	6.8	0.6	5.7	8.5
16	1.0	6.7	1.6	1.2	0.7	2.5	2.9	2.6	1.8
17	1.5	2.7	1.3	0.1	0.7		1.3		
18	1.0	1.3	0.0	0.5	0.6		2.6		
19	0.5	4.5	0.0	1.9	1.3		3.5		
20	0.0	2.0	0.0	1.3	0.5		1.4		
21	1.0	2.6	0.0	0.4	1.6		4.5		
22	0.2	3.3	0.0	0.3	0.9				
23		7.9	3.1	0.6	0.2				
24		1.9	0.0	1.1	0.4				
25		3.0	1.5	0.6	2.0				
26		2.4	0.0	1.8	0.2				
27		1.1	0.0		**2.5**				
28		3.1	0.0		0.9				
29					1.2				
30									

*, [#], [+], ‘, °, ^, [§] matching tags indicate identical peak annotation. Hence a comparison across different experiments is possible.

4.2.2 Adherent versus suspension growth

Both, the use of adherent and suspension cells has pros and cons [6, 61]. For instance, easier passaging without the need of trypsin treatment and no requirement of tissue-culture treated vessels or microcarriers represent clear advantages of suspension cell lines. In contrast easier washing of microcarrier-associated adherent cells during stirred cultivations are among the advantages of adherent cell cultivations. Therefore, the impact of an adaptation of adherent MDCK cells to serum-free suspension growth on the HA *N*-glycosylation pattern was characterized [6]. Briefly, an adherently growing, serum-dependent MDCK cell line was adapted to serum-free suspension growth in two independent biological duplicates in cooperation with Prof. Dr. K. Scharfenberg (University of Applied Sciences, Emden, Germany, [142]). Two suspension cell lines namely MDCK.SUS1 and MDCK.SUS3 resulted. The suspension cell line MDCK.SUS1 was further adapted to improve growth kinetics in suspension, resulting in the MDCK.SUS2 cell line [142]. These four cell lines (MDCK, MDCK.SUS1, MDCK.SUS2, MDCK.SUS3) were used for RKI-strain production and HA *N*-glycosylation was characterized by CGE-LIF-based glycoanalysis. The comparison of all four fingerprints revealed significant differences between MDCK, MDCK.SUS1 and MDCK.SUS3 cell lines regarding the types of glycans attached as well as their relative abundances (figure 17). MDCK cell-derived fingerprints exhibit a higher number of high abundant peaks than all MDCK.SUS cells. Furthermore, 100 % of all high abundant peaks of MDCK cells have normalized migration times above 320 MTU', whereas 17 % to 50 % of all high abundant peaks (number 25) of any MDCK.SUS cell line have migration times below 320 MTU' (figure 17A). A detailed comparison of fingerprints reveals that peaks 1, 7, 22 and 24 were only present in MDCK, peak 15 was only present in MDCK.SUS1 and MDCK.SUS2, whereas peaks 2 and 8 were only present in MDCK.SUS3 (figure 17B). The maximum difference of RPH was 18.1 % (peak 18, figure 17C). The comparison of RPH of MDCK.SUS1 and

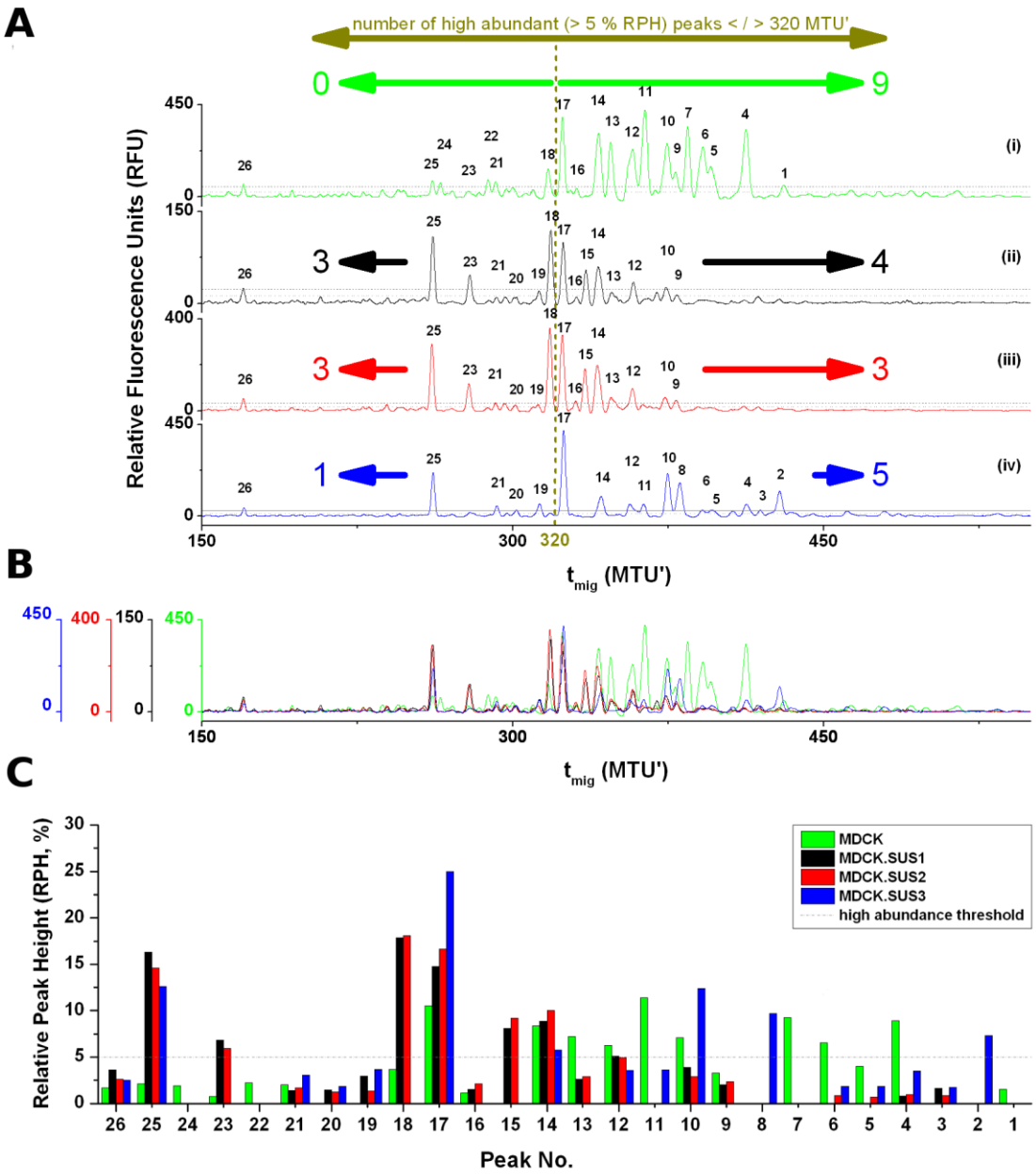


figure 17: Impact of host cell adaptation to serum-free suspension growth on the HA *N*-glycosylation pattern of the RKI-strain.
(A) HA *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). All peaks exceeding the 10x baseline noise threshold in at least one fingerprint are annotated. Serum-requiring adherent MDCK cell line (i), MDCK cell line adapted to serum-free suspension growth (ii, MDCK.SUS1) and further adapted MDCK.SUS1 cell line to better growth characteristics (iii, MDCK.SUS2) [142]. The first adaptation step was performed in biological duplicates (iv, MDCK.SUS3). The number of high abundant peaks (RPH > 5 %) with migration times below or above 320 MTU' is indicated. (B) Overlay of all four *N*-glycosylation fingerprints. (C) Relative *N*-glycan structure abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks < 5 % RPH (---) are defined low abundant.

MDCK.SUS2-derived fingerprints revealed extensive analogy (max. |ΔRPH| = 1.9 %, peak 17; figure 17C). Interestingly, the first adaptation step to serum-free suspension growth resulted for both adaptation series (MDCK to MDCK.SUS1 and MDCK to MDCK.SUS3) in a totally new cell line, not only with respect to their growth characteristics, e.g. reduced doubling times and smaller aggregates [142], but also to their HA *N*-glycosylation characteristics. Therefore, figure S 1 (extended figure 15) shows the impact of host system (including MDCK.SUS2 and MDCK.SUS3) on the HA *N*-glycan fingerprint.

These results were consistent with data obtained from NIBSC-strain production in MDCK, MDCK.SUS1, MDCK.SUS2 and MDCK.SUS3 cells (section 12.2, figure S 2).

4.3 Virus strain

Most manufacturers have their own, approved production cell line, in which viruses are propagated for all seasonal or pandemic influenza vaccines. In order to investigate the impact of the virus strain on the HA *N*-glycosylation pattern *N*-glycan fingerprints from different MDCK cell-derived IVA were analyzed: two strains, namely RKI-strain and IVA-California, belong to the H1N1 subtype and two, namely IVA-Uruguay and IVA-Victoria, belong to the H3N2 subtype. The comparison of fingerprints demonstrated that most peaks were present for all tested viruses (except for low abundant peaks 2, 5, 15, 16, 24 and 30, figure 18A, B). The relative peak abundance varied with a maximum difference of 25.2 % (peak 22, figure 18C). Separate consideration of H1N1 and H3N2 strains resulted in much smaller maximum differences in RPH of 12.4 % (H1N1-strains) and 9.8 % (H3N2-strains), respectively (section 12.3, figure S 4).

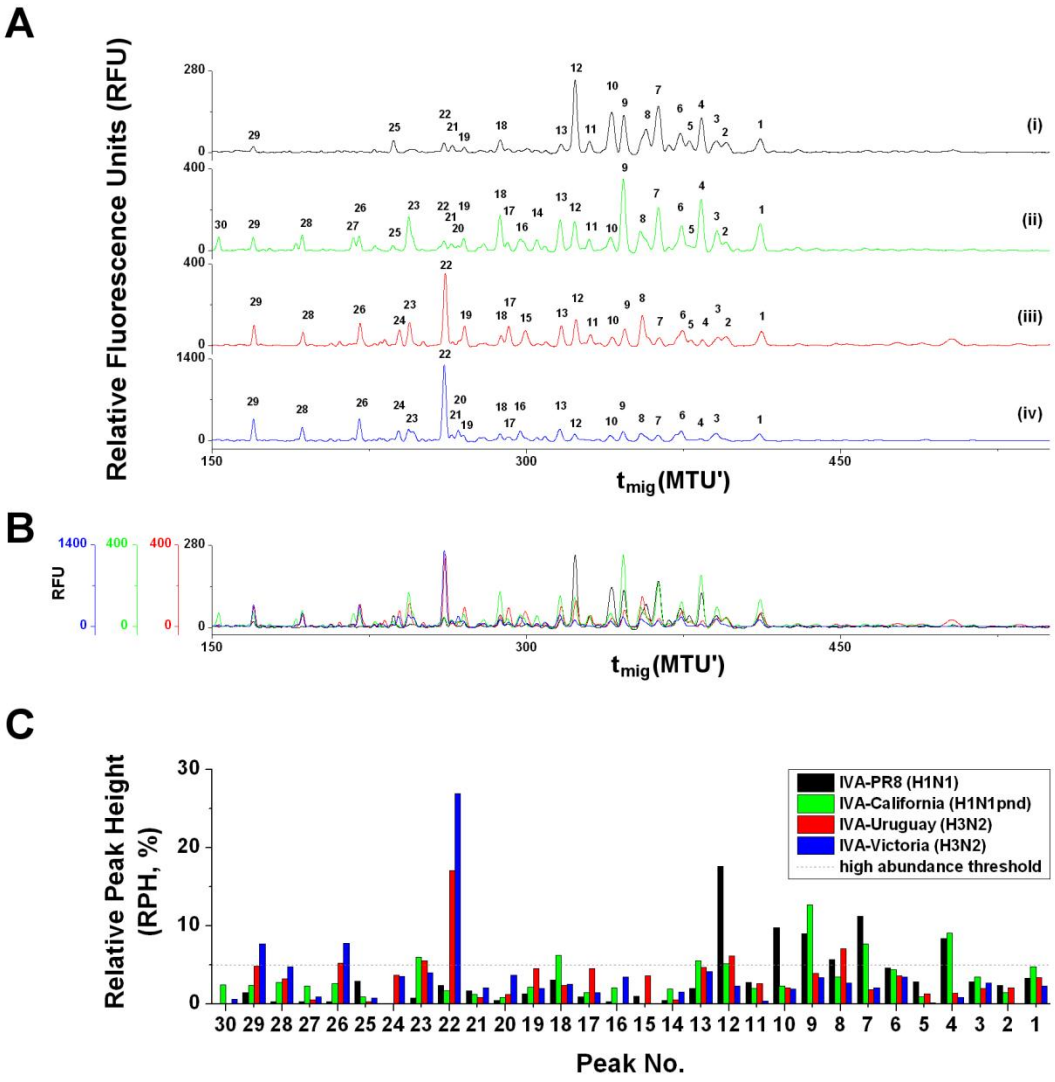


figure 18: Impact of virus strain on MDCK cell-derived HA *N*-glycosylation patterns.
(A) Shifted overlay of HA *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). IVA-PR8 (i), IVA-California (ii), IVA-Uruguay (iii) and IVA-Victoria (iv) were produced in MDCK cell culture. All peaks exceeding in at least one of the fingerprints (i-iv) the 10x baseline noise threshold are annotated (no. 1 - 30). (B) Direct overlay of HA *N*-glycan fingerprints. C) Relative *N*-glycan structure abundances (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined high abundant if RPH > 5 % (---).

4.4 Virus passage

The stability of HA *N*-glycan fingerprints during multiple virus passages in one cell line is one essential condition for the characterization of HA *N*-glycosylation patterns using different cultivation conditions in influenza virus production. Therefore, in this study 10 consecutive passages of the RKI-strain were

produced in MDCK cells using roller bottles. 24 hpi HA-titers of these 10 virus passages ranged between 2 and 2.4 HAU (data not shown). Moreover, throughout all passages HA *N*-glycosylation fingerprints all featured the same 15 main peaks (no. 1-4, 6-16, figure 19A). Maximum differences of RPH were 3.5 % (peak 15, table 3, figure 19B, [4]). Overall, these results demonstrate a high stability of the HA *N*-glycosylation pattern over 10 successive virus passages in the same host cell system [4].

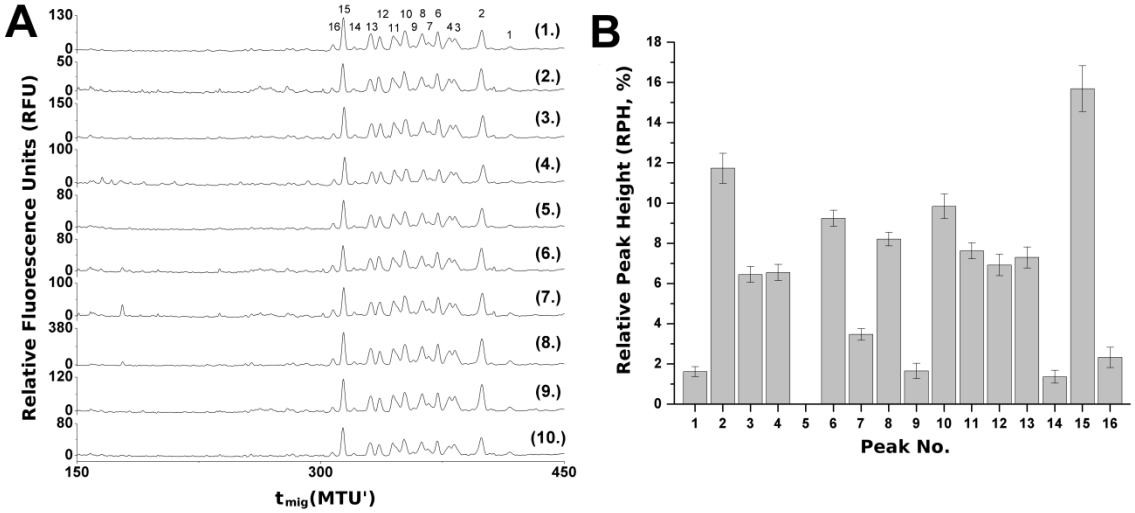


figure 19: Impact of virus passage on the HA *N*-glycosylation patterns of the RKI-strain. (A) Shifted overlays of *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time in normalized migration time units (MTU'). HA *N*-glycosylation patterns are reproducible over 10 successive virus passages in MDCK cells. All 10 patterns exhibit the same 15 numbered main peaks (no. 1 – 4, 6 – 16, numbering according to virus adaptation from MDCK to Vero cells and back, section 4.6.2, table 3 indicated by *) between 300 MTU' and 420 MTU'. (B) Relative peak heights (RPH) of the 15 main peaks. Standard deviations (error bars) for 10 successive virus passages, range between 0.25 % and 1.14 %. Modified and reprinted with permission [4].

4.5 Virus supplier

Next, two IVA-PR8 seed viruses were compared. One was provided from the Robert Koch Institute (RKI-strain) and the other was purchased from NIBSC (NIBSC-strain). Both strains differ in seven amino acid positions within the HA molecule [4]. Both HA *N*-glycan fingerprints exhibited the set of 16 adherent MDCK cell-specific peaks (1-16) between 300 MTU' and 450 MTU', (figure 20A). For the RKI strain a total number of eight high abundant *N*-glycan peaks (2, 4, 5, 7, 9, 12, 15 and 16) was identified (figure 20B). For the NIBSC-strain, the total number of high abundant peaks was reduced as peak 7 dropped below 5 % RPH

(figure 20B). The minimum and maximum differences of RPH of the two strains were calculated with 0.1 % and 5 %, respectively ($|\Delta RPH_{max}|$: peak 2, table 3, [6]).

Producing the RKI- and the NIBSC-strain in MDCK.SUS2 cells resulted in similar findings. Interestingly, the minimal and maximum differences of RPH increased to 0.6 % and 11.0 %, respectively ($|\Delta RPH_{max}|$: peak 9, section 12.3, figure S 3).

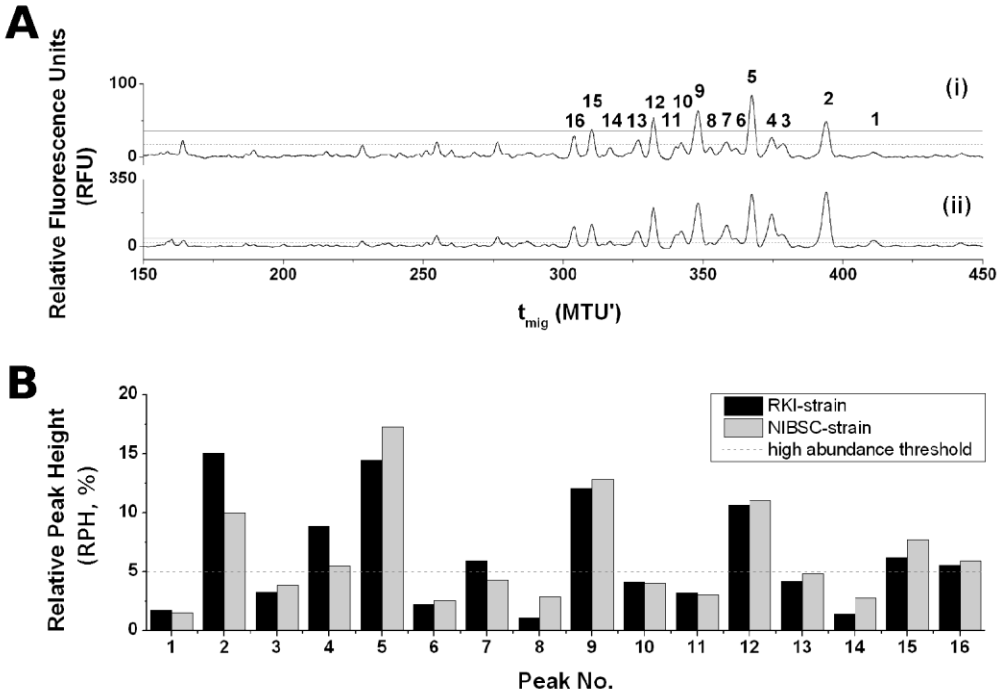


figure 20: Comparison of two IVA-PR8 strains from different suppliers with respect to HA *N*-glycosylation patterns.
(A) Shifted overlay of HA *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). MDCK cell-specific peaks between 300 MTU' and 450 MTU' exceeding the 10x baseline noise threshold (---) in at least one of the direct comparable fingerprints (table 3, +) are annotated. The virus was purchased from either the Robert Koch Institute (i) or NIBSC (ii). (B) Relative peak abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined low abundant if RPH < 5 % (- - -).Modified and reprinted with permission [6].

4.6 Virus adaptation

4.6.1 Virus replication dynamics

Usually, viruses require adaptation to new host cells to optimize yields and time of harvest [72]. In order to characterize this adaptation process and to investigate possible biological mechanisms MDCK cell-adapted virus was propagated for 24 h in MDCK cells (passage 1). As a result HA-titers of 1.9 HAU and 2.2 HAU

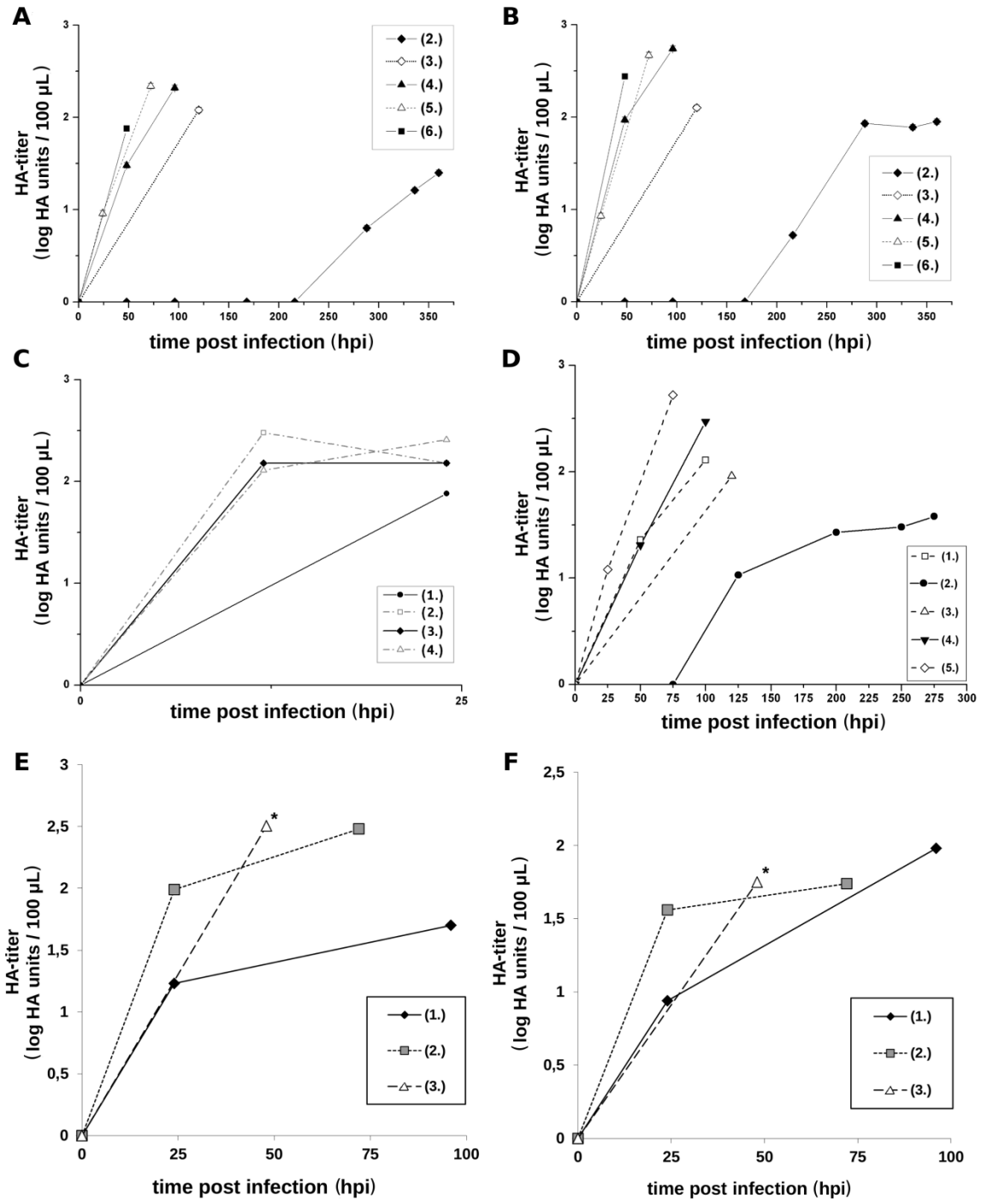


figure 21: HA-titers during IVA adaptation to different host cells.
MDCK cell-adapted RKI-strain (A) or NIBSC-strain (B) served as seed for infection of five consecutive passages of Vero cells (no. 2 to 6). Reprinted with permission [4]. (C) MDCK cell-adapted RKI-strain was adapted to replication in AGE1.CR.pIX cells during 4 consecutive passages (no. 1 to 4). (D) MDCK cell-adapted IVA-Uruguay was adapted during 5 consecutive passages (no. 1 to 5) to replication in Vero cells. Egg-adapted IVA-Viktoria (E) and IVA-California (F) were adapted during 3 consecutive passages (no. 1 to 3) to replication in MDCK cells. * the mean value of two measurements is plotted. The 95 % confidence interval of the HA-assay is maximum 15 %. All titers at 0 hpi are defined as zero.

were obtained for the strains from RKI and NIBSC, respectively (see supplementary, table S 1). Supernatant of this passage served as the virus seed for a first infection of Vero cells (passage 2). For the RKI-strain virus-release was not detected before 288 hpi (0.8 HAU, figure 21A) and the maximal HA-titer of 1.4 HAU was finally reached 360 hpi. For the NIBSC-strain virus release was first detected 216 hpi and virus replication was continued until 360 hpi, exhibiting a final HA-titer of 1.95 HAU (figure 21B). An aliquot of these supernatants served as virus seeds for the next infections of Vero cells (passage 3) and so on (passage 4 to 6). During this virus adaptation to Vero cells, viral fitness of both strains, RKI and NIBSC, improved: The time required to achieve specific HAU-values (≥ 1.4 HAU) decreased, whereas maximum HA-titers increased from passage 2 to 4. No significant differences were detected for passage 4 to 6 of the RKI-strain, indicating the completion of the adaptation process (figure 21A). For the NIBSC-strain, the second passage in Vero cells (passage 3) reached 2.1 HAU at 120 hpi, and passage 4 reached 2.7 HAU at 96 hpi. For all subsequent passages (numbered 5 and 6) the time required to achieve a HA-titer of at least 1.8 HAU as well as the maximal titers were more or less the same. This indicated the completion of the adaptation process (figure 21B). The following five virus passages (7 to 11) were again performed in MDCK cells to monitor virus back-adaptation. Here, for the RKI- as well as the NIBSC-strain all titers ranged between 2 HAU and 2.5 HAU at 48 hpi and 96 hpi (see supplementary, table S 1). In contrast to the adaptation to Vero cells, no impact on HA-titer level and virus release dynamics was observed during back-adaptation [4].

Similar results, though less distinct, were obtained for the adaptation of the RKI-strain from MDCK to duck-derived AGE1.CR.pIX cells (passage 1 to 4, figure 21C): Virus adaptation to the new avian host cell line allowed an increment of virus titers (figure 21C). Accordingly, virus fitness in the new host system has improved during the adaptation process. Nevertheless, the virus titer of passage 1 with 1.9 HAU suggests that the duck-derived AGE1.CR.pIX cell line allowed rather efficient virus replication from the beginning [7].

The efficient replication from the beginning in duck cells is consistent with data obtained from adaptation of the MDCK cell-derived RKI-strain to replication in embryonated chicken eggs, which was performed in cooperation with Dr. B. Hundt (IDT Biologika GmbH, Dessau-Rosslau, Germany). For each passage, three different virus dilutions (10^0, 10^{-1}, 10^{-2}) were used to infect three sets of five eggs (altogether 15 eggs). The virus was harvested 96 hpi and pooled. The titer of the first passage pool already reached 3.0 HAU (two eggs of the 10^{-2}- dilution died before harvest at 48 hpi and 72 hpi; data not shown). An aliquot of this first passage pool served for the infection of three further sets of embryonated chicken eggs. Again, 96 hpi the virus from 13 surviving eggs (two eggs died: dilution 10^{-1}/72 hpi and 10^{-2}/72 hpi) was pooled. The titer again reached 3.0 HAU (data not shown). The pool of the third passage from altogether 12 eggs (three eggs died: dilution 10^0/72 hpi, 10^0/96 hpi and 10^{-1}/72 hpi) reached a titer of 3.3 HAU (96 hpi; data not shown). Interestingly, the titers reached (3.0 HAU to 3.3 HAU) during the adaptation from MDCK cells to embryonated chicken eggs demonstrated that embryonated chicken eggs allowed rather efficient virus replication from the very beginning and further increase of virus yield by adaptation was not detected. The RKI-strain - as it origins from an egg-adapted ancestor - seems to have retained its ability to replicate rather well in avian cells.

Interestingly, during adaptation of MDCK cell-derived IVA-Uruguay to replication in Vero cells the HA-titer of the second passage in Vero cells was even lower than the titer of the first passage but in passages 3, 4 and 5 titers increased (figure 21D). Taken all data together, the adaptation resulted in a final titer increase from 2.1 HAU (passage 1, 96 hpi) to 2.7 HAU (passage 5, 72 hpi).

Furthermore two egg-derived vaccine strains IVA-Victoria and IVA-California were adapted to MDCK cells. Therefore, the egg-adapted virus lyophilisate was resuspended in sterile H_2O_{MQ}, and was used to infect a first passage of MDCK cells. 24 hpi an aliquot of the supernatant was used to infect the second passage of MDCK cells and so on. The virus replicated well from the beginning. The adaptation resulted in increased maximum virus titers for IVA-Viktoria (2.5 HAU, 48 hpi, figure 21E). In contrast final titers of IVA-California did not significantly

increase by adaptation reaching 1.75 HAU 48 hpi in the third passage (figure 21F).

In a next step, it was tested if improved growth characteristics after virus adaptation were maintained, even if no selection pressure persisted. Therefore, after the adaptation from MDCK to Vero cells (in 5 passages) and back to MDCK cells (in another 5 passages) the RKI- as well as the NIBSC-strain were adapted again to replication in Vero cells (in 3 passages, numbered passages 12 to 14, figure 22). This further adaptation to Vero cells highlighted that the forth- and back-adapted virus has somehow acquired and kept (throughout MDCK-passaging, passages 7-11) the ability to replicate efficiently in Vero cells (figure 22): HA-titers already increased after one (NIBSC-strain) or two days (RKI-strain) in passages 12 and even earlier for the RKI-strain in passages 13 and 14. In contrast, a Vero cell-passage (passage 2') infected with supernatant of non-adapted virus from passage 1 showed titer increase not before 94 hpi (NIBSC-strain) and 160 hpi (RKI-strain). This clearly demonstrated an improved viral fitness in Vero cells after virus adaptation.

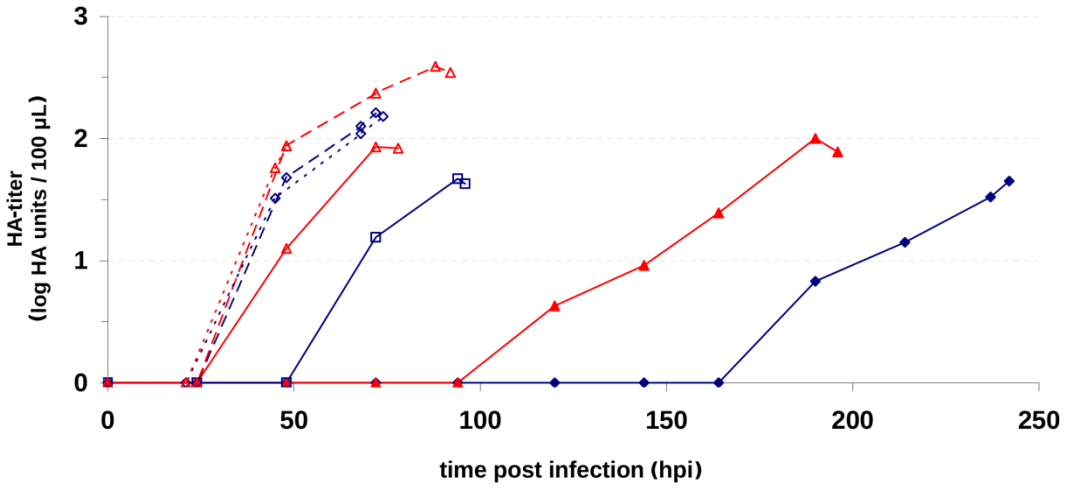


figure 22: HA-titers of IVA-PR8 from RKI (blue) and NIBSC (red) during first (▲, ♦) and second (Δ, ◊) adaptation to Vero cells.
MDCK cell-derived virus from passage 1 (no Vero cell contact before) and from passage 11 (Vero cell contact during passages 2- 6) was used to infect a passage of Vero cells (passage 2': filled symbol, --- and passage 12: empty symbol, ——), respectively. An aliquot of the supernatant of passage 12 was used to infect another passage of Vero cells (passage 13, empty symbol, – – –) and so on (passage 14, empty symbol, ----). Viruses, which had Vero cell contact before replicated faster in Vero cells than viruses without previous Vero cell contact. Viruses have retained the ability to replicate in Vero cells during several passages (7-11) in MDCK cell culture.

4.6.2 Host cell-specificity of HA N-glycosylation patterns during virus adaptation

Next, *N*-glycan fingerprints were analyzed for all adaptation passages (see also section 4.6.1) to investigate the impact of virus adaptation on HA *N*-glycosylation. All HA *N*-glycosylation patterns of MDCK cell-derived IVA-PR8 samples were similar. The same applied to all IVA-PR8 Vero cell-derived virus samples (figure 23A, B). In agreement with earlier studies the HA *N*-glycosylation pattern was strictly host cell-specific and changed significantly with the switch to the new host cells [72, 73]. However, of all Vero cell-derived HA *N*-glycosylation patterns for the RKI- as well as the NIBSC-strain passage 2 revealed the biggest differences in RPH (figure 23, figure 24). Here, the RPH of peak 5 and 7 was almost twice as high as for all subsequent passages in Vero cells. This is in agreement with the time series in Vero cells (also see section 3.6.21.1.1; figure 12) where these RPH almost doubled until 360 hpi. Furthermore, the low abundant glycan structure represented by peak 11 was missing in passage 2 of both IVA-PR8 strains (figure 23, figure 24). This is most likely due to a drop below the detection limit. During the adaptation of the RKI-strain the height of peak 16 decreased by a factor of two in passage 2 (figure 23A, figure 24A), which was in agreement with the steady decrease of peak 16 during time course experiments (also see section 1.1.1; figure 12). Besides, peaks 17 and 18 were missing in passage 2 of the RKI-strain. In the time series in Vero cells these peaks represented low abundant structures with only 0.2 – 2.7 % RPH (also see section 3.6.2; figure 12). This probably indicates again a drop below the detection limit [4].

In table 3 differences of RPH in controls (different virus passages, section 4.4 and different harvest time points, section 3.6.2) and during adaptation experiments are compared for each peak of both host-specific glycosylation patterns. The standard deviations (SD) are given in the supplementary (table S 3, page 6). A more than 3-fold higher SD compared to the controls was considered significant. For the RKI-strain, during adaptation to Vero cells three peaks (11, 14 and 18) showed a more than 3-fold higher SD of RPH (table S 3; Adaptation series H1N1, RKI) compared to the same peaks of the time series experiment (also see section 1.1.1; figure 12). Closer examination of the 15 main peaks of all

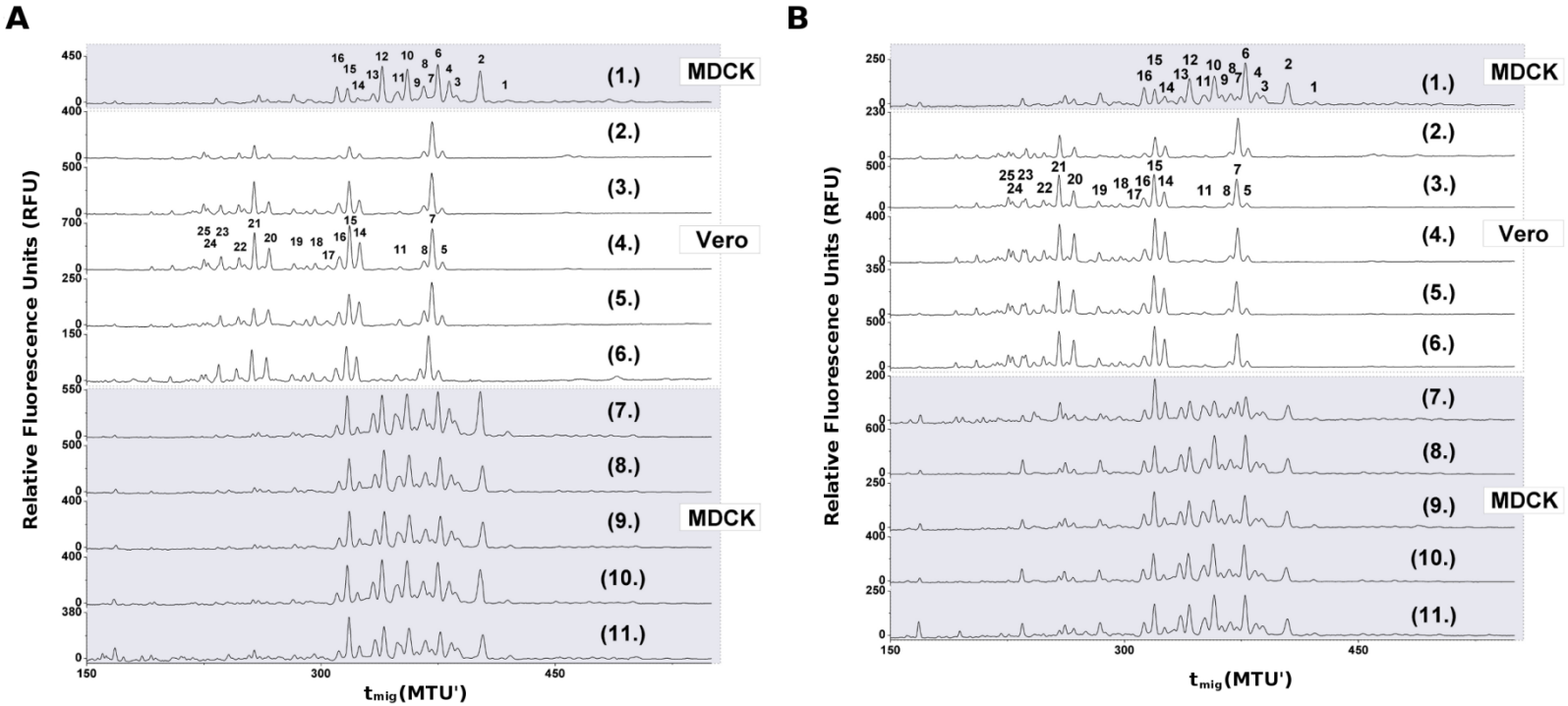


figure 23: HA *N*-glycosylation patterns during IVA-PR8 adaptation from MDCK to Vero cells and back to MDCK cells.

Shifted overlay of HA *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). Fingerprints represent 11 virus passages during virus adaptation of the RKI- (A) and the NIBSC-strain (B): (1.) MDCK cell-adapted virus seed, (2.) to (6.) adaptation to replication in Vero cells, (7.) to (11.) back-adaptation to replication in MDCK cells. Modified and reprinted with permission [4].

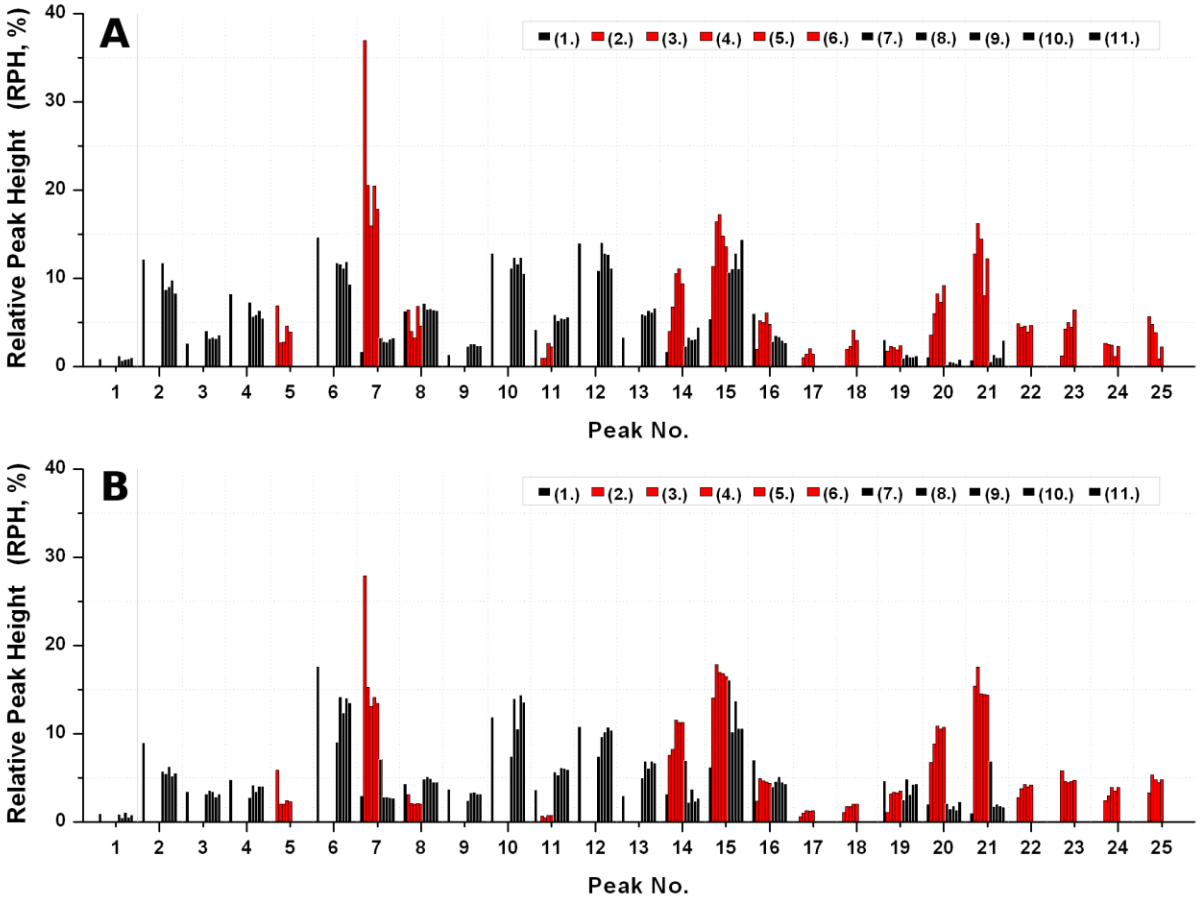


figure 24: Relative peak height of IVA-PR8 HA *N*-glycosylation patterns during virus adaptation.
(A) RKI-strain: In passages (1.) and (7.) to (11.) virus was propagated in MDCK cell culture (■). In passages (2.) to (6.) virus was propagated in Vero cell culture (■). Most of the 25 different major peaks are host cell-specific. Only the HA *N*-glycan structures represented by peak no. 7, 8, 11, 14 to 16 and 19 to 21 are present in virus samples from both host cells. (B) NIBSC-strain: In passage (1.) and (7.) to (11.) virus was propagated in MDCK cell culture (■). In passage (2.) to (6.) virus was propagated in Vero cell culture (■). Most of the 25 different major peaks are host cell-specific. The HA *N*-glycan structures represented by peak no. 7, 8, 11, 14 to 16 and 19 to 21 are present in virus samples from both host cells. Modified and reprinted with permission [4].

MDCK cell-derived HA *N*-glycosylation patterns during back-adaptation revealed a more than 3-fold higher SD in RPH for peaks 6 and 14 (table S 3; Adaptation series H1N1, RKI) compared to the time series in MDCK cells (section 3.6.2; figure 12) and biological reproducibility experiments (section 4.4, figure 19). Of these peaks, only number 6, with an average RPH of 12.1 %, represents a high abundant glycan structure. For the NIBSC-strain, during forward adaptation, a 3-fold higher SD was found for the high abundant glycan structure represented by peak 14, as well as for another low abundant glycan structure represented by peak 19. Regarding the back-adaptation to MDCK cells the low abundant glycan structures of peaks 7 and 14 as well as the high abundant glycan structures of peak 6, 10 and 13 exhibited a more than 3-fold higher SD compared to controls.
Overall, however, a clear trend during the forward and backward adaptation process was not evident (figure 23; figure 24). Furthermore, good reproducibility of the host cell-specific HA *N*-glycosylation pattern during virus adaptation to different host cell lines was demonstrated for both IVA-PR8 strains [4].
Data from adapting MDCK cell-derived RKI-strain to replication in AGE1.CR.pIX cells are consistent with the previous data demonstrating host cell-specificity. The HA *N*-glycan fingerprint of the MDCK cell-derived seed virus (figure 25A) was dominated by 14 peaks exhibiting at least 20% of the RFU of the highest peak. These peaks are numbered 1–3, 5–8, 10–12, 14, 15, 18, and 20. With the first passage in AGE1.CR.pIX cells (passage 1, figure 25A), the HA *N*-glycan fingerprint changed significantly. Now, the fingerprint was dominated by the 14 peaks (no. 4, 6, 9, 13–17, 19, and 21–25). Peaks 6, 13–15, 17–19, and 23 were present in both cell line-derived HA *N*-glycan fingerprint, whereas peaks 1–3, 5, 7, 8, 10–12, and 20 were MDCK cell-specific, and peaks 4, 9, 16, 21, 22, 24, and 25 were AGE1.CR.pIX cell-specific. During all four consecutive virus passages of the adaptation process the HA *N*-glycan fingerprint was more or less stable. However, peaks 4, 9, 16–18, and 24 dropped below the 20 % threshold of the highest peak in the fingerprints of passages 2 to 4. Peak 13 dropped below this threshold only in passages 2 and 3 and peak 25 dropped below this threshold only in passage 2 (figure 25). For relative quantification, RPH was calculated and

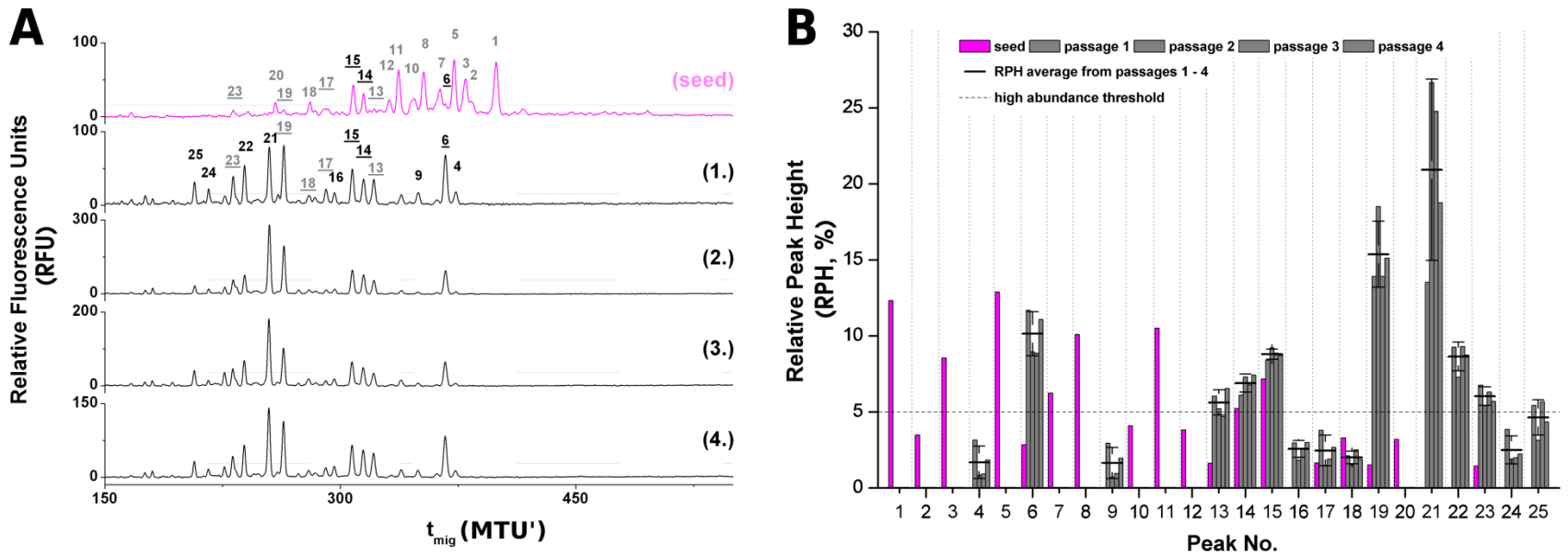


figure 25: Adaptation of RKI-strain from MDCK (pink) to AGE1.CR.pIX cells (dark grey).

(A) Shifted overlay of HA *N*-glycosylation fingerprints, in which relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). Threshold (- - -) indicates peaks exceeding 20 % of the highest peak. MDCK cell-derived virus seed was adapted over four passages (1.) to (4.) to replicate in AGE1.CR.pIX cells. Peaks annotated in grey are only present in MDCK cell-derived HA, whereas peaks annotated in black are only present in AGE1.CR.pIX-derived HA. Annotated peaks in black underlined are present in both cell lines above the 20 % threshold, whereas peaks annotated in grey underlined are present in both cell lines, though in one only below the 20 °% threshold. (B) Relative peak abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined low abundant if RPH < 5 % (- - -). The average RPH from all adaptation samples (1.) to (4.) is indicated by black bars (—), and the standard deviations of the RPH of each peak of passages (1.) to (4.) are indicated by black error bars. Modified and reprinted with permission [7].

plotted over the peak number (figure 25B). This highlights the low abundance of peaks 2, 6, 10, 12, 13, 17, 18, 19, 20, and 23 as well as the high abundance of peaks 1, 3, 5, 7, 8, 11, 14, and 15 in the MDCK cell-specific HA *N*-glycan fingerprints. In contrast, in the AGE1.CR.pIX cell-specific fingerprints, peaks 4, 9, 16–18, and 24 were low abundant in all analyzed passages, whereas peak 13 dropped to low abundance only in passage 3 and peak 25 dropped to low abundance only in passages 2 and 4. During adaptation, all peaks, except peaks 19 and 21, showed differences in relative *N*-glycan structure abundances of ≤ 2.8 %. Peaks 19 and 21 varied stronger with maximum differences for RPH of 4.6 % and 13.1 %, respectively. The RPH of peak 21 almost doubled from passage 1 to passage 2 and decreased again in passages 3 and 4. This suggests that the structure represented by peak 21 is not as stable as other structures or more susceptible to minor variations in the culture and infection process. Either its synthesis or its degradation seems to vary for all consecutive AGE1.CR.pIX cell-derived virus passages. For all peaks no clear trend during the adaptation process was detected [7].

In line with so far observed host cell-specificity during virus adaptations, the data from adapting MDCK cell-derived RKI-strain to replication in embryonated chicken eggs demonstrated strict host cell-specificity, too (figure 26). The fingerprints of the viral HA changed with the switch of host cell completely (figure 26A, B). Thereafter, the fingerprint remained more or less robust during the following three passages in embryonated chicken eggs: relative peak abundances differed with maximal 8.2 % (peak 11, figure 26C).

Furthermore, the data from the adaptation of MDCK cell-derived IVA-Uruguay to Vero cells were consistent with the previous results demonstrating that first and foremost the host cell determines the HA *N*-glycan fingerprint. However, also the choice of the IVA strain has an impact on the HA *N*-glycan fingerprint (as shown in section 4.3). In contrast to the 16 MDCK cell-specific peaks of IVA-PR8-derived HA, the fingerprint of MDCK cell-derived IVA-Uruguay seed exhibited 25 different characteristic peaks in the range of 160 MTU' to 400 MTU'. Of these, a total number of 11 peaks representing large glycans (275 MTU' - 400 MTU'; 12,

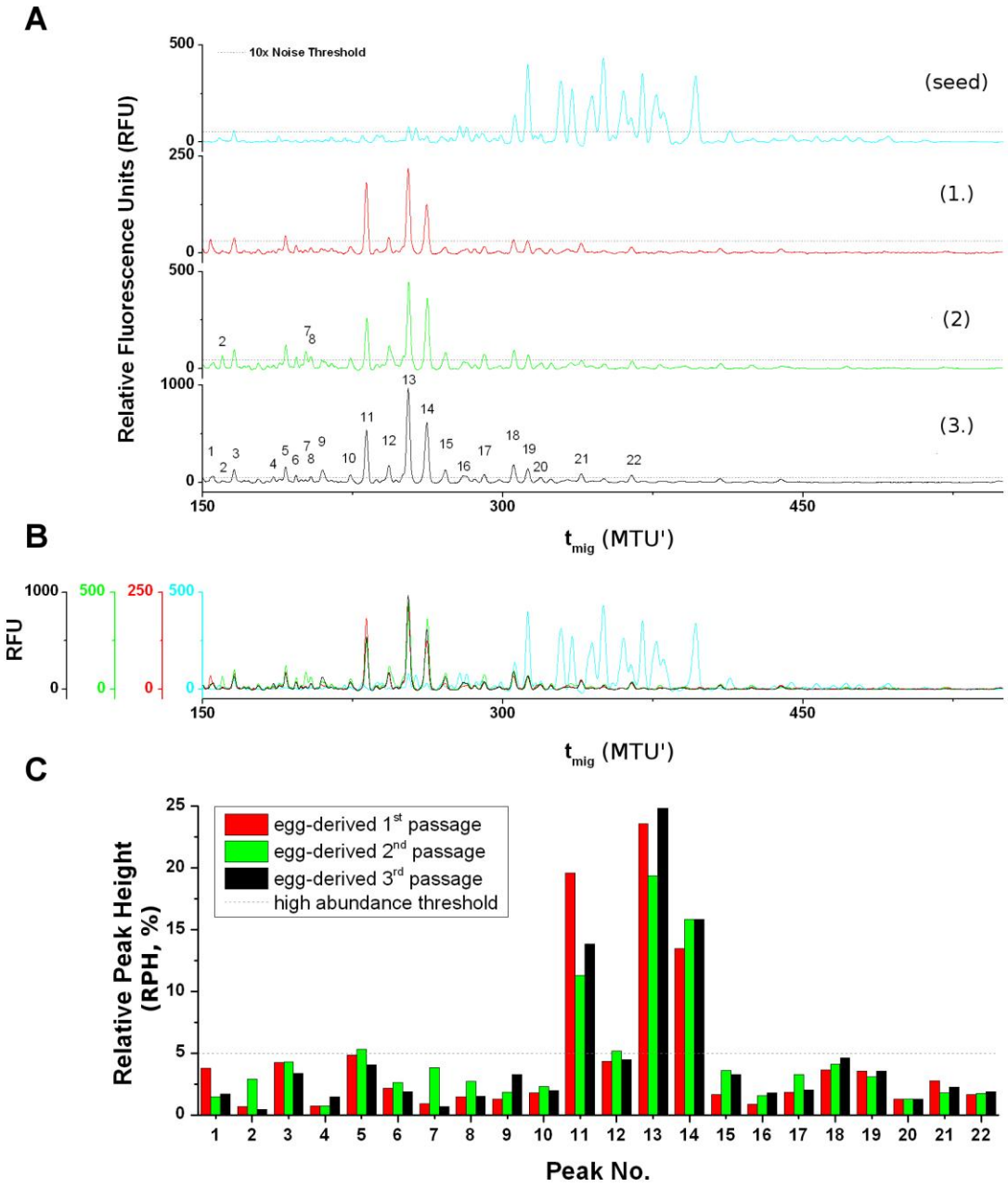


figure 26: Adaptation of the RKI-strain from MDCK cell- to embryonated chicken egg-based replication.

(A) Shifted overlay of HA *N*-glycan fingerprints, relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). Egg-specific peaks between 150 MTU' and 450 MTU' exceeding the 10x baseline noise threshold (---) in at least one of the fingerprints are annotated. MDCK cell-derived seed virus (seed) was adapted over three passages (1.) to (3.) to replicate in embryonated chicken eggs. (B) Direct overlay of all fingerprints. (C) Relative peak abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined low abundant, if RPH < 5 % (---).

13, 18 - 22, 24, 26 - 28) were unique to MDCK cell-derived virus (figure 27A, seed). The HA *N*-glycosylation pattern changed significantly with the first passage in Vero cells. Here, 15 different peaks between 150 MTU' and 380 MTU' characterized the Vero cell-specific HA *N*-glycan fingerprint. Four peaks (no. 3, 5, 23 and 25) were unique to Vero cell-derived virus (figure 27A, first passage). In comparison to MDCK cell-derived HA, the Vero cell-derived antigen showed a tendency towards smaller glycan structures. This is in agreement with the tendency towards smaller glycan structures attached to HA of Vero cell-derived viruses from the RKI- and the NIBSC-strain. The relative abundance of each peak over all Vero passages only varied marginally with standard deviations (SD) ≤ 2.1 % and maximal differences in RPH ≤ 5.1 % ($|\Delta RPH_{max}|$, peak 10, table 3 and figure 27B, [5]).

Consistent with all previous results, also for IVA-Viktoria as well as for IVA-California host cell-specificity of the HA *N*-glycosylation during virus adaptation was demonstrated: with the first passage in MDCK cells, the HA *N*-glycosylation pattern changed completely (figure 28A - D). During further adaptation the fingerprint stayed more or less robust. Differences of relative peak abundance were ≤ 2.5 % ($|\Delta RPH|_{max}$ for peak 27, figure 28E) and ≤ 3.1 % ($|\Delta RPH|_{max}$ for peak 7, figure 28F), indicating no significant further impact of the adaptation process on the HA *N*-glycan fingerprint of IVA-Viktoria and IVA-California, respectively.

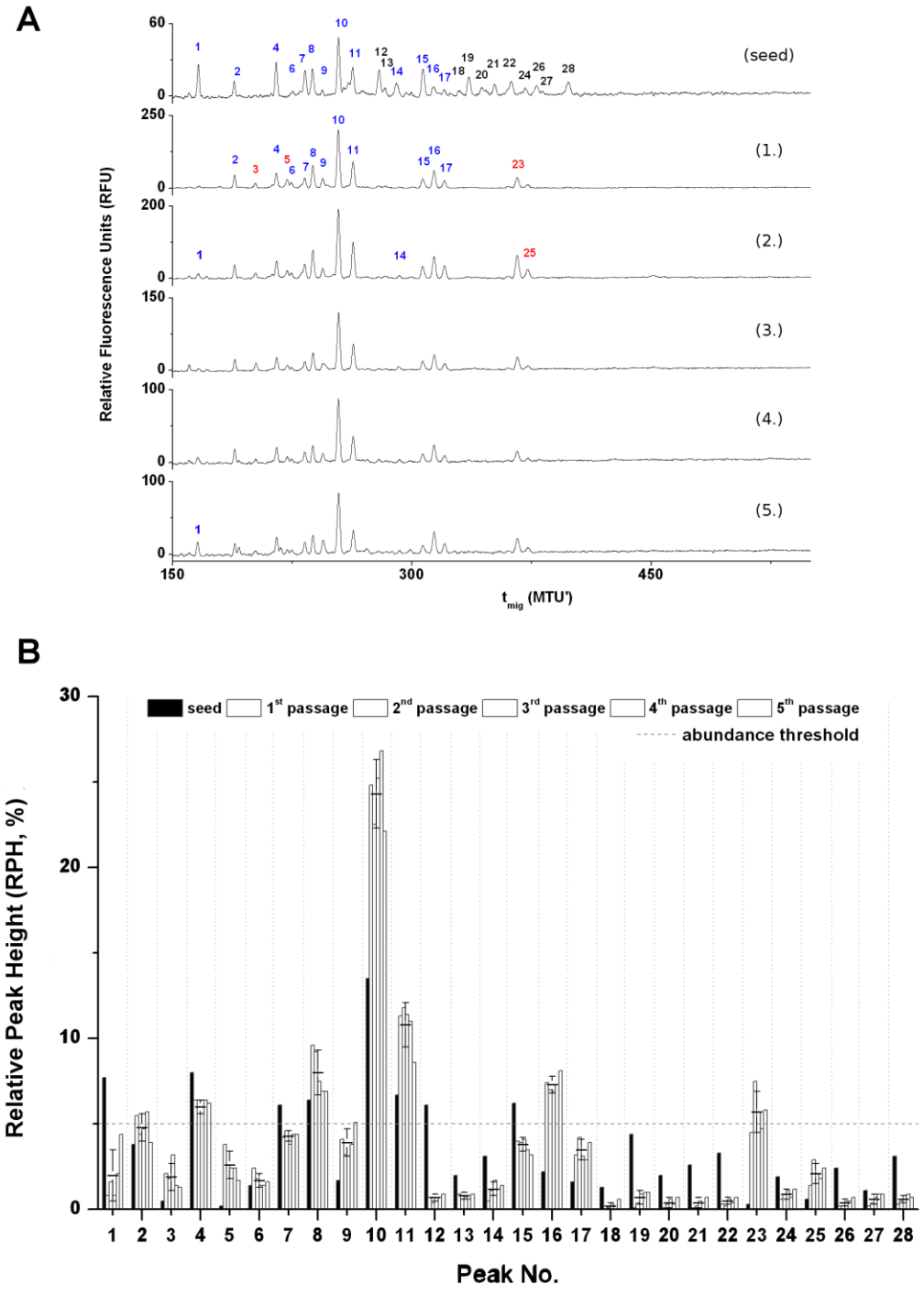


figure 27: Adaptation of IVA-Uruguay from MDCK to Vero cells.
(A) HA *N*-glycan fingerprints. Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). (A) MDCK cell-derived virus (seed) was consecutively passaged in Vero cells over five passages (1.) to (5.). Peaks annotated in blue are present in MDCK as well as Vero cell-derived HA; peaks annotated in black are MDCK cell-specific, and red annotation indicates Vero cell-specific peaks. Modified and reprinted with permission [5]. (B) Relative peak height (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined low abundant if RPH < 5 % (- - -, [5]).

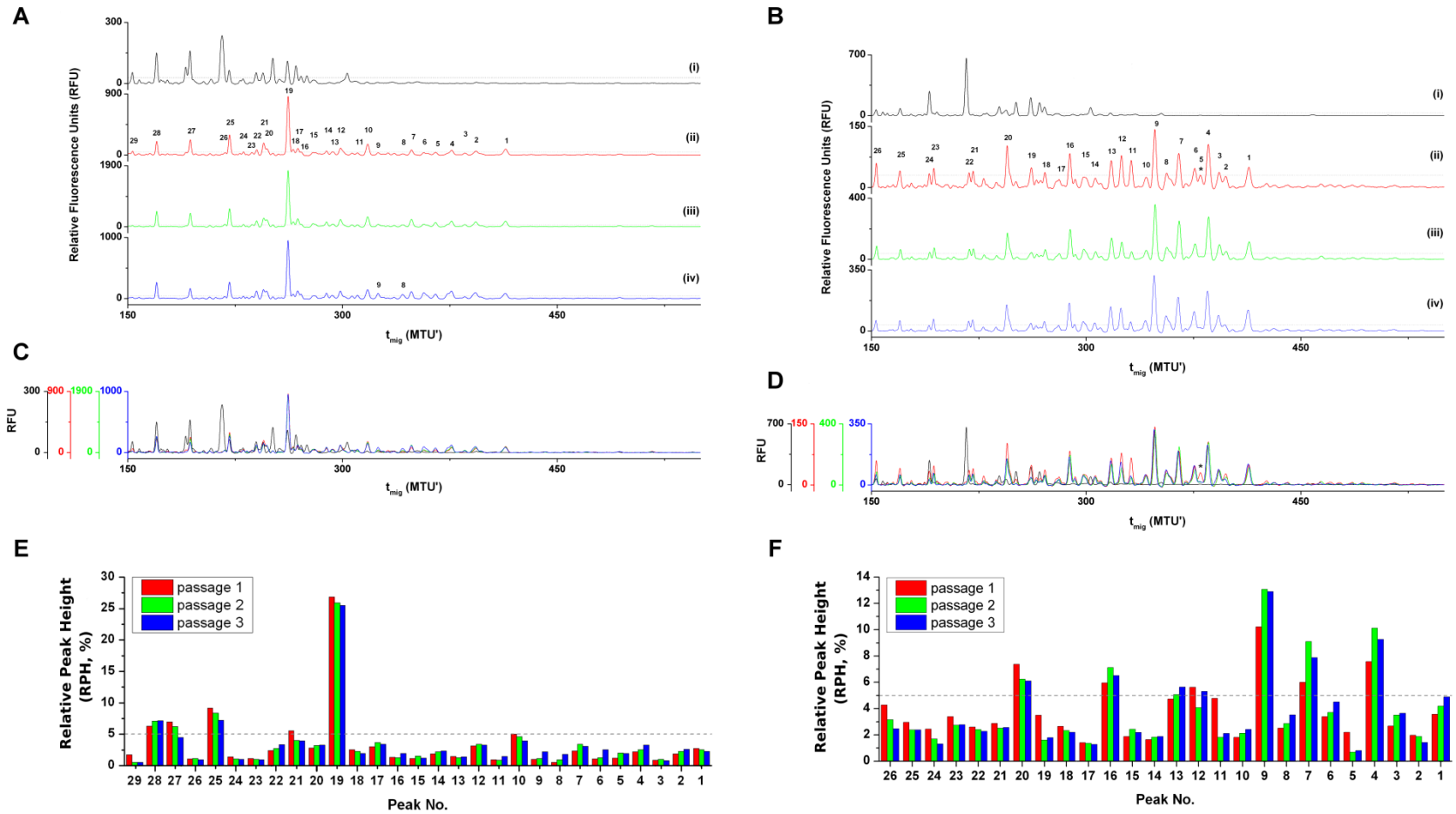


figure 28: Adaptation of IVA-Viktoria (A, C, E) and IVA-California (B, D, F) from embryonated chicken eggs to MDCK cells.
(A, B) Shifted overlay of HA *N*-glycan fingerprints, relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). MDCK cell-specific peaks between 150 MTU' and 450 MTU' exceeding the 10x baseline noise threshold in at least one of the fingerprints compared are annotated. Egg-derived virus (i) was adapted in three subsequent passages (ii) to (iv) to replicate in MDCK cells. (C, D) Direct overlay of all four fingerprints. (E, F) Relative peak abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks) during the adaptation process. Peaks are defined low abundant if RPH < 5 % (---).

4.6.3 Changes in quasispecies composition during virus adaptation from MDCK to Vero cells and back

An impact of HA *N*-glycosylation patterns on virus' properties, such as virus replication, has been described before [104, 161-164]. Interestingly, during all analyzed adaptations the *N*-glycan fingerprint stabilized soon after the first passage in the new host. This clearly suggests that further increase in HA-titer, e.g. as observed for both IVA-PR8 strains during adaptation from MDCK to Vero cells, was not driven by further changes in the HA *N*-glycosylation pattern. Other factors must contribute to higher virus yields. Therefore, changes in the viral genome during virus adaptation were investigated in more detail. Each viral gene contributes to the virus' interaction with the host and to resulting infection. Understanding the adaptation of each single viral gene to the new host may hence provide new insights in the mechanisms of virus adaptation. For this purpose, samples of non-adapted (passages 1, section 4.6.1, paragraph 1), Vero cell-adapted (passages 6 as well as seed virus for later immunogenicity studies [72]) and back-adapted (passages 11) IVA-PR8, either the RKI- or the NIBSC-strain, were sequenced by massive parallel pyrosequencing.

All consensus cDNA sequences from passages 1 of the RKI-strain and the NIBSC-strain as well as translated and aligned consensus amino acid sequences from passages 1 are compiled in the supplementary (sections 12.7 and 12.8). The composition of the quasispecies for the RKI- and the NIBSC-strain during virus adaptation for all virus segments is presented in table 4 and table 5, respectively.

table 4: Overview of changes in quasispecies composition of the RKI-strain during virus adaptation to Vero cells.

Segment	Coded Protein	cDNA / AA Different from Seed Virus Consensus Sequence (Consensus Position Substitution)						Ratio of cDNA / AA Differing from Seed Virus Consensus Sequence (%)			
		Base-Substitution			AA-Substitution			Adaptation IVA-PR/8 (RKI)			Adaptation IVA-PR/8 (RKI)
								Passage 1	Passage 6	Passage11	Genzel *et al.* 2010 (Passage 5)
1	PB2	-	333	A	frameshift			0	0	0	12
		T	343	C	Y	115	H	0	13	0	0
		-	357	A	frameshift			0	0	0	11
		C	588	G	C	196	W	0	7§	**50**	0
		ACA	670-672	C	frameshift			0	11	0	0
		T	1212	C	F	404	F	0	27	0	10§
		T	1227	C	C	409	C	0	6§	1§	12
		T	1338	-	frameshift			0	16§	0	21
		G	1351	A	V	451	I	0	20	0	0
		C	1362	T	I	454	I	0	15	0	0
2	PB1	ACAAAGA	524-530	CAAG	NKE	175-177	TR	0	16	0	0
		ACA	1037-1039	C	frameshift			0	47	0	0
		-	1043	A	frameshift			0	0	0	15
		A	1286	-	frameshift			0	**87**	0	0
		-	1445	A	frameshift			10	0	0	0
		T	2000	CTA	frameshift			12	1§	15	2§
		TAT	2084-2086	A	frameshift			0	36	0	0
3	PA	T	150	C	D	50	D	0	4§	35	0
		G	301	T	E	101	stop	0	0	0	18
		-	303	C	frameshift			0	0	0	17
		G	585	A	E	195	E	0	0	38	0
		G	954	A	K	318	K	0	0	0	23
		G	1053	A	E	351	E	0	10	0	0
		A	1087	-	frameshift			0	54§	0	28
		ATG	1728	T	frameshift			0	40	0	0
4	HA	CAT	1011-1113	-	I	338	-	0	2§	1§	46
		A	1258	-	frameshift			0	0	0	37
		A	1268	-	frameshift			0	0	0	18
		-	1284	A	frameshift			0	0	0	19
		C	1370	T	S	457	L	0	19*	9*	0
		A	1378	G	K	460	E	0	**80***	**81***	0
		-	1420	A	frameshift			0	0	0	10
		initial seed virus			no AA-substitutions			**100**	few reads#	10	n.d.
5	NP	-	34	G	frameshift			**0**	0	0	15
		AA	354-355	-	frameshift			0	10	0	0
		G	359	A	W	120	stop	0	10	0	0
		G	417	A	W	139	stop	11	0	0	0
		A	859	C	S	287	R	0	17	54	0
		G	882	T	E	294	D	0	47	9§	36
		A	926	G	N	309	S	0	20	0	0
		T	1260	-	frameshift			0	0	0	12
		G	1324	A	A	442	T	0	12	0	0
		G	1414	C	A	472	P	0	42	0	12
		G	1418	A	S	473	N	0	38	0	9§
		last bp missing			last AA missing			4 AA	4 AA	3 AA	0
6	NA	A	21	G	I	7	M	0	15	**85**	0
		A	607	-	frameshift			20§	24§	17§	19
		-	610	A	frameshift			0	0	0	11
		A	622	-	frameshift			12	16§	5§	17
		A	1258	-	frameshift			0	0	0	53
7	M1	-	730	G	frameshift			0	0	0	15
	M2	last bp missing			last AA missing			1 AA	2 AA	1 AA	0
8	NS1	T	307	C	S	103	P	0	**100+**	**100**	0
	NEP	C	27	T	F	9	F	0	0	0	**100+**

Passage 1 represents the MDCK cell-adapted seed virus. Passage 6 represents the last of five consecutive virus passages in Vero cells. Passage 11 represents the last of five subsequent passages in MDCK cells, i.e. after final back-adaptation. S457L and K460E in segment 4 are uncoupled in passages 6 and 11.
Genzel *et al.* 2010 (Passage 5, [72]) represents the Vero cell-adapted virus seed used for the immunogenicity studies. 0% values correspond to below the detection limit. Table content for segment 4 is reprinted with permission [4].
in Passage 6 only few reads of the original sequence of segment 4 was detected;
+ a few reads with the original sequence were detected (difference was detected with a frequency of ≥ 1 %; no HCD quality was reached);
§ detected frequency of difference was ≥ 1 % and did not reach high confidence quality.

table 5: Overview of changes in quasispecies composition of the NIBSC-strain during virus adaptation to Vero cells.

Segment	Coded Protein	cDNA / AA Different from Seed Virus Consensus Sequence (Consensus Position Substitution)						Ratio of cDNA / AA Differing from Seed Virus Consensus Sequence (%)		
		Base-Substitution			AA-Substitution			Adaptation IVA-PR/8 (NIBSC)		
								Passage 1	Passage 6	Passage11
1	PB2	T	1287	C	N	429	N	0	8§	15
2	PB1	G	627	A	K	209	K	13	1§	1§
		G	747	A	R	249	R	10	1§	1§
		-	1479	T	frameshift			0	0	13
		T	2000	CTA	frameshift			8§	12	9§
		last bp missing			last AA missing			0	5 AA	5 AA
3	PA	A	399	G	E	133	E	0	5§	11
		C	1365	A	A	455	A	0	6§	13
		C	1374	T	Y	458	Y	2§	6§	12
		T	1687	C	L	563	L	6§	28	12
4	HA	T	70	C	Y	24	H	22	0	0
		G	1183	A	V	395	M	0	41.5	11.3
		A	1189	G	T	397	A	1.3	0	0
		A	1189	T	T	397	S	0.6	0	5.4
		G	1363	T	D	455	Y	21.4	6.1	3.2
		G	1363	C	D	455	H	0	**52**	44.1
		A	1375	G	K	459	E	0	0	44.2
		A	1378	G	N	460	D	12.2	41.1	10.5
5	NP	first bp missing			first AA missing			1 AA	0	0
		A	59	C	Q	20	P	12	0	0
		G	684	A	G	228	G	0	**80**	**64**
		A	859	C	S	287	R	0	**86**	**77**
		A	881	T	E	294	V	0	11	3§
		T	1191	A	N	397	K	0	0	15
		G	1323	A	R	441	R	0	0	12
		last bp missing			last AA missing			4 AA	2 AA	4 AA
6	NA	T	20	C	I	7	T	0	**85**	**65**
		A	904	G	I	302	V	0	0	12
		G	940	A	E	314	K	0	0	12
		-	999	T	frameshift			0	20	0
		A	1032	-	frameshift			0	11	0
		G	1038	A	R	346	R	5§	20	1§
		G	1046	A	S	349	N	0	5§	26
		T	1269	C	T	423	T	0	0	10
		A	1300	G	N	434	D	0	0	11
		G	1352	C	S	451	T	0	0	10

Passage 1 represents the MDCK cell-adapted seed virus. Passage 6 represents the last of five consecutive virus passages in Vero cells. Passage 11 represents the last of five subsequent passages in MDCK cells, i.e. after final back-adaptation. 0% values correspond to below the detection limit. In Passage 1, segment 4 D455Y and N460D are uncoupled. Incomplete coding sequences are indicated. Table content for segment 4 is reprinted with permission [4].

* if coupled GTC->GCA: V -> A; # if coupled CGT->CAC: R -> H; °substitutions are coupled;

“ if coupled: GTT->ACT: V -> T; § detected frequency of difference was ≥ 1 % and did not reach high confidence quality (not defined as HCD).

table 5 continued

Segment	Coded Protein	cDNA / AA Different from Seed Virus Consensus Sequence (Consensus Position Substitution)						Ratio of cDNA / AA Differing from Seed Virus Consensus Sequence (%)		
		Base-Substitution			AA-Substitution			Adaptation IVA-PR/8 (NIBSC)		
								Passage 1	Passage 6	Passage11
7	M1	T	33	A	V	11	V	15	30	25
		G	75	A	A	25	A	13	31	30
		A	88	G	N	30	D	13	28	27
		A	105	G	K	35	K	14	28	28
		AG	215-216	GA	Q	72	R	14	0	0
		T	255	C	N	85	N	13	6§	0
		G	280	A	D	94	N	7§	36	45
		G	289	A	V	97	I	1§	13	6§
		T	309	C	L	103	L	10	23	21
		T	328	C	Y	110	H	10	24	23
		T	336	C	A	112	A	11	24	22
		A	345	C	I	115	I	11°	23	25
		G	346	T	A	116	S		25	27
		C	360	T	S	120	S	12	27	27
		T	376	A	C	126	S	11	27	33
		A	409	G	T	137	A	12	29	34
		T	417	C	T	139	T	12	27	34
		G	418	A	A	140	T	12	28	34
		T	428*	C	V	143	A	12°	28	30
		G	429*	A	V	143	V		28	29
		G	499	A	A	167	T	15	26	24
		G	618	T	A	206	A	16	33	30
		T	620	G	I	207	S	16	33	29
		G	623	A	R	208	Q	16	34	30
		G	630	A	R	210	R	15	34	29
		G	642	A	Q	214	Q	15	35	32
		A	645	G	A	215	A	15	35	32
		G	691	A	D	231	N	10	27	21
		G	767#	A	R	256	H	13	36	22
		T	768#	C	R	256	R	13	34	22
		T	776	G	L	259	R	13	33	21
		G	780	A	K	260	K	12	28	21
		T	804	C	D	268	D	12	30	20
		T	849	G	S	283	S	12	28	20
		TT	862-863	AC	F	288	T	12	28	21
		A	869	G	K	290	R	12	28	21
		A	918	G	S	306	S	13	26	21
	M2	G	79"	A	V	27	I	13	36	22
		T	80"	C	V	27	A	13	34	22
		T	88	G	S	30	A	13	33	21
		G	92	A	S	31	N	12	28	21
		T	116	C	I	39	T	12	30	20
		T	161	G	L	54	R	12	28	20
		TT	174-175	AC	GL	58-59	GL	12	28	21
		A	181	G	R	61	G	12	28	21
		A	230	G	Q	77	R	13	26	21
		last bp missing			last AA missing			0	1 AA	1 AA
8	NS1	C	240	A	T	80	T	0	0	13
		G	301	A	D	101	N	0	0	15
		G	306	A	W	102	stop	0	49	0
		C	320	A	P	107	H	1§	13	2§
		G	547	A	G	183	R	0	7§	25
		G	551	A	G	184	E	0	4§	24
		G	565	A	D	189	N	0	5§	23
		C	778	T	L	260	F	1§	5§	29
	NEP (NS2)	G	75	A	S	25	S	0	0	26
		G	79	A	D	27	N	0	0	21
		G	93	A	M	31	I	0	0	24
		C	306	T	A	102	A	0	0	29

4.6.3.1 Segment 4 coding for HA

At first, particular focus was laid on the sequence coding for HA (segment 4) and on possible changes in potential HA *N*-glycosylation sites: Regarding the consensus amino acid sequences of HA, the initial virus seeds from RKI and NIBSC differed in 7 amino acid positions. The differences comprise K147-, A156E, E158K, I208L, R269M, F309Y and S398T (RKI versus NIBSC, figure S 9 in the supplementary). For the RKI-strain, sequencing analysis clearly indicated that the virus population from the first initial MDCK passage 1 was uniform concerning RNA segment 4, i.e. only one virus variant was detected above the detection limit. In contrast, segment 4 of the initial virus seed from NIBSC (passage 1) already comprised several virus variants. The following variants different from the consensus were detected with the specified frequencies, (table 5): Y24H (22 %), T397A (1.3 %), T397S (0.6 %), D455Y (21.4 %) and N460D (12.2 %, [4]).

After complete adaptation to Vero cells (passage 6), 80% of the RKI-strain virus population carried amino acid substitution K460E, where a positively charged lysine was replaced by the negatively charged glutamic acid. Another population of 19 % carried the S457L substitution, where polar serine was replaced by non-polar leucine, (table 4). These two substitutions were uncoupled. No single read was detected, which carried both substitutions. Hence, 99 % of sequenced viruses carried either one of these two substitutions, indicating a crucial region of the HA for adaptation to efficient virus growth in Vero cells. Only single reads were detected that carried the original sequence of the first passage. This suggests that on the one hand, the initial virus did replicate in Vero cells, but only poorly. Considering the dilutions of the initial virus seed during the five Vero passages without any replication, its concentration would have been below 0.1 virion/mL in passage 6 and would most likely not have been detected at all. On the other hand, only a change in very few single amino acids in this HA-region is necessary to increase virus fitness for sufficient growth in Vero cells. In passage 6 of the adaptation of the NIBSC-strain a new variant carrying a V395M

substitution was detected with a frequency of 41.5 %, which introduced an additional sulfur containing residue into the HA_2 peptide chain. Furthermore, the D455Y virus seed subpopulation dropped to 6.1 %, but a new variant D455H, in which aspartic acid was substituted by the basic amino acid histidine, was detected that dominated the virus population with 52 %. In addition, the N460D subpopulation increased to 41.1 % (table 5). Sequencing analysis of the Vero-adapted virus seed [72] for later immunogenicity studies revealed a different set of subpopulations within the HA-coding sequence: Except from a few frameshift-causing and silent mutations, one deletion (I338-: 46 %) as well as one substitution (V459M: 11 %) was detected [4].

After back-adaptation of the RKI-strain to MDCK cells the dominating virus population (81%) still carried the K460E substitution in the HA. The minor subpopulation (19 % after Vero-adaptation) carrying the S457L substitution decreased to 9 % after back-adaptation. The remaining population of 10 % represented the initial virus seed sequence from passage 1, (table 4). This demonstrates that the K460E variant allows good virus replication in both cell lines. In contrast, the S457L variant seems to be less efficiently replicating in MDCK cells. These results, particularly the fitness of the K460E variant in MDCK as well as in Vero cells, strongly suggests that a mutation in this region was acquired, rather than an already existing virus subpopulation of the initial virus seed was selected. This mutation finally allowed sufficient virus replication in passage 2 of the adaptation series to reach an HA-titer of 1.4 HAU at 288 hpi. After back-adaptation of the NIBSC-strain to MDCK cells the V395M, the D455Y, the D455H, and the N460D variants decreased to 11.3 %, 3.2 %, 44.1 %, and 10.5 %, respectively. In contrast, the T397S variant that was not detected in passage 6, came-up again and made up for 5.4 % in passage 11, (table 5). This T397S substitution abolished the only difference of the consensus sequences in the HA_2 chain from the two IVA-PR8 strains (RKI- and NIBSC-strain, figure S 9 in the supplementary). Furthermore a new variant carrying the K459E substitution was detected with a frequency of 44.2 % [4].

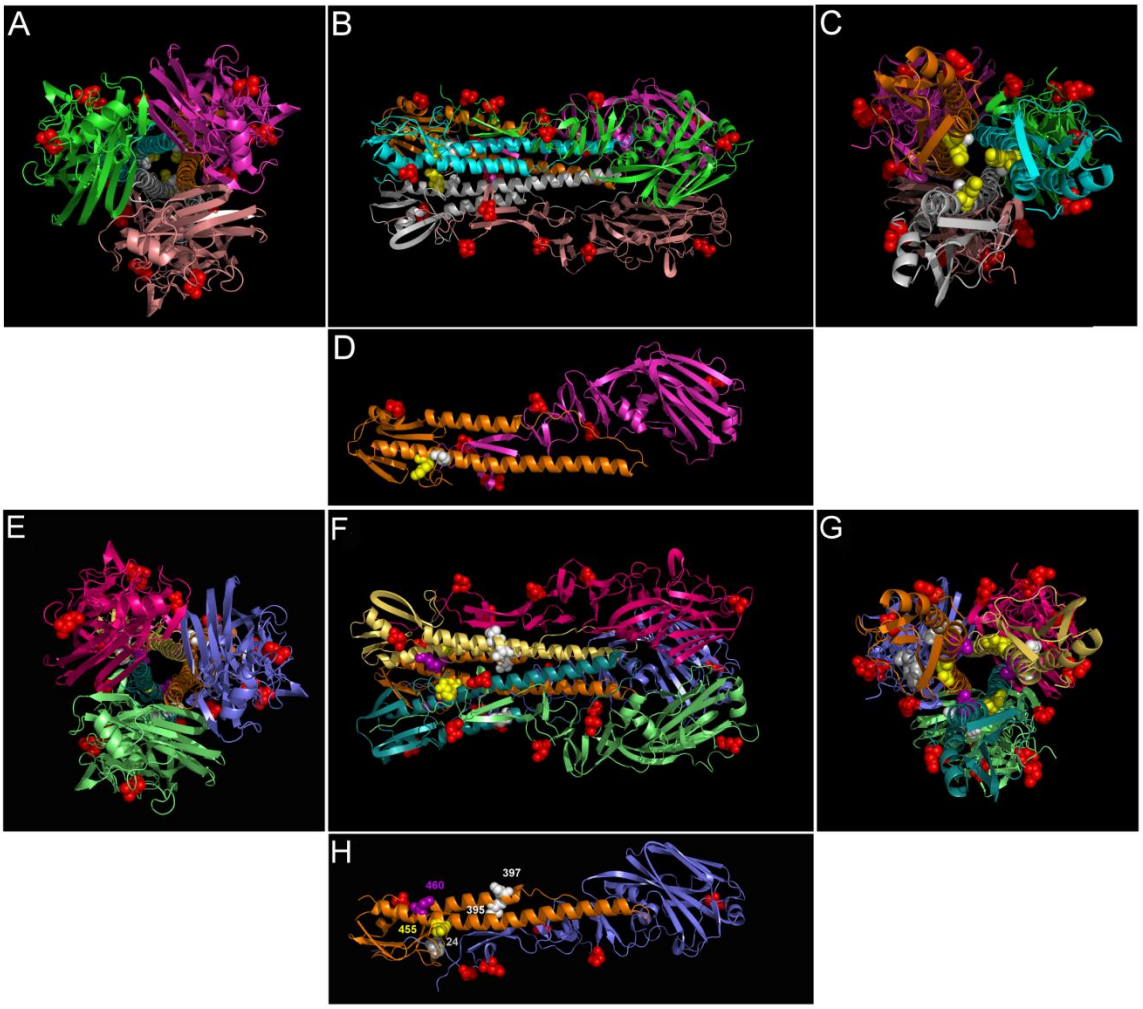


figure 29: Localization of substitutions during adaptation from MDCK to Vero and back to MDCK cells within the 3D HA-structure.
Structures are displayed in a cartoon diagram with potential HA *N*-glycosylation sites highlighted by red space filled residues. (A, B, C, D) RKI-strain, (E, F, G, H) NIBSC-strain . Trimeric (A-C, E-G) and monomeric (D, H) HA molecules. (A-D) The HA_1 chains are colored in pink, green and brown; the HA_2 chains are colored in blue, grey and orange. The K460E mutation is highlighted in yellow, the S457L substitution by white space-filled residues. (A) Bottom (B) side and (C) top view. (D) indicates the close proximity within the monomer of these two substitutions, which are one helix turn apart from each other. (E-H) For the isolate from NIBSC the HA_1 chains are colored in pink, purple and green; the HA_2 chains are colored in turquoise, yellow and orange. The substitutions already present in the virus seed are highlighted by grey (Y24H), yellow (D455Y) and pink (N460D) space-filled residues. Substitutions occurring during virus adaptation are indicated by white (V395M, T397S) or yellow (D455H) space-filled residues. (E) Top, (F, H) side, (G) bottom view. The PDB entry 1RU7 and PyMOL (v0.99, DeLano Scientific LLC, California, USA) software was used for structure display. Reprinted with permission [4].

All substitutions detected during forward and backward adaptation of both IVA-PR8 strains are located in the HA_2 chain, neither inside nor in close proximity of any *N*-glycosylation site. They are, however, located in the inside of the HA

trimer within or in close proximity to the fusion peptide pocket: within the subunits' contact site for the RKI-strain (figure 29A-D) and within the subunits' and monomer contact sites for the NIBSC-strain (figure 29E-H). This position belongs to the fusion subdomain [165]. Interestingly, the Vero-adapted virus seed [72] used for immunogenicity studies revealed beside a few frameshift-causing and silent mutations, one deletion (I338-) in the HA_1 just before the HA_1-HA_2 cleavage site at residues 344-345 as well as one substitution in the HA_2 (V459M), just between the two mutations observed before for the RKI strain (S457L and K460E), located in the inside of the HA trimer within or in close proximity to the fusion peptide pocket (figure 30, [4]).

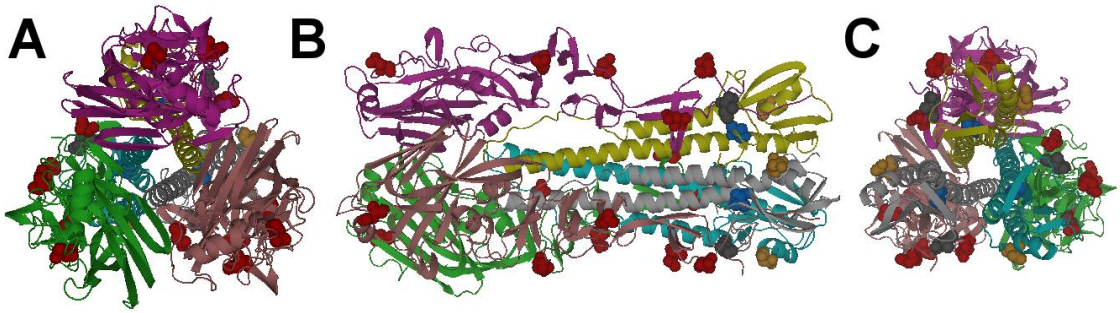


figure 30: Localization of substitution/deletion within the 3D HA-structure in the Vero-adapted seed virus [72] for the immunogenicity studies.
Trimeric HA molecules are displayed in a cartoon diagram with potential HA *N*-glycosylation sites highlighted by red (HA_1) or orange (HA_2) space filled residues. (A-C) The HA_1 chains are colored in pink, green and rosé; the HA_2 chains are colored in yellow, turquoise and grey. The I338-deletion is highlighted in grey, the V459M substitution by blue space-filled residues. (A) top (B) side and (C) top view. The PDB entry 1RU7 and PyMOL (v0.99, DeLano Scientific LLC, California, USA) software was used for structure display.

4.6.3.2 Virus segments 1 - 3, 5 - 8 coding for all other virus proteins

Sequence analysis of segment 4 coding for HA revealed no differences with respect to potential *N*-glycosylation sites, neither during the adaptation process nor between the two tested IVA-PR8 strains RKI and NIBSC. In contrast, the comparison of potential *N*-glycosylation sites in NA (segment 6) between the RKI- and the NIBSC-strain seed virus revealed one potential *N*-glycosylation site more in the RKI-sequence at position 131. In NIBSC, N at position 131 is replaced by S, disrupting the *N*-glycosylation sequon NGT (section 12.8, figure S 11). Altogether, five (positions 44, 58, 73, 131, 220) and four (positions 44, 58, 73, 220) potential *N*-glycosylation sites were predicted for the NA of the RKI- and

the NIBSC-strain seed virus, respectively. During the adaptation process one additional potential *N*-glycosylation site was formed by the S349N substitution detected in passage 11 of the NIBSC-strain (table 5).

Furthermore, the sequence information of all other segments (1 - 3, 5 – 8) confirmed the findings for segment 4 (HA). The RKI-seed virus (passage 1) was uniform for all segments: Only one virus variant was detected except from a few stop- and frameshift substitutions (W139stop [NP, segment 5], frameshift at nucleotides 622 [NA, segment 6], 1445, 2000 [PB1, segment 2], table 4). On the basis of the obtained sequencing data it is difficult to decide whether the detected stop codon and frameshifts are sequencing artifacts or non-infective virus particles. Different sequencing techniques, e.g. Illumina sequencing, are recommended to verify this. In contrast to the uniform RKI-seed, the initial NIBSC-seed virus (passage 1) comprised multiple virus variants: silent substitutions were found in segments 2 (PB1), 3 (PA), 6 (NA) and 8 (NS1, NS2/NEP), although the substitutions in segments 3, 6 and 8 did not reach HCD quality before later time points of the adaptation processes (table 5). Segment 5 (NP) of the NIBSC seed virus comprised one base substitution, whereas segment 7 (M1, M2) comprised 39 base substitutions, resulting in one amino acid substitution for NP and in 22 amino acid substitutions in the M1 and M2 proteins (table 5).

After complete adaptation to Vero cells (passage 6), different subpopulations were detected for the RKI-strain in segments 1 (PB2), 2 (PB1), 3 (PA), 5 (NP) and 6 (NA, table 4). Populations making up for more than 50 % of the quasispecies were detected in segment 2 (PB1, frameshift: 87 %) and in segment 8 (NS1, NS2/NEP; S103P: 100 %). For the NIBSC-strain, the initially inhomogeneous segments 2 (PB1), 5 (NP) and 7 (M1, M2) remained rather inhomogeneous during adaptation to Vero cells. However, the subpopulations of the quasispecies often varied in percentages from the initial ones. Additionally, in passage 6, also segments 3 (PA), 6 (NA) and 8 (NS1, NS2/NEP) revealed various subpopulations. Some newly derived populations in segment 5 (NP;

G228G: 80 %, S287R: 86 %) and segment 6 (NA; I7T:85 %) dominated the virus population by passage 6 with ratios above 50 %.

Interestingly, the substitution S287R in segment 5 (NP) was detected for the RKI- (17 %) as well as the NIBSC-strain (86 %). Furthermore, in segment 5 (NP) of the RKI-strain the substitution E294D (47 %) was detected, whereas 11 % of the NIBSC sequences carried the substitution E294V. The Vero-adapted virus seed [72], which was later used for immunogenicity studies, also comprised a subpopulation of 36 % carrying the E294D substitution in segment 5 (NP). Also in segment 6 (NA) the seventh amino acid was partly substituted in the RKI- (I7M: 15 %) as well as in the NIBSC-strain (I7T: 85 %). The substitution A472P in segment 5 (NP) of the RKI-strain was detected in passage 6 of the forward and backward adaptation [4] as well as in the Vero cell-adapted seed virus [72] for the immunogenicity studies (table 4, table 5).

After back-adaptation of the RKI-strain (passage 11), most subpopulations dropped below the detection limit, some remained (table 4; segment 5, NP, S287R: 54 %; segment 6, NA, I7M: 85 %; segment 8, NS1, NS2/NEP, S103P: 100 %), though with different frequencies. Other substitutions even newly emerged (table 4; segment 1, PB2, C196W: 50 %). A similar dynamic was observed after back-adaptation of the NIBSC-strain (passage 11): some subpopulations dropped below the detection limit, others remained (e.g. segment 5, NP, S287R: 77 %; segment 6, NA, I7T: 65 %) though often with different frequencies and yet others newly emerged. The drop below the detection limit of the E294D- (RKI) and the E294V- (NIBSC) variants suggests less efficient replication of these variants in MDCK cells. In contrast, the S287R variant of the Vero cell-adapted RKI (17 %) as well as NIBSC strain (86 %) persisted well over five passages in MDCK cells (RKI: 46 %; NIBSC: 77 %). In agreement with the significantly delayed HA-titer increase for Vero-inexperienced viruses (figure 22), these results again, strongly suggest that mutations were acquired to allow sufficient replication in the new host rather than already existing virus variants in the initial seed virus were selected.

4.7 Cultivation scale and vessel

In a next step the impact of different cultivation scales and vessels on the HA *N*-glycan fingerprint was investigated. As reviewed by Genzel *et al.* [62] a large range of influenza vaccine manufacturing processes up to the 6000 L scale have been established. In this study, the RKI-strain was produced in MDCK cell culture using T75-, T175-flasks, roller bottles and 1 L-STR. Again, all resulting HA *N*-glycan fingerprints exhibited the 16 MDCK cell-specific peaks between 300 MTU' and 450 MTU' (figure 31A). Relative quantification (figure 31B) revealed eight high abundant peaks (2, 4, 5, 7, 9, 12, 15 and 16) for both T-flask cultivations. A total of nine high abundant peaks were observed when virus was cultivated in roller bottles (2, 4, 5, 7, 9, 11 - 13 and 15) and 1 L-STR (2 - 5, 7, 9, 11, 12 and 15). For all scales and vessels tested, the differences in RPH were in the range of 0.2 % (peak 8) - 7.3 % (peak 2, table 3, [6]).

These results were confirmed for a H3N2 virus subtype, IVA-Uruguay (figure 31C, D): A comparison of MDCK cell-derived IVA-Uruguay revealed similar robust HA *N*-glycosylation fingerprints (figure 31C) for virus production in a T75-flask and a 5 L-STR, showing differences in relative *N*-glycan structure abundances between 0.1 % (peak 3) and 6.6 % (peak 6, figure 31D).

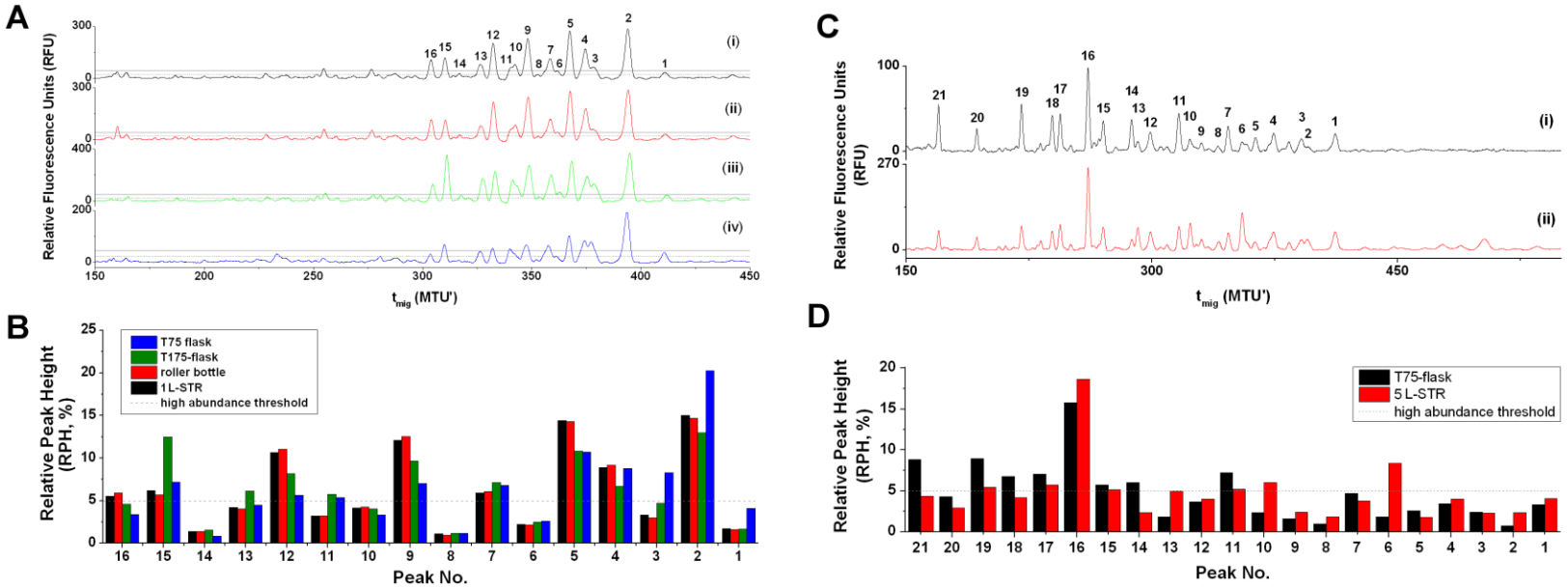


figure 31: Impact of cultivation scale and vessel on the HA *N*-glycosylation pattern of IVA.
(A, B) RKI-strain was produced in T75-flask (i), T175-flask (ii), roller bottle (iii) or 1 L-stirred tank reactor (STR, iv) using MDCK cell culture. Modified and reprinted with permission [6]. (C, D) IVA-Uruguay was produced in T75-flask (i) and 5 L-STR (ii) using MDCK cell culture. (A, C) Shifted overlays of HA *N*-glycan fingerprints, relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). Peaks exceeding the 10x baseline noise threshold (---) in at least one of the direct comparable fingerprints are annotated. (B, D) Relative peak height (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined high abundant if RPH > 5 % (---).

4.8 Virus production media

4.8.1 Media composition

To investigate the impact of media composition on the HA *N*-glycosylation pattern of MDCK cell-derived RKI-strain, three different virus production media were tested: GMEM, Episerf and the chemically defined, protein- and peptide-free SMIF8 medium (for details see table 1). Again all *N*-glycan fingerprints exhibited the 16 MDCK cell-specific peaks between 300 MTU' and 450 MTU' figure 32A). Relative quantification by RPH comparison revealed seven high abundant peaks (2, 4, 5, 7, 9, 12 and 15) for all media used and two additional high abundant peaks when Episerf was used (10 and 16, figure 32B). Differences of RPHs for each peak ranged between 0.2 % and 5.7 % ($I\Delta RPH_{max}I$: peak 12, table 3, [6]).

4.8.2 Trypsin activities

Generally, during virus production trypsin is added for increased virus yields. Hence, in a next step the effect of trypsin additions on the HA *N*-glycosylation pattern of MDCK cell-derived RKI-strain was characterized. Therefore, a standard trypsin activity of 5 U/mL was added [72] once (at time of infection) or every 24 h. As a control, cells were infected without trypsin. Controls reached maximum titers of 2.1 - 2.2 HAU (120 hpi), cultivations infected with trypsin at time of infection achieved up to 2.5 – 2.7 HAU (96 hpi). Daily addition of trypsin resulted in 2.6 – 2.7 HAU (96 hpi, data not shown). As before, the same MDCK cell-specific *N*-glycan fingerprints, consisting of 16 peaks between 300 MTU' and 450 MTU' were obtained for all conditions, (figure 32C). Relative quantification by RPH revealed eight high abundant peaks (2, 5, 7, 9, 10, 12, 13 and 15, figure 32D) for all trypsin containing samples (standard and daily addition) during virus production. The control samples showed a slightly different set of eight high abundant peaks (7, 9, 10, 11, 12, 13, 14 and 15, figure 32D). For all samples differences of relative *N*-glycan structure abundances ranged from 0.6 % -

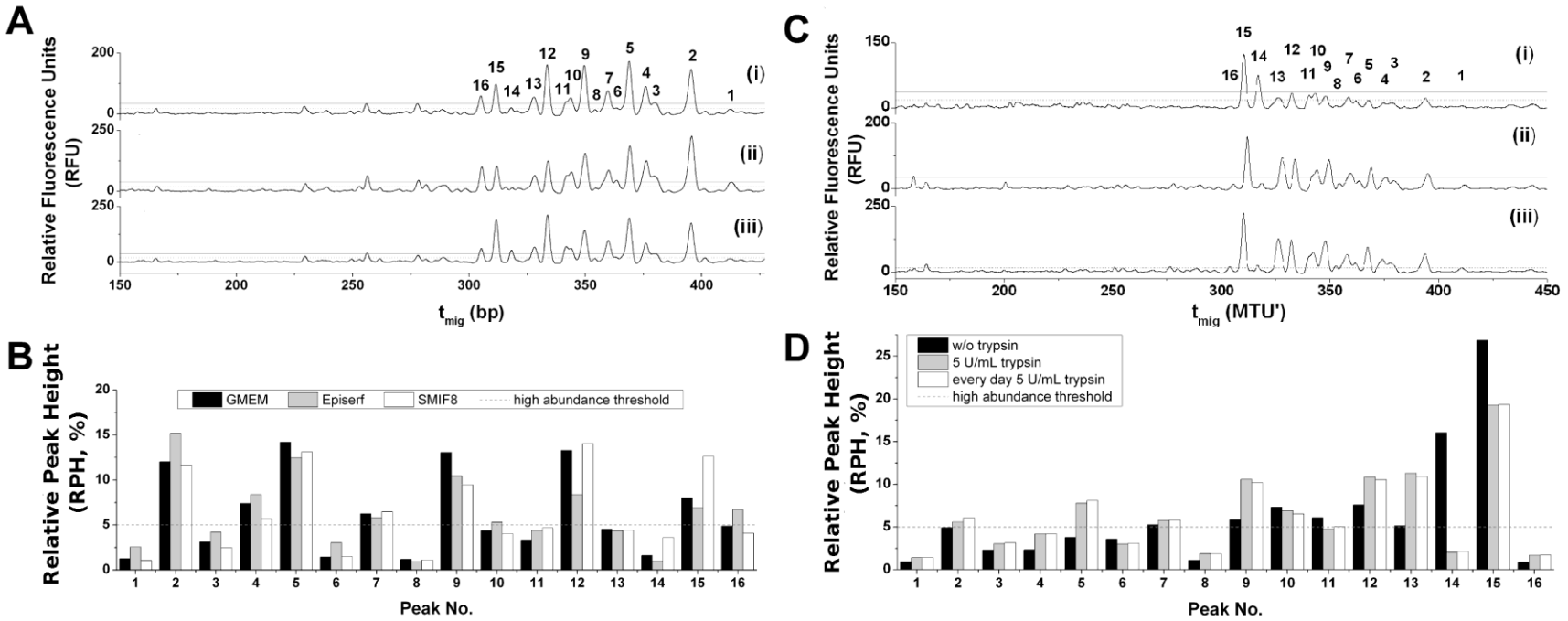


figure 32: Impact of virus production media compositions (A, B) and trypsin activities (C, D) on the HA *N*-glycosylation pattern of IVA-PR8.

(A, C) HA *N*- glycan fingerprints, relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). MDCK cell-specific peaks between 300 MTU' and 450 MTU' exceeding the 10x baseline noise threshold (---) in at least one of the direct comparable fingerprints (table 3, +) were annotated. (A) The virus was produced in GMEM (i), Episerf (ii) and SMIF8 (iii) medium. (C) Virus was produced without (w/o) trypsin (i), with 5 U/mL trypsin (ii) or with an addition of 5 U/mL trypsin every 24 h (iii). Representatives from duplicate samples are displayed (B, D) Relative peak abundance (RPH) in % of the total peak height (TPH, sum of all annotated peaks). Peaks are defined low abundant if RPH < 5 % (---). Modified and reprinted with permission [6].

14.0 % ($I\Delta RPH_{max}I$: peak 14, table 3). These results clearly indicate that the addition of trypsin has no significant impact on the HA *N*-glycosylation pattern with respect to HA *N*-glycan structure presence. However, highest impact on relative *N*-glycan structure abundance has trypsin presence or absence since the differences of relative *N*-glycan structure abundances for all trypsin supplemented samples is ≤ 1 % ($I\Delta RPH_{max}I$: peak 15,data not shown, [6]).

4.9 Virus N-glycosylation and immunogenicity

So far, results demonstrated that some cultivation conditions impact HA *N*-glycosylation of produced virus significantly (e.g. host cell), while other conditions hardly affect HA *N*-glycosylation (e.g. cultivation scale and vessel). It is known that carbohydrates are involved in antigen uptake, processing, presentation, and act in an adjuvant manner as described for saponin from *Quillaja saponaria* [166] or phosphatidylinositol mannosides from *Mycobacterium tuberculosis* [7, 167]. Furthermore, few studies have addressed the impact of HA *N*-glycosylation on anti- and immunogenicity [1-3, 88]. Therefore in cooperation with Dr. B. Lepenies and J. Hütter (Glycoimmunology, MPI-KG, Potsdam-Golm, Germany)) immunogenicity studies were designed in order to directly address the impact of differential *N*-glycosylation on the immunogenicity of IVA preparations. In the following the distribution of work is briefly summarized as well as the results are presented (sections 4.9.1 to 4.9.2), which were published in 2013 [1].

Since the influence of HA *N*-glycosylation on virus immunogenicity and hence on vaccine efficiency would make or break the relevance and value of this work, I started early during my thesis to think of possible ways to address the impact of *N*-glycosylation on immunogenicity. After discussion with a former colleague, Prof. Dr. Th. Schüler, a cooperation with Dr. E. Rapp and myself (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany), Dr. B. Lepenies and J. Hütter (MPI-KG, Potsdam-Golm, Germany) as well as Dr. B. Hundt (IDT Biologika GmbH, Dessau-Rosslau, Germany), was initiated. Within this cooperation the impact of *N*-glycosylation on immunogenicity of different

virus preparations was investigated during immunogenicity studies using a TCR-HA transgenic mouse model and in addition a BALB/c *wildtype* (*wt*) mouse model. T cells from these transgenic mice only express a TCR specific for the $HA_{110-120}$ peptide presented by MHC class-II molecules [168, 169]. The great advantage of this transgenic model is that it allows the characterization of HA-specific $CD4^+$ T cell stimulation by differentially *N*-glycosylated whole virus preparations. The work was distributed as followed:

Multiple differently glycosylated virus preparations were produced by myself. This included cell culture-based virus production, β-PL-inactivation of the virus containing culture broth, cell culture-based inactivation testing, virus quantification by HA-assay, sterility testing, virus isolation and washing. Furthermore, I established and performed native enzyme-based deglycosylation procedures for the generation of natively folded, non-glycosylated virus preparations. Finally, I quantified all virus preparations for their protein content by BCA-assay and characterized each sample's HA *N*-glycosylation pattern by CGE-LIF-based glycoanalysis. Furthermore, seed viruses were sequenced by next-generation pyrosequencing in cooperation with Dr. D. Höper (FLI, Greifswald - Insel Riems, Germany) to confirm congruence of potential *N*-glycosylation sites as well as presence of $HA_{110-120}$ peptide, for which transgenic HA-TCR T cells are specific. Dr. D. Höper and colleagues performed sequencing analysis, processed, sorted, assembled and mapped raw data, whereas the analysis and interpretation of sequence information was generally conducted by myself.

J. Hütter (PhD-student, MPI-KG, Potsdam-Golm, Germany) established and performed *in vitro* whole spleen cell assays with TCR-HA transgenic mice or BALB/c *wt* mice (controls). She used magnetic assisted cell sorting- (MACS-) isolated spleen cell subsets ($CD11c^+$, $CD19^+$) to identify responsible APC population. She analyzed T cell activation by flow cytometry (CD69, CD25) as well as ELISA (IL-2, IL-4, IFNγ) using the cell culture supernatant. Furthermore, she performed *in vivo* boost immunizations (day 0 and 14) with BALB/c *wt* mice, quantifying HA-specific IgG/IgM Ab-titers of mice sera by ELISA (14 days post

immunization, dpi; 28 dpi) and performed an adoptive transfer of labeled TCR-HA transgenic T cells, analyzing their proliferation in the spleen (4 dpi) by flow cytometry of spleen cells. For *in vivo* experiments IL-2 and IFNγ production of $CD4^+$ spleen T cells was quantified after restimulation by ELISPOT. The results of these immunogenicity studies will also be part of J. Hütter's PhD-thesis.
HAI assays to determine titers of Ab in immunized mice sera, which are able to inhibit hemagglutination, were performed by Dr. B. Hundt (IDT Biologika GmbH, Dessau-Rosslau, Germany) and colleagues.

4.9.1 In vitro studies

The following section 4.9.1 contains text, analogous text content, structure, figures and figure legends or parts of figure legends taken from a paper published together with J. Hütter, D. Höper, P.H. Seeberger, E. Rapp and B. Lepenies, January 2013 in *J. Immunol.* [1]. As described before (section 1, last two paragraphs), scentences and/or paragraphs containing quotations are not indicated specifically. The reference will only be given after the phrase or paragraph by the number of the quoted reference.
MDCK and Vero cell-derived HA *N*-glycosylation patterns of the RKI-strain differ significantly (figure 33A(i), B(i), section 4.1 and [73]). The impact of HA *N*-glycosylation on immunogenicity was investigated by stimulating TCR-HA transgenic spleen cells with either β-propiolactone inactivated MDCK or Vero cell-derived virus preparations [1]. Stimulations with Vero cell-derived virus preparations demonstrated significantly increased frequencies of the activation marker CD69 for splenic T cells than with MDCK cell-derived preparations (figure 34A). Consistent with the enhanced CD69 expression, IL-2 levels in splenocyte supernatants were significantly higher when incubated with the Vero than with the MDCK cell-derived virus preparations (figure 34B). The same tendency, i.e. an increased cytokine secretion after stimulation with Vero cell- in comparison to MDCK cell-derived virus, was observed for IFN-γ. Whereas IL-4 secretion was comparable upon stimulation with the Vero and the MDCK cell-derived virus preparations (figure 34C, D). These findings demonstrate that HA *N*-glycosylation

had a marked impact on CD69 expression as well as on IL-2 secretion and to a smaller extent on IFN-γ production, suggesting a marked influence of differential HA *N*-glycosylation on the initiated T cell response [1].

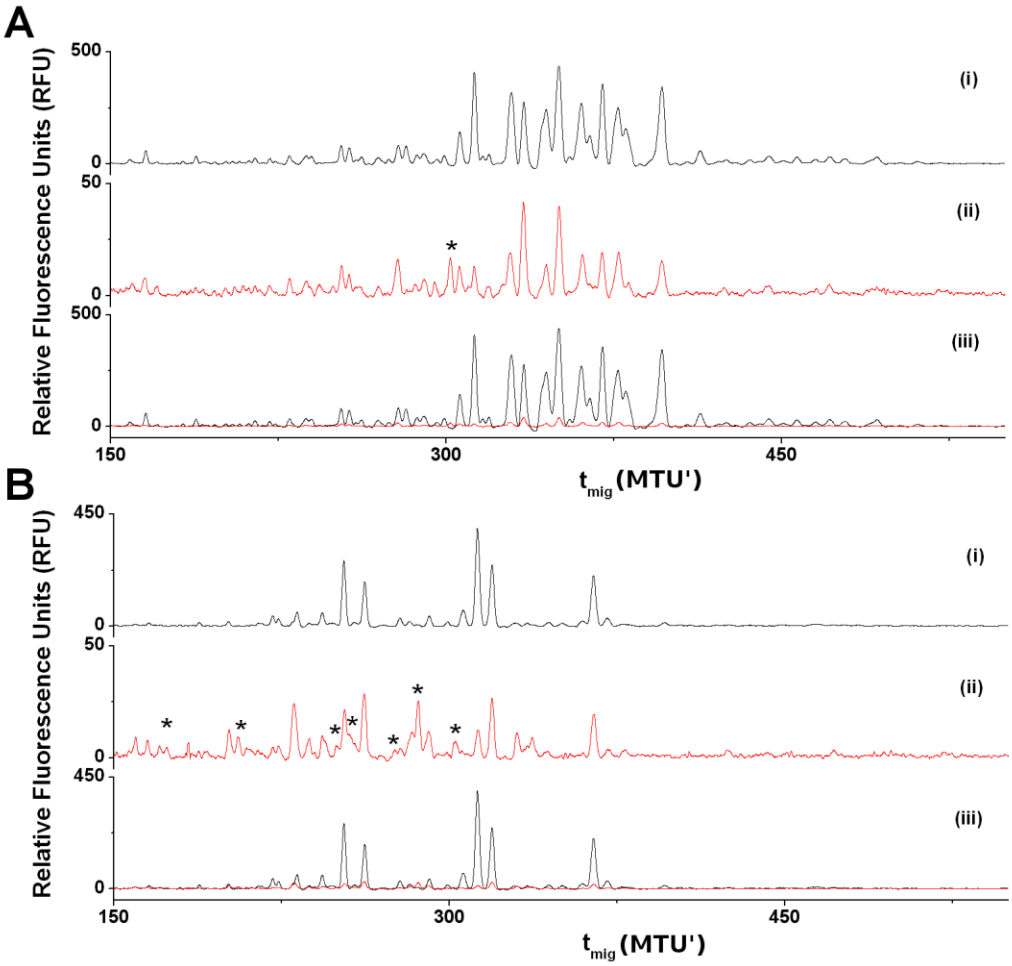


figure 33: HA *N*-glycan fingerprints of glycosylated (i) and deglycosylated (ii) MDCK (A) and Vero (B) cell-derived RKI-strain.

Relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). Shifted overlays (i, ii) and direct overlays (iii) of fully *N*-glycosylated (i) and native deglycosylated HA (ii) show efficient but not complete deglycosylation (note the different scale in i and ii). Glycoanalysis indicated that at least about 90 % of HA *N*-glycan structures were cut off. New glycan structures detected after deglycosylation on the Vero (B (ii)) or MDCK (A (ii)) cell-derived HA are marked with an asterisk (*). Reprinted with permission [1]. Inactivated virus samples, figure and respective data were generated and analyzed by myself.

Moreover, T cell activation kinetics using the two influenza glycovariants indicated that Vero cell- in comparison to MDCK cell-derived virus induces faster T cell activation, thus enhancing T cell proliferation: higher CD69 expression 24 h post stimulation (hps) but not 72 hps with Vero cell-derived virus and higher IL-2

production 72 hps but not 24 hps after stimulation with the Vero cell-derived virus preparations (data not shown, [1]). This differential T cell activation by Vero and MDCK cell-derived virus preparations was demonstrated to be mediated by $CD11c^+$ dendritic cells (data not shown, [1]).

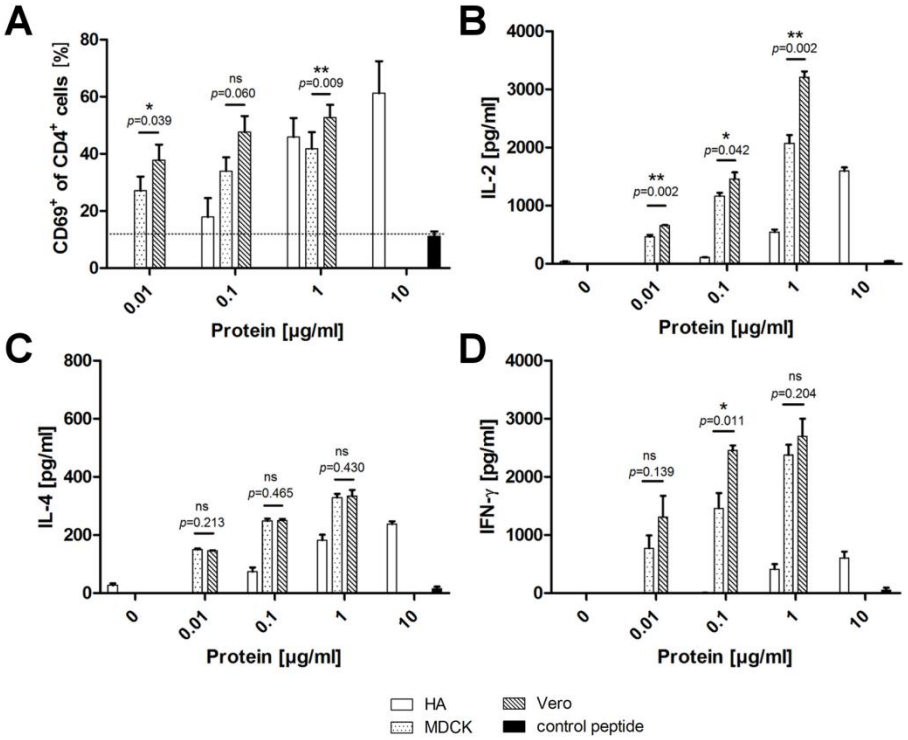


figure 34: Increased T cell activation upon stimulation with Vero cell-derived RKI-strain in TCR-HA transgenic whole spleen cell assays.
Cells were stimulated with a HA peptide ($HA_{110-120}$, positive control, white), with the MDCK (dotted) or Vero cell-derived RKI-strain (striped) or with an ovalbumin peptide ($OVA_{323-339}$, negative control, black). (A) As measured by flow cytometry, a higher frequency of TCR-HA transgenic splenic T cells expressed CD69 (very early T cell activation marker) after stimulation with Vero than with MDCK cell-derived virus. Background frequency of $CD4^+CD69^+$ cells is indicated (•••). (B, C, D) Bar diagrams represent results from three independent experiments. Cytokine levels as measured by ELISA in cell culture supernatants. Stimulation with Vero cell-derived virus induced significantly higher IL-2 secretion by splenocytes, indicating an impact of HA *N*-glycosylation on T cell proliferation. Data are representative of four independent experiments. All data are expressed as mean + SEM. The p-values were determined using paired Student's t-test (*p<0.05, **p<0.01) for MDCK vs. Vero. Significance is indicated by asterisks (*), ns= no significance. Modified and reprinted with permission [1].
Figure and respective data were generated and analyzed by J. Hütter (MPI-KG, Potsdam-Golm, Germany). Inactivated virus samples (MDCK, Vero) were generated by myself.

To further investigate the effect of HA *N*-glycosylation on T cell stimulation, the MDCK and Vero cell-derived glycovariants of the RKI-strain were natively deglycosylated using a variety of endo- and exoglycosidases (section 3.7). HA bands were shifted to lower molecular weights in the SDS-PAGE indicating

successful deglycosylation (data not shown). Bands of fully glycosylated HA monomers were found at about 70 kDa (MDCK cell-derived just above and Vero cell-derived just below). In contrast, deglycosylated preparations exhibited a more diffuse and broader band pattern just below the fully glycosylated band between about 55 kDa and 70 kD (data not shown). *N*-glycosylation pattern analysis confirmed that both variants were deglycosylated for the most part (figure 33A(ii-iii), B(ii-iii)). In particular, the significant reduction of signal intensity from about 450 RFU - 500 RFU to below 50 RFU indicates efficient protein deglycosylation (this conclusion is only possible, if same samples were used, if whole sample preparation was performed in the same batch and if same material batches were used for *N*-glycan analysis). Though, no complete deglycosylation was achieved, glycan analysis showed a reduction of *N*-glycosylation by at least a factor of 10. Moreover, glycan analysis suggested that Vero cell-derived *N*-glycan structures were removed more efficiently (figure 33B(ii)) than MDCK cell-derived structures (figure 33A(ii)) (max. RFU_{Vero} < max. RFU_{MDCK}). Interestingly, deglycosylation resulted in multiple truncated glycan structures of lower molecular weight on the Vero cell-derived HA (marked with an asterisk (*) in figure 33B (ii)). In contrast, on the MDCK cell-derived HA only one shorter glycan structure above 300 MTU' was detected after deglycosylation (marked with an asterisk (*) in figure 33A (ii), [1]).

The impact of HA deglycosylation on immunogenicity was investigated by stimulating TCR-HA transgenic splenocytes with MDCK or Vero cell-derived virus preparations as well as with the deglycosylated control preparations in a whole spleen cell assay (figure 35). Particularly at lower protein concentrations, deglycosylation led for both glycovariants to reduced frequencies of CD69 expression in the $CD4^+$ T cell population as well as to reduced cytokine production levels (figure 35A-D). Though, differences between the glycosylated and deglycosylated virus preparations were more pronounced for the Vero than for the MDCK cell-derived virus (figure 35A-D, [1]). This might be due to the fact that deglycosylation of the MDCK cell-derived virus preparation reduced the HA *N*-glycosylation level by about 90% but without modifying most *N*-glycan

structures. In contrast, HA *N*-glycans of the Vero cell-derived virus preparation were truncated during deglycosylation leading to multiple new glycan structures in addition to the reduced level of glycosylation (figure 33B(ii), [1]).

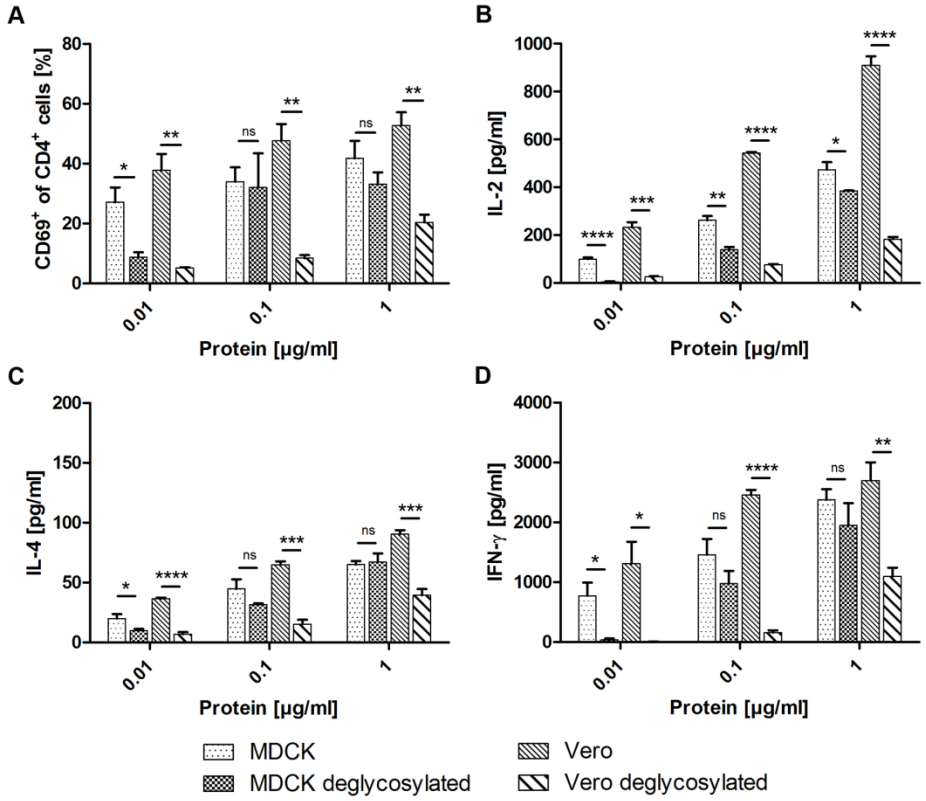


figure 35: Deglycosylation of the RKI-strain reduces T cell activation significantly *in vitro*.
TCR-HA transgenic splenocytes were stimulated for 48 h with glycosylated and natively deglycosylated MDCK or Vero cell-derived virus preparations. (A) CD69 (early T cell activation marker) frequency of $CD4^+$ T cells (duplicates each). Upon deglycosylation frequencies of $CD69^+$ T cells decreased markedly. Also cytokine levels of (B) IL-2, (C) IL-4 and (D) IFN-γ, in particular for the Vero cell-derived virus decreased after deglycosylation. Bar diagrams are representatives of triplicates and represent mean + SEM. Significance was tested by unpaired student's t-test (ns= no significance, (*p<0.05, **p<0.01, ***p<0.001, ****p<0.0001) for glycosylated versus natively deglycosylated virus preparations. Reprinted with permission [1].
Figure and respective data were generated and analyzed by J. Hütter (MPI-KG, Potsdam-Golm, Germany). Inactivated virus samples (MDCK, Vero, MDCK deglycosylated, Vero deglycosylated) were generated by myself.

4.9.2 In vivo studies in mice

The following section 4.9.2 contains analogous text content, structure and modified figures taken from a paper published together with J. Hütter, D. Höper, P.H. Seeberger, E. Rapp and B. Lepenies, January 2013 in *J. Immunol. [1].* As described before (section 1, last two paragraphs), scentences and/or paragraphs

containing quotations are not indicated specifically. The reference will only be given after the phrase or paragraph by the number of the quoted reference.

To investigate the impact of *N*-glycosylation on immunogenicity *in vivo* and in a *wt* background, BALB/c mice were prime-boost immunized intraperitoneal (i.p.) on day 1 and 14 with fully glycosylated Vero or MDCK cell-derived virus preparations or with the deglycosylated controls. IgG/IgM titers of HA-specific Ab in mice sera were determined on day 14 and 28 by ELISA. In agreement with the *in vitro* results, deglycosylation led to dramatically reduced HA-specific Ab levels in sera of immunized mice for the MDCK as well as for the Vero cell-derived virus 14 dpi and 28 dpi (figure 36A). Interestingly, immunization with MDCK cell-derived virus induced significantly higher levels of HA-specific Ab on day 14 than immunization with Vero cell-derived virus (figure 36A, left). This effect was no more evident on day 28 (figure 36A, right). Consistent with the ELISA, the ability of mice sera of day 28 to inhibit hemagglutination in an HAI assay decreased significantly when mice were immunized with the deglycosylated virus preparations (figure 36B). Interestingly, on day 28 Ab induced by immunization with MDCK cell-derived virus still significantly better inhibited hemagglutination than the Ab induced by immunization with Vero cell-derived virus (figure 36B, [1]).

Altogether, these results suggest that the *N*-glycosylation pattern of the MDCK cell-derived virus preparation more promotes the B cell-mediated humoral immune response whereas the Vero cell-derived virus preparation slightly more stimulates the T cell-mediated cellular immune response [1].

Furthermore, T cell proliferation in the spleen of wild-type BALB/c mice of labeled and adoptively transferred TCR-HA transgenic T cells was comparable after immunization with Vero and MDCK cell-derived virus (figure 36C). But consistent with the *in vitro* findings the frequency of IL-2 producing splenocytes was higher after immunization with the Vero than with the MDCK cell-derived virus preparations (figure 36D, [1]).

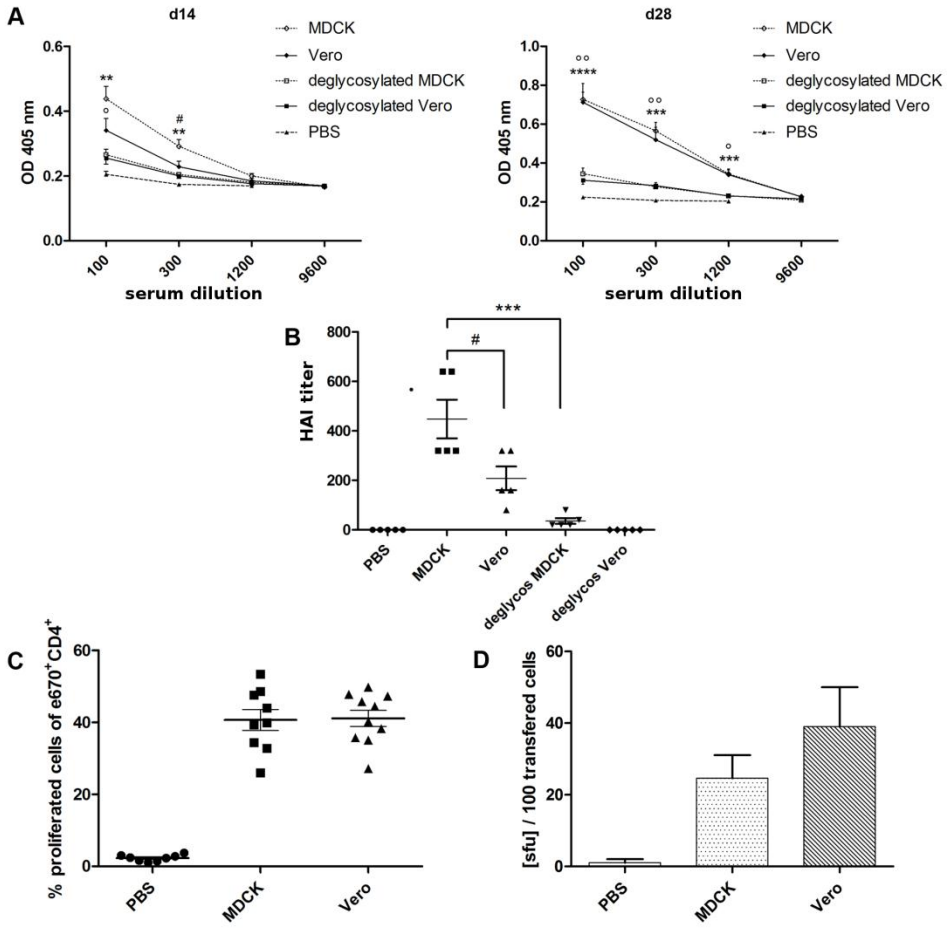


figure 36: *In-vivo* immunogenicity of glycovariants of the RKI-strain.
(A) HA-specific Ab titers in BALB/c *wt* mice sera after prime- (day 0) boost- (day 14) immunization (i.p.) with 10 µg virus preparation (fully glycosylated or deglycosylated M- or V-variants) as measured by ELISA (in triplicates). Data on day 14 and 28 are derived from 5 BALB/c *wt* mice, respectively. Data represent means + SEM. Significance was tested by unpaired student's t-test for glycosylated versus natively deglycosylated virus variants (MDCK cell-derived: *; Vero cell-derived: °) and for glycosylated MDCK versus Vero cell-derived virus variants (#; ns= no significance, */°/$^{\#}$p<0.05, **/°°p<0.01, ***p<0.001, ****p<0.0001). (B) Inhibition of hemagglutination (HAI) titers of immunized mice sera were measured on day 28. One symbol represents one mouse. HAI for Vero-deglycosylated was 0, hence no significance could be determined for Vero-glycosylated versus -deglycosylated. (C) MACS-purified TCR-HA transgenic T cells were labeled with e670 (cell-proliferation dye) on day 0 were i. v. adoptively transferred into BALB/c *wt* mice. On day 1 mice were immunized with 50 µg of glycovariants (MDCK or Vero) or with PBS. Proliferation of TCR-HA transgenic T cells and T cell activation were analyzed on day 5. Gating on $e670^{+}CD4^{+}$ cells was performed for flow cytometric analysis. Frequency of transferred and labeled TCR-HA transgenic T cells, which had proliferated. Each symbol represents one mouse, derived from altogether three independent experiments. (D) Frequency of IL-2 producing splenocytes as analyzed by ELISpot in triplicates, after re-stimulation of isolated splenocytes with $HA_{110-120}$-peptide. Data are normalized to the number of adoptively transferred TCR-HA transgenic T cells based on flow cytometry data. Bar diagram represents mean + SEM. Modified and reprinted with permission [1].
Figure and respective data were generated (except B; Dr. B. Hundt, IDT Biologika GmbH, Dessau-Rosslau, Germany) and analyzed by J. Hütter (MPI-KG, Potsdam-Golm, Germany). Inactivated virus samples (MDCK, Vero, deglycosylated MDCK, deglycosylated Vero) were generated by myself.

In conclusion, these studies indicate that differential *N*-glycosylation has a marked impact on IVA-PR8 immunogenicity *in vitro* and *in vivo*, in TCR-HA transgenic as well as in *wt* BALB/c mice [1].

5 Discussion

5.1 Impact of host cells and host cell adaptation on the HA N-glycosylation pattern

This study confirms that host cell choice significantly impacts HA *N*-glycosylation patterns of IVA. Previously, Schwarzer *et al.* demonstrated strict host cell-specificity for HA *N*-glycosylation [73]. They showed by exoglycosidase digestions, that all *N*-glycan structures attached to MDCK cell-derived HA are of the complex type with either terminal α- or β-galactose. On the other hand most *N*-glycan structures attached to Vero cell-derived HA are of the complex type with terminal β-galactose but some other structures are of the high mannose type. Interestingly, AGE1.CR cell-derived HA also carried mostly complex *N*-glycans with terminal β-galactose and some high mannose structures, but additionally hybrid structures were detected [73]. The present work extends the data of Schwarzer *et al.* [73] for a variety of additional cell lines such as AGE1.CR.pIX cells, Cap cells and embryonated hens' eggs (section 4.1, figure 15). Strict host cell-specificity was demonstrated and is in agreement with Raju *et al.*, who showed significant differences in *N*-glycosylation for IgG produced in host cells from 13 different species [170]. Host cell-specificity of glycosylation is caused by species-specific differences, e.g. in the complex glycosylation machinery with respect to expressed enzyme repertoires and enzyme activities [31]. Also, the hosts' cellular metabolism associated with glycosylation was reported to impact glycosylation outcomes due to differences in the availability of intracellular sugar-nucleotide donors [171]. Another factor influencing the *N*-glycosylation of proteins is the transit time in the Golgi, which may also differ from cell line to cell line [31, 171].

Interestingly, taking all different host cell-derived HA *N*-glycan fingerprints obtained in this work into account, MDCK cell-derived fingerprints show the largest differences compared to all other hosts in terms of migration times (tendency towards bigger structures), number of peaks (tendency towards higher variety of different structures), and peak intensities (tendency towards higher

numbers of different, more or less high abundant *N*-glycan structures). MDCK cell-derived *N*-glycan structures are in average bigger. This may be either due to longer oligosaccharide chains or due to higher numbers of antennae, which remains to be investigated e.g. by MS-based *N*-glycan structure elucidation.

The comparison of the AGE1.CR.pIX- with the AGE1.CR-derived HA *N*-glycan fingerprint published by Schwarzer *et al.* [73] shows no significant differences. This suggests that no significant changes to the host cell's glycosylation machinery were induced by the stable integration of the adenoviral pIX gene into the AGE1.CR genome. The pIX gene was integrated into the genome to promote capsid stabilization of cognate adenovirus in particular and may also alter further cell line properties [7, 67, 73, 172].

The adaptation of adherent MDCK cells from serum-containing to serum-free cell growth in Episerf caused no significant changes in peak presence but resulted in differences in relative peak abundance (maximum $|\Delta RPH|$ of 8.4 %, table 3) of the HA *N*-glycan fingerprints (section 4.2.1, figure 16). Therefore, this host cell adaptation had no significant impact on the host cell's glycosylation machinery. The differences in relative *N*-glycan structure abundances may be caused by differing nutrient concentrations in the media. In particular, it was shown for recombinant protein production processes that glucose and glutamine concentrations in the media can affect *N*-glycosylation [173-175]. Furthermore, glucose or glutamine limitations can lead to decreased sialylation and increased presence of hybrid and high-mannose glycan types [176], which is possibly due to reduced intracellular UDP-*N*-acetylgalactosamine concentrations [177]. Moreover, metabolic by-product accumulation of lactate and ammonia, influenced by glucose and glutamine concentrations in the medium [6, 178] have been previously described to have an impact on *N*-glycosylation [179-181] influencing antennary-, sialylation- [182] or galactosylation- [183] levels of glycoproteins. Finally, adaptation-induced changes of enzyme activities of the glycosylation machinery or within the hosts' cellular metabolism associated with intracellular sugar-nucleotide donors may also have contributed to differences in relative *N*-glycan structure abundances [6].

In contrast, adaptation of adherent MDCK cells to serum-free suspension growth [142] changed peak presence as well as relative peak abundances in HA *N*-glycan fingerprints (maximum │ΔRPH│ of 18.1 %, table 3, section 4.2.2, figure 17). This emphasizes that every repetition of host cell adaptation - as for MDCK.SUS1 and MDCK.SUS3 - may result in a totally new HA *N*-glycosylation pattern, showing a distinct fingerprint. This may be explained by adaptation-induced, altered presence or absence of oligosaccharide-processing enzymes in the host cell's proteome or by host cell-specific differences in the relative activities of competitive enzymes involved. For example, glycoengineering allows to modulate, knock out or newly introduce specific glycosyltransferases in established producer cells to achieve a more suitable glycosylation pattern for therapeutic glycoproteins [184-187]. Concerning the effect of enzyme activities, e.g. Kobata described transformational changes of sugar chain moieties of the human luteinizing hormone due to an enhancement of the fucosyltransferase [188]. Such changes in the glycosylation machinery may occur due to spontaneous mutations in the host cell genome as Stanley reported for CHO cell mutants, which acquired multiple glycosylation defects leading to altered glycosylation of endogenous as well as recombinant glycoproteins [189, 190]. The occurrence of such mutations is also suggested by proteomic studies, revealing significant changes in the host cell's proteome, acquired after adaptation to suspension growth (manuscript in preparation by Kluge *et al.*, MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany). In contrast, further adaptation of MDCK.SUS1 to MDCK.SUS2 for better growth characteristics did not impact the host cell's *N*-glycosylation machinery any further as indicated by rather stable HA *N*-glycosylation fingerprints (maximum │ΔRPH│ of 1.9 %, table 3, [6]).

5.2 Impact of virus strain, virus supplier, virus passage and virus adaptation status on the HA N-glycosylation pattern

First and foremost the host cell determines the HA *N*-glycosylation pattern. In addition, the virus strain may influence the HA *N*-glycosylation. In this study,

CGE-LIF-based glycoanalysis revealed distinct *N*-glycan fingerprints for two H1N1 and two H3N2 strains produced in MDCK cell culture (section 4.3, figure 18). However, mostly RPH, representing relative *N*-glycan structure abundances (maximum $|\Delta RPH|$ of 25.2 %, table 3), vary between the strains. The differences in RPH were smaller, if two strains of the same subtype (IVA-PR8 and IVA-Victoria, H1N1; or IVA-Uruguay and IVA-California, H3N2) were compared (maximum $|\Delta RPH_{H1N1}|$ of 12.4 %, maximum $|\Delta RPH_{H3N2}|$ of 9.8 %, table 3). This highlights the closer relation of the H1N1-strains and the H3N2-strains. A few differences with respect to peak presence were detected, when all strains were compared: however, these represented mostly low abundant structures (e.g. peak 24, not present in both H1N1 strains, figure 18). These differences in *N*-glycan structure abundance as well as presence are most probably due to slightly varying three-dimensional conformations between all four analyzed IVA strains. Interestingly, Schwarzer *et al.* showed for MDCK cell-specific glycovariants by sequential exoglycosidase digestions that all glycans attached to IVA-PR8-derived HA were of the complex type with terminal α- and β-galactose, while IVA-WSN/67/2005- (H3N2) derived HA only possessed few complex *N*-glycan structures with terminal α- and β-galactose and the highest abundant structures were of the high mannose type [73]. Similar digestion patterns were found for influenza virus B/Mal/2506/2004-derived HA: highest abundant structures were of the high mannose type, but all detected complex *N*-glycan structures were terminated by β-galactose [73]. Already in 1997 Mir-Shekari *et al.* highlighted the importance of the three-dimensional structure of HA for *N*-glycosylation. They demonstrated that each of the four *N*-glycosylation sites within the HA_1 subunit was occupied by conserved *N*-glycan structures depending on specific *N*-glycosylation site characteristic: in loop regions bi-, tri- and tetra-antennary complex *N*-glycans were present. In contrast, the glycosylation site buried in the α-helix was occupied by high mannose structures, indicating that these structures are no more accessible for glycosylation modulating enzymes after proper protein folding and trimerization of the HA molecule [31, 191]. However, beside the three-dimensional structure of the

protein, also specific glycosylation characteristics have been described to influence further glycan processing: e.g. Harpaz and Schachter demonstrated that the presence of bisecting GlcNAc inhibited GlcNAc transferases, which is responsible for further glycan branching [192].

In addition to different virus strains, the impact of IVA-PR8 origin with respect to suppliers (i.e. RKI- versus NIBSC-strain) on HA *N*-glycosylation was investigated (section 4.5, figure 20). The choice of virus supplier did only slightly effect relative peak heights (maximum $|\Delta RPH|$ of 5.0 %, table 3). Overall, this suggests that *N*-glycan fingerprints of strains from different suppliers, which may as well slightly vary in their HA genome sequence (as shown for RKI- and NIBSC-strain), can also show differences regarding relative peak abundance, i.e. probably with low abundance. This was expected as these two closely related IVA-PR8 virus seeds do not differ in any potential glycosylation site of the HA molecule [4, 6]. Still, the two virus seeds differed significantly concerning maximum virus titers obtained, interferon and apoptosis induction as well as the activation of general host cell responses [193-196]. These variations may be caused by different ratios of defective interfering virus particles in seed viruses [197]. Furthermore, Andersen *et al.* suggested in 2000, that the ratio of cells in G0/G1 phase determines the glycosylation efficiency [198]. 20 years before, Hakimi *et al.* reported that virus isolated from rapidly growing fibroblasts exhibited a higher number of larger glycans compared to virus replicated in non-growing cells [199]. All this suggests, that the cell status can affect *N*-glycosylation and that the slight differences in RPH in the HA *N*-glycan fingerprints between these two tested virus seeds may be due to differing virus-induced physiological statuses of the host cells rather than caused by sequential differences in the HA [6].

The characterization of HA *N*-glycan fingerprints derived from multiple virus passages demonstrated a good reproducibility and an overall high biological stability of HA *N*-glycosylation (maximum $|\Delta RPH|$ of 3.5 %, table 3, section 4.4, figure 19). This stability makes further investigations possible, e.g. during virus adaptation. However, HA *N*-glycosylation pattern stability does not allow any conclusions concerning occurred mutation events, since slight differences in

virus genome sequences do not necessarily lead to significantly altered HA *N*-glycan fingerprints, as observed during the comparison of the RKI- and the NIBSC-strain.

Next, the impact on HA *N*-glycosylation of seed virus adaptation to various host cells was investigated (section 4.6.1/2, figure 21 to figure 24). Virus seed adaptation is often necessary to optimize virus yields in production cells. The results clearly show that predominantly the host cell line determines the HA *N*-glycosylation pattern of a specific virus strain (details in section 5.1). For NIBSC- as well as RKI-strain-derived HA *N*-glycan fingerprints, the variation between the first passage in Vero cells (passage 2) and the patterns of all subsequent Vero cell adaptation passages (passages 3 – 6) cannot completely be explained by applied harvest time point. This indicates that in addition to the harvest time point (maximum $|\Delta RPH|$ of 14.2 % for RKI-strain in Vero cells, table 3), the viral adaptation status (maximum $|\Delta RPH|$ of 21.0 % for RKI-strain to Vero cells, table 3) may impact relative *N*-glycan structure abundances. Interestingly, for all tested virus strains and host cells, the glycan pattern stabilized at the latest after the first passage in the new cell line. Moreover, the quicker release of virus particles and to begin with the increase in HA-titers - as observed for most seed virus adaptations – indicates that other factors are involved in the adaptation process resulting in improved fitness of virus subpopulations. Such factors include the acquisition of changes in the virus genome sequence [4].

Taking together data from all adaptation series, the results clearly indicate that the impact of virus adaptation status on HA *N*-glycan fingerprints and virus replication dynamics depends on chosen virus strains as well as on selected cell lines for virus replication (section 4.6.2, figure 21 to figure 28). This is caused by strain- and/or host cell-specific characteristics: e.g. the passage history of a seed virus may have already selected fitter virus variants for the new host system or the host cells express favoured receptors. For instance, the expression of both α2,3 and α2,6 linked sialic acids on cell surfaces may also support the propagation of avian (binding α2,3 sialic acids) as well as human (binding α2,6 sialic acids) influenza strains, as observed for AGE1.CR.pIX cells [200, 201].

5.3 Impact of virus adaptation on quasispecies composition

5.3.1 Characterization of virus seeds

The fact, that cDNA sequences are homogenous in the virus seed of the RKI-strain (passage 1), clearly suggests the presence of only one virus population (table 4). In contrast, the initial virus seed population (passage 1) of the NIBSC-strain comprises multiple virus variants (table 5) regarding segments 2 (silent), 3 (silent, non-HCD quality), 4, 5, 6 (silent, non-HCD quality), 7 and 8 (non-HCD quality): in segment 4 four substituted AA residues (Y24H, T397A/S, D455Y, N460D) were detected, whereas in segment 5 one subpopulation carrying the substitution Q20P was detected. Q20P is located within a domain interacting with cellular proteins, in particular with the splicing factor BAT1 [202, 203]. In 2011 Ping *et al.* characterized this Q20P substitution as a mouse adaptive mutation [204]. Its presence may be explained by the NIBSC-strain passage history, which is not clear. However, in segment 7 coding for M1 and M2 proteins a total number of 22 AA substitutions have been detected in the NIBSC seed virus (passage 1). These comprise the substitutions N30D, Q72R, A167 T207S, R208Q and D231N, which are located within the domain of M1, interacting with the ribonucleocapsid protein (RNP) [204]. The sequence conflict SQ207/208IR was characterized before by comparing the sequences published by Winter *et al.* in 1980 and by Ghedin *et al.* in the course of the National Institute of Allergy and Infectious Diseases (NIAD) influenza genome sequencing project [205, 206]. Further subpopulations detected for M1 carried the substitutions Y110H, A116S, C126S, T137A, A140T and V143A, which are all located within α-helices, involved in membrane-binding [158, 207]. Subpopulations detected for M2 comprised the substitutions V27A/I, S30A, S31N and I39T, which are located within the ion channel [204]. Of these V27A and S31N have been associated with amantadine resistance [208, 209]. Furthermore, a comparison of available sequence information for IVA-PR8 reveals the sequence conflicts A27I, L54R, R61G, Q77R as well as characterizes I39T as a natural variant of the IVA-PR8/Mount Sinai-strain [158, 205, 206, 210].

5.3.2 Comparison of two IVA-PR8 virus seeds, RKI- vs. NIBSC-strain

The comparison of consensus sequences of both IVA-PR8 seed viruses (passages 1) – the RKI- and the NIBSC-strain – reveals four AA substitutions for segment 1 coding for PB2 (table 6, figure S 5), six AA substitutions for segment 2 coding for PB1 (table 6, figure S 6), three AA substitutions for segment 2 coding for PB1-F2 (table 6, figure S 7), three AA substitutions for segment 3 coding for PA (table 6, figure S 8), six AA substitutions and one deletion for segment 4 coding for HA (table 6, figure S 9), four AA substitutions for segment 5 coding for NP (table 6, figure S 10), nine AA substitutions for segment 6 coding for NA (table 6, figure S 11), twelve AA substitutions for segment 7 coding for M1 (table 6, figure S 12), six AA substitutions for segment 7 coding for M2 (table 6, figure S 13), two AA substitutions for segment 8 coding for NS1 (table 6, figure S 14) and two AA substitutions for segment 8 coding for NS2/NEP (table 6, figure S 15). Interestingly, the RKI-sequence for segment 7 matches the subpopulations of the NIBSC-sequence (carrying the HCD) perfectly on the AA-level. Even on the cDNA-level the 33 differences are reduced to one silent substitution (A438G). Similar findings were observed for segment 6, though to a lesser extend: the five subpopulations (cDNA/AA: A904G/I302V, G940A/E314K, T1269C/T423T, A1300G/N434D, G1352C/S451T) arising during the adaptation processes reduce sequential differences between the RKI- and the NIBSC-strain from nine to five and from fourteen to nine positions at the AA- and cDNA-level, respectively. Hence, these data may indicate a possible contamination and subsequent reassortment of segment 7 of the NIBSC-seed with the RKI-strain. Or if the original NIBSC-sequence matches the RKI-sequence of segment 7, the NIBSC-seed may have been contaminated with some other virus strain, which already dominates the virus population in passage 1. Since the sequences from other NIBSC-strain seed virus variants, also comprised a mixture of RKI-specific (e.g. G620/S207) and NIBSC-specific (T620/I207) residues in segment 7 (data not shown, data obtained by Dr. T. Frensing and B. Heynisch (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) in cooperation with Dr. D. Höper (FLI, Greifswald - Insel Riems, Germany) a possible contamination of the

table 6: Sequence differences of viral genomes between the RKI- and the NIBSC-strain.
Differing AAs for both strains are indicated by the AA found in the RKI-strain, followed by the AA position, followed by the AA in the NIBSC-strain. Consensus suquences were used for the alignment.

		number of mutations											
segment	**protein**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
segment 1	**PB2**	M105I	V259L	V504I	R702K								
segment 2	**PB1**	A155T	V195M	I205M	R208K	I336V	P394S						
segment 2	**PB1-F2**	P43Q	K59R	Q60R									
segment 3	**PA**	K237E	V354I	L550I									
segment 4	**HA**	K147-	A156E	E158K	I208L	R269M	F309Y	S398T					
segment 5	**NP**	V353L	V425I	T430N	A442T								
segment 6	**NA**	M15L	R128K	N131S	V302I	K314E	E371K	I403M	D434N	T451S			
segment 7	**M1**	D30N	R72Q	H110Y	S116A	S126C	A137T	T140A	A143V	T167A	S207I	Q208R	N231D
segment 7	**M2**	A27V	A30S	N31S	R54L	G61R	R77Q						
segment 8	**NS1**	E55K	E101D										
segment 8	**NEP (NS2)**	R88K	V89I										

NIBSC-strain with the RKI-strain during these studies can be ruled out. Another possibility is the accumulation of mutations within segment 7 of the NIBSC-strain. Previously, the theoretically identical RKI- and the NIBSC-strain have been described to differ significantly in infection characteristics such as IFN response, apoptosis induction, final virus yields [194, 195] and the activation of general host cell response [193, 196]: The NIBSC-strain was characterized to induce higher levels of IFN, to express more Mx proteins, to induce apoptosis earlier, and to reach lower final titers than the RKI-strain. Seitz *et al.* hypothesized that two amino acid substitutions in the non-structural protein 1 (NS1) might be related to these differences. Another approach correlates higher amounts of defective interfering virus particles (DIPS) in seed virus preparations to higher apoptosis rates, IFN expression and to decreased final virus titers (personal communication Pflugmacher, Frensing, MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany). Our present findings suggest that lower average virus yields, as well as increased IFN responses and apoptosis rates may at least be partly due to a broadened quasispecies of the NIBSC-strain virus seed comprising low-yield-virus variants [4].

5.3.3 Quasispecies of segments 1 to 3

During virus adaptation the quasispecies of the RKI-strain broadened in dependence of the adaptation pressure. The substitutions within segment 1 coding for PB2 Y115H (located within PB1-binding site), C196W (located within

NP-binding-site) and V451I (located within nuclear localization signal, NLS) are all located within regions important for protein-protein contact or protein transport, respectively [204]. To the author's knowledge, these mutations have not yet been described before. For the NIBSC strain no AA substitutions were detected during the adaptation processes within segments 1, 2 and 3.

5.3.4 Quasispecies of segment 4

During all performed IVA-PR8 adaptations from MDCK to Vero cells, additional virus variants with mutations in the HA stem region were generated. While the quasispecies of the homogenous seed virus of the RKI-strain broadened, the heterogeneity of the seed NIBSC-strain persisted during the whole adaptation processes: from passages 1 to 11 samples comprised multiple virus subpopulations. Thereby, sizes of subpopulations as well as selection and extinction of specific virus variants varied in dependence of adaptation pressure, i.e. the host. However, no potential HA *N*-glycosylation sites or AA residue in their close neighborhood were affected [4].

A comparison of results obtained for both adaptations from MDCK to Vero cells and back using the RKI- and NIBSC-strain revealed one interesting similarity: the substitution of lysine by glutamic acid at position 459/460 (RKI-/NIBSC-strain). The K459E variant of the NIBSC-strain that carries a deletion in the HA_1 chain (K147-) corresponds to the K460E variant of the RKI-strain. Moreover, all substitutions detected during forward and backward adaptation of both IVA-PR8 strains are located in the HA_2 chain, neither inside nor in close proximity of any *N*-glycosylation site. They are, however, located in the inside of the HA trimer within or in close proximity to the fusion peptide pocket: within the subunits' contact site for the RKI-strain (figure 29A-D) and within the subunits and monomer contact sites for the NIBSC-strain (figure 29E-H). The second adaptation of the RKI-strain by Genzel *et al.* [72] resulted, besides some frameshift mutations, in the deletion of AA 338 (I338-). This deletion is located within the fusion subdomain [165] of the HA molecule (figure 30). Substitutions in the region of HA subunit contacts have been described to be crucial for the

stability of the structure of the native protein. Among other factors, the optimal stability depends on pH or temperature [22, 211-213]. In this regard lower pH environments for instance require higher native structure stability, whereas elevated pH-values require less stable structure conformations to mediate membrane fusion. In 2010 Reed *et al.* demonstrated using recombinant H5N1 influenza viruses that substitutions within the fusion peptide pocket and the α-helix of HA_2 alter the pH of activation of HA, which in term effects influenza virus pathogenicity as well as transmissibility in mallards [214]. Furthermore, Thoennes *et al.* showed that different substitutions at HA_2 position 111 of a H3N2 influenza virus strain significantly effected fusion pH, suggesting, a key role of this residue for neutral pH structure stability [215]. This is consistent with findings of Kawaoka's group in 2012. They demonstrated, that the N460D (HA_2N117D as published by Murakami *et al.*) substitution, we detected after NIBSC-strain adaptation to Vero cells, is responsible for improved replication in Vero cells without inhibiting growth in MDCK cells [216]. Furthermore, they showed that the N460D substitution resulted in an increased or boarded optimal pH range for viral membrane fusion in comparison to the *wt* virus without the N460D substitution. Based on two fluorescent dyes – one pH stable, the other pH sensitive – they demonstrated a higher intensity ratio for Vero cells then for MDCK cells, suggesting higher pH values in the early endosome of Vero cells [216]. Also Nakowitsch *et al.* described mutations in the HA_1 and HA_2 after passaging a H3N2 IVA on Vero cells, which led to impaired virus stability. Interestingly, these mutations also led to decreased immunogenicity in ferrets [217]. Other authors also reported substitutions within the HA stalk region, altering HA stability and hence pH of membrane fusion [22, 214, 218, 219]. Therefore, all substitutions within the HA molecule in this study, occurring during virus adaptation from MDCK to Vero cell-based replication of the RKI- or NIBSC strain most probably modulate the electrostatic balance of the HA molecule and hence alter stability of its three-dimensional conformation. Interestingly, Rott *et al.* [220] described the occurrence of virus variants after adaptation to MDCK cells, which exhibited elevated fusion pH caused by substitutions within the HA_1 chain. However, these

experiments indicate a crucial role of the HA stem region for virus adaptation from MDCK to Vero cells. Here, inter-monomer or inter-subunit contact is mediated, suggesting HA-initiated fusion as a driving factor of adaptation pressure. Zaraket *et al.* reported that altered HA acid stability of a H5N1-strain impacts virus growth in the upper respiratory tract and contact transmission in ferrets [221]. Furthermore they described an impact on virus replication and pathogenesis in mice [4, 222].

In this study, all three virus adaptations show a significant change of the HA *N*-glycosylation pattern with the change of host cell system as recently published by Genzel *et al.* as well as Roedig *et al.* [4, 72]. Additionally, all three experiments demonstrate that only a small sequence adaptation is required for successful infection and fast growth to high titers in new host cell lines. These findings are in agreement with results of Wagner *et al.* and Klenk *et al.* [163, 164] who reported that *N*-glycans attached to the stem domain of HA efficiently regulate influenza A virus replication. The authors showed that a loss of *N*-glycans in the stem region results in increased pH-sensitivity of the virus and that these viruses are also temperature sensitive. However, for most detected substitutions (except for N460D during NIBSC-strain adaptation), the exact functions remain to be investigated, including the question whether these substitutions alter acid stability of HA [223], pH of activation, or membrane fusion [224]. Or, whether they simply counteract steric hindrance [164] caused by Vero cell-specific changes in HA *N*-glycosylation, e.g. on residues 28 and/or 40 in the stem region of HA to achieve low-pH conformation required for membrane fusion [4].

5.3.5 Quasispecies of segment 5

In segment 5 coding for NP several substitutions were detected during adaptation processes within functional domains of the RKI-strain: the substitution S287R is located within the domain for NP-binding, whereas substitutions E294D (E294V during NIBSC-strain adaptation), N309S, A442T, A472P and S473N are located in a region involved in NP-binding as well as PB2-binding [204]. Although mutations occurred in both regions of interaction - in PB2 (C196W) as well as in

NP - a correlation between the PB2 and the NP adaptation is not probable due to strongly differing mutant population sizes (NP-binding site of PB2: 7 % *versus* NP- and PB2-binding-site of NP: 47 %, 20 %, 12 %, 42 %, 38 %). Furthermore, Chen *et al.* defined the residue 442 as a host-specific genetic signature, where avian strains usually carry a T and human strains carry an A [225]. On the one hand the A442T subpopulation may have already been present below the detection limit in passage 1. The presence of A442T in the virus seed is then most likely due to the virus' passage history: the virus was isolated from humans in 1934, passaged in embryonated chicken eggs as all viruses were in former times and mostly still are, before the virus was adapted in our laboratory to MDCK cell culture-based virus replication. If this was the case, its detection in passage 6 is most probably due to selection in Vero cell-based virus replication. On the other hand, the fact that a second adaptation of the RKI-strain to Vero cells by Genzel *et al.* did not result in a A442T subpopulation strongly indicates that no selection but rather a mutation led to the A442T substitution. However, the accumulation of substitutions between AA 294 and 309 (47 % and 20 %) and between AA 442 and 473 (12 % - 42 %) suggests that a structural change in this region was necessary for successful virus replication in Vero cells. This is supported by the data from the second adaptation by Genzel *et al.* [72] showing comparable accumulations of AA substitutions within these two regions. Interestingly, all substitutions in these two regions were first detected in passage 6 and decreased to under the detection limit in passage 11, which strongly indicates a significant advantage of these substitutions for Vero cell-based, but a significant disadvantage for MDCK cell-based virus replication. In contrast the S287R substitution, which was also detected for the NIBSC-strain, was first detected in passage 6 (RKI: 17 %, NIBSC: 86 %) and further persisted until passage 11 (RKI: 54 %, NIBSC: 77 %) finally representing the major virus population in both strains. For segment 5 of the NIBSC strain in passage 11 one additional substitution (N397K, region involved in NP-binding as well as PB2-binding) to the ones also observed for the RKI-strain (S287R and E294V) was detected. Naffakh *et al.* hypothesized a co-evolution of PB2, PA and NP due to

strong physical as well as functional interactions [202]. This is in agreement with the occurrence of substitutions within the PB2 and NP molecules observed during the adaptation processes of the RKI-strain.

5.3.6 Quasispecies of segment 6

For segment 6 coding for NA one subpopulation (I7M) was generated during virus adaptation processes of the RKI- as well as the NIBSC-strain. This I7M subpopulation replicated rather well in Vero as well as in MDCK cells (RKI/NIBSC: 0 %/0 % passage 1; 15 %/85 % passage 6; 85 %/65 % passage 11). Position 7 is located in the transmembrane domain (7 - 35) of the NA molecule and was described as a potential signal-anchor for type-II membrane proteins [158, 226]. In contrast to the RKI-strain, further substitutions were detected for the NIBSC-strain in passage 11. These comprised the substitutions I302K, E314K, S349N, N434D and S451T, which are all located within the head of the neuraminidase protein [204]. The comparison of sequencing data from Fields *et al.* with data derived from the NIAID influenza genome sequencing project reveals the E314K and S451T sequence conflict [205, 227], which may suggest these residue substitutions as frequently occurring variants. Interestingly, the substitutions I302V, E314K, N434D and S451T result in matching AA sequences at these positions for the RKI- and NIBSC-strain.

5.3.7 Quasispecies of segment 7

With respect to segment 7 coding for M1 and M2 proteins the RKI-strain stayed homogeneous throughout the virus adaptation processes, i.e. no additional subpopulation was detected. This is consistent with the persistence of the rather broad quasispecies of the NIBSC-strain from passage 1. Only the subpopulation Q72R decreased to under the detection limit in passages 6 and 11. These findings suggest that the selection pressure during virus adaptation from MDCK to Vero cells and vice versa hardly acts on the two matrix proteins M1 and M2.

5.3.8 Quasispecies of segment 8

Differential splicing results in two proteins encoded by segment 8 namely NS1 and NEP (NS2). During virus adaptations different substitutions within the NS1 protein have been detected. Interestingly, after virus adaptation from IFN-competent MDCK to IFN-deficient Vero cells (passage 6) approximately 100 % of the viruses of the RKI-strain carried the substitution S103P. Only a few sequence reads could be detected carrying the original sequence. This substitution persisted and made up for 100 % of the virus population in passage 11 after back-adaptation to MDCK cells. This clearly suggests a fitness gain by S103P for replication in Vero and definitively no fitness loss for replication in MDCK cells. In contrast, after the adaptation of the NIBSC-strain to Vero cells, a subpopulation P107H was detected, which decreased again during back-adaptation to MDCK cells. Altogether, this suggests a fitness gain for Vero cell-based virus replication but a fitness loss for MDCK cell-based replication. Other substitutions, namely D101N, G183R, G184E and D189N were detected with high confidence after back-adaptation of the NIBSC-strain to MDCK cells. The substitutions S103P (RKI-strain), P107H and D101N (NIBSC-strain) are located in the effector domain of the NS1 protein within the region involved in interaction with eukaryotic initiation factor 4 G1 (eIF4G1) as well as the cleavage and polyadenylation specificity factor (CPSF), which is involved in the 3'-polyadenylation of cellular mRNAs [204]. So far, it is assumed that the ability of an NS1 molecule to interact with eIF4G1 enhances viral mRNA translation [228, 229], whereas the ability to interact with CPSE-30 downregulates host gene expression including IFN expression [230, 231]. Recently, Forbes *et al.* demonstrated that the mutation F103L in this region of mouse-adapted IVA-HongKong/1/1968 (H3N2) bound the F2F3 domain of CPSF with significantly decreased affinities [232]. Furthermore, it was reported that, a hydrophilic S at position 103 of the IVA-PR8 eliminates CPSF30 binding affinity, but an unknown viral strategy maintained suppression of IFN-β mRNA production [231, 233]. Structural analysis indicated that the aromatic side chain of residue 103 interacts with hydrophobic residues of the

F2F3 domain of CPSF and contributes to the stability of tetrameric NS1A-F2F3 complex [233]. Furthermore, the F103L mutation was shown to increase significantly virus yields in mouse kidney epithelium cells as well as to increase early viral growth (12-24 hpi) in IFN-β primed cells in contrast to *wt* viruses [232]. Interestingly, not only the antagonism to IFN-β was enhanced for the F103L mutant but it also induced significantly lower IFN-β expression 1 dpi in the lungs of mice [232]. Since MDCK cells are IFN-competent whereas Vero cells are INF-deficient, it may be assumed that the virus adapted according to IFN competence of the host cells. However, this is unlikely since recent studies by Seitz *et al.* demonstrated that IFN expression is no limiting factor for IVA replication in MDCK cells [195]. Rather IFN-expression independent functions contribute to NS1 mutations during the adaptation processes. The appearance and further increase of the G183R, G184E and D189N virus variants during the NIBSC-strain adaptation may indicate a fitness gain by tuning CPSF interactions. Residues 183 and 184 are located in a domain (144 - 188) interacting with CPSF, whereas residue 189 is located just outside this domain [204]. Structural analysis demonstrated that residues 183/184 belong to the largely hydrophobic F2F3 binding pocket within the NS1 molecule interacting with the F3 zinc finger of the CPSF complex contributing to CPSF-binding [233]. A strong influence of G184 by an unknown mechanism on IVA-PR8 virulence was described, which is independent from the IFN system [234]. However, residues F103 and M106 have been identified as critical for CPSF binding. Some viruses, e.g. IVA-PR8, vary in these positions. As a result, attenuation or even complete loss of CPSF-binding was described [231, 235, 236].

Altogether, the emergence of NS1 variants with all substitutions located in the elF4G1- (AA81 - 113) and/or the CPSF- (AA81 -113, AA147 - 188) binding site or nearby (AA189) suggests a fitness gain for all variants in Vero cells and for most variants also in MDCK cells, whenever a further increase during back-adaptation to MDCK cells was observed. Since Vero cells are IFN-deficient and even in MDCK cells IFN is not a limiting factor for virus replication another, IFN-independent function probably drives NS1 adaptation from MDCK to Vero cells

and back. Such a factor may be involved in (i) the suppression of RNAi induction [237-240], (ii) the inhibition of host cell mRNA processing by interacting with CPSF and the suppression of nuclear export of poly-A-tailed transcripts [233, 241-244], (iii) suppressing apoptosis in infected cells [245, 246] or (iv) the stimulation of the translation of viral transcripts by interacting with eIF4G1 [229, 247].

The fact that 100 % of detected viruses of the RKI-strain carried the S103P substitution in passages 6 and 11 suggests that either a mutation was essential for sufficient virus replication in Vero cells or that this variant had an enormous fitness-gain, having outcompeted other virus variants by passage 6. The latter is more likely, since after the adaptation by Genzel *et al.* from MDCK to Vero cells no NS1 variants were detected. Consistent with this is the accumulation of some substitutions in similar regions during the adaptation of the NIBSC strain. Whether the detected substitutions modify CPSF30-, eIF4G1-binding or other, so far unknown interactions located in this region, remains to be further investigated.

Regarding the NEP protein, no AA substitutions were detected during both RKI-strain adaptations, while after back-adaptation of the NIBSC-strain to MDCK cells two subpopulations carrying the substitutions D27N and M31I were detected. To the author's knowledge, so far, no specific function was mapped to this region.

5.3.9 General remarks

In general, next-generation pyrosequencing was performed to address whether sequence changes of the viral genome were required to ensure efficient virus replication in the new host system (section 4.6.3). It should be mentioned that on the basis of obtained sequencing data, it was difficult to decide whether the detected frameshifts are sequencing artifacts or represent non-infectious and/or truncated virus particles. Further sequencing by another method e.g. Illumina sequencing would be necessary to verify this. Hence, in the previous, only sense-mutations were discussed, whereas introduced stop codons and nonsense-mutations such as frameshifts were ignored. Moreover, several silent substitutions were detected. Some may have occurred by chance, others may

have been selected to better match the new host's codon usage bias, i.e. favoured base triplets coding for a specific AA [248]. However, in the previous, the focus was laid on substitutions changing the AA sequence.

5.4 Impact of cultivation scale, vessel and virus production media on the HA N-glycosylation pattern

So far, the host cell line, the host cell line's adaptation status to adherent or serum-free suspension growth and to some extend the virus strain have been shown to have an impact on the HA *N*-glycan fingerprint with respect to *N*-glycan structure presence as well as relative structure abundances. In contrast, different harvest time points, different virus passages, different IVA-PR8 virus suppliers as well as different virus adaptation statuses only slightly impacted HA *N*-glycan fingerprints, i.e. only affecting *N*-glycan structure abundances. In a next step, for robustness during scale-up, the impact of different cultivation vessels (T75-flasks up to microcarrier-based 1L-STR) on the HA *N*-glycosylation pattern was assessed. Again, the use of different cultivation vessels had no significant impact on *N*-glycan structure presence in HA *N*-glycan fingerprints. Only relative *N*-glycan structure abundances were affected, (RKI-strain: maximal lΔRPHI of 7.3 % for peak 2; IVA-Uruguay: maximal lΔRPHI of 6.6 % for peak 6, table 3). Nevertheless, one possible cause for such variations can be found in the different concentrations of dissolved oxygen (DO) in T-flasks, roller bottles and DO-/pH-controlled bioreactors. For example, Kunkel *et al.* observed different galactosylation levels for different DO concentrations for monoclonal antibodies [249, 250], and Restelli *et al.* described DO concentration-dependent fucosylation efficiencies for recombinant human erythropoietin [251]. Beside DO concentration, also time course of pH differs between pH-controlled STR cultivations, T-flasks and roller bottles (both uncontrolled). Variations of the extracellular pH in the range of 6.1 to 8.7 have been described, e.g. to influence significantly the glycosylation pattern of CHO-derived recombinant mouse placental lactogen-I. At low and high pH values a decreased extent of glycosylation was found [252]. Furthermore, Zanghi *et al.* showed that changes in

pH as well as partial pressures of CO_2 can alter protein glycosylation in CHO cells [253]. Other possible causes for variations in RPH are different shear forces in T-flasks, roller bottles and STRs. For example, Senger *et al.* reported shear-dependent ratios of partially and fully glycosylated recombinant tissue-type plasminogen activator protein [254]. Besides, growth in suspension or in microcarrier culture can have an impact. Nam *et al.* observed decreased fucosylation and increased sialylation of the recombinant model glycoprotein secreted human placental alkaline phosphatase produced in CHO microcarrier in contrast to suspension culture [6, 255].

The choice of virus production medium only slightly affected relative peak abundances (media composition: maximal IΔRPHI = 5.7 %; trypsin activity: maximal IΔRPHI = 14.0 %, table 3) of the HA *N*-glycan fingerprints. As addressed before, nutrient concentrations in the medium and accumulation of metabolic by-products may cause such minor variations of the RPH [6].

In summary, except for the choice of host cells, the adaptation of host cells to serum-free suspension growth and potentially the selection of virus strain, all other investigated process modifications resulted only in minor differences regarding relative *N*-glycan structure abundances. This brings up questions concerning the possible impact of such variations on the quality of antigens, i.e. immunogenicity or efficacy. Recently, de Vries *et al.* (2012) investigated the impact of *N*-glycosylation on the immunogenicity of recombinant HA, showing that HA antigens carrying terminal mannose residues induced significantly lower HAI Ab titers than HA modified by complex glycan structures or single *N*-acetylglucosamine side chains. However, using a HA_1 antigen microarray they demonstrated a comparable breadth of Ab response for all tested recombinant HA glycovariants [3, 6].

5.5 Impact of HA N-glycosylation on immunogenicity

The following section 5.5 contains analogous text content and structure taken from a paper published together with J. Hütter, D. Höper, P.H. Seeberger, E. Rapp and B. Lepenies, January 2013 in *J. Immunol*. [1]. As described before

(section 1, last two paragraphs), scentences and/or paragraphs containing quotations are not indicated specifically. The reference will only be given after the phrase or paragraph by the number of the quoted reference.

To investigate the impact of *N*-glycosylation on immunogenicity of virus preparations *in vitro* as well as *in vivo*, mouse studies were performed in cooperation with Dr. B. Lepenies and J. Hütter (MPI-KG, Potsdam-Golm, Germany). Within these studies it was demonstrated that differences in HA *N*-glycosylation have a marked impact on immunogenicity *in vitro* as well as *in vivo* [1]. Studies were performed with β-propiolactone inactivated MDCK and Vero cell-derived as well as natively deglycosylated virus preparations. MDCK and Vero cells were selected for virus production due to their industrial relevance [1].

Whole spleen cell assays demonstrated higher $CD69^+$ frequencies of $CD4^+$ TCR-HA transgenic T cells as well as increased secretion of IL-2 after stimulation with Vero cell-derived virus. These higher levels of CD69 and IL-2 indicated faster T cell activation after stimulation with the Vero than with the MDCK cell-derived virus. IL-2 is generally produced very early upon Ag stimulation of Th0 cells, promoting division, differentiation and survival of T cells, but may also be produced by Th1 cells [1, 256]. Interestingly, secretion levels of IFNγ were less affected whereas secretion levels of IL-4 were hardly affected at all. Comparable Th1/Th2 effector cytokine levels (IFNγ, IL-4) for both glycovariants and higher IL-2 levels for Vero cell-derived virus suggest faster recognition and Ag uptake by APC of the Vero cell-derived virus preparation. However, a tendency towards higher IFNγ-secretion upon splenocyte stimulation with the Vero cell-derived virus was observed, although mostly not significant. IFNγ is predominantly produced by stimulated Th1 cells, which are essential for viral clearance [1, 257].

The differences observed in T cell activation were predominantly mediated by $CD11c^+$ DCs as shown in a DC/T cell co-cultivation assay by Hütter *et al.* [1]. Typically DCs, which are part of the innate immunity, take-up the Ag in the periphery, at the site of infection and transport it to the regional lymph node or the spleen, where they present specific peptides of the processed Ag, and activate specific $CD4^+$ and $CD8^+$ T cells (via cross-presentation) by multiple

interactions (see section 2.11 and [258]). Usually $CD4^+$ T cells induce optimal T cell activation as well as stimulate Ab-producing B plasma cells [259]. However, in this study, DCs were demonstrated being crucial for distinct T cell activation by means of $CD69^+$ expression (MDCK *vs.* Vero cell-derived virus) or IL-2, IL-4 and IFNγ secretion (glycosylated *vs.* deglycosylated virus preparations, [1]). So far, different DC subsets have been described, being involved in IVA infections. These may cause differential T cell activation induced by MDCK and the Vero cell-derived virus preparations. For instance $CD103^+$ $CD11b^{low/neg}$ DCs efficiently transport Ag to posterior mediastinal lymph nodes, efficiently load viral peptides onto MHCI complexes as well as efficiently present them to $CD8^+$ T cells [260]. In contrast, although $CD103^+$ $CD11b^{high}$ DCs take up Ag even more efficiently, they mainly remain in the lung tissue, where they produce proinflammatory cytokines [260]. Another subset is represented by $CD8\alpha^+$ DCs, which play a crucial role for CTL priming during various infections including IVA infections [261]. However, these data may be mouse-specific and not relevant for human vaccine design. In a recent human challenge study preexisting $CD4^+$ T cells, but not $CD8^+$ T cells correlated with protection and virus clear before Ab responses were stimulated [262], thus, highlighting the need for further experiments in human DC and T cell models [1].

Furthermore, DCs express pathogen pattern recognition receptors, in particular C-type lectin receptors (CLRs), such as the macrophage mannose receptor (MMR, [263]). The MMR was recently shown to be involved in macrophage infection by IVA [264]. In our study the high mannose *N*-glycan structures attached to the Vero, but not to the MDCK cell-derived virus may be specifically recognized by such CLRs on splenic $CD11c^+$ DCs and therefore lead to higher immunogenicity *in vitro* with regard to CD69 expression on $CD4^+$ T cells and IL-2 secretion. Other receptors identified to interact with glycan moieties of IVA include macrophage galactose-type lectin (MGL, [265]) and dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN, [266]), allowing sialic acid independent virus entry. Also soluble expressed lectins were described to interact with glycan moieties of IVA, such as surfactant protein D

(SP-D), which is suggested to contribute to viral clearance via IVA-aggregation (reviewed in [267]). Another potential cause of differential immunogenicity of differentially glycosylated virus preparations observed in this study may also be attributed to masking of antigenic epitopes by *N*-glycan side chains [1, 268, 269].

However, in most studies so far discussed only live influenza viruses were used. In order to exclude virus replication in mouse cells and resulting changes in *N*-glycosylation of virus preparations, β-propiolactone inactivated virus preparations were used in this study. Moreover, β-propiolactone treatment was demonstrated an ideal IVA-PR8 inactivation method, hardly affecting HA *N*-glycosylation at all (maximal IΔRPHI = 0.8, [6]). Thus, solely the impact of HA *N*-glycosylation on immunogenicity could be addressed [1].

Furthermore, in this study virus preparations of MDCK as well as Vero cell-derived virus were natively deglycosylated. T cell activation was dramatically reduced after deglycosylation of both glycovariants with respect to CD69 expression of $CD4^+$ TCR-HA transgenic T cells as well as to IL-2, IL-4 and IFNγ secretion. Interestingly, the decrease of T cell activation after deglycosylation was more pronounced for the Vero than for the MDCK cell-derived virus. This may be explained by differential deglycosylation results: deglycosylation efficiency of at least about 90 % was achieved for the MDCK as well as the Vero cell-derived virus preparation as indicated by signal intensities of the HA *N*-glycan fingerprints. However, while the remaining HA *N*-glycans on the deglycosylated MDCK cell-derived virus still represented predominantly MDCK cell-specific structures, the HA *N*-glycans detected for the deglycosylated Vero cell-derived virus also comprised various new, most probably truncated *N*-glycan structures with reduced migration times. Altogether, the marked impact of removed glycan moieties, e.g. mannose and β-galactose, on T cell activation, confirms the dramatic impact of *N*-glycosylation on immunogenicity *in vitro* [1].

This is in agreement with obtained *in vivo* data, demonstrating significantly reduced HA-specific Ab levels for the deglycosylated virus variants in contrast to the glycosylated ones. Thus, HA *N*-glycosylation markedly impacts the B cell-mediated, humoral immunity. Interestingly, 14 days after prime immunization HA-

specific Ab levels induced by the MDCK cell-derived virus were significantly higher than levels induced by the Vero cell-derived virus. HAI-titers in mouse sera after prime-boost immunization with the MDCK cell-derived virus were still markedly higher at 28 days than HAI-titers induced by the Vero cell-derived virus. This is consistent with findings by de Vries *et al.*, demonstrating higher HAI activities for HA carrying complex *N*-glycans or single GlcNAc residues than for HA carrying high mannose structures [3]. In agreement with that, Lin *et al.* showed very recently that complex terminally sialated and asialyated-galactose type *N*-glycans induced higher-quality antibodies (i.e. higher neutralizing titers) after immunization in mice than pauci-mannose or high-mannose type *N*-glycans did [2]. Interestingly, Lin et al. observed higher total HA-specific Ab titers and stronger T cell responses for HA carrying pauci-mannose or high-mannose type *N*-glycans [2]. The second is consistent with our findings of stronger T cell activation after immunization with Vero (complex β-galactose and high-mannose type *N*-glycans) than with MDCK cell-derived virus (complex α- and β-galactose type *N*-glycans). Altogether these results indicated differential impact of HA *N*-glycosylation on B cell and T cell mediated immune responses [1].

Furthermore, TCR-HA splenic T cells were purified, labeled and adoptively transferred into *wt* mice. Although splenic T cell proliferation upon immunization was comparable for the MDCK and the Vero cell-derived virus, the frequency of IL-2 producing, transferred T cells was higher upon immunization with the Vero cell-derived virus. This is consistent with the *in vitro* data indicating a faster T cell activation upon immunization with Vero cell-derived virus. On the other hand the MDCK cell-derived virus induced higher HA-specific Ab titers 14 days after prime immunization. This clearly indicates that MDCK cell-specific HA *N*-glycosylation triggers a pronounced humoral immune response [1].

In summary, these data suggest that Vero cell-specific HA *N*-glycosylation stronger promotes cellular immunity whereas MDCK cell-specific HA *N*-glycosylation stronger promotes humoral immunity. Although a dramatic effect of HA *N*-glycosylation on immunogenicity was demonstrated *in vitro* as well as *in vivo*, further investigations are required to address the impact of HA *N*-

glycosylation on vaccine’s efficacy and safety and to finally identify critical quality attributes for animal and human vaccines [1].

6 Conclusion and outlook

6.1 Impact of cultivation conditions on HA N-glycosylation

In this study, multiple process conditions were investigated for their impact on HA *N*-glycosylation. Thereby, most tested process modifications resulted only in minor differences in HA *N*-glycan fingerprints regarding relative *N*-glycan structure abundances. Hence, HA *N*-glycosylation seems a rather robust protein modification, which is only slightly affected with respect to RPH for most process variations tested including variation of virus passages, virus seed suppliers, virus adaptation statuses, composition as well as trypsin activity of virus production media, virus production vessels and time points of virus harvest. Moreover, Rödig *et al.* demonstrated that also different virus production temperatures (33 °C, 37 °C, 39 °C) only slightly impacted relative *N*-glycan structure abundances and interestingly final β-propiolactone inactivation did basically not change HA *N*-glycan fingerprints at all, making it an ideal inactivation method for manufacturing influenza vaccines [6]. However, this study identified a few cultivation conditions, which affect HA *N*-glycosylation significantly with respect to *N*-glycan structure presence as well as relative *N*-glycan structure abundances. These included the use of different host cell lines, the adaptation status of the host cell line to serum-requiring adherent or serum-free suspension growth and potentially the selected virus strain.

Overall, these results demonstrate that the HA *N*-glycosylation pattern of IVA is remarkably stable, regarding changes in the production process. In particular, small variations in a production process are unlikely to change the HA *N*-glycosylation dramatically and therefore, it can be assumed that wanted and unwanted process variations in influenza vaccine manufacturing have only a minor or no impact on product quality. However, the impact of changes in glycan profiles, i.e. *N*-glycan structure presence as well as relative *N*-glycan structure abundance, on properties of live and dead vaccines, e.g. safety and efficacy as well as its mechanisms should be further characterized in more detail [4, 6].

However, other factors may impact HA *N*-glycosylation and remain to be investigated. These include the host cell status (cell cycle: G1,G0, etc. as well as physiological cell status: virus produced in different time slots: e.g. 0-12 hpi, 12-24 hpi, 24-48 hpi, …), cell density, multiplicity of infection for the virus production phase, other media compositions and media supplements, methods for virus harvest (e.g. gradient step centrifugation, filtration), virus inactivation by formalin and Triton X-100, methods for virus purification and concentration (e.g. sucrose density gradient, affinity and pseudo-affinity chromatography, hydrophobic interaction chromatography, SEC, ion-exchange chromatography), detergent treatment for HA/NA isolation as used for split and subunit vaccines, benzonase treatment for DNA reduction and the dilution in formulation buffer. Moreover, the impact of cultivation conditions on NA *N*-glycosylation should be investigated. Furthermore, *N*-glycan structure elucidation is of special importance for further data interpretation and project development and could be achieved by parallel MS/MS- and CGE-LIF-based approaches.

In conclusion, monitoring HA *N*-glycosylation patterns during vaccine production processes allows not only to investigate the impact of process modifications on antigen quality, but also offers a sensitive tool to evaluate consequences of unwanted process variations or process failures [4].

6.2 Fitness gain by virus adaptation and identification of key mutations

Taken together, this study demonstrated that the adaptation of virus seeds to new host cell lines was often necessary for sufficient high virus titers. However, the degree of titer enhancement by virus adaptation depends on host cell characteristics as well as on viruses. For two IVA-PR8 strains significantly increased final titers within shorter cultivation times were achieved by virus adaptation from MDCK cells to Vero cells. Interestingly, for all virus strains investigated, the HA *N*-glycan fingerprint stabilized latest with the second passage in a new cell line. The faster release of virus particles and the further increase in HA-titers, which are mostly observed during the first passages in the new host indicate that other factors are as well involved in the adaptation

process, i.e. changes in the sequence of the viral genome resulting in improved fitness of virus subpopulations [4].

Amino acid substitutions within the stem region of HA and/or within the NP- and/or PB2-binding domain of NP and/or the transmembrane region of NA and/or the CPSF30- and/or eIF4G1-binding domain of NS1 rescued the virus population and ensured efficient virus replication in the new host [4]. The fitness of these adapted virus variants during backward adaptation to MDCK cells varies – some variants as the K460E (RKI) replicate well and infect either cell line while others grow only poorly in one cell line (e.g. S457L (RKI) in MDCK cells, [4]). This fitness gain of adapted viruses was also shown in the direct comparison of non-adapted virus seeds, Vero cell-adapted viruses as well as Vero cell-adapted and back-adapted viruses. The non-adapted viruses hardly replicated during the first days after infection, whereas Vero cell-adapted as well as back-adapted viruses replicated well from the very beginning. The improved virus fitness of adapted viruses is consistent with the sequencing data, since several substitutions occurring after adaptation to Vero cells, persisted during back-adaptation to MDCK cells. This persistence of substitutions strongly indicates no or only small disadvantages for replication in MDCK cells. Moreover, quasispecies analysis during virus adaptation of MDCK cell-derived virus seeds identified potential key mutations allowing sufficient virus replication in Vero cells. In particular mutations within the stem region of the HA molecule, tuning the HA pH-stability, seemed to be necessary to adapt to Vero cells' endosomal pH. Furthermore, substitutions within the NP- and PB2-binding sites of the NP molecule may significantly contribute to Vero cell adaptation of different IVA-PR8 strains. For the NA molecule the N-terminal transmembrane region was identified as potentially contributing to successful virus adaptation. Last but not least, substitutions within the CPSF30- and/or the eIF4G1-binding domains of the NS1 molecule were identified as potential key factors for efficient virus replication in the new host cell line. In a next step, each substitution's contribution to successful virus adaptation should be characterized in more detail using reverse genetics. In general, due to the time frame required for efficient virus replication in the new host (RKI- and

NIBSC-strain adaptation from MDCK to Vero cells and back; passage 1; HA-titer increase not before 216 hpi and 168 hpi, respectively), we can certainly assume that most substitutions detected during virus adaptation were caused by mutations which were then selected for, rather than solely selection events.

6.3 Impact of HA N-glycosylation on immunogenicity

It was demonstrated that the Vero cell-specific HA *N*-glycosylation shows a bias towards the cellular immune response, whereas MDCK cell-specific HA *N*-glycosylation more promotes the humoral immune response. In order to identify key characteristics for optimal HA *N*-glycosylation other IVA-PR8 glycovariants, including egg-, AGE1.CR/.pIX-, MDCK.SUS2-derived virus preparations should be characterized with respect to their immunogenicity using *in vitro* and *in vivo* approaches and (transgenic) mouse or other animal models (e.g. ferret, pig). Moreover, screening of virus preparations derived from glycosylation-deficient cell lines such as HEK293S and CHO 15B cells (both deficient in *N*-acetylglucosaminyltransferase I) would complement this approach. However, glycosylation-deficient cell lines should be selected with care, since some have been described as not very susceptible to IVA infections [270, 271]. Another approach is the generation of (conditional) knock-out cell lines as well as knock-in cell lines. This would allow to directly design optimal product glycosylation, i.e. HA *N*-glycosylation for optimal vaccine efficacy [272-274]. Moreover, the mechanisms by which the Vero cell-derived virus shows a bias towards cellular immunity and the MDCK cell-derived virus towards humoral immunity should further be investigated. Therefore, *in vitro* studies using TCR-HA transgenic and *wt* mouse models would allow focusing on MHCI- and MHCII-mediated responses, whereas *in vivo* studies using *wt* models would address the complete range of immune responses. Within such studies, CD40 expression would identify levels of mature/immature DCs, its expression also indicates cell adhesion, cell proliferation and signal transduction in B cells [275]. Investigating the expression of the co-stimulatory ligands B7.1 (CD80) and B7.2 (CD86) and its interaction with CD28 would further characterize the mechanisms how

glycosylation impacts T-, B cell and DC activation, co-stimulation and immunoregulation. Monitoring CD74 expression would give more insights regarding the regulation of T- and B-cell interactions, their development, activation, growth and motility. Furthermore, the expression of MHCI and MHCII levels on DCs would give more insights into DC Ag processing and presentation. Also, T cells ($CD3^+$) and B cells ($CD19^+$, $B220^+$) from other lymphatic organs than the spleen should be characterized for CD4, CD8, CD69 expression and CD22 (B cell activation marker) respectively. However, especially memory CD8 T cells and CD4 T cells have been described to mediate heterosubtypic immunity, i.e. protection of severe disease by prior infection with a virus of a different subtype [276, 277]. Therefore, the ability of the different virus variants (in particular MDCK and Vero cell-derived virus) to cross-protect should be investigated for example in challenge studies. The most robust protection may be provided by simultaneous induction of multiple immune pathways of the cellular as well as the humoral immunity. Thus optimal HA *N*-glycosylation for B cell-, $CD4^+$ and $CD8^+$ T cell-mediated immune responses should be identified with respect to heterosubtypic (cross-protection) and homotypic (no cross protection between different virus subtypes) immunity. Moreover, the relevance of HA *N*-glycosylation in humans should be addressed in more detail, using human DCs and peripheral blood mononuclear cell (PBMCs). Finally, the role of *N*-glycosylation could be evaluated by specifically designed and synthesized glycopeptides (most advisable in cooperation with experts in the field). Altogether, this would finally allow the definition of ideal cultivation conditions for influenza virus vaccine production, the development of an optimally glyco-engineered production cell line, improved vaccine formulations substituted with specific glycan structures and/or directed chemical *N*-glycan synthesis onto viral peptides.

7 References

1. Hütter J, Rödig JV, Höper D, Reichl U, Seehofer PH, Rapp E, Lepenies B: **Toward animal cell culture-based influenza vaccine design: viral hemagglutinin N-glycosylation markedly impacts immunogenicity**. *J Immunol* 2013, **190**(1):220-230.
2. Lin SC, Jan JT, Dionne B, Butler M, Huang MH, Wu CY, Wong CH, Wu SC: **Different immunity elicited by recombinant H5N1 hemagglutinin proteins containing pauci-mannose, high-mannose, or complex type N-glycans**. *PLoS One* 2013, **8**(6):e66719.
3. de Vries RP, Smit CH, de Bruin E, Rigter A, de Vries E, Cornelissen LA, Eggink D, Chung NP, Moore JP, Sanders RW *et al*: **Glycan-dependent immunogenicity of recombinant soluble trimeric hemagglutinin**. *J Virol* 2012, **86**(21):11735-11744.
4. Roedig JV, Rapp E, Höper D, Genzel Y, Reichl U: **Impact of Host Cell Line Adaptation on Quasispecies Composition and Glycosylation of Influenza A Virus Hemagglutinin**. *PLos ONE* 2011, **6(12)**.
5. Roedig JV, Rapp E, Genzel Y, Reichl U: **Impact of different influenza cultivation conditions on HA N-Glycosylation**. *BMC proceedings* 2011, **5 Suppl 8**:P113.
6. Rödig JV, Rapp E, J. B, Kampe M, Kaffka H, Bock A, Genzel Y, Reichl U: **Impact of cultivation conditions on N-glycosylation of influenza virus A hemagglutinin produced in MDCK cell culture**. *Biotechnology and Bioengineering* 2013, **110**(6):1691-1703.
7. Rödig J, Rapp E, Djeljadini S, Lohr V, Genzel Y, Jordan I, Sandig V, Reichl U: **Impact of Influenza Virus Adaptation status on HA N-Glycosylation Patterns in Cell Culture-Based Vaccine Production**. *Journal of Carbohydrate Chemistry* 2011, **30**:281-290.
8. Genzel Y, Rödig JV, Rapp E, Reichl U: **Vaccine production - upstream processing with adherent or suspension cell lines**. In: *Animal cell biotechnology - methods and protocols.* Edited by Pörtner R, vol. 25: Humana Press; 3rd ed. 2013.
9. Genzel Y, Behrendt I, Rödig J, Rapp E, Kueppers C, Kochanek S, Schiedner G, Reichl U: **CAP, a new human suspension cell line for influenza virus production**. *Appl Microbiol Biotechnol* 2012.
10. Cox NJ, Neumann G, Donis RO, Kawaoka Y: **Orthomyxoviruses: Influenza**. In: *Topley & Wilson's Microbiology and Microbial Infections.* John Wiley & Sons, Ltd; 2010: 634-698.
11. WHO: **A revision of the system of nomenclature for influenza viruses: a WHO memorandum**. *Bull World Health Organ* 1980, **58**(4):585-591.
12. Lamb RA, Krug RM: **Orthomyxooviridae: The Viruses and Their Replication**. In: *Fields Virology.* Edited by Knipe DM, Howley PM, vol. 1, 4th edn. Philadelphia: LIPPINCOTT WILLIAMS & WILKINS; 2001: 1487-1531.
13. Chen W, Calvo PA, Malide D, Gibbs J, Schubert U, Bacik I, Basta S, O'Neill R, Schickli J, Palese P *et al*: **A novel influenza A virus**

mitochondrial protein that induces cell death. *Nat Med* 2001, **7**(12):1306-1312.

14. Wise HM, Foeglein A, Sun J, Dalton RM, Patel S, Howard W, Anderson EC, Barclay WS, Digard P: **A complicated message: Identification of a novel PB1-related protein translated from influenza A virus segment 2 mRNA**. *J Virol* 2009, **83**(16):8021-8031.
15. Wise HM, Hutchinson EC, Jagger BW, Stuart AD, Kang ZH, Robb N, Schwartzman LM, Kash JC, Fodor E, Firth AE *et al*: **Identification of a novel splice variant form of the influenza A virus M2 ion channel with an antigenically distinct ectodomain**. *PLoS Pathog* 2012, **8**(11):e1002998.
16. **Influenza Type A Viruses and Subtypes** [http://www.cdc.gov/flu/avianflu/influenza-a-virus-subtypes.htm]
17. **Influenzavirus A** [http://viralzone.expasy.org/all_by_species/6.html]
18. Scholtissek C: **Molecular evolution of influenza viruses**. *Virus Genes* 1995, **11**(2-3):209-215.
19. Steinhauer DA, Holland JJ: **Rapid evolution of RNA viruses**. *Annu Rev Microbiol* 1987, **41**:409-433.
20. Domingo E, Martin V, Perales C, Grande-Perez A, Garcia-Arriaza J, Arias A: **Viruses as quasispecies: biological implications**. *Curr Top Microbiol Immunol* 2006, **299**:51-82.
21. Lauring AS, Andino R: **Quasispecies theory and the behavior of RNA viruses**. *PLoS Pathog* 2010, **6**(7):e1001005.
22. Lin YP, Wharton SA, Martin J, Skehel JJ, Wiley DC, Steinhauer DA: **Adaptation of egg-grown and transfectant influenza viruses for growth in mammalian cells: selection of hemagglutinin mutants with elevated pH of membrane fusion**. *Virology* 1997, **233**(2):402-410.
23. Webster RG, Laver WG, Air GM, Schild GC: **Molecular mechanisms of variation in influenza viruses**. *Nature* 1982, **296**(5853):115-121.
24. Hensley SE, Das SR, Bailey AL, Schmidt LM, Hickman HD, Jayaraman A, Viswanathan K, Raman R, Sasisekharan R, Bennink JR *et al*: **Hemagglutinin Receptor Binding Avidity Drives Influenza A Virus Antigenic Drift**. *Science* 2009, **326**(5953):734-736.
25. Rogers GN, Paulson JC: **Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin**. *Virology* 1983, **127**(2):361-373.
26. Webster RG, Bean WJ, Gorman OT, Chambers TM, Kawaoka Y: **Evolution and ecology of influenza A viruses**. *Microbiol Rev* 1992, **56**(1):152-179.
27. Hirst GK: **Studies on the Mechanism of Adaptation of Influenza Virus to Mice**. *J Exp Med* 1947, **86**(5):357-366.
28. Domingo E, Baranowski E, Ruiz-Jarabo CM, Martin-Hernandez AM, Saiz JC, Escarmis C: **Quasispecies structure and persistence of RNA viruses**. *Emerg Infect Dis* 1998, **4**(4):521-527.

29. Domingo E, Sabo D, Taniguchi T, Weissmann C: **Nucleotide sequence heterogeneity of an RNA phage population**. *Cell* 1978, **13**(4):735-744.
30. Biebricher CK, Eigen M: **What is a quasispecies?** *Curr Top Microbiol Immunol* 2006, **299**:1-31.
31. Rudd PM, Dwek RA: **Glycosylation: heterogeneity and the 3D structure of proteins**. *Critical reviews in biochemistry and molecular biology* 1997, **32**(1):1-100.
32. Sidorenko Y, Reichl U: **Structured model of influenza virus replication in MDCK cells**. *Biotechnol Bioeng* 2004, **88**(1):1-14.
33. Heldt FS, Frensing T, Reichl U: **Modeling the intracellular dynamics of influenza virus replication to understand the control of viral RNA synthesis**. *J Virol* 2012, **86**(15):7806-7817.
34. Medina RA, Garcia-Sastre A: **Influenza A viruses: new research developments**. *Nat Rev Microbiol* 2011, **9**(8):590-603.
35. Klenk HD, Rott R, Orlich M, Blodorn J: **Activation of influenza A viruses by trypsin treatment**. *Virology* 1975, **68**(2):426-439.
36. Enami M, Luytjes W, Krystal M, Palese P: **Introduction of site-specific mutations into the genome of influenza virus**. *Proc Natl Acad Sci U S A* 1990, **87**(10):3802-3805.
37. Gubareva LV, Kaiser L, Hayden FG: **Influenza virus neuraminidase inhibitors**. *Lancet* 2000, **355**(9206):827-835.
38. Martin K, Helenius A: **Nuclear transport of influenza virus ribonucleoproteins: the viral matrix protein (M1) promotes export and inhibits import**. *Cell* 1991, **67**(1):117-130.
39. Boulo S, Akarsu H, Ruigrok RW, Baudin F: **Nuclear traffic of influenza virus proteins and ribonucleoprotein complexes**. *Virus Res* 2007, **124**(1-2):12-21.
40. Guilligay D, Tarendeau F, Resa-Infante P, Coloma R, Crepin T, Sehr P, Lewis J, Ruigrok RW, Ortin J, Hart DJ *et al*: **The structural basis for cap binding by influenza virus polymerase subunit PB2**. *Nat Struct Mol Biol* 2008, **15**(5):500-506.
41. Li ML, Rao P, Krug RM: **The active sites of the influenza cap-dependent endonuclease are on different polymerase subunits**. *EMBO J* 2001, **20**(8):2078-2086.
42. Ruigrok RW, Crepin T, Hart DJ, Cusack S: **Towards an atomic resolution understanding of the influenza virus replication machinery**. *Curr Opin Struct Biol* 2010, **20**(1):104-113.
43. Dias A, Bouvier D, Crepin T, McCarthy AA, Hart DJ, Baudin F, Cusack S, Ruigrok RW: **The cap-snatching endonuclease of influenza virus polymerase resides in the PA subunit**. *Nature* 2009, **458**(7240):914-918.
44. Yuan P, Bartlam M, Lou Z, Chen S, Zhou J, He X, Lv Z, Ge R, Li X, Deng T *et al*: **Crystal structure of an avian influenza polymerase PA(N) reveals an endonuclease active site**. *Nature* 2009, **458**(7240):909-913.

45. Biswas SK, Nayak DP: **Mutational analysis of the conserved motifs of influenza A virus polymerase basic protein 1**. *J Virol* 1994, **68**(3):1819-1826.
46. Poch O, Sauvaget I, Delarue M, Tordo N: **Identification of four conserved motifs among the RNA-dependent polymerase encoding elements**. *EMBO J* 1989, **8**(12):3867-3874.
47. Min JY, Krug RM: **The primary function of RNA binding by the influenza A virus NS1 protein in infected cells: Inhibiting the 2'-5' oligo (A) synthetase/RNase L pathway**. *Proc Natl Acad Sci U S A* 2006, **103**(18):7100-7105.
48. Wandinger-Ness A, Bennett MK, Antony C, Simons K: **Distinct transport vesicles mediate the delivery of plasma membrane proteins to the apical and basolateral domains of MDCK cells**. *J Cell Biol* 1990, **111**(3):987-1000.
49. Nayak DP, Hui EK, Barman S: **Assembly and budding of influenza virus**. *Virus Res* 2004, **106**(2):147-165.
50. Rossman JS, Lamb RA: **Influenza virus assembly and budding**. *Virology* 2011, **411**(2):229-236.
51. Palese P, Tobita K, Ueda M, Compans RW: **Characterization of temperature sensitive influenza virus mutants defective in neuraminidase**. *Virology* 1974, **61**(2):397-410.
52. Lin D, Lan J, Zhang Z: **Structure and function of the NS1 protein of influenza A virus**. *Acta biochimica et biophysica Sinica* 2007, **39**(3):155-162.
53. Grienke U, Schmidtke M, von Grafenstein S, Kirchmair J, Liedl KR, Rollinger JM: **Influenza neuraminidase: A druggable target for natural products**. *Nat Prod Rep* 2011.
54. Gerdil C: **The annual production cycle for influenza vaccine**. *Vaccine* 2003, **21**(16):1776-1779.
55. Robertson JS, Nicolson C, Harvey R, Johnson R, Major D, Guilfoyle K, Roseby S, Newman R, Collin R, Wallis C: **The development of vaccine viruses against pandemic A(H1N1) influenza**. *Vaccine* 2011, **29**(9):1836-1843.
56. **Pandemic influenza vaccine manufacturing process and timeline**
57. Huang DB, Wu JJ, Tyring SK: **A review of licensed viral vaccines, some of their safety concerns, and the advances in the development of investigational viral vaccines**. *J Infect* 2004, **49**(3):179-209.
58. **Complete List of Vaccines Licensed for Immunization and Distribution in the US**
59. Lambe T: **Novel viral vectored vaccines for the prevention of influenza**. *Mol Med* 2012.
60. Hickling J, D'Hondt E: **A review of production technologies for influenza virus vaccines, and their suitability for deployment in developing countries for influenza pandemic preparedness**. In*: 20. December 2004; Geneva, Switzerland*. http://who.int/vaccine_research/diseases/influenza/Flu_vacc_manuf_tech_

report.pdf: World Health Organization, Initiative for Vaccine Research 2006.
61. Gregersen JP, Schmitt HJ, Trusheim H, Broker M: **Safety of MDCK cell culture-based influenza vaccines**. *Future Microbiol* 2011, **6**(2):143-152.
62. Genzel Y, Reichl U: **Continuous cell lines as a production system for influenza vaccines**. *Expert Rev Vaccines* 2009, **8**(12):1681-1692.
63. Howard MK, Kistner O, Barrett PN: **Pre-clinical development of cell culture (Vero)-derived H5N1 pandemic vaccines**. *Biol Chem* 2008, **389**(5):569-577.
64. Kistner O, Barrett PN, Mundt W, Reiter M, Schober-Bendixen S, Dorner F: **Development of a mammalian cell (Vero) derived candidate influenza virus vaccine**. *Vaccine* 1998, **16**(9-10):960-968.
65. Kistner O, Howard MK, Spruth M, Wodal W, Bruhl P, Gerencer M, Crowe BA, Savidis-Dacho H, Livey I, Reiter M *et al*: **Cell culture (Vero) derived whole virus (H5N1) vaccine based on wild-type virus strain induces cross-protective immune responses**. *Vaccine* 2007, **25**(32):6028-6036.
66. **Vepacel**
67. Lohr V, Rath A, Genzel Y, Jordan I, Sandig V, Reichl U: **New avian suspension cell lines provide production of influenza virus and MVA in serum-free media: studies on growth, metabolism and virus propagation**. *Vaccine* 2009, **27**(36):4975-4982.
68. Pau MG, Ophorst C, Koldijk MH, Schouten G, Mehtali M, Uytdehaag F: **The human cell line PER.C6 provides a new manufacturing system for the production of influenza vaccines**. *Vaccine* 2001, **19**(17-19):2716-2721.
69. Cox RJ, Madhun AS, Hauge S, Sjursen H, Major D, Kuhne M, Hoschler K, Saville M, Vogel FR, Barclay W *et al*: **A phase I clinical trial of a PER.C6 cell grown influenza H7 virus vaccine**. *Vaccine* 2009, **27**(13):1889-1897.
70. Smith KA, Colvin CJ, Weber PS, Spatz SJ, Coussens PM: **High titer growth of human and avian influenza viruses in an immortalized chick embryo cell line without the need for exogenous proteases**. *Vaccine* 2008, **26**(29-30):3778-3782.
71. Seo SH, Goloubeva O, Webby R, Webster RG: **Characterization of a porcine lung epithelial cell line suitable for influenza virus studies**. *J Virol* 2001, **75**(19):9517-9525.
72. Genzel Y, Dietzsch C, Rapp E, Schwarzer J, Reichl U: **MDCK and Vero cells for influenza virus vaccine production: a one-to-one comparison up to lab-scale bioreactor cultivation**. *Appl Microbiol Biotechnol* 2010, **88**(2):461-475.
73. Schwarzer J, Rapp E, Hennig R, Genzel Y, Jordan I, Sandig V, Reichl U: **Glycan analysis in cell culture-based influenza vaccine production: influence of host cell line and virus strain on the glycosylation pattern of viral hemagglutinin**. *Vaccine* 2009, **27**(32):4325-4336.
74. Robertson JS, Bootman JS, Newman R, Oxford JS, Daniels RS, Webster RG, Schild GC: **Structural changes in the haemagglutinin which**

accompany egg adaptation of an influenza A(H1N1) virus. *Virology* 1987, **160**(1):31-37.

75. Schild GC, Oxford JS, de Jong JC, Webster RG: **Evidence for host-cell selection of influenza virus antigenic variants**. *Nature* 1983, **303**(5919):706-709.
76. Sandig V, Jordan, I: **Immortalized cell lines for virus production**. In.; 2005.
77. Brown SW, Mehtali M: **The Avian EB66(R) Cell Line, Application to Vaccines, and Therapeutic Protein Production**. *PDA J Pharm Sci Technol* 2010, **64**(5):419-425.
78. Nicholson K: **Report of meeting on the development of influenza vaccines with broad spectrum and long-lasting immune responses,, ,** . In*: 26–27 February 2004; Geneva, Switzerland*. http://www.who.int/vaccine_research/documents/Report_Feb_04.pdf: World Health Organization 2004: 26-36.
79. (WHO) WHO: **Recommendations for the production and control of influenza vaccine (inactivated)**. In: *Technical Report Series.* vol. No. 927. http://www.who.int/biologicals/publications/trs/areas/vaccines/influenza/ANNEX%203%20InfluenzaP99-134.pdf; 2005: 99-134.
80. Nakamura K, Bhown AS, Compans RW: **Glycosylation sites of influenza viral glycoproteins. Tryptic glycopeptides from the A/WSN (H0N1) hemagglutinin glycoprotein**. *Virology* 1980, **107**(1):208-221.
81. Keil W, Niemann H, Schwarz RT, Klenk HD: **Carbohydrates of influenza virus. V. Oligosaccharides attached to individual glycosylation sites of the hemagglutinin of fowl plague virus**. *Virology* 1984, **133**(1):77-91.
82. Teillaud JL: **Antibody dependent Cellular Cytotoxicity (ADCC)**. *eLS John Wiley & Sons Ltd, Chichester* 2012.
83. Abe Y, Takashita E, Sugawara K, Matsuzaki Y, Muraki Y, Hongo S: **Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin**. *Journal of virology* 2004, **78**(18):9605-9611.
84. Saito T, Nakaya Y, Suzuki T, Ito R, Saito T, Saito H, Takao S, Sahara K, Odagiri T, Murata T *et al*: **Antigenic alteration of influenza B virus associated with loss of a glycosylation site due to host-cell adaptation**. *Journal of Medical Virology* 2004, **74**(2):336-343.
85. Chen Z, Aspelund A, Jin H: **Stabilizing the glycosylation pattern of influenza B hemagglutinin following adaptation to growth in eggs**. *Vaccine* 2008, **26**(3):361-371.
86. Opitz L, Zimmermann A, Lehmann S, Genzel Y, Lubben H, Reichl U, Wolff MW: **Capture of cell culture-derived influenza virus by lectins: strain independent, but host cell dependent**. *J Virol Methods* 2008, **154**(1-2):61-68.
87. de Vries RP, de Vries E, Bosch BJ, de Groot RJ, Rottier PJM, de Haan CAM: **The influenza A virus hemagglutinin glycosylation state affects receptor-binding specificity**. *Virology* 2010, **403**(1):17-25.

88. Sun X, Jayaraman A, Maniprasad P, Raman R, Houser KV, Pappas C, Zeng H, Sasisekharan R, Katz JM, Tumpey TM: **N-Linked Glycosylation of the Hemagglutinin Protein Influences Virulence and Antigenicity of the 1918 Pandemic and Seasonal H1N1 Influenza A Viruses**. *J Virol* 2013, **87**(15):8756-8766.
89. Bohne-Lang A, von der Lieth CW: **GlyProt: in silico glycosylation of proteins**. *Nucleic Acids Res* 2005, **33**(Web Server issue):W214-219.
90. Helenius A, Aebi M: **Roles of N-linked glycans in the endoplasmic reticulum**. *Annu Rev Biochem* 2004, **73**:1019-1049.
91. Larkin A, Imperiali B: **The expanding horizons of asparagine-linked glycosylation**. *Biochemistry* 2011, **50**(21):4411-4426.
92. Maeda Y, Kinoshita T: **Dolichol-phosphate mannose synthase: structure, function and regulation**. *Biochim Biophys Acta* 2008, **1780**(6):861-868.
93. Burda P, Aebi M: **The ALG10 locus of Saccharomyces cerevisiae encodes the alpha-1,2 glucosyltransferase of the endoplasmic reticulum: the terminal glucose of the lipid-linked oligosaccharide is required for efficient N-linked glycosylation**. *Glycobiology* 1998, **8**(5):455-462.
94. Bause E: **Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes**. *Biochem J* 1983, **209**(2):331-336.
95. Imperiali B, Shannon KL, Unno M, Rickert KW: **A Mechanistic Proposal for Asparagine-Linked Glycosylation**. *J Am Chem Soc* 1992, **114**:7944-7945.
96. Gavel Y, von Heijne G: **Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering**. *Protein Eng* 1990, **3**(5):433-442.
97. Kelleher DJ, Karaoglu D, Mandon EC, Gilmore R: **Oligosaccharyltransferase isoforms that contain different catalytic STT3 subunits have distinct enzymatic properties**. *Molecular cell* 2003, **12**(1):101-111.
98. Ellgaard L, Helenius A: **Quality control in the endoplasmic reticulum**. *Nature reviews Molecular cell biology* 2003, **4**(3):181-191.
99. Helenius A, Aebi M: **Intracellular functions of N-linked glycans**. *Science* 2001, **291**(5512):2364-2369.
100. Basak S, Tomana M, Compans RW: **Sialic acid is incorporated into influenza hemagglutinin glycoproteins in the absence of viral neuraminidase**. *Virus Res* 1985, **2**(1):61-68.
101. Klenk HD, Compans RW, Choppin WP: **An electron microscopic study of the presence or absence of neuraminic acid in enveloped viruses**. *Virology* 1970, **42**(4):1158-1162.
102. **NetNGlyc 1.0 Server** [http://www.cbs.dtu.dk/services/NetNGlyc/]
103. Blom N, Sicheritz-Ponten T, Gupta R, Gammeltoft S, Brunak S: **Prediction of post-translational glycosylation and phosphorylation of**

proteins from the amino acid sequence. *Proteomics* 2004, **4**(6):1633-1649.

104. Wang CC, Chen JR, Tseng YC, Hsu CH, Hung YF, Chen SW, Chen CM, Khoo KH, Cheng TJ, Cheng YS *et al*: **Glycans on influenza hemagglutinin affect receptor binding and immune response**. *Proc Natl Acad Sci U S A* 2009, **106**(43):18137-18142.
105. Veits J, Weber S, Stech O, Breithaupt A, Graber M, Gohrbandt S, Bogs J, Hundt J, Teifke JP, Mettenleiter TC *et al*: **Avian influenza virus hemagglutinins H2, H4, H8, and H14 support a highly pathogenic phenotype**. *Proc Natl Acad Sci U S A* 2012, **109**(7):2579-2584.
106. Patton WF: **Fluorescence detection of glycoproteins in gels and on electroblots**. *Current protocols in cell biology / editorial board, Juan S Bonifacino [et al]* 2002, **Chapter 6**:Unit 6 8.
107. Sharon N, Lis H: **History of lectins: from hemagglutinins to biological recognition molecules**. *Glycobiology* 2004, **14**(11):53R-62R.
108. Pilobello KT, Krishnamoorthy L, Slawek D, Mahal LK: **Development of a lectin microarray for the rapid analysis of protein glycopatterns**. *Chembiochem : a European journal of chemical biology* 2005, **6**(6):985-989.
109. Mulloy B, Hart GW, Stanley P: **Structural Analysis of Glycans**. In: *Essentials of Glycobiology 2nd edition.* Edited by Varki A. CRD, Esko J.D., et al., editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2009.
110. Cataldi TR, Campa C, De Benedetto GE: **Carbohydrate analysis by high-performance anion-exchange chromatography with pulsed amperometric detection: the potential is still growing**. *Fresenius' journal of analytical chemistry* 2000, **368**(8):739-758.
111. Ruhaak LR, Zauner G, Huhn C, Bruggink C, Deelder AM, Wuhrer M: **Glycan labeling strategies and their use in identification and quantification**. *Analytical and bioanalytical chemistry* 2010, **397**(8):3457-3481.
112. Schwarzer J, Rapp E, Reichl U: **N-glycan analysis by CGE-LIF: profiling influenza A virus hemagglutinin N-glycosylation during vaccine production**. *Electrophoresis* 2008, **29**(20):4203-4214.
113. Callewaert N, Geysens S, Molemans F, Contreras R: **Ultrasensitive profiling and sequencing of N-linked oligosaccharides using standard DNA-sequencing equipment**. *Glycobiology* 2001, **11**(4):275-281.
114. Bigge JC, Patel TP, Bruce JA, Goulding PN, Charles SM, Parekh RB: **Nonselective and efficient fluorescent labeling of glycans using 2-amino benzamide and anthranilic acid**. *Anal Biochem* 1995, **230**(2):229-238.
115. Ruhaak LR, Steenvoorden E, Koeleman CA, Deelder AM, Wuhrer M: **2-picoline-borane: a non-toxic reducing agent for oligosaccharide labeling by reductive amination**. *Proteomics* 2010, **10**(12):2330-2336.

116. Calabro A, Benavides M, Tammi M, Hascall VC, Midura RJ: **Microanalysis of enzyme digests of hyaluronan and chondroitin/dermatan sulfate by fluorophore-assisted carbohydrate electrophoresis (FACE)**. *Glycobiology* 2000, **10**(3):273-281.
117. Suzuki H, Muller O, Guttman A, Karger BL: **Analysis of 1-aminopyrene-3,6,8-trisulfonate-derivatized oligosaccharides by capillary electrophoresis with matrix-assisted laser desorption/ionization time-of-flight mass spectrometry**. *Anal Chem* 1997, **69**(22):4554-4559.
118. Suzuki S, Kuwahara Y, Makiura K, Honda S: **Preparation of various silica-based columns for capillary electrochromatography by in-column derivatization**. *J Chromatogr A* 2000, **873**(2):247-256.
119. Zhong H, El Rassi Z: **Neutral polar methacrylate-based monoliths for normal phase nano-LC and CEC of polar species including N-glycans**. *Journal of separation science* 2009, **32**(1):10-20.
120. Manzi AE, Hvan Halbeek H: **Principles of Structural Analysis and Sequencing of Glycans**. In: *Essentials of Glycobiology.* Edited by Varki A CR, Esko J, et al., editors. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 1999.
121. Pabst M, Altmann F: **Glycan analysis by modern instrumental methods**. *Proteomics* 2011, **11**(4):631-643.
122. Wuhrer M, de Boer AR, Deelder AM: **Structural glycomics using hydrophilic interaction chromatography (HILIC) with mass spectrometry**. *Mass spectrometry reviews* 2009, **28**(2):192-206.
123. Royle L, Campbell MP, Radcliffe CM, White DM, Harvey DJ, Abrahams JL, Kim YG, Henry GW, Shadick NA, Weinblatt ME *et al*: **HPLC-based analysis of serum N-glycans on a 96-well plate platform with dedicated database software**. *Anal Biochem* 2008, **376**(1):1-12.
124. Takahashi N: **Three-dimensional mapping of N-linked oligosaccharides using anion-exchange, hydrophobic and hydrophilic interaction modes of high-performance liquid chromatography**. *J Chromatogr A* 1996, **720**(1-2):217-225.
125. Ruhaak LR, Deelder AM, Wuhrer M: **Oligosaccharide analysis by graphitized carbon liquid chromatography-mass spectrometry**. *Analytical and bioanalytical chemistry* 2009, **394**(1):163-174.
126. Townsend RR, Lipniunas PH, Bigge C, Ventom A, Parekh R: **Multimode high-performance liquid chromatography of fluorescently labeled oligosaccharides from glycoproteins**. *Anal Biochem* 1996, **239**(2):200-207.
127. Rapp E, Schwarzer J, Reichl U, Bohne C: **Method for automated high throughput identification of carbohydrates and carbohydrate mixture composition patterns as well as systems therefore**. In. Edited by patent E, vol. EP 08007887.6; 2008.
128. Rapp E, Schwarzer J, Reichl U, Bohne C: **Method for automated high throughput identification of carbohydrates and carbohydrate mixture composition patterns as well as systems therefore**. In. Edited by patent U, vol.: US 12/428,003; 2009.

129. Franca LT, Carrilho E, Kist TB: **A review of DNA sequencing techniques**. *Quarterly reviews of biophysics* 2002, **35**(2):169-200.
130. Mardis ER: **Next-generation sequencing platforms**. *Annual review of analytical chemistry* 2013, **6**:287-303.
131. Janeway CA Jr TP, Walport M, et al. : **Immunobiology: The Immune System in Health and Disease. 5th edition.** New York: : Garland Science; 2001.
132. Racine R, Chatterjee M, Winslow GM: **CD11c expression identifies a population of extrafollicular antigen-specific splenic plasmablasts responsible for CD4 T-independent antibody responses during intracellular bacterial infection**. *J Immunol* 2008, **181**(2):1375-1385.
133. Metlay JP, Witmer-Pack MD, Agger R, Crowley MT, Lawless D, Steinman RM: **The distinct leukocyte integrins of mouse spleen dendritic cells as identified with new hamster monoclonal antibodies**. *J Exp Med* 1990, **171**(5):1753-1771.
134. Lai L, Alaverdi N, Maltais L, Morse HC, 3rd: **Mouse cell surface antigens: nomenclature and immunophenotyping**. *J Immunol* 1998, **160**(8):3861-3868.
135. Janeway C: **The adaptive immune response**. In: *Immunology.* Edited by M. Robertson RW, E. Lawrence, R. Palmer, E. Dorey. London/UK, New York/USA: Current Biology Ltd./Garland Publishing Inc.; 1994: 1-7.
136. Sancho D, Gomez M, Sanchez-Madrid F: **CD69 is an immunoregulatory molecule induced following activation**. *Trends in immunology* 2005, **26**(3):136-140.
137. Broere F, Aspasov SG, Sitkovsky MV, van Eden W: **T cell subsets and T cell-mediated immunity**. In: *Principles of Immunopharmacology: 3rd revised and extended edition.* Edited by Parnham FPNaMJ: Springer Basel AG; 2011: 15-27.
138. McGee HS, Agrawal DK: **TH2 cells in the pathogenesis of airway remodeling: regulatory T cells a plausible panacea for asthma**. *Immunologic research* 2006, **35**(3):219-232.
139. Bertram EM, Dawicki W, Watts TH: **Role of T cell costimulation in anti-viral immunity**. *Seminars in immunology* 2004, **16**(3):185-196.
140. McHeyzer-Williams MG: **B cells as effectors**. *Current opinion in immunology* 2003, **15**(3):354-361.
141. Pierce SK: **Lipid rafts and B-cell activation**. *Nature reviews Immunology* 2002, **2**(2):96-105.
142. Lohr V, Genzel Y, Behrendt I, Scharfenberg K, Reichl U: **A new MDCK suspension line cultivated in a fully defined medium in stirred-tank and wave bioreactor**. *Vaccine* 2010, **28**(38):6256-6264.
143. Jordan I, Northoff S, Thiele M, Hartmann S, Horn D, Howing K, Bernhardt H, Oehmke S, von Horsten H, Rebeski D *et al*: **A chemically defined production process for highly attenuated poxviruses**. *Biologicals* 2011, **39**(1):50-58.
144. Kalbfuss B, Knochlein A, Krober T, Reichl U: **Monitoring influenza virus content in vaccine production: precise assays for the quantitation of**

hemagglutination and neuraminidase activity. *Biologicals* 2008, **36**(3):145-161.
145. Kalbfuss B, Genzel Y, Wolff M, Zimmermann A, Morenweiser R, Reichl U: **Harvesting and concentration of human influenza A virus produced in serum-free mammalian cell culture for the production of vaccines**. *Biotechnol Bioeng* 2007, **97**(1):73-85.
146. Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, Fujimoto EK, Goeke NM, Olson BJ, Klenk DC: **Measurement of protein using bicinchoninic acid**. *Anal Biochem* 1985, **150**(1):76-85.
147. Rödig J, Rapp E, Hennig R, Schwarzer J, Reichl U: **Optimized CGE-LIF-Based Glycan Analysis for High-Throughput Applications**. In: *Proceedings of the 21st Annual Meeting of the European Society for Animal Cell Technology (ESACT), June 7–10, 2009.* Dublin, Ireland: Springer Science+Business Media B.V.; 2011.
148. Rapp E, Hennig R, Borowiak M, Kottler R, Reichl U: **High-Throughput Glycosylation Pattern Analysis of Glycoproteins Utilizing a Multiplexing Capillary-DNA-Sequencer**. *Glycoconjugate Journal* 2011, **28**:234-235.
149. Frey T: Studienarbeit: **Weiterentwicklung der Analysemethode zur Charakterisierung des N-Glykosylierungsmusters des Influenza Proteins Hämagglutinin mittels CGE-LIF**. Magdeburg: Otto von Guericke University; 2009.
150. Ruhaak LR, Hennig R, Huhn C, Borowiak M, Dolhain RJ, Deelder AM, Rapp E, Wuhrer M: **Optimized workflow for preparation of APTS-labeled N-glycans allowing high-throughput analysis of human plasma glycomes using 48-channel multiplexed CGE-LIF**. *J Proteome Res* 2010, **9**(12):6655-6664.
151. Eilers PH, Boelens HFM: **Baseline Correction with Assymetric Least Squares Smoothing**. *Leiden University Medical Centre report* 2005.
152. Tree JA, Richardson C, Fooks AR, Clegg JC, Looby D: **Comparison of large-scale mammalian cell culture systems with egg culture for the production of influenza virus A vaccine strains**. *Vaccine* 2001, **19**(25-26):3444-3450.
153. Aggarwal K, Jing F, Maranga L, Liu J: **Bioprocess optimization for cell culture based influenza vaccine production**. *Vaccine* 2011, **29**(17):3320-3328.
154. Hoper D, Hoffmann B, Beer M: **A comprehensive deep sequencing strategy for full-length genomes of influenza A**. *PLoS One* 2011, **6**(4):e19075.
155. Hoper D, Hoffmann B, Beer M: **Simple, sensitive, and swift sequencing of complete H5N1 avian influenza virus genomes**. *J Clin Microbiol* 2009, **47**(3):674-679.
156. Leifer I, Hoffmann B, Hoper D, Bruun Rasmussen T, Blome S, Strebelow G, Horeth-Bontgen D, Staubach C, Beer M: **Molecular epidemiology of current classical swine fever virus isolates of wild boar in Germany**. *J Gen Virol* 2010, **91**(Pt 11):2687-2697.

157. Wiley G, Macmil S, Qu C, Wang P, Xing Y, White D, Li J, White JD, Domingo A, Roe BA: **Methods for generating shotgun and mixed shotgun/paired-end libraries for the 454 DNA sequencer**. *Curr Protoc Hum Genet* 2009, **Chapter 18**:Unit18 11.
158. Consortium TU: **Reorganizing the protein space at the Universal Protein Resource (UniProt)**. *Nucleic Acids Research* 2012, **40**(Database D71-D75):5.
159. Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E *et al*: **ExPASy: SIB bioinformatics resource portal**. *Nucleic Acids Research* 2012, **40**(W1):W597-W603.
160. Gupta R, Jung E, Brunak S: **Prediction of N-glycosylation sites in human proteins.** . *in preparation* 2004.
161. Tsuchiya E, Sugawara K, Hongo S, Matsuzaki Y, Muraki Y, Li ZN, Nakamura K: **Effect of addition of new oligosaccharide chains to the globular head of influenza A/H2N2 virus haemagglutinin on the intracellular transport and biological activities of the molecule**. *J Gen Virol* 2002, **83**(Pt 5):1137-1146.
162. Deshpande KL, Fried VA, Ando M, Webster RG: **Glycosylation affects cleavage of an H5N2 influenza virus hemagglutinin and regulates virulence**. *Proc Natl Acad Sci U S A* 1987, **84**(1):36-40.
163. Klenk HD, Wagner R, Heuer D, Wolff T: **Importance of hemagglutinin glycosylation for the biological functions of influenza virus**. *Virus Res* 2002, **82**(1-2):73-75.
164. Wagner R, Heuer D, Wolff T, Herwig A, Klenk HD: **N-Glycans attached to the stem domain of haemagglutinin efficiently regulate influenza A virus replication**. *J Gen Virol* 2002, **83**(Pt 3):601-609.
165. Gamblin SJ, Haire LF, Russell RJ, Stevens DJ, Xiao B, Ha Y, Vasisht N, Steinhauer DA, Daniels RS, Elliot A *et al*: **The structure and receptor binding properties of the 1918 influenza hemagglutinin**. *Science* 2004, **303**(5665):1838-1842.
166. Deng K, Adams MM, Damani P, Livingston PO, Ragupathi G, Gin DY: **Synthesis of QS-21-xylose: establishment of the immunopotentiating activity of synthetic QS-21 adjuvant with a melanoma vaccine**. *Angewandte Chemie* 2008, **47**(34):6395-6398.
167. Boonyarattanakalin S, Liu X, Michieletti M, Lepenies B, Seeberger PH: **Chemical synthesis of all phosphatidylinositol mannoside (PIM) glycans from Mycobacterium tuberculosis**. *J Am Chem Soc* 2008, **130**(49):16791-16799.
168. Hansen W, Westendorf AM, Reinwald S, Bruder D, Deppenmeier S, Groebe L, Probst-Kepper M, Gruber AD, Geffers R, Buer J: **Chronic antigen stimulation in vivo induces a distinct population of antigen-specific Foxp3 CD25 regulatory T cells**. *J Immunol* 2007, **179**(12):8059-8068.
169. Kirberg J, Baron A, Jakob S, Rolink A, Karjalainen K, von Boehmer H: **Thymic selection of CD8+ single positive cells with a class II major**

histocompatibility complex-restricted receptor. *J Exp Med* 1994, **180**(1):25-34.

170. Raju TS, Briggs JB, Borge SM, Jones AJ: **Species-specific variation in glycosylation of IgG: evidence for the species-specific sialylation and branch-specific galactosylation and importance for engineering recombinant glycoprotein therapeutics**. *Glycobiology* 2000, **10**(5):477-486.
171. Butler M: **Optimisation of the cellular metabolism of glycosylation for recombinant proteins produced by Mammalian cell systems**. *Cytotechnology* 2006, **50**(1-3):57-76.
172. Jordan I, Vos A, Beilfuss S, Neubert A, Breul S, Sandig V: **An avian cell line designed for production of highly attenuated viruses**. *Vaccine* 2009, **27**(5):748-756.
173. Hooker AD, Goldman MH, Markham NH, James DC, Ison AP, Bull AT, Strange PG, Salmon I, Baines AJ, Jenkins N: **N-glycans of recombinant human interferon-gamma change during batch culture of chinese hamster ovary cells**. *Biotechnol Bioeng* 1995, **48**(6):639-648.
174. Hayter PM, Curling EM, Baines AJ, Jenkins N, Salmon I, Strange PG, Tong JM, Bull AT: **Glucose-limited chemostat culture of Chinese hamster ovary cells producing recombinant human interferon-gamma**. *Biotechnol Bioeng* 1992, **39**(3):327-335.
175. Gershman H, Robbins PW: **Transitory effects of glucose starvation on the synthesis of dolichol-linked oligosaccharides in mammalian cells**. *J Biol Chem* 1981, **256**(15):7774-7780.
176. Chee Furng Wong D, Tin Kam Wong K, Tang Goh L, Kiat Heng C, Gek Sim Yap M: **Impact of dynamic online fed-batch strategies on metabolism, productivity and N-glycosylation quality in CHO cell cultures**. *Biotechnol Bioeng* 2005, **89**(2):164-177.
177. Nyberg GB, Balcarcel RR, Follstad BD, Stephanopoulos G, Wang DI: **Metabolic effects on recombinant interferon-gamma glycosylation in continuous culture of Chinese hamster ovary cells**. *Biotechnol Bioeng* 1999, **62**(3):336-347.
178. Ljunggren J, Haggstrom L: **Catabolic control of hybridoma cells by glucose and glutamine limited fed batch cultures**. *Biotechnol Bioeng* 1994, **44**(7):808-818.
179. Andersen DC, Goochee CF: **The effect of cell-culture conditions on the oligosaccharide structures of secreted glycoproteins**. *Curr Opin Biotechnol* 1994, **5**(5):546-549.
180. Borys MC, Linzer DI, Papoutsakis ET: **Ammonia affects the glycosylation patterns of recombinant mouse placental lactogen-I by chinese hamster ovary cells in a pH-dependent manner**. *Biotechnol Bioeng* 1994, **43**(6):505-514.
181. Yang M, Butler M: **Effects of ammonia on CHO cell growth, erythropoietin production, and glycosylation**. *Biotechnol Bioeng* 2000, **68**(4):370-380.

182. Yang M, Butler M: **Effect of ammonia on the glycosylation of human recombinant erythropoietin in culture**. *Biotechnol Prog* 2000, **16**(5):751-759.
183. Gillmeister MP, Tomiya N, Jacobia SJ, Lee YC, Gorfien SF, Betenbaugh MJ: **An HPLC-MALDI MS method for N-glycan analyses using smaller size samples: application to monitor glycan modulation by medium conditions**. *Glycoconj J* 2009, **26**(9):1135-1149.
184. Aumiller JJ, Mabashi-Asazuma H, Hillar A, Shi X, Jarvis DL: **A new glycoengineered insect cell line with an inducibly mammalianized protein N-glycosylation pathway**. *Glycobiology* 2012, **22**(3):417-428.
185. Schuster M, Umana P, Ferrara C, Brunker P, Gerdes C, Waxenecker G, Wiederkum S, Schwager C, Loibner H, Himmler G *et al*: **Improved effector functions of a therapeutic monoclonal Lewis Y-specific antibody by glycoform engineering**. *Cancer Res* 2005, **65**(17):7934-7941.
186. Viswanathan K, Lawrence S, Hinderlich S, Yarema KJ, Lee YC, Betenbaugh MJ: **Engineering sialic acid synthetic ability into insect cells: identifying metabolic bottlenecks and devising strategies to overcome them**. *Biochemistry* 2003, **42**(51):15215-15225.
187. Walsh G: **Biopharmaceutical benchmarks 2010**. *Nat Biotechnol* 2010, **28**(9):917-924.
188. Kobata A: **Structures, function, and transformational changes of the sugar chains of glycohormones**. *J Cell Biochem* 1988, **37**(1):79-90.
189. Stanley P: **Glycosylation mutants of animal cells**. *Annu Rev Genet* 1984, **18**:525-552.
190. Stanley P: **Chinese hamster ovary cell mutants with multiple glycosylation defects for production of glycoproteins with minimal carbohydrate heterogeneity**. *Mol Cell Biol* 1989, **9**(2):377-383.
191. Mir-Shekari SY, Ashford DA, Harvey DJ, Dwek RA, Schulze IT: **The glycosylation of the influenza A virus hemagglutinin by mammalian cells. A site-specific study**. *J Biol Chem* 1997, **272**(7):4027-4036.
192. Harpaz N, Schachter H: **Control of glycoprotein synthesis. Processing of asparagine-linked oligosaccharides by one or more rat liver Golgi alpha-D-mannosidases dependent on the prior action of UDP-N-acetylglucosamine: alpha-D-mannoside beta 2-N-acetylglucosaminyltransferase I**. *J Biol Chem* 1980, **255**(10):4894-4902.
193. Heynisch B, Frensing T, Heinze K, Seitz C, Genzel Y, Reichl U: **Differential activation of host cell signalling pathways through infection with two variants of influenza A/Puerto Rico/8/34 (H1N1) in MDCK cells**. *Vaccine* 2010, **28**(51):8210-8218.
194. Schulze-Horsel J, Schulze M, Agalaridis G, Genzel Y, Reichl U: **Infection dynamics and virus-induced apoptosis in cell culture-based influenza vaccine production-Flow cytometry and mathematical modeling**. *Vaccine* 2009, **27**(20):2712-2722.
195. Seitz C, Frensing T, Hoper D, Kochs G, Reichl U: **High yields of influenza A virus in Madin-Darby canine kidney cells are promoted**

by an insufficient interferon-induced antiviral state. *J Gen Virol* 2010, **91**(Pt 7):1754-1763.
196. Vester D, Rapp E, Kluge S, Genzel Y, Reichl U: **Virus-host cell interactions in vaccine production cell lines infected with different human influenza A virus variants: a proteomic approach**. *J Proteomics* 2010, **73**(9):1656-1669.
197. Frensing T, Heldt FS, Pflugmacher A, Behrendt I, Jordan I, Flockerzi D, Genzel Y, Reichl U: **Continuous influenza virus production in cell culture shows a periodic accumulation of defective interfering particles**. *PLoS One* 2013, **8**(9):e72288.
198. Andersen DC, Bridges T, Gawlitzek M, Hoy C: **Multiple cell culture factors can affect the glycosylation of Asn-184 in CHO-produced tissue-type plasminogen activator**. *Biotechnology and Bioengineering* 2000, **70**(1):25-31.
199. Hakimi J, Atkinson PH: **Growth-dependent alterations in oligomannosyl glycopeptides expressed in Sindbis virus glycoproteins**. *Biochemistry* 1980, **19**(24):5619-5624.
200. Ogorek C: **Charakterisierung des N-Glykosylierungsprofils der neuen Entenzelllinie AGE1.CR und die gerichtete Modifikation der Fucosylierung**. *dissertation.* Freie Universität Berlin; 2011.
201. Lohr V, Genzel Y, Jordan I, Katinger D, Mahr S, Sandig V, Reichl U: **Live attenuated influenza viruses produced in a suspension process with avian AGE1.CR.pIX cells**. *BMC biotechnology* 2012, **12**:79.
202. Naffakh N, Tomoiu A, Rameix-Welti MA, van der Werf S: **Host restriction of avian influenza viruses at the level of the ribonucleoproteins**. *Annual review of microbiology* 2008, **62**:403-424.
203. Momose F, Basler CF, O'Neill RE, Iwamatsu A, Palese P, Nagata K: **Cellular splicing factor RAF-2p48/NPI-5/BAT1/UAP56 interacts with the influenza virus nucleoprotein and enhances viral RNA synthesis**. *J Virol* 2001, **75**(4):1899-1908.
204. Ping J, Keleta L, Forbes NE, Dankar S, Stecho W, Tyler S, Zhou Y, Babiuk L, Weingartl H, Halpin RA *et al*: **Genomic and protein structural maps of adaptive evolution of human influenza A virus to increased virulence in the mouse**. *PLoS One* 2011, **6**(6):e21740.
205. Ghedin E, Spiro D, Miller N, Zaborsky J, Feldblyum T, Subbu V, Shumway M, Sparenborg J, Groveman L, Halpin R *et al*: **The NIAID Influenza Genome Sequencing Project**. In*.* EMBL/GenBank/DDBJ databases; 2006.
206. Winter G, Fields S: **Cloning of influenza cDNA ino M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein**. *Nucleic Acids Res* 1980, **8**(9):1965-1974.
207. Consortium TU: **Matrix protein 1**. In*.* Protein Knowledgebase (UniProtKB); Last modified November 28, 2012. Version 100.
208. Hay AJ, Wolstenholme AJ, Skehel JJ, Smith MH: **The molecular basis of the specific anti-influenza action of amantadine**. *EMBO J* 1985, **4**(11):3021-3024.

209. Furuse Y, Suzuki A, Kamigaki T, Oshitani H: **Evolution of the M gene of the influenza A virus in different host species: large-scale sequence analysis**. *Virology journal* 2009, **6**:67.
210. Consortium TU: **Matrix Protein 2**. In. Protein Knowledgebase (UniProtKB): Reviewed, UniProtKB/Swiss-Prot Last modified November 28, 2012. Version 100.
211. Ruigrok RW, Wrigley NG, Calder LJ, Cusack S, Wharton SA, Brown EB, Skehel JJ: **Electron microscopy of the low pH structure of influenza virus haemagglutinin**. *EMBO J* 1986, **5**(1):41-49.
212. Ruigrok RW, Martin SR, Wharton SA, Skehel JJ, Bayley PM, Wiley DC: **Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes**. *Virology* 1986, **155**(2):484-497.
213. Doms RW, Gething MJ, Henneberry J, White J, Helenius A: **Variant influenza virus hemagglutinin that induces fusion at elevated pH**. *J Virol* 1986, **57**(2):603-613.
214. Reed ML, Bridges OA, Seiler P, Kim JK, Yen HL, Salomon R, Govorkova EA, Webster RG, Russell CJ: **The pH of activation of the hemagglutinin protein regulates H5N1 influenza virus pathogenicity and transmissibility in ducks**. *J Virol* 2010, **84**(3):1527-1535.
215. Thoennes S, Li ZN, Lee BJ, Langley WA, Skehel JJ, Russell RJ, Steinhauer DA: **Analysis of residues near the fusion peptide in the influenza hemagglutinin structure for roles in triggering membrane fusion**. *Virology* 2008, **370**(2):403-414.
216. Murakami S, Horimoto T, Ito M, Takano R, Katsura H, Shimojima M, Kawaoka Y: **Enhanced growth of influenza vaccine seed viruses in vero cells mediated by broadening the optimal pH range for virus membrane fusion**. *J Virol* 2012, **86**(3):1405-1410.
217. Nakowitsch S, Wolschek M, Morokutti A, Ruthsatz T, Krenn BM, Ferko B, Ferstl N, Triendl A, Muster T, Egorov A *et al*: **Mutations affecting the stability of the haemagglutinin molecule impair the immunogenicity of live attenuated H3N2 intranasal influenza vaccine candidates lacking NS1**. *Vaccine* 2011, **29**(19):3517-3524.
218. Xu R, Wilson IA: **Structural characterization of an early fusion intermediate of influenza virus hemagglutinin**. *J Virol* 2011, **85**(10):5172-5182.
219. Reed ML, Yen HL, DuBois RM, Bridges OA, Salomon R, Webster RG, Russell CJ: **Amino acid residues in the fusion peptide pocket regulate the pH of activation of the H5N1 influenza virus hemagglutinin protein**. *J Virol* 2009, **83**(8):3568-3580.
220. Rott R, Orlich M, Klenk HD, Wang ML, Skehel JJ, Wiley DC: **Studies on the adaptation of influenza viruses to MDCK cells**. *EMBO J* 1984, **3**(13):3329-3332.
221. Zaraket H, Bridges OA, Duan S, Baranovich T, Yoon SW, Reed ML, Salomon R, Webby RJ, Webster RG, Russell CJ: **Increased acid stability of the hemagglutinin protein enhances H5N1 influenza virus**

growth in the upper respiratory tract but is insufficient for transmission in ferrets. *J Virol* 2013.

222. Zaraket H, Bridges OA, Russell CJ: **The pH of activation of the hemagglutinin protein regulates H5N1 influenza virus replication and pathogenesis in mice**. *J Virol* 2013, **87**(9):4826-4834.
223. Brown JD, Swayne DE, Cooper RJ, Burns RE, Stallknecht DE: **Persistence of H5 and H7 avian influenza viruses in water**. *Avian Dis* 2007, **51**(1 Suppl):285-289.
224. Reed ML, Yen HL, DuBois RM, Bridges OA, Salomon R, Webster RG, Russell CJ: **Amino Acid Residues in the Fusion Peptide Pocket Regulate the pH of Activation of the H5N1 Influenza Virus Hemagglutinin Protein**. *Journal of Virology* 2009, **83**(8):3568-3580.
225. Chen GW, Chang SC, Mok CK, Lo YL, Kung YN, Huang JH, Shih YH, Wang JY, Chiang C, Chen CJ *et al*: **Genomic signatures of human versus avian influenza A viruses**. *Emerg Infect Dis* 2006, **12**(9):1353-1360.
226. Consortium TU: **Neuraminidase**. In. Protein KnowledgebaseKB (UniProtKB): Reviewed, UniProtKB/Swiss-Prot; Last modified November 28, 2012. Version 101.
227. Fields S, Winter G, Brownlee GG: **Structure of the neuraminidase gene in human influenza virus A/PR/8/34**. *Nature* 1981, **290**(5803):213-217.
228. Aragon T, de la Luna S, Novoa I, Carrasco L, Ortin J, Nieto A: **Eukaryotic translation initiation factor 4GI is a cellular target for NS1 protein, a translational activator of influenza virus**. *Mol Cell Biol* 2000, **20**(17):6259-6268.
229. Burgui I, Aragon T, Ortin J, Nieto A: **PABP1 and eIF4GI associate with influenza virus NS1 protein in viral mRNA translation initiation complexes**. *J Gen Virol* 2003, **84**(Pt 12):3263-3274.
230. Chen Z, Li Y, Krug RM: **Influenza A virus NS1 protein targets poly(A)-binding protein II of the cellular 3'-end processing machinery**. *EMBO J* 1999, **18**(8):2273-2283.
231. Kochs G, Garcia-Sastre A, Martinez-Sobrido L: **Multiple anti-interferon actions of the influenza A virus NS1 protein**. *J Virol* 2007, **81**(13):7011-7021.
232. Forbes NE, Ping J, Dankar SK, Jia JJ, Selman M, Keleta L, Zhou Y, Brown EG: **Multifunctional adaptive NS1 mutations are selected upon human influenza virus evolution in the mouse**. *PLoS One* 2012, **7**(2):e31839.
233. Das K, Ma LC, Xiao R, Radvansky B, Aramini J, Zhao L, Marklund J, Kuo RL, Twu KY, Arnold E *et al*: **Structural basis for suppression of a host antiviral response by influenza A virus**. *Proc Natl Acad Sci U S A* 2008, **105**(35):13093-13098.
234. Steidle S, Martinez-Sobrido L, Mordstein M, Lienenklaus S, Garcia-Sastre A, Staheli P, Kochs G: **Glycine 184 in nonstructural protein NS1 determines the virulence of influenza A virus strain PR8 without**

affecting the host interferon response. *J Virol* 2010, **84**(24):12761-12770.

235. Hayman A, Comely S, Lackenby A, Hartgroves LC, Goodbourn S, McCauley JW, Barclay WS: **NS1 proteins of avian influenza A viruses can act as antagonists of the human alpha/beta interferon response**. *J Virol* 2007, **81**(5):2318-2327.
236. Kuo RL, Krug RM: **Influenza a virus polymerase is an integral component of the CPSF30-NS1A protein complex in infected cells**. *J Virol* 2009, **83**(4):1611-1616.
237. de Vries W, Haasnoot J, Fouchier R, de Haan P, Berkhout B: **Differential RNA silencing suppression activity of NS1 proteins from different influenza A virus strains**. *J Gen Virol* 2009, **90**(Pt 8):1916-1922.
238. Li WX, Li H, Lu R, Li F, Dus M, Atkinson P, Brydon EW, Johnson KL, Garcia-Sastre A, Ball LA *et al*: **Interferon antagonist proteins of influenza and vaccinia viruses are suppressors of RNA silencing**. *Proc Natl Acad Sci U S A* 2004, **101**(5):1350-1355.
239. Bucher E, Hemmes H, de Haan P, Goldbach R, Prins M: **The influenza A virus NS1 protein binds small interfering RNAs and suppresses RNA silencing in plants**. *J Gen Virol* 2004, **85**(Pt 4):983-991.
240. Delgadillo MO, Saenz P, Salvador B, Garcia JA, Simon-Mateo C: **Human influenza virus NS1 protein enhances viral pathogenicity and acts as an RNA silencing suppressor in plants**. *J Gen Virol* 2004, **85**(Pt 4):993-999.
241. Lu Y, Qian XY, Krug RM: **The influenza virus NS1 protein: a novel inhibitor of pre-mRNA splicing**. *Genes & development* 1994, **8**(15):1817-1828.
242. Fortes P, Beloso A, Ortin J: **Influenza virus NS1 protein inhibits pre-mRNA splicing and blocks mRNA nucleocytoplasmic transport**. *EMBO J* 1994, **13**(3):704-712.
243. Qiu Y, Krug RM: **The influenza virus NS1 protein is a poly(A)-binding protein that inhibits nuclear export of mRNAs containing poly(A)**. *J Virol* 1994, **68**(4):2425-2432.
244. Satterly N, Tsai PL, van Deursen J, Nussenzveig DR, Wang Y, Faria PA, Levay A, Levy DE, Fontoura BM: **Influenza virus targets the mRNA export machinery and the nuclear pore complex**. *Proc Natl Acad Sci U S A* 2007, **104**(6):1853-1858.
245. Xing Z, Cardona CJ, Adams S, Yang Z, Li J, Perez D, Woolcock PR: **Differential regulation of antiviral and proinflammatory cytokines and suppression of Fas-mediated apoptosis by NS1 of H9N2 avian influenza virus in chicken macrophages**. *J Gen Virol* 2009, **90**(Pt 5):1109-1118.
246. Ehrhardt C, Wolff T, Pleschka S, Planz O, Beermann W, Bode JG, Schmolke M, Ludwig S: **Influenza A virus NS1 protein activates the PI3K/Akt pathway to mediate antiapoptotic signaling responses**. *J Virol* 2007, **81**(7):3058-3067.

247. Enami K, Sato TA, Nakada S, Enami M: **Influenza virus NS1 protein stimulates translation of the M1 protein**. *J Virol* 1994, **68**(3):1432-1437.
248. Goni N, Iriarte A, Comas V, Sonora M, Moreno P, Moratorio G, Musto H, Cristina J: **Pandemic influenza A virus codon usage revisited: biases, adaptation and implications for vaccine strain development**. *Virology journal* 2012, **9**:263.
249. Kunkel JP, Jan DC, Butler M, Jamieson JC: **Comparisons of the glycosylation of a monoclonal antibody produced under nominally identical cell culture conditions in two different bioreactors**. *Biotechnol Prog* 2000, **16**(3):462-470.
250. Kunkel JP, Jan DC, Jamieson JC, Butler M: **Dissolved oxygen concentration in serum-free continuous culture affects N-linked glycosylation of a monoclonal antibody**. *J Biotechnol* 1998, **62**(1):55-71.
251. Restelli V, Wang MD, Huzel N, Ethier M, Perreault H, Butler M: **The effect of dissolved oxygen on the production and the glycosylation profile of recombinant human erythropoietin produced from CHO cells**. *Biotechnol Bioeng* 2006, **94**(3):481-494.
252. Borys MC, Linzer DI, Papoutsakis ET: **Culture pH affects expression rates and glycosylation of recombinant mouse placental lactogen proteins by Chinese hamster ovary (CHO) cells**. *Biotechnology (N Y)* 1993, **11**(6):720-724.
253. Zanghi JA, Schmelzer AE, Mendoza TP, Knop RH, Miller WM: **Bicarbonate concentration and osmolality are key determinants in the inhibition of CHO cell polysialylation under elevated pCO(2) or pH**. *Biotechnol Bioeng* 1999, **65**(2):182-191.
254. Senger RS, Karim MN: **Effect of shear stress on intrinsic CHO culture state and glycosylation of recombinant tissue-type plasminogen activator protein**. *Biotechnol Prog* 2003, **19**(4):1199-1209.
255. Nam JH, Zhang F, Ermonval M, Linhardt RJ, Sharfstein ST: **The effects of culture conditions on the glycosylation of secreted human placental alkaline phosphatase produced in Chinese hamster ovary cells**. *Biotechnol Bioeng* 2008, **100**(6):1178-1192.
256. O'Garra A: **Cytokines induce the development of functionally heterogeneous T helper cell subsets**. *Immunity* 1998, **8**(3):275-283.
257. Gardner EM, Beli E, Clinthorne JF, Duriancik DM: **Energy intake and response to infection with influenza**. *Annual review of nutrition* 2011, **31**:353-367.
258. Legge KL, Braciale TJ: **Accelerated migration of respiratory dendritic cells to the regional lymph nodes is limited to the early phase of pulmonary infection**. *Immunity* 2003, **18**(2):265-277.
259. Castellino F, Germain RN: **Cooperation between CD4+ and CD8+ T cells: when, where, and how**. *Annual review of immunology* 2006, **24**:519-540.
260. Ho AW, Prabhu N, Betts RJ, Ge MQ, Dai X, Hutchinson PE, Lew FC, Wong KL, Hanson BJ, Macary PA *et al*: **Lung CD103+ dendritic cells**

efficiently transport influenza virus to the lymph node and load viral antigen onto MHC class I for presentation to CD8 T cells. *J Immunol* 2011, **187**(11):6011-6021.

261. Belz GT, Smith CM, Eichner D, Shortman K, Karupiah G, Carbone FR, Heath WR: **Cutting edge: conventional CD8 alpha+ dendritic cells are generally involved in priming CTL immunity to viruses**. *J Immunol* 2004, **172**(4):1996-2000.
262. Wilkinson TM, Li CK, Chui CS, Huang AK, Perkins M, Liebner JC, Lambkin-Williams R, Gilbert A, Oxford J, Nicholas B *et al*: **Preexisting influenza-specific CD4+ T cells correlate with disease protection against influenza challenge in humans**. *Nat Med* 2012, **18**(2):274-280.
263. Figdor CG, van Kooyk Y, Adema GJ: **C-type lectin receptors on dendritic cells and Langerhans cells**. *Nature reviews Immunology* 2002, **2**(2):77-84.
264. Reading PC, Miller JL, Anders EM: **Involvement of the mannose receptor in infection of macrophages by influenza virus**. *J Virol* 2000, **74**(11):5190-5197.
265. Upham JP, Pickett D, Irimura T, Anders EM, Reading PC: **Macrophage receptors for influenza A virus: role of the macrophage galactose-type lectin and mannose receptor in viral entry**. *J Virol* 2010, **84**(8):3730-3737.
266. Hillaire ML, Nieuwkoop NJ, Boon AC, de Mutsert G, Vogelzang-van Trierum SE, Fouchier RA, Osterhaus AD, Rimmelzwaan GF: **Binding of DC-SIGN to the Hemagglutinin of Influenza A Viruses Supports Virus Replication in DC-SIGN Expressing Cells**. *PLoS One* 2013, **8**(2):e56164.
267. Hillaire ML, Haagsman HP, Osterhaus AD, Rimmelzwaan GF, van Eijk M: **Pulmonary Surfactant Protein D in First-Line Innate Defence against Influenza A Virus Infections**. *Journal of innate immunity* 2013.
268. Abe Y, Takashita E, Sugawara K, Matsuzaki Y, Muraki Y, Hongo S: **Effect of the addition of oligosaccharides on the biological activities and antigenicity of influenza A/H3N2 virus hemagglutinin**. *J Virol* 2004, **78**(18):9605-9611.
269. Wang W, Lu B, Zhou H, Suguitan AL, Jr., Cheng X, Subbarao K, Kemble G, Jin H: **Glycosylation at 158N of the hemagglutinin protein and receptor binding specificity synergistically affect the antigenicity and immunogenicity of a live attenuated H5N1 A/Vietnam/1203/2004 vaccine virus in ferrets**. *J Virol* 2010, **84**(13):6570-6577.
270. Chu VC, Whittaker GR: **Influenza virus entry and infection require host cell N-linked glycoprotein**. *Proc Natl Acad Sci U S A* 2004, **101**(52):18153-18158.
271. de Vries E, de Vries RP, Wienholts MJ, Floris CE, Jacobs MS, van den Heuvel A, Rottier PJ, de Haan CA: **Influenza A virus entry into cells lacking sialylated N-glycans**. *Proc Natl Acad Sci U S A* 2012, **109**(19):7457-7462.

272. Fant X, Samejima K, Carvalho A, Ogawa H, Xu Z, Yue Z, Earnshaw WC, Ruchaud S: **Use of DT40 conditional-knockout cell lines to study chromosomal passenger protein function**. *Biochem Soc Trans* 2010, **38**(6):1655-1659.
273. Baumeister H, Goletz S: **A Matter of Cell Line Development**. *European BioPharmaceutical Review* 2010:54-58.
274. Ono T, Nishijima H, Adachi N, Iiizumi S, Morohoshi A, Koyama H, Shibahara K: **Generation of tetracycline-inducible conditional gene knockout cells in a human Nalm-6 cell line**. *J Biotechnol* 2009, **141**(1-2):1-7.
275. Ma DY, Clark EA: **The role of CD40 and CD154/CD40L in dendritic cells**. *Seminars in immunology* 2009, **21**(5):265-272.
276. La Gruta N, Kelso A, Brown LE, Chen W, Jackson DC, Turner SJ: **Role of CD8(+) T-cell immunity in influenza infection: potential use in future vaccine development**. *Expert review of respiratory medicine* 2009, **3**(5):523-537.
277. McKinstry KK, Strutt TM, Swain SL: **Hallmarks of CD4 T cell immunity against influenza**. *Journal of internal medicine* 2011, **269**(5):507-518.

8 Index of figures

9 Index of supplementary figures

10 Index of tables

11 Index of supplementary tables

12 Supplementary

12.1 Impact of different IVA production cells on HA N-glycosylation (MDCK.SUS2 and MDCK.SUS3 cells included)

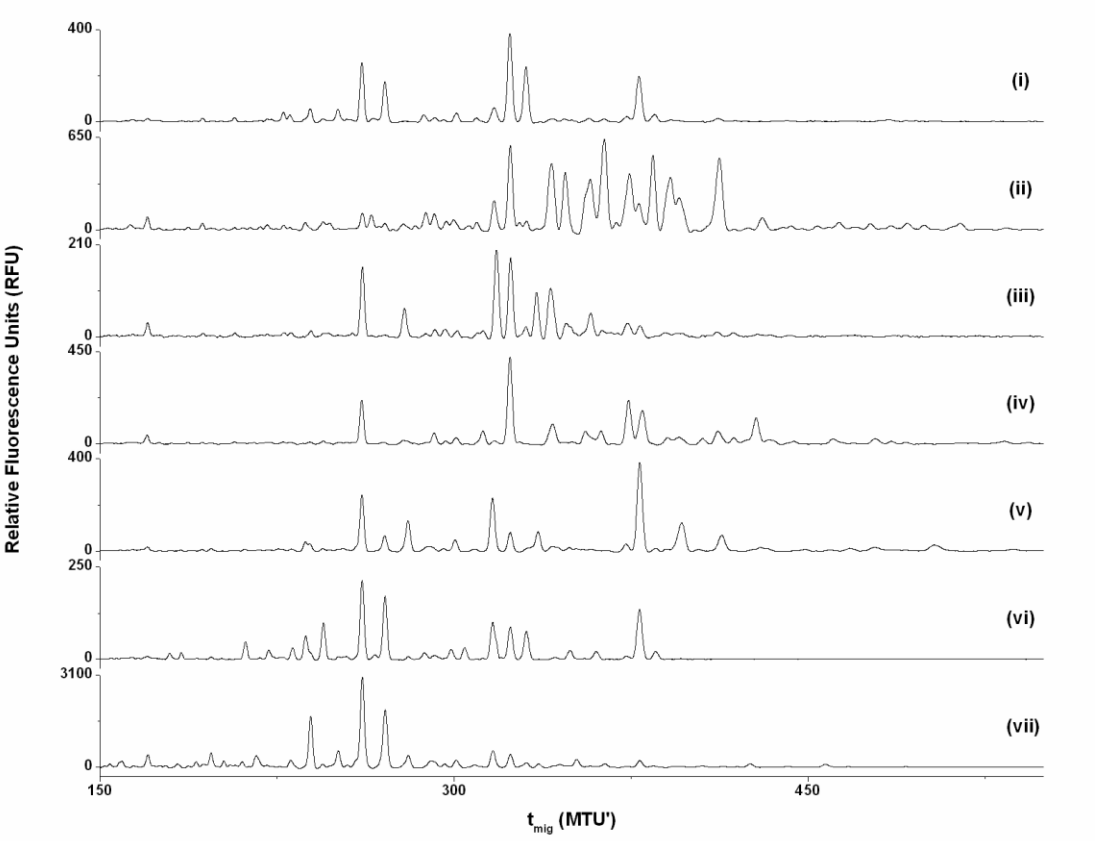


figure S 1: HA *N*-glycosylation fingerprints derived from viruses produced in different host cells.
The RKI-strain (IVA-PR8) was produced in adherently growing Vero cells (i) of African green monkey origin or MDCK (ii) cells of canine origin. Furthermore, the adherently growing MDCK cell line was adapted to suspension growth in two biological independent settings generating the in suspension growing MDCK.SUS2 (iii) and MDCK.SUS3 (iv) cell line. Additional suspension cells such as human-derived Cap cells (v) or duck-derived AGE1.CR.pIX cells (vi) were used for virus production. Finally, cell culture-derived virus was compared with virus produced in embryonated hen's eggs (vii).

12.2 Host cell adaptation to serum-free suspension growth

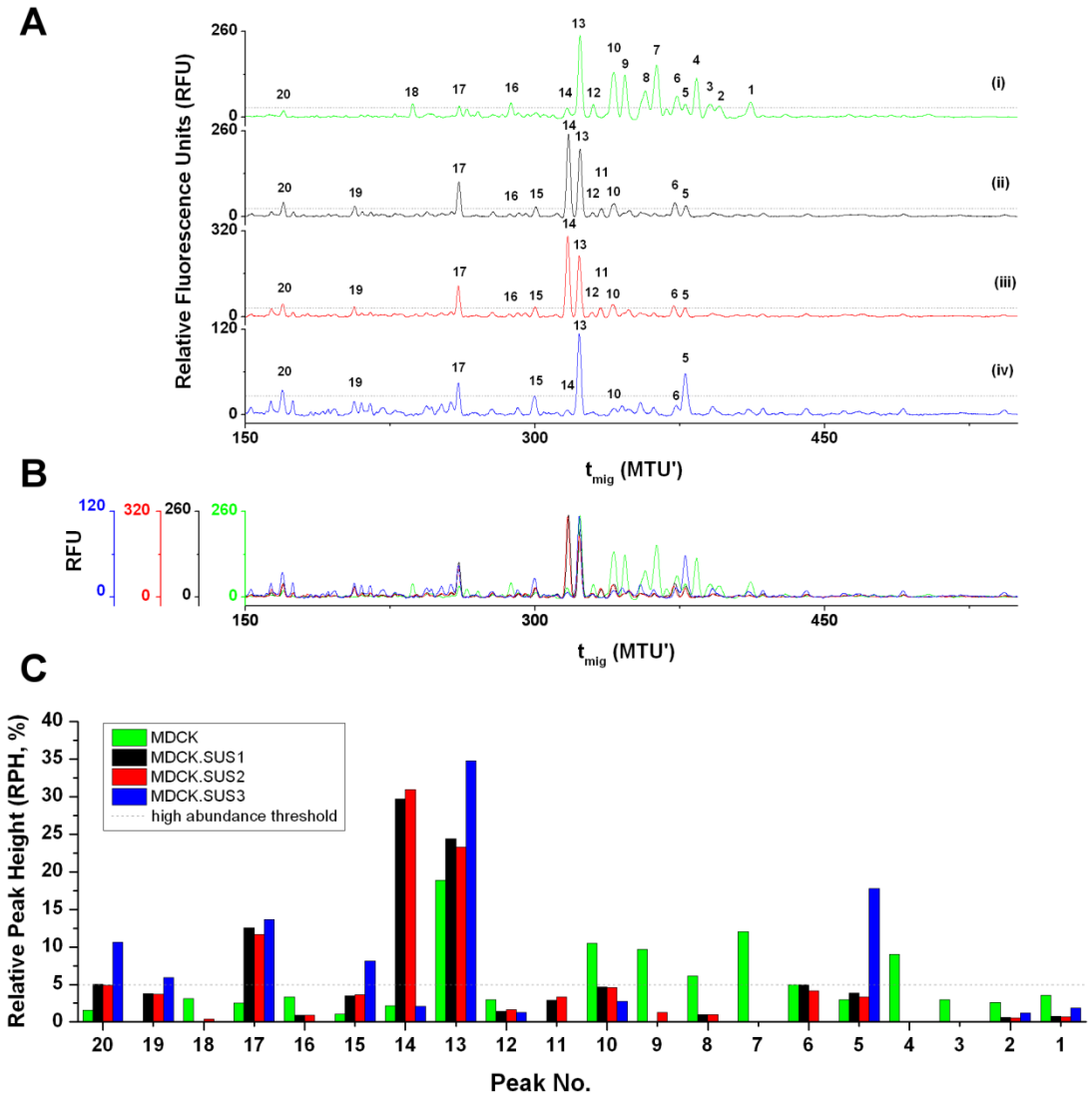


figure S 2: Impact of host cell adaptation to serum-free suspension growth on the HA *N*-glycosylation pattern of the NIBSC-strain.
(A) HA *N*-glycan fingerprints, relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). [6] All peaks exceeding the 10x baseline noise threshold (···) in at least one fingerprint are annotated. The non-adapted, serum requiring, adherent MDCK cell line (i) was adapted to serum-free suspension growth (ii, MDCK.SUS1) and further adapted to better growth characteristics (iii, MDCK.SUS2) [142]. The first adaptation step was performed in biological duplicates (iv, MDCK.SUS3). (B) Overlay of all four *N*-glycosylation fingerprints. (C) Relative peak height (RPH) in % of the total peak height (TPH, sum of all annotated peaks) displaying relative peak abundance. Peaks are defined high abundant if RPH > 5 % (---).

12.3 Impact of different virus suppliers in MDCK.SUS2 cells

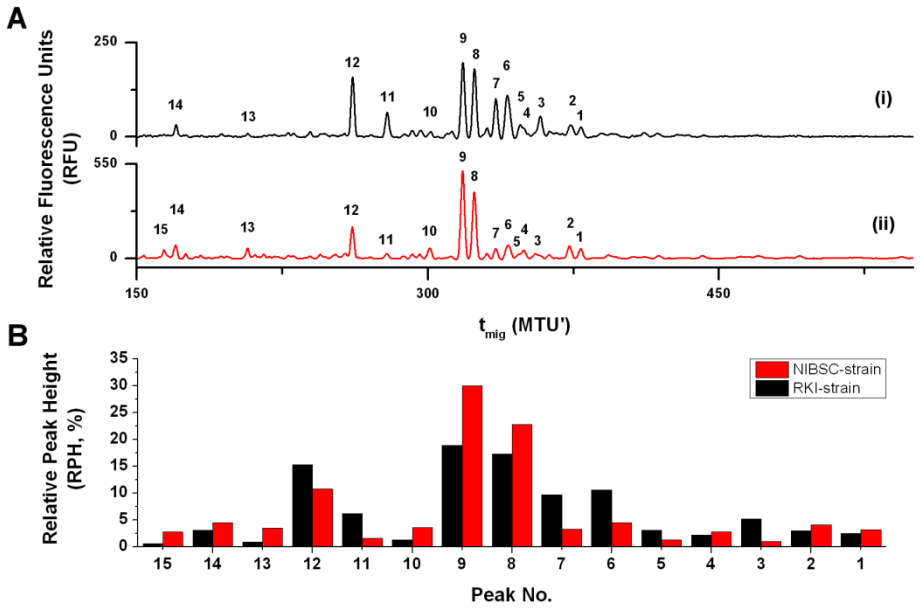


figure S 3: Comparison of HA *N*-glycosylation pattern of two IVA-PR8 strains from different suppliers.
(A) HA *N*-glycan fingerprints, relative fluorescence units (RFU) are plotted over the migration time (t_{mig}) in normalized migration time units (MTU'). All peaks exceeding in at least one of the two fingerprints the 10x baseline noise threshold are annotated (1 - 15). The virus was purchased from the Robert Koch Institute (RKI, i) or from NIBSC (ii). (B) Relative peak height (RPH) in % of the total peak height (TPH, sum of all annotated peaks) displays relative peak abundance.

12.4 Similarity of HA N-glycan fingerprints derived from different viruses produced in MDCK cells

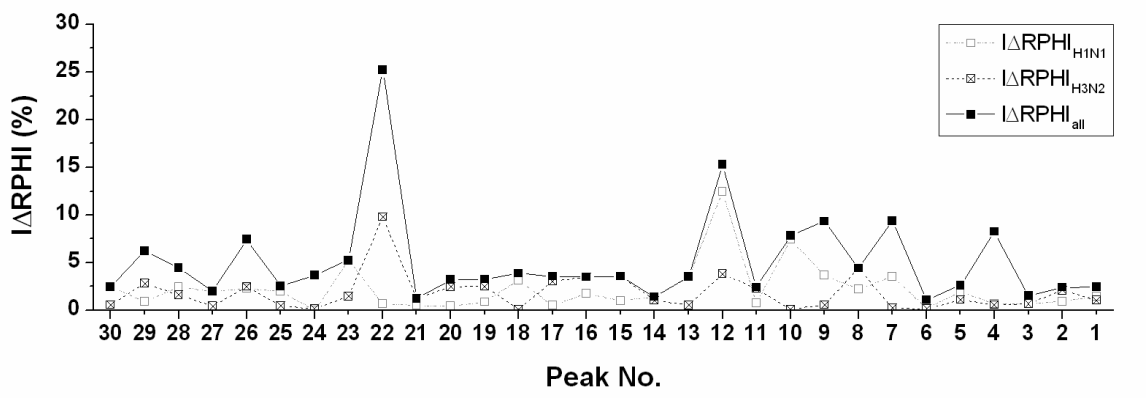


figure S 4: Differences of relative peak height (|ΔRPH|) for different viruses produced in MDCK cells.
The differences of RPH between two different H1N1 virus strains (□) and two different H3N2 virus strains (■) are smaller than the differences over all tested viruses (■): the variation of relative peak abundance within a virus subtype is smaller than over different subtypes.

12.5 HA-Titers in MDCK cells during virus adaptation

table S 1: MDCK cell-derived HA-titers during virus adaptation.
The RKI- and the NIBSC-strain were adapted from MDCK (seed virus, passage 1) to Vero (passages 2-6) and back to MDCK cells (passages 7-11).

RKI-strain			NIBSC-strain		
passage (no.)	hours post infection (hpi)	HA-titer (HAU)	passage (no.)	hours post infection (hpi)	HA-titer (HAU)
1	24	1.9	1	24	2.2
7	96	2.5	7	96	2.1
8	48	2.0	8	48	2.4
8	72	2.0	8	72	2.4
9	96	2.2	9	96	2.4
10	48	2.3	10	48	2.6
10	72	2.2	10	72	2.5
11	96	2.2	11	96	2.4

12.6 Supplementary tables

table S 2: Accession numbers for IVA-PR8 sequences during virus adaptation.
(1)-(6) Adaptation from MDCK to Vero cells and back or from MDCK to Vero cells (7). The original virus seed was either purchased from the National Institute for Biological Standards and Control (NIBSC) or the Robert Koch Institute (RKI). (1, 4) The first virus passages, produced in MDCK cell culture (M1), served as virus seed for the first passages in Vero cells. (2, 5) The last of five consecutive Vero cell-derived virus passages (V5) served as seed for five consecutive MDCK cell-derived virus passages, of which M6 (3, 6) represents the last [4]. (7) The fourth virus passage in Vero cells represents the Vero-adapted seed virus generated by Genzel *et al.* [72]. This seed was used within the immunogenicity studies in cooperation with Dr. B. Lepenies and J. Hütter (MPI-KG, Potsdam-Golm, Germany). All sequences were generated by pyrosequencing and are deposited in the GISAID EpiFlu database (www.gisaid.org, [4]).

no.	Segment_Ids	Isolate_Name	Passage History	Seq_Id (HA)
1	EPI304412	A/MDCK/Germany/[A/Puerto Rico/8/1934]-RKI-M1 (H1N1)	P5	utf_rki-m1-ges_unambig_HA_Segment4_00001 length=1769 numreads=9121
2	EPI304420	A/MDCK/Germany/[A/Puerto Rico/8/1934]-RKI-V5 (H1N1)	P10	utf_rki-v5-ges_unambig_HA_Segment4_00001 length=1763 numreads=3748
3	EPI304428	A/MDCK/Germany/[A/Puerto Rico/8/1934]-RKI-M6 (H1N1)	P15	mutf_rki-m6-ges_unambig_HA_Segment4_00001, 1..1761 length=1760 numreads=2824
4	EPI304436	A/MDCK/Germany/[A/Puerto Rico/8/1934]-NIBSC-M1 (H1N1)	P4	mutf_nibsc-m1-ges_unambig_HA_Segment4_00001 length=1762 numreads=4279
5	EPI304444	A/MDCK/Germany/[A/Puerto Rico/8/1934]-NIBSC-V5 (H1N1)	P9	ctf_nibsc-v5-ges_unambig_HA_Segment4_00001 length=1760 numreads=3493
6	EPI304452	A/MDCK/Germany/[A/Puerto Rico/8/1934]-NIBSC-M6 (H1N1)	P14	mutf_nibsc-m6-ges_unambig_HA_Segment4_00001, 1..1759 length=1760 numreads=8589
7	EF.190974.1	A/MDCK/Germany/[A/Puerto Rico/8/1934]-RKI-V4genzel2010 (H1N1)	P7	segment 4 contig00006 gi|145322834|gb|EF190974.1|, 1..1778 length=1778 numreads=23028

table S 3: Overview of relative peak height (RPH) averages and according standard (SD) and relative standard deviations (RSD).
The average RPH and the respective SD and RSD of each peak (no. 1 - 25) within each experiment (control 1: pattern stability for ten consecutive virus passages; control 2: reproducibility in Vero or MDCK time series; adaptation series of H1N1 from RKI and adaptation series of H1N1 from NIBSC) are listed. The factors indicate x–fold increase (> 1) or decrease (< 1) of respective deviations observed during adaptation compared to the maximal deviation during controls. Factors of > 3 are defined as significantly influenced during adaptation and are highlighted in blue bold numbers. Reprinted with permission [4].

Host Cell Line	Peak No.	Pattern Stability (Control 1)*			Reproducibility in Time Series (Control 2)°			max. deviation of controls		Adaptation Series *H1N1 (Amp. 3138,* RKI)#					Adaptation Series *H1N1* (#06/114, NIBSC)#				
		$RPH_{average}$ [%]	SD [%]	RSD [%]	$RPH_{average}$ [%]	SD [%]	RSD [%]	SD [%]	RSD [%]	$RPH_{average}$ [%]	SD [%]	RSD [%]	$factor_{RSD}$	$factor_{SD}$	$RPH_{average}$ [%]	SD [%]	RSD [%]	$factor_{RSD}$	$factor_{SD}$
Vero	5	n.a.	n.a.	n.a.	4.01	0.99	24.83	0.99	24.83	4.18	1.72	41.21	1.66	1.73	2.92	1.65	56.73	2.29	1.66
	7	n.a.	n.a.	n.a.	21.76	6.06	27.83	6.06	27.83	22.34	8.40	37.62	1.35	1.39	16.77	6.29	37.52	1.35	1.04
	8	n.a.	n.a.	n.a.	2.63	0.71	27.08	0.71	27.08	5.01	1.54	30.79	1.14	2.17	2.28	0.48	20.93	0.77	0.67
	11	n.a.	n.a.	n.a.	1.09	0.26	23.97	0.26	23.97	1.35	1.06	78.51	**3.28**	**4.05**	0.50	0.30	59.93	2.50	1.16
	14	n.a.	n.a.	n.a.	12.56	0.54	4.30	0.54	4.30	8.35	2.94	35.23	**8.19**	**5.44**	9.99	1.91	19.13	**4.45**	**3.54**
	15	n.a.	n.a.	n.a.	16.41	1.51	9.19	1.51	9.19	14.66	2.33	15.89	1.73	1.55	16.42	1.40	8.53	0.93	0.93
	16	n.a.	n.a.	n.a.	3.41	1.11	32.41	1.11	32.41	4.60	1.56	33.98	1.05	1.41	4.18	1.06	25.27	0.78	0.96
	17	n.a.	n.a.	n.a.	1.00	0.52	51.99	0.52	51.99	1.17	0.75	63.93	1.23	1.43	1.08	0.28	26.14	0.50	0.54
	18	n.a.	n.a.	n.a.	2.29	0.34	15.04	0.34	15.04	2.26	1.52	66.98	**4.45**	**4.41**	1.71	0.40	23.47	1.56	1.17
	19	n.a.	n.a.	n.a.	1.79	0.14	7.76	0.14	7.76	2.09	0.27	12.73	1.64	1.92	2.88	1.03	35.90	**4.63**	**7.45**
	20	n.a.	n.a.	n.a.	9.57	1.81	18.93	1.81	18.93	6.85	2.18	31.81	1.68	1.20	9.53	1.78	18.63	0.98	0.98
	21	n.a.	n.a.	n.a.	13.00	1.53	11.75	1.53	11.75	12.72	3.06	24.08	2.05	2.01	15.27	1.34	8.80	0.75	0.88
	22	n.a.	n.a.	n.a.	3.24	0.92	28.33	0.92	28.33	4.49	0.36	8.02	0.28	0.39	3.78	0.58	15.37	0.54	0.63
	23	n.a.	n.a.	n.a.	3.74	1.34	35.77	1.34	35.77	4.27	1.89	44.33	1.24	1.41	4.83	0.54	11.20	0.31	0.40
	24	n.a.	n.a.	n.a.	1.66	0.77	46.51	0.77	46.51	2.19	0.61	27.90	0.60	0.79	3.34	0.65	19.33	0.42	0.84
	25	n.a.	n.a.	n.a.	1.83	0.95	51.96	0.95	51.96	3.47	1.96	56.43	1.09	2.06	4.52	0.76	16.74	0.32	0.80
MDCK	1	1.62	0.25	15.50	1.55	0.24	15.75	0.25	15.75	0.93	0.18	19.14	1.22	0.71	0.85	0.26	30.39	1.93	1.03
	2	11.73	0.75	6.37	10.45	0.57	5.48	0.75	6.37	10.29	1.68	16.31	2.56	2.25	6.76	1.49	22.09	**3.47**	2.00
	3	6.45	0.40	6.14	5.66	0.03	0.59	0.40	6.14	3.42	0.47	13.83	2.25	1.19	3.57	0.30	8.44	1.38	0.76
	4	6.55	0.41	6.20	6.48	0.53	8.13	0.53	8.13	6.69	1.14	17.00	2.09	2.16	4.21	0.71	16.90	2.08	1.35
	6	9.24	0.40	4.37	9.56	0.57	6.00	0.57	6.00	12.13	1.83	15.13	2.52	**3.20**	14.64	2.90	19.79	**3.30**	**5.05**
	7	3.47	0.28	8.16	3.72	0.01	0.30	0.28	8.16	2.89	0.59	20.43	2.50	2.09	3.88	2.03	52.29	**6.41**	**7.16**
	8	8.21	0.34	4.16	8.40	0.03	0.33	0.34	4.16	6.75	0.26	3.89	0.94	0.77	5.13	0.37	7.15	1.72	1.07
	9	1.65	0.38	22.91	2.22	0.19	8.72	0.38	22.91	2.31	0.46	19.71	0.86	1.20	3.47	0.41	11.80	0.52	1.08
	10	9.84	0.61	6.24	9.89	0.10	1.03	0.61	6.24	12.21	0.92	7.56	1.21	1.50	13.02	2.80	21.53	**3.45**	**4.57**
	11	7.63	0.40	5.20	7.20	0.24	3.32	0.40	5.20	5.45	0.58	10.60	2.04	1.46	5.96	1.04	17.44	**3.35**	2.62
	12	6.92	0.53	7.65	7.04	0.09	1.27	0.53	7.65	13.02	1.45	11.16	1.46	2.74	10.74	1.24	11.57	1.51	2.35
	13	7.30	0.52	7.15	7.92	0.53	6.69	0.53	7.15	5.87	1.23	20.88	2.92	2.31	6.26	1.66	26.48	**3.71**	**3.12**
	14	1.36	0.32	23.64	1.56	0.06	3.79	0.32	23.64	3.08	0.99	32.01	1.35	**3.07**	3.86	2.05	53.22	2.25	**6.38**
	15	15.69	1.14	7.28	14.90	2.34	15.69	2.34	15.69	11.27	3.18	28.22	1.80	1.36	12.28	3.85	31.32	2.00	1.64
	16	2.33	0.51	21.69	3.46	0.82	23.69	0.82	23.69	3.68	1.32	35.73	1.51	1.61	5.36	1.14	21.26	0.90	1.39

* n = 10
°n = 4 (Vero); n = 2 (MDCK)
n = 5 (Vero); n = 6 (MDCK)

12.7 cDNA consensus sequences of the RKI- and the NIBSC-strain from passage 1

12.7.1 >Segment_1_RKI_PB2

aTGGAAAGAATAAAAGAACTAAGAAATCTAATGTCGCAGTCTCGCACCCGCG
AGATACTCACAAAAACCACCGTGGACCATATGGCCATAATCAAGAAGTACAC
ATCAGGAAGACAGGAGAAGAACCCAGCACTTAGGATGAAATGGATGATGGC
AATGAAATATCCAATTACAGCAGACAAGAGGATAACGGAAATGATTCCTGAG
AGAAATGAGCAAGGACAAACTTTATGGAGTAAAATGAATGATGCCGGATCA
GACCGAGTGATGGTATCACCTCTGGCTGTGACATGGTGGAATAGGAATGGA
CCAATGACAAATACAGTTCATTATCCAAAAATCTACAAAACTTATTTTGAAAG
AGTCGAAAGGCTAAAGCATGGAACCTTTGGCCCTGTCCATTTTAGAAACCAA
GTCAAAATACGTCGGAGAGTTGACATAAATCCTGGTCATGCAGATCTCAGTG
CCAAGGAGGCACAGGATGTAATCATGGAAGTTGTTTTCCCTAACGAAGTGG
GAGCCAGGATACTAACATCGGAATCGCAACTAACGATAACCAAAGAGAAGA
AAGAAGAACTCCAGGATTGCAAAATTTCTCCTTTGATGGTTGCATACATGTT
GGAGAGAGAACTGGTCCGCAAAACGAGATTCCTCCCAGTGGCTGGTGGAA
CAAGCAGTGTGTACATTGAAGTGTTGCATTTGACTCAAGGAACATGCTGGG
AACAGATGTATACTCCAGGAGGGGAAGTGAGGAATGATGATGTTGATCAAA
GCGTGATTATTGCTGCTAGGAACATAGTGAGAAGAGCTGCAGTATCAGCAG
ACCCACTAGCATCTTTATTGGAGATGTGCCACAGCACACAGATTGGTGGAAT
TAGGATGGTAGACATCCTTAGGCAGAACCCAACAGAAGAGCAAGCCGTGGA
TATATGCAAGGCTGCAATGGGACTGAGAATTAGCTCATCCTTCAGTTTTGGT
GGATTTACATTTAAGAGAACAAGCGGATCATCAGTCAAGAGAGAGGAAGAG
GTGCTTACGGGCAATCTTCAAACATTGAAGATAAGAGTGCATGAGGGATAT
GAAGAGTTCACAATGGTtGGGAGAAGAGCAACAGCCATACTCAGAAAAGCA
ACCAGGAGATTGATTCAGCTGATAGTGAGTGGGAGAGACGAACAGTCGATT
GCCGAAGCAATAATTGTGGCCATGGTATTTTCACAAGAGGATTGTATGATAA
AAGCAGTTAGAGGTGATCTGAATTTCGTCAATAGGGCGAATCAGCGACTGA
ATCCTATGCATCAACTTTTAAGACATTTTCAGAAGGATGCGAAAGTGCTTTTT
CAAAATTGGGGAGTTGAACCTATCGACAATGTGATGGGAATGATTGGGATAT

TGCCCGACATGACTCCAAGCATCGAGATGTCAATGAGAGGAGTGAGAATCA
GCAAAATGGGTGTAGATGAGTACTCCAGCACGGAGAGGGTAGTGGTGAGC
ATTGACCGGTTTTTGAGAGTCCGGGACCAACGAGGAAATGTACTACTGTCT
CCCGAGGAGGTCAGTGAAACACAGGGAACAGAGAAACTGACAATAACTTAC
TCATCGTCAATGATGTGGGAGATTAATGGTCCTGAATCAGTGTTGGTCAATA
CCTATCAATGGATCATCAGAAACTGGGAAACTGTTAAAATTCAGTGGTCCCA
GAACCCTACAATGCTATACAATAAAATGGAATTTGAACCATTTCAGTCTTTAG
TACCTAAGGCCATTAGAGGCCAATACAGTGGGTTTGTAAGAACTCTGTTCCA
ACAAATGAGGGATGTGCTTGGGACATTTGATACCGCACAGATAATAAAACTT
CTTCCCTTCGCAGCCGCTCCACCAAAGCAAAGTAGAATGCAGTTCTCCTCAT
TTACTGTGAATGTGAGGGGATCAGGAATGAGAATACTTGTAAGGGGCAATT
CTCCTGTATTCAACTACAACAAGGCCACGAAGAGACTCACTGTTCTCGGAAA
GGATGCTGGCACTTTAACCGAAGACCCAGATGAAGGCACTGCTGGAGTGG
AGTCCGCTGTTCTGAGGGGATTCCTCATTCTGGGCAAAGAAGACAGGAGAT
ATGGGCCAGCATTAAGCATCAATGAACTGAGCAACCTTGCGAAAGGAGAGA
AGGCTAATGTGCTAATTGGGCAAGGAGACGTGGTGTTGGTAATGAAACGGA
AACGGGACTCTAGCATACTTACTGACAGCCAGACAGCGACCAAAAGAATTC
GGATGGCCATCAATTAGTGTCGAAtaGTT

12.7.2 >Segment_1_NIBSC_PB2

CGAAAGACAGGTaCAAATAaTATTCAATATGGAAAGAATAAAAGAACTAAGAA
ATCTAATGTCGCAGTCTCGCACCCGCGAGATACTCACAAAAACCACCGTGG
ACCATATGGCCATAATCAAGAAGTACACATCAGGAAGACAGGAGAAGAACC
CAGCACTTAGGATGAAATGGATGATGGCAATGAAATATCCAATTACAGCAGA
CAAGAGGATAACGGAAATGATTCCTGAGAGAAATGAGCAAGGACAAACTTT
ATGGAGTAAAATGAATGATGCCGGATCAGACCGAGTGATGGTATCACCACT
GGCTGTGACATGGTGGAATAGGAATGGACCAATAACAAATACAGTTCATTAT
CCAAAAATCTACAAAACTTATTTTGAAAGAGTCGAAAGGCTAAAGCATGGAA
CCTTTGGCCCTGTCCATTTTAGAAACCAAGTCAAAATACGTCGGAGAGTTGA
CATAAATCCTGGTCATGCAGATCTCAGTGCCAAGGAGGCACAGGATGTAAT
CATGGAAGTTGTTTTCCCTAACGAAGTGGGAGCCAGGATACTAACATCGGA

ATCGCAACTAACGATAACCAAAGAGAAGAAAGAAGAACTCCAGGATTGCAA
AATTTCTCCTTTGATGGTTGCATACATGTTGGAGAGAGAACTGGTCCGCAAA
ACGAGATTCCTCCCAGTGGCTGGTGGAACAAGCAGTGTGTACATTGAAGTG
TTGCATTTGACTCAAGGAACATGCTGGGAACAGATGTATACTCCAGGAGGG
GAAGTGAGGAATGATGATGTTGATCAAAGCTTGATTATTGCTGCTAGGAACA
TAGTGAGAAGAGCTGCAGTATCAGCAGATCCACTAGCATCTTTATTGGAGAT
GTGCCACAGCACACAGATTGGTGGAATTAGGATGGTAGACATCCTTAGGCA
GAACCCAACAGAAGAGCAAGCCGTGGATATATGCAAGGCTGCAATGGGACT
GAGAATTAGCTCATCCTTCAGTTTTGGTGGATTCACATTTAAGAGAACAAGC
GGATCATCAGTCAAGAGAGAGGAAGAGGTGCTTACGGGCAATCTTCAAACA
TTGAAGATAAGAGTGCATGAGGGATATGAAGAGTTCACAATGGTtGGGAGAA
GAGCAACAGCCATACTCAGAAAAGCAACCAGGAGATTGATTCAGCTGATAG
TGAGTGGGAGAGACGAACAGTCGATTGCCGAAGCAATAATTGTGGCCATGG
TATTTTCACAAGAGGATTGTATGATAAAAGCAGTCAGAGGTGATCTGAATTT
CGTCAATAGGGCGAATCAGCGATTGAATCCTATGCATCAACTTTTAAGACAT
TTTCAGAAGGATGCGAAAGTGCTTTTTCAAAATTGGGGAGTTGAACCTATCG
ACAATGTGATGGGAATGATTGGGATATTGCCCGACATGACTCCAAGCATCG
AGATGTCAATGAGAGGAGTGAGAATCAGCAAAATGGGTGTAGATGAGTACT
CCAGCACGGAGAGGGTAGTGGTGAGCATTGACCGTTTTTTGAGAATCCGGG
ACCAACGAGGAAATGTACTACTGTCTCCCGAGGAGGTCAGTGAAACACAGG
GAACAGAGAAACTGACAATAACTTACTCATCGTCAATGATGTGGGAGATTAA
TGGTCCTGAATCAGTGTTGGTCAATACCTATCAATGGATCATCAGAAACTGG
GAAACTGTTAAAATTCAGTGGTCCCAGAACCCTACAATGCTATACAATAAAA
TGGAATTTGAACCATTTCAGTCTTTAGTACCTAAGGCCATTAGAGGCCAATA
CAGTGGGTTTGTAAGAACTCTGTTCCAACAAATGAGGGATGTGCTTGGGAC
ATTTGATACCGCACAGATAATAAAACTTCTTCCCTTCGCAGCCGCTCCACCA
AAGCAAAGTAGAATGCAGTTCTCCTCATTTACTGTGAATGTGAGGGGATCAG
GAATGAGAATACTTGTAAGGGGCAATTCTCCTGTATTCAACTATAACAAGGC
CACGAAGAGACTCACAGTTCTCGGAAAGGATGCTGGCACTTTAACTGAAGA
CCCAGATGAAGGCACAGCTGGAGTGGAGTCCGCTGTTCTGAGGGGATTCC
TCATTCTGGGCAAAGAGGACAAGAGATATGGGCCAGCACTAAGCATCAATG

AACTGAGCAACCTTGCGAAAGGAGAGAAGGCTAATGTGCTAATTGGGCAAG
GAGACGTGGTGTTGGTAATGAAACGGAAACGGGACTCTAGCATACTTACTG
ACAGCCAGACAGCGACCAAAAGAaTTCGGATGGCCATCAATTAGTGTCGAAt
AGTTTAAAacgaCc

12.7.3 >Segment_2_RKI_PB1_PB1-F2

AcCAttGaATGGATGTCAATCCGACCtTACTTTTCTTAAAAGTGCCAGCACAAA
ATGCTATAAGCACAACTTTCCCTTATACTGGAGACCCTCCTTACAGCCATGG
GACAGGAACAGGATACACCATGGATACTGTCAACAGGACACATCAGTACTC
AGAAAAGGGAAGATGGACAACAAACACCGAAACTGGAGCACCGCAACTCAA
CCCGATTGATGGGCCACTGCCCGAAGACAATGAACCAAGTGGTTATGCCCA
AACAGATTGTGTATTGGAAGCAATGGCTTTCCTTGAGGAATCCCATCCTGGT
ATTTTTGAAAACTCGTGTATTGAAACGATGGAGGTTGTTCAGCAAACACGAG
TAGACAAGCTGACACAAGGCCGACAGACCTATGACTGGACTCTAAATAGAA
ACCAGCCTGCTGCAACAGCATTGGCCAACACAATAGAAGTGTTCAGATCAA
ATGGCCTCGCGGCCAATGAGTCTGGAAGGCTCATAGACTTCCTTAAGGATG
TAATGGAGTCAATGAACAAAGAAGAAATGGGGATCACAACTCATTTCCAGAG
AAAGAGACGGGTGAGAGACAATGTGACTAAGAAAATGATAACACAGAGAAC
AATAGGTAAAAGGAAGCAGAGATTGAACAAAAGGAGTTATCTAATTAGAGCA
TTGACCCTGAACACAATGACCAAAGATGCTGAGAGAGGGAAGCTAAAACGG
AGAGCAATTGCAACCCCAGGGATGCAAATAAGGGGGTTTGTATACTTTGTT
GAGACACTGGCAAGGAGTATATGTGAGAAACTTGAACAATCAGGGTTGCCA
GTTGGAGGCAATGAGAAGAAAGCAAAGTTGGCAAATGTTGTAAGGAAGATG
ATGACCAATTCTCAGGACACCGAACTTTCTTTCACCATCACTGGAGATAACA
CCAAATGGAACGAAAATCAGAATCCTCGGATGTTTTTGGCCATGATCACATA
TATGACCAGAAATCAGCCCGAATGGTTCAGAAATATTCTAAGTATTGCTCCA
ATAATGTTCTCAAACAAAATGGCGAGACTGGGAAAAGGGTATATGTTTGAGA
GCAAGAGTATGAAACTTAGAACTCAAATACCTGCAGAAATGCTAGCAAGCAT
TGATTTGAAATATTTCAATGATTCAACAAGAAAgaagATTGAAAAAATCCGACC
GCTCTTAATAGAGGGGACTGCATCATTGAGCCCTGGaATGATGATGGGCAT
GTTCAATATGTTAAGCACTGTATTAGGCGTCTCCATCCTGAATCTTGGACAA

AAGaGATACACCAAGACTACTTACTGGTGGGATGGTCTTCAATCCTCTGACG
ATTTTGCTCTGATTGTGAATGCACCCAATCATGAAGGGATTCAAGCCGGAGT
CGACAGGTTTTATCGAACCTGTAAGCTACTTGGAATCAATATGAGCAAGAAA
AAGTCTTACATAAACAGAACAGGTACATTTGAATTCACAAGTTTTTTCTATCG
TTATGGGTTTGTTGCCAATTTCAGCATGGAGCTTCCCAGTTTTGGGGTGTCT
GGGATCAACGAGTCAGCGGACATGAGTATTGGAGTTACTGTCATCAAAAAC
AATATGATAAACAATGATCTTGGTCCAGCAACAGCTCAAATGGCCCTTCAGT
TGTTCATCAAAGATTACAGGTACACGTACCGATGCCATAGAGGTGACACACA
AATACAAACCCGAAGATCATTTGAAATAAAGAAACTGTGGGAGCAAACCCGT
TCCAAAGCTGGACTGTTGGTCTCCGACGGAGGCCCAAATTTGTACAACATT
AGAAATCTCCACATTCCTGAAGTCTGCCTAAAATGGGAATTGATGGATGAGG
ATTACCAGGGGCGTTTATGCAACCCACTGAACCCATTTGTCAGCCATAAAGA
AATTGAATCAATGAACAATGCAGTGATGATGCCAGCACATGGTCCAGCCAAA
AACATGGAGTATGATGCTGTTGCAACAACACACTCCTGGATCCCCAAAAGAA
ATCGATCCATCTTGAATACAAGTCAAAGAGGAGTACTTGAAGATGAACAAAT
GTACCAAAGGTGCTGCAATTTATTTGAAAAATTCTTCCCCAGCAGTTCATAC
AGAAGACCAGTCGGGATATCCAGTATGGTGGAGGCTATGGTTTCCAGAGCC
CGAATTGATGCACGGATTGATTTCGAATCTGGAAGGATAAAGAAAGAAGAGT
TCACTGAGATCATGAAGATCTGTTCCACCATTGAAGAGCTCAGACGGCAAAA
A

12.7.4 >Segment_2_NIBSC_PB1_PB1-F2

GGcAaCCATTtGAATGGaTGTCAATCCGACCTTACTTTTCTTAAAaGTGCCAG
CACAAAATGCTATAAGCACAACTTTCCCTTATACTGGAGACCCTCCTTACAG
CCATGGGACAGGAACAGGATACACCATGGATACTGTCAACAGGACACATCA
GTACTCAGAAAAGGGAAGATGGACAACAAACACCGAAACTGGAGCACCGCA
ACTCAACCCGATTGATGGGCCACTGCCAGAAGACAATGAACCAAGTGGTTA
TGCCCAAACAGATTGTGTATTGGAGGCGATGGCTTTCCTTGAGGAATCCCAT
CCTGGTATTTTTGAAAACTCGTGTATTGAAACGATGGAGGTTGTTCAGCAAA
CACGAGTAGACAAGCTGACACAAGGCCGACAGACCTATGACTGGACTCTAA
ATAGAAACCAACCTGCTGCAACAGCATTGGCCAACACAATAGAAGTGTTCA

GATCAAATGGCCTCACGGCCAATGAGTCTGGAAGGCTCATAGACTTCCTTA
AGGATGTAATGGAGTCAATGAACAAAGAAGAAATGGGGATCACAACTCATTT
TCAGAGAAAGAGACGGGTGAGAGACAATATGACTAAGAAAATGATAACACA
GAGAACAATGGGTAAAAAGAAGCAGAGATTGAACAAAAGGAGTTATCTAATT
AGAGCATTGACCCTGAACACAATGACCAAAGATGCTGAGAGAGGGAAGCTA
AAACGGAGAGCAATTGCAACCCCAGGGATGCAAATAAGGGGGTTTGTATAC
TTTGTTGAGACACTGGCAAGGAGTATATGTGAGAAACTTGAACAATCAGGGT
TGCCAGTTGGAGGCAATGAGAAGAAAGCAAAGTTGGCAAATGTTGTAAGGA
AGATGATGACCAATTCTCAGGACACCGAACTTTCTTTCACCATCACTGGAGA
TAACACCAAATGGAACGAAAATCAGAATCCTCGGATGTTTTTGGCCATGATC
ACATATATGACCAGAAATCAGCCCGAATGGTTCAGAAATGTTCTAAGTATTG
CTCCAATAATGTTCTCAAACAAAATGGCGAGACTGGGAAAAGGGTATATGTT
TGAGAGCAAGAGTATGAAACTTAGAACTCAAATACCTGCAGAAATGCTAGCA
AGCATCGATTTGaaaTATTTCAATGATTCAACAAGAAAGAAGATTGAAAAAAT
CCGATCGCTCTTAATAGAGGGGACTGCATCATTGAGCCCTGGAATGATGAT
GGGCATGTTCAATATGTTAAGCACTGTATTAGGCGTCTCCATCCTGAATCTT
GGACAAAAGaGATACACCAAGACTACTTACTGGTGGGATGGTCTTCAATCCT
CTGACGATTTTGCTCTGATTGTGAATGCACCCAATCATGAAGGGATTCAAGC
CGGAGTCGACAGGTTTTATCGAACCTGTAAGCTACTTGGAATCAATATGAGC
AAGAAAAAGTCTTACATAAACAGAACAGGTACATTTGAATTCACAAGTTTTTT
CTATCGTTATGGGTTTGTTGCCAATTTCAGCATGGAGCTTCCCAGTTTTGGG
GTGTCTGGGATCAACGAGTCAGCGGACATGAGTATTGGAGTTACTGTCATC
AAAAACAATATGATAAACAATGATCTTGGTCCAGCAACAGCTCAAATGGCCC
TTCAGTTGTTCATCAAAGATTACAGGTACACGTACCGATGCCATAGAGGTGA
CACACAAATACAAACCCGAAGATCATTTGAAATAAAGAAACTGTGGGAGCAA
ACCCGTTCCAAAGCTGGACTGCTGGTCTCCGACGGAGGCCCAAATTTATAC
AACATTAGAAATCTCCACATTCCTGAAGTCTGCCTAAAATGGGAATTGATGG
ATGAGGATTACCAGGGGCGTTTATGCAACCCACTGAACCCATTTGTCAGCC
ATAAAGAAATTGAATCAATGAACAATGCAGTGATGATGCCAGCACATGGTCC
AGCCAAAAACATGGAGTATGATGCTGTTGCAACAACACACTCCTGGATCCC
CAAAAGAAATCGATCCATCTTGAATACAAGTCAAAGAGGAGTACTTGAGGAT

GAACAAATGTACCAAAGGTGCTGCAATTTATTTGAAAAATTCTTCCCCAGCA
GTTCATACAGAAGACCAGTCGGGATATCCAGTATGGTGGAGGCTATGGTTT
CCAGAGCCCGAATTGATGCACGGATTGATTTCGAATCTGGAAGGATAAAGA
AAGAAGAGTTCACTGAGATCATGAAGATCTGTTCCACCATTGAAGAGCTCAG
ACGGCAAAAATGTGAATTTAGCttt

12.7.5 >Segment_3_RKI_PA

CAaaaTGGAAGATTTTGTGCGACAATGCTTCAATCCGATGATTGTCGAGCTT
GCGGAAAAAACAATGAAAGAGTATGGGGAGGACCTGAAAATCGAAACAAAC
AAATTTGCAGCAATATGCACTCACTTGGAAGTATGCTTCATGTATTCAGACTT
TCACTTCATCAATGAGCAAGGCGAGTCAATAATCGTGGAACTTGGTGATCCA
AATGCACTTTTGAAGCACAGATTTGAAATAATCGAGGGAAGAGATCGCACAA
TGGCCTGGACAGTAGTAAACAGTATTTGCAACACTACAGGGGCTGAGAAAC
CAAAGTTTCTACCAGATTTGTATGATTACAAGGAGAATAGATTCATCGAAATT
GGAGTAACAAGGAGAGAAGTTCACATATATTATCTGGAAAAGGCCAATAAAA
TTAAATCTGAGAAAACACACATCCACATTTTCTCGTTCACTGGGGAAGAAAT
GGCCACAAAGGCAGACTACACTCTCGATGAAGAAAGCAGGGCTAGGATCAA
AACCAGACTATTCACCATAAGACAAGAAATGGCCAGCAGAGGCCTCTGGGA
TTCCTTTCGTCAGTCCGAAAGAGGGGAAGAGACAATTGAAGAAAGGTTTGA
AATCACAGGAACAATGCGCAAGCTTGCCGACCAAAGTCTCCCGCCGAACTT
CTCCAGCCTTGAAAATTTTAGAGCCTATGTGGATGGATTCAAACCGAACGGC
TACATTGAGGGCAAGCTGTCTCAAATGTCCAAAGAAGTAAATGCTAGAATTG
AACCTTTTTTGAAAACAACACCACGACCACTTAGACTTCCGAATGGGCCTCC
CTGTTCTCAGCGGTCCAAATTCCTGCTGATGGATGCCTTAAAATTAAGCATT
GAGGACCCAAGTCATGAAGGAGAGGGAATACCGCTATATGATGCAATCAAA
TGCATGAGAACATTCTTTGGATGGAAGGAACCCAATGTTGTTAAACCACACG
AAAAGGGAATAAATCCAAATTATCTTCTGTCATGGAAGCAAGTACTGGCAGA
ACTGCAGGACATTGAGAATGAGGAGAAAGTTCCAAAGACTAAAAATATGAAG
AAAACAAGTCAGCTAAAGTGGGCACTTGGTGAGAACATGGCACCAGAAAAG
GTAGACTTTGACGACTGTAAAGATGTAGGTGATTTGAAGCAATATGATAGTG
ATGAACCAGAATTGAGGTCGCTTGCAAGTTGGATTCAGAATGAGTTTAACAA

GGCATGCGAACTGACAGATTCAAGCTGGATAGAGCTCGATGAGATTGGAGA
AGATGTGGCTCCAATTGAACACATTGCAAGCATGAGAAGGAATTATTTCACA
TCAGAGGTGTCTCACTGCAGAGCCACAGAATACATAATGAAGGGGGTGTAC
ATCAATACTGCCTTGCTTAATGCATCTTGTGCAGCAATGGATGATTTCCAATT
AATTCCAATGATAAGCAAGTGTAGAACTAAGGAGGGAAGGCGAAAGACCAA
CTTGTATGGTTTCATCATAAAAGGAAGATCCCACTTAAGGAATGACACCGAC
GTGGTAAACTTTGTGAGCATGGAGTTTTCTCTCACTGACCCAAGACTTGAAC
CACACAAATGGGAGAAGTACTGTGTTCTTGAGATAGGAGATATGCTTCTAAG
AAGTGCCATAGGCCAGGTTTCAAGGCCCATGTTCTTGTATGTGAGAACAAAT
GGAACCTCAAAAATTAAAATGAAATGGGGAATGGAGATGAGGCGTTGCCTC
CTCCAGTCACTTCAACAAATTGAGAGTATGATTGAAGCTGAGTCCTCTGTCA
AAGAGAAAGACATGACCAAAGAGTTCTTTGAGAACAAATCAGAAACATGGCC
CATTGGAGAGTCCCCCAAAGGAGTGGAGGAAAGTTCCATTGGGAAGGTCTG
CAGGACTTTATTAGCAAAGTCGGTATTCAACAGCTTGTATGCATCTCCACAA
CTAGAAGGATTTTCAGCTGAATCAAGAAAACTGCTTCTTATCGTTCAGGCTC
TTAGGGACAACCTGGAACCTGGGACCTTTGATCTTGGGGGGCTATATGAAG
CAATTGAGGAGTGCCTGATTAATGATCCCTGGGTTTTGCTTAATGCTTCTTG
GTTCAACTCCTTCCTTACACATGCATTGAGTTAGTTGTGGCAGTGCTACTATT
TGCTATCCATACTGT

12.7.6 >Segment_3_NIBSC_PA

CTAaaaTGGAAGATTTTGTGCGACAATGCTTCAATCCGATGATTGTCGAGCTT
GCGGAAAAAACAATGAAAGAGTATGGGGAGGACCTGAAAATCGAAACAAAC
AAATTTGCAGCAATATGCACTCACTTGGAAGTATGCTTCATGTATTCAGATTT
TCACTTCATCAATGAGCAAGGCGAGTCAATAATCGTAGAACTTGGTGATCCA
AATGCACTTTTGAAGCACAGATTTGAAATAATCGAGGGAAGAGATCGCACAA
TGGCCTGGACAGTAGTAAACAGTATTTGCAACACTACAGGGGCTGAGAAAC
CAAAGTTTCTACCAGATTTGTATGATTACAAGGAGAATAGATTCATCGAAATT
GGAGTAACAAGGAGAGAAGTTCACATATACTATCTGGAAAAGGCCAATAAAA
TTAAATCTGAGAAAACACACATCCACATTTTCTCGTTCACTGGGGAAGAAAT
GGCCACAAAGGCAGACTACACTCTCGATGAAGAAAGCAGGGCTAGGATCAA

AACCAGACTATTCACCATAAGACAAGAAATGGCCAGCAGAGGCCTCTGGGA
TTCCTTTCGTCAGTCCGAGAGAGGAGAAGAGACAATTGAAGAAAGGTTTGA
AATCACAGGAACAATGCGCAAGCTTGCCGACCAAAGTCTCCCGCCGAACTT
CTCCAGCCTTGAAAATTTTAGAGCCTATGTGGATGGATTCGAACCGAACGG
CTACATTGAGGGCAAGCTGTCTCAAATGTCCAAAGAAGTAAATGCTAGAATT
GAACCTTTTTTGAAAACAACACCACGACCACTTAGACTTCCGAATGGGCCTC
CCTGTTCTCAGCGGTCCAAATTCCTGCTGATGGATGCCTTAAAATTAAGCAT
TGAGGACCCAAGTCATGAAGGAGAGGGAATACCGCTATATGATGCAATCAA
ATGCATGAGAACATTCTTTGGATGGAAGGAACCCAATGTTGTTAAACCACAC
GAAAAGGGAATAAATCCAAATTATCTTCTGTCATGGAAGCAAGTACTGGCAG
AACTGCAGGACATTGAGAATGAGGAGAAAATTCCAAAGACTAAAAATATGAA
GAAAACAAGTCAGCTAAAGTGGGCACTTGGTGAGAACATGGCACCAGAAAA
GGTAGACTTTGACGACTGTAAAGATGTAGGTGATTTGAAGCAATATGATAGT
GATGAACCAGAATTGAGGTCGCTTGCAAGTTGGATTCAGAATGAGTTTAACA
AGGCATGCGAACTGACAGATTCAAGCTGGATAGAGCTCGATGAGATTGGAG
AAGATGTGGCTCCAATTGAACACATTGCAAGCATGAGAAGGAATTATTTCAC
ATCAGAGGTGTCTCACTGCAGAGCCACAGAATACATAATGAAGGGGGTGTA
CATCAATACTGCCTTGCTTAATGCATCTTGTGCAGCAATGGATGATTTCCAAT
TAATTCCAATGATAAGCAAGTGTAGAACTAAGGAGGGAAGGCGAAAGACCA
ACTTGTATGGTTTCATCATAAAAGGAAGATCCCACTTAAGGAATGACACCGA
CGTGGTAAACTTTGTGAGCATGGAGTTTTCTCTCACTGACCCAAGACTTGAA
CCACATAAATGGGAGAAGTACTGTGTTCTTGAGATAGGAGATATGCTTATAA
GAAGTGCCATAGGCCAGGTTTCAAGGCCCATGTTCTTGTATGTGAGAACAA
ATGGAACCTCAAAAATTAAAATGAAATGGGGAATGGAGATGAGGCGTTGCC
TCCTCCAGTCACTTCAACAAATTGAGAGTATGATTGAAGCTGAGTCCTCTGT
CAAAGAGAAAGACATGACCAAAGAGTTCTTTGAGAACAAATCAGAAACATGG
CCCATTGGAGAGTCCCCCAAAGGAGTGGAGGAAAGTTCCATTGGGAAGGT
CTGCAGGACTTTATTAGCAAAGTCGGTATTCAACAGCTTGTATGCATCTCCA
CAACTAGAAGGATTTTCAGCTGAATCAAGAAAACTGCTTCTTATCGTTCAGG
CTCTTAGGGACAACCTTGAACCTGGGACCTTTGATCTTGGGGGGCTATATG
AAGCAATTGAGGAGTGCCTGATTAATGATCCCTGGGTTTTGCTTAATGCTTC

TTGGTTCAACTCCTTCCTTACACATGCATTGAGTTAGTTGTGGCAGTGCTAC
TATTTGCTATCCATACTGtaCCAAAAAAaG

12.7.7 >Segment_4_RKI_HA

cGTAGCGAAAGCAGGggAAAaTAAAAACAACCAAAATGAAGGCAAACCTACT
GGTCCTGTTATGTGCACTTGCAGCTGCAGATGCAGACACAATATGTATAGG
CTACCATGCGAACAATTCAACCGACACTGTTGACACAGTGCTCGAGAAGAA
TGTGACAGTGACACACTCTGTTAACCTGCTCGAAGACAGCCACAACGGAAA
ACTaTGTAGATTAAAAGGAaTAGCCCCACTACAATTGGGGAAATGTAACATC
GCCGGATGGCTCTTGGGAAACCCAGAATGCGACCCACTGCTTCCAGTGAG
ATCATGGTCCTACATTGTAGAAACACCAAACTCTGAGAATGGAATATGTTAT
CCAGGAGATTTCATCGACTATGAGGAGCTGAGGGAGCAATTGAGCTCAGTG
TCATCATTCGAAAGATTCGAAATATTTCCCAAAGAAAGCTCATGGCCCAACC
ACAACACAAACAAAGGAGTAACGGCAGCATGCTCCCATGCGGGGgAAAGCA
GTTTTTACAGAAATTTGCTATGGCTGACGGAGAAGGAGGGCTCATACCCAA
AGCTGAAAAATTCTTATGTGAACAAGAAAGGGAAAGAAGTCCTTGTACTGTG
GGGTATTCATCACCCGTCTAACAGTAAGGAGCAACAGAATATCTATCAGAAT
GAAAATGCTTATGTCTCTGTAGTGACTTCAAATTATAACAGGAGATTTACCCC
GGAAATAGCAGAAAGACCCAAAGTAAGAGATCAAGCTGGGAGGATGAACTA
TTACTGGACCTTGCTAAAACCCGGAGACACAATAATATTTGAGGCAAATGGA
AATCTAATAGCACCAAGGTATGCTTTCGCACTGAGTAGAGGCTTTGGGTCC
GGCATCATCACCTCAAACGCATCAATGCATGAGTGTAACACGAAGTGTCAAA
CACCCCTGGGAGCTATAAACAGCAGTCTCCCTTTCCAGAATATACACCCAGT
CACAATAGGAGAGTGCCCAAAATACGTCAGGAGTGCCAAATTGAGGATGGT
TACAGGACTAAGGAACATTCCGTCCATTCAATCCAGAGGTCTATTTGGAGCC
ATTGCCGGTTTTATtGAAGGGGGATGGACTGGAATGATAGATGGaTGGTACG
GTTATCATCATCAGAATGAACAGGGATCAGGCTATGCAGCGGATCAAAAAA
GCACACAAAATGCCATTAACGGGATTACAAACAAGGTGAACTCTGTTATCGA
GAAAATGAACATTCAATTCACAGCTGTGGGTAAAGAATTCAACAAATTAGAA
AAAAGGATGGAAAATTTAAATAAAAAAGTTGATGATGGATTTCTGGACATTTG
GACATATAATGCAGAATTGTTAGTTCTACTGGAAAATGAAAGGACTCTGGAT

TTCCATGACTCAAATGTGAAAAATCTGTATGAGAAAGTAAAAAGCCAATTAAA
GAATAATGCCAAAGAAATCGGAAATGGATGTTTTGAGTTCTACCACAAGTGT
GACAATGAATGCATGGAAAGTGTAAGAAATGGGACTTATGATTATCCCAAAT
ATTCAGAAGAGTCAAAGTTGAACAGGGAAAAGGTAGATGGAGTGAAATTGG
AATCAATGGGGATCTATCAGATTCTGGCGATCTACTCAACTGTCGCCAGTTC
ACTGGTGCTTTTGGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGTTCTAAT
GGATCTTTGCAGTGCAGAATATGCATCTGAGATTAGAATTTCAGAAATATGA
GGAAAAACACC

12.7.8 >Segment_4_NIBSC_HA

GTaGCgAAAGCAGGGgAAAATAAAAACAACCAAAATGAAGGCAAACCTACTG
GTCCTGTTATGTGCACTTGCAGCTGCAGATGCAGACACAATATGTATAGGCt
ACCATGCGAACAATTCAACCGACACTGTTGACACAGTACTCGAGAAGAATGT
GACAGTGACACACTCTGTTAACCTGCTCGAAGACAGCCACAACGGAAAACT
aTGTAGATTAAAAGGAaTAGCCCCACTACAATTGGGGAAATGTAACATCGCC
GGATGGCTCTTGGGAAACCCAGAATGCGACCCACTGCTTCCAGTGAGATCA
TGGTCCTACATTGTAGAAACACCAAACTCTGAGAATGGAATATGTTATCCAG
GAGATTTCATCGACTATGAGGAGCTGAGGGAGCAATTGAGCTCAGTGTCAT
CATTCGAAAGATTCGAAATATTTCCCAAAGAAAGCTCATGGCCCAACCACAA
CACAAACGGAGTAACGGCAGCATGCTCCCATGAGGGGAAAAGCAGTTTTTA
CAGAAATTTGCTATGGCTGACGGAGAAGGAGGGCTCATACCCAAAGCTGAA
AAATTCTTATGTGAACAAAAAaGGGAAAGAAGTCCTTGTACTGTGGGGTATT
CATCACCCGTCTAACAGTAAGGAACAACAGAATCTCTATCAGAATGAAAATG
CTTATGTCTCTGTAGTGACTTCAAATTATAACAGGAGATTTACCCCGGAAATA
GCAGAAAGACCCAAAGTAAGAGATCAAGCTGGGAGGATGAACTATTACTGG
ACCTTGCTAAAACCCGGAGACACAATAATATTTGAGGCAAATGGAAATCTAA
TAGCACCAATGTATGCTTTCGCACTGAGTAGAGGCTTTGGGTCCGGCATCA
TCACCTCAAACGCATCAATGCATGAGTGTAACACGAAGTGTCAAACACCCCT
GGGAGCTATAAACAGCAGTCTCCCTTACCAGAATATACACCCAGTCACAATA
GGAGAGTGCCCAAAATACGTCAGGAGTGCCAAATTGAGGATGGTTACAGGA
CTAAGGAACATTCCGTCCATTCAATCCAGAGGTCTATTtGGAGCCATTGCCG

GTTTTATtGAAGGGGGatGGACTGGAATGATAGATGGATGGTATGGTTATCAT
CATCAGAATGAACAGGGATCAGGCTATGCAGCGGATCAAAAAAGCACACAA
AATGCCATTAACGGGATTACAAACAAGGTGAACACTGTTATCGAGAAAATGA
ACATTCAATTCACAGCTGTGGGTAAAGAATTCAACAAATTAGAAAAAAGGAT
GGAAAATTTAAATAAAAAaGTTGATGATGGATTTCTGGACATTTGGACATATA
ATGCAGAATTGTTAGTTCTACTGGAAAATGAAAGGACTCTGGATTTCCATGA
CTCAAATGTGAAGAATCTGTATGAGAAAGTAAAAAGCCAATTAAAGAATAAT
GCCAAAGAAaTCGGAAATGGATGTTTTGAGTTCTACCACAAGTGTGACAATG
AATGCATGGAAAGTGTAAGAAATGGGACTTATGATTATCCCAAATATTCAGA
AGAGTCAAAGTTGAACAGGgAAAaGGTAGATGGAGTGAAATTGGAATCAATG
GGGATCTATCAGATTCTGGCGATCTACTCAACTGTCGCCAGTTCACTGGTG
CTTTTGGTCTCCCTGGGGGCAATCAGTTTCTGGATGTGTTCTAATGGATCTT
TGCAGTGCAGAATATGCATCTGAGATTAGAATTTCAGAGATATGAGGAAAAA
Cn

12.7.9 >Segment_5_RKI_NP

CAACATCATGGCGTCCCAAGGCACCAAACGGTCTTACGAACAGATGGAGAC
TGATGGAGAACGCCAGAATGCCACTGAAATCAGAGCATCCGTCGGAAAAAT
GATTGGTGGAATTGGACGATTCTACATCCAAATGTGCACAGAACTTAAACTC
AGTGATTATGAGGGACGGTTGATCCAAAACAGCTTAACAATAGAGAGAATG
GTGCTCTCTGCTTTTGACGAAAGGAGAAATAAATACCTGGAAGAACATCCCA
GTGCGGGGAAAGATCCTAAGAAAACTGGAGGACCTATATACAGAAGAGTAA
ACGGAAAGTGGATGAGAGAACTCATCCTTTATGACAAAGAAGAAATAAGGC
GAATCTGGCGCCAAGCTAATAATGGTGACGATGCAACGGCTGGTCTGACTC
ACATGATGATCTGGCATTCCAATTTGAATGATGCAACTTATCAGAGGACAAG
AGCTCTTGTTCGCACTGGAATGGATCCCAGGATGTGCTCTCTGATGCAAGG
TTCAACTCTCCCTAGGAGGTCTGGAGCCGCAGGTGCTGCAGTCAAAGGAGT
TGGAACAATGGTGATGGAATTGGTCAGGATGATCAAACGTGGGATCAATGA
TCGGAACTTCTGGAGGGGTGAGAATGGACGAAAAACAAGAATTGCTTATGA
AAGAATGTGCAACATTCTCAAAGGGAAATTTCAAACTGCTGCACAAAAAGCA
ATGATGGATCAAGTGAGAGAGAGCCGGAACCCAGGGAATGCTGAGTTCGA

AGATCTCACTTTTCTAGCACGGTCTGCACTCATATTGAGAGGGTCGGTTGCT
CACAAGTCCTGCCTGCCTGCCTGTGTGTATGGACCTGCCGTAGCCAGTGGG
TACGACTTTGAAAGAGAGGGATACTCTCTAGTCGGAATAGACCCTTTCAGAC
TGCTTCAAAACAGCCAAGTGTACAGCCTAATCAGACCAAATGAGAATCCAGC
ACACAAGAGTCAACTGGTGTGGATGGCATGCCATTCTGCCGCATTTGAAGA
TCTAAGAGTATTGAGCTTCATCAAAGGGACGAAGGTGGTCCCAAGAGGGAA
GCTTTCCACTAGAGGAGTTCAAATTGCTTCCAATGAAAATATGGAGACTATG
GAATCAAGTACACTTGAACTGAGAAGCAGGTACTGGGCCATAAGGACCAGA
AGTGGAGGAAACACCAATCAACAGAGGGCATCTGCGGGCCAAATCAGCATA
CAACCTACGTTCTCAGTACAGAGAAATCTCCCTTTTGACAGAACAACCGTTA
TGGCAGCATTCACTGGGAATACAGAGGGGAGAACATCTGACATGAGGGCC
GAAATCATAAGGATGATGGAAAGTGCAAGACCAGAAGATGTGTCTTTCCAG
GGGCgGGGAGTCTTCGAGCTCTCGGACGAAAAGGCAGCGAGCCCGATCGT
GCCTTCCTTTGACATGAGTAAtGAAGGATCTTATTTCTTCGGAGACAATGCAG
AGGA

12.7.10 >Segment_5_NIBSC_NP

TGGCGTCTCAAGGCACCAAACGATCTTACGAACAGATGGAGACTGATGGAG
AACGCCAGAATGCCACTGAAATCAGAGCATCCGTCGGAAAAATGATTGGTG
GAATTGGACGATTCTACATCCAAATGTGCACCGAACTCAAACTCAGTGATTA
TGAGGGACGGTTGATCCAAAACAGCTTAACAATAGAGAGAATGGTGCTCTC
TGCTTTTGACGAAAGGAGAAATAAATACCTTGAAGAACATCCCAGTGCGGG
GAAAGATCCTAAGAAAACTGGAGGACCTATATACAGGAGAGTAAACGGAAA
GTGGATGAGAGAACTCATCCTTTATGACAAAGAAGAAATAAGGCGAATCTG
GCGCCAAGCTAATAATGGTGACGATGCAACGGCTGGTCTGACTCACATGAT
GATCTGGCATTCCAATTTGAATGATGCAACTTATCAGAGGACAAGAGCTCTT
GTTCGCACCGGAATGGATCCCAGGATGTGCTCTCTGATGCAAGGTTCAACT
CTCCCTAGGAGGTCTGGAGCCGCAGGTGCTGCAGTCAAAGGAGTTGGAAC
AATGGTGATGGAATTGGTCAGAATGATCAAACGTGGGATCAATGATCGGAA
CTTCTGGAGGGGTGAGAATGGACGAAAAACAAGAATTGCTTATGAAAGAAT
GTGCAACATTCTCAAAGGGAAATTTCAAACTGCTGCACAAAAAGCAATGATG

GATCAAGTGAGAGAGAGCCGGAACCCAGGGAATGCTGAGTTCGAAGATCT
CACTTTTCTAGCACGGTCTGCACTCATATTGAGAGGGTCGGTTGCTCACAAG
TCCTGCCTGCCTGCCTGTGTGTATGGACCTGCCGTAGCCAGTGGGTACGAC
TTTGAAAGGGAGGGATACTCTCTAGTCGGAATAGACCCTTTCAGACTGCTTC
AAAACAGCCAAGTGTACAGCCTAATCAGACCAAATGAGAATCCAGCACACA
AGAGTCAACTGGTGTGGATGGCATGCCATTCTGCCGCATTTGAAGATCTAA
GAGTATTAAGCTTCATCAAAGGGACGAAGGTGCTCCCAAGAGGGAAGCTTT
CCACTAGAGGAGTTCAAATTGCTTCCAATGAAAATATGGAGACTATGGAATC
AAGTACACTTGAACTGAGAAGCAGGTACTGGGCCATAAGGACCAGAAGTGG
AGGAAACACCAATCAACAGAGGGCATCTGCGGGCCAAATCAGCATACAACC
TACGTTCTCAGTACAGAGAAATCTCCCTTTTGACAGAACAACCATTATGGCA
GCATTCAATGGGAATACAGAGGGGAGAACATCTGACATGAGGACCGAAATC
ATAAGGATGATGGAAAGTGCAAGaCCAGAAGATGTGTCTTTCCAGGGGCgG
GGAGTCTTCGAGCTCTCGGACGAAAAGGCAGCGAGCCCGATCGTGCCTTC
CtTTGACATGAGTAATGAAGGATCTTATTTCTTCGGAGACAATGCAGAGGA

12.7.11 >Segment_6_RKI_NA

CGTAGCGaaAgcagGGGTTTaAAATGAATCCAAACCAGaAAATAATAACCATTG
GATCAATCTGTATGGTAGTCGGACTAATTAGCCTAATATTGCAAATAGGGAA
TATAATCTCAATATGGATTAGCCATTCAATTCAAACTGGAAGTCAAAACCATA
CTGGAATATGCAACCAAAACATCATTACCTATAAAAATAGCACCTGGGTAAA
GGACACAACTTCAGTGATATTAACCGGCAATTCATCTCTTTGTCCCATCCGT
GGGTGGGCTATATACAGCAAAGACAATAGCATAAGAATTGGTTCCAAAGGA
GACGTTTTTGTCATAAGAGAGCCCTTTATTTCATGTTCTCACTTGGAATGCAG
GACCtTTTTTCTGACCCAAGGTGCCTTACTGAATGACAGGCATTCAAATGGG
ACTGTTAAGGACAGAAGCCCTTATAGGGCCTTAATGAGCTGCCCTGTCGGT
GAAGCTCCGTCCCCGTACAATTCAAGATTTGAATCGGTTGCTTGGTCAGCAA
GTGCATGTCATGATGGCATGGGCTGGCTAACAATCGGAATTTCAGGTCCAG
ATAATGGAGCAGTGGCTGTATTAAAATACAACGGCATAATAACTGAAACCAT
AAAAAGTTGGAGGAAGAAAATATTGAGGACACAAGAGTCTGAATGTGCCTG
TGTAAATGGTTCATGTTTTACTATAATGACTGATGGCCCGAGTGATGGGCTG

GCCTCGTACAAAATTTTCAAGATCGAAAAGGGGAAGGTTACTAAATCAATAG
AATTGAATGCACCTAATTCTCACTATGAGGAATGTTCCTGTTACCCTGATACC
GGCAAAGTGATGTGTGTGTGCAGAGACAATTGGCATGGTTCGAACCGGCCA
TGGGTGTCTTTCGATCAAAACCTGGATTATCAAATAGGATACGTCTGCAGTG
GGGTTTTCGGTGACAACCCGCGTCCCAAAGATGGAACAGGCAGCTGTGGT
CCAGTGTATGTTGATGGAGCAAACGGAGTAAAGGGATTTTCATATAGGTATG
GTAATGGTGTtTGGATAGGAAGGACCAAAAGTCACAGTTCCAGGCATGGGTT
TGAGATGATTTGGGATCCTAATGGATGGACAGAGACTGATAGTGAGTTCTCT
GTGAGGCAAGATGTTGTGGCAATGACTGATTGGTCAGGGTATAGCGGAAGT
TTCGTTCAACATCCTGAGCTAACAGGGCTAGACTGTATAAGGCCGTGCTTCT
GGGTTGAATTAATCAGGGGACGACCTAAAGaAAAAACAATCTGGACCAGTG
CGAGCAGCATTTCtTTTTGTGGCGTGGATAGTGATACTGTAGATTGGTCTTG
GCCAGACGGTGCTGAGTTGCCATTCACCATTGACAAGTAGTCTGTTCAAAAA
AACtT

12.7.12 >Segment_6_NIBSC_NA

GGGGtTTaAAATGAATCCAAATCAGAAAATAATAACCATTGGATCAATCTGTC
TGGTAGTCGGACTAATTAGCCTAATATTGCAAATAGGGAATATAATCTCAATA
TGGATTAGCCATTCAATTCAAACTGGAAGTCAAAACCATACTGGAATATGCA
ACCAAAACATCATTACCTATAAAAATAGCACCTGGGTAAAGGACACAACTTC
AGTGATATTAACCGGCAATTCATCTCTTTGTCCCATCCGTGGGTGGGCTATA
TACAGCAAAGACAATAGCATAAGAATTGGTTCCAAAGGAGACGTTTTTGTCA
TAAGAGAGCCCTTTATTTCATGTTCTCACTTGGAATGCAGGACCtTTTTTCTG
ACCCAAGGTGCCTTACTGAATGACAAGCATTCAAGTGGGACTGTTAAGGAC
AGAAGCCCTTATAGGGCCTTAATGAGCTGCCCTGTCGGTGAAGCTCCGTCC
CCGTACAATTCAAGATTTGAATCGGTTGCTTGGTCAGCAAGTGCATGTCATG
ATGGCATGGGCTGGCTAACAATCGGAATTTCAGGTCCAGATAATGGAGCAG
TGGCTGTATTAAAATACAACGGCATAATAACTGAAACCATAAAAAGTTGGAG
GAAGAAAATATTGAGGACACAAGAGTCTGAATGTGCCTGTGTAAATGGTTCA
TGTTTTACTATAATGACTGATGGCCCGAGTGATGGGCTGGCCTCGTACAAAA
TTTTCAAGATCGAAAAGGGGAAGGTTACTAAATCAATAGAGTTGAATGCACC

TAATTCTCACTATGAGGAATGTTCCTGTTACCCTGATACCGGCAAAGTGATG
TGTGTGTGCAGAGACAATTGGCATGGTTCGAACCGGCCATGGGTGTCTTTC
GATCAAAACCTGGATTATCAAATAGGATACATCTGCAGTGGGGTTTTCGGTG
ACAACCCGCGTCCCGAAGATGGAACAGGCAGCTGTGGTCCAGTGTATGTTG
ATGGAGCAAACGGAGTAAAGGGAtTTTCATATAGGTATGGTAATGGTGtTTGG
ATAGGAAGGACCAAAAGTCACAGTTCCAGACATGGGTTTGAGATGATTTGG
GATCCTAATGGATGGACAGAGACTGATAGTAAGTTCTCTGTGAGGCAAGAT
GTTGTGGCAATGACTGATTGGTCAGGGTATAGCGGAAGtTTCGTTCAACATC
CTGAGCTGACAGGGCTAGACTGTATGAGGCCGTGCTTCTGGGTTGAATTAA
TCAGGGGACGACCTAAAGAAAAAACAATCTGGACTAGTGCGAGCAGCATTT
CtTTTTGTGGCGTGAATAGTGATACTGTAGATTGGTCTTGGCCAGACGGTGC
TGAGTTGCCATTCAGCATTGACAAGTAGTC

12.7.13 >Segment_7_RKI_M1_M2

GTTGAAAGATGAGTCTTCTAACCGAGGTCGAAACGTACGTTCTCTCTATCAT
CCCGTCAGGCCCCCTCAAAGCCGAGATCGCACAGAGACTTGAAGATGTCTT
TGCAGGGAAGAACACCGATCTTGAGGTTCTCATGGAATGGCTAAAGACAAG
ACCAATCCTGTCACCTCTGACTAAGGGGATTTTAGGATTTGTGTTCACGCTC
ACCGTGCCCAGTGAGCGAGGACTGCAGCGTAGACGCTTTGTCCAAAATGC
CCTTAATGGGAACGGGGATCCAAATAACATGGACAAAGCAGTTAAACTGTAT
AGGAAGCTCAAGAGGGAGATAACATTCCATGGGGCCAAAGAAATCTCACTC
AGTTATTCTGCTGGTGCACTTGCCAGTTGTATGGGCCTCATATACAACAGGA
TGGGGGCTGTGACCACTGAAGTGGCATTTGGCCTAGTATGTGCAACCTGTG
AACAGATTGCTGACTCCCAGCATCGGTCTCATAGGCAAATGGTGACAACAA
CCAATCCACTAATCAGACATGAGAACAGAATGGTTTTAGCCAGCACTACAGC
TAAGGCTATGGAGCAAATGGCTGGATCGAGTGAGCAAGCAGCAGAGGCCA
TGGAGGTTGCTAGTCAGGCTAGGCAAATGGTGCAAGCGATGAGAACCATTG
GGACTCATCCTAGCTCCAGTGCTGGTCTGaAAAATGATCTTCTTGAAAATTT
GCAGGCCTATCAGAAACGAATgGGGGTGCAGATGCAACGGTTCAAGTGATC
CTCTCGCTATTGCCGCAAATATCATTGGGATCTTGCACTTGATATTGTGGAT
TCTTGATCGTCTTTTTTTCAAATGCATTTACCGTCGCTTTAAATACGGACTGA

AAGGAGGGCCTTCTACGGAAGGAGTGCCAAAGTCTATGAGGGAAGAATATC
GAAAGGAACAGCAGAGTGCTGTGGATGCTGACGATGGTCATTTTGtCAGCAT
AGAGCT

12.7.14 >Segment_7_NIBSC_M1_M2

gtTGaAAGaTGAGTCTTCTAACCGAGGTCGAAACGTACGTTCTCTCTATCATC
CCGTCAGGCCCCCTCAAAGCCGAGATCGCGCAGAGACTTGAAAATGTCTTT
GCAGGGaAAAACACCGATCTTGAGGTTCTCATGGAATGGCTAAAGACAAGA
CCAATCCTGTCACCTCTGACTAAGGGGATTTTAGGATTTGTGTTCACGCTCA
CCGTGCCCAGTGAGCAGGGACTGCAGCGTAGACGCTTTGTCCAAAATGCC
CTTAATGGGAACGGGGATCCAAATAACATGGACAAAGCAGTTAAACTGTATA
GGAAGCTTAAGAGGGAGATAACATTCTATGGGGCTAAAGAAATAGCACTCA
GTTATTCCGCTGGTGCACTTGCCTGTTGTATGGGCCTCATATACAACAGGAT
GGGGACTGTGACTGCTGAAGTGGTGTTTGGCCTGGTATGTGCAACCTGTGA
ACAGATTGCTGACTCCCAGCATCGGTCTCATAGGCAAATGGTGGCAACAAC
CAATCCACTAATCAGACATGAGAACAGAATGGTTTTAGCCAGCACTACAGCT
AAGGCTATGGAGCAAATGGCTGGATCGAGTGAGCAAGCAGCAGAGGCCAT
GGAGGTTGCGATTCGGGCTAGGCAAATGGTGCAGGCAATGAGAACCATTG
GGACTCATCCTAGCTCCAGTGCTGGTCTGAAAGATGATCTTCTTGAAAATTT
GCAGGCCTATCAGAAACGAATGGGGGTGCAGATGCAACGGTTCAAGTGATC
CTCTCGTTATTGCCTCAAGTATCATTGGGATCTTGCACTTGATATTGTGGATT
CTTGATCGTCTTTTTTTCAAATGCATTTACCGTCTCTTTAAATACGGTTTGAAA
AGAGGGCCTTCTACGGAAGGAGTGCCAAAGTCTATGAGGGAAGAATATCAA
AAGGAACAGCAGAGTGCTGTGGATGCTGACGATGGTCATTTTGTCAGCATA
GAGCTGGAGTAAAA

12.7.15 >Segment_8_RKI_NS1_NS2

GTGACAAAGACATAATGGATCCAAACACTGTGTCAAGCTTTCAGGTAGATTG
CTTTCTTTGGCATGTCCGCAAACGAGTTGCAGACCAAGAACTAGGTGATGC
CCCATTCCTTGATCGGCTTCGCCGAGATCAGAAATCCCTAAGAGGAAGGGG
CAGCACTCTCGGTCTGGACATCGAGACAGCCACACGTGCTGGAAAGCAGAT
AGTGGAGCGGATTCTGAAAGAAGAATCCGATGAGGCACTTAAAATGACCAT

GGCCTCTGTACCTGCGTCGCGTTACCTAACTGACATGACTCTTGAGGAAAT
GTCAAGGGAATGGTCCATGCTCATACCCAAGCAGAAAGTGGCAGGCCCTCT
TTGTATCAGAATGGACCAGGCGATCATGGATAAGAACATCATACTGAAAGC
GAACTTCAGTGTGATTTTTGACCGGCTGGAGACTCTAATATTGCTAAGGGCT
TTCACCGAAGAGGGAGCAATTGTTGGCGAAATTTCACCATTGCCTTCTCTTC
CAGGACATACTGCTGAGGATGTCAAAAATGCAGTTGGAGTCCTCATCGGAG
GACTTGAATGGAATGATAACACAGTTCGAGTCTCTGAAACTCTACAGAGATT
CGCTTGGAGAAGCAGTAATGAGAATGGGAGACCTCCACTCACTCCAAAACA
GAAACGAGAAATGGCGGGAACAATTAGGTCAGAAGTTTGAAGAAATAAGAT
GGTTGATTGAAGAAGTGAGACACAAACTGAGGGTAACAGAGAATAGTTTTG
AGCAAATAACATTTATGCAAGCCTTACATCTATTGCTTGAAGTGGAGCAAGA
GATAAGAACTTTCTCATTTCAGCTTATTTAATAA

12.7.16 >Segment_8_NIBSC_NS1_NS2

GGgtGaCaGAcaTAaTGGATCCAAACACTGTGTCAAGCTTTCAGGTAGATTGC
TTTCTTTGGCATGTCCGCAAACGAGTTGCAGACCAAGAACTAGGTGATGCC
CCATTCCTTGATCGGCTTCGCCGAGATCAGAAATCCCTAAGAGGAAGGGGC
AGTACTCTCGGTCTGGACATCAAGACAGCCACACGTGCTGGAAAGCAGATA
GTGGAGCGGATTCTGAAAGAAGAATCCGATGAGGCACTTAAAATGACCATG
GCCTCTGTACCTGCGTCGCGTTACCTAACTGACATGACTCTTGAGGAAATGT
CAAGGGACTGGTCCATGCTCATACCCAAGCAGAAAGTGGCAGGCCCTCTTT
GTATCAGAATGGACCAGGCGATCATGGATAAGAACATCATACTGAAAGCGA
ACTTCAGTGTGATTTTTGACCGGCTGGAGACTCTAATATTGCTAAGGGCTTT
CACCGAAGAGGGAGCAATTGTTGGCGAAATTTCACCATTGCCTTCTCTTCCA
GGACATACTGCTGAGGATGTCAAAAATGCAGTTGGAGTCCTCATCGGAGGA
CTTGAATGGAATGATAACACAGTTCGAGTCTCTGAAACTCTACAGAGATTCG
CTTGGAGAAGCAGTAATGAGAATGGGAGACCTCCACTCACTCCAAAACAGA
AACGAGAAATGGCGGGAACAATTAGGTCAGAAGTTTGAAGAAATAAGATGG
TTGATTGAAGAAGTGAGACACAAACTGAAGATAACAGAGAATAGTTTTGAGC
AAATAACATTTATGCAAGCCTTACATCTATTGCTTGAAGTGGAGCAAGAGAT
AAGAACTTTCTCGTTTCAGCTTATTTAGTACTAA

12.8 Alignment of amino acid consensus sequences of the RKI- and the NIBSC-strain from passage 1

```
1    MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD   60  Segment_1_RKI_PB2
1    MERIKELRNLMSQSRTREILTKTTVDHMAIIKKYTSGRQEKNPALRMKWMMAMKYPITAD   60  Segment_1_NIBSC_PB2
     ************************************************************

61   KRITEMIPERNEQGQTLWSKMNDAGSDRVMVSPLAVTWWNRNGPMTNTVHYPKIYKTYFE  120  Segment_1_RKI_PB2
61   KRITEMIPERNEQGQTLWSKMNDAGSDRVMVSPLAVTWWNRNGPITNTVHYPKIYKTYFE  120  Segment_1_NIBSC_PB2
     ******************************************:***************

121  RVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE  180  Segment_1_RKI_PB2
121  RVERLKHGTFGPVHFRNQVKIRRRVDINPGHADLSAKEAQDVIMEVVFPNEVGARILTSE  180  Segment_1_NIBSC_PB2
     ************************************************************

181  SQLTITKEKKEELQDCKISPLMVAYMLERELVRKTRFLPVAGGTSSVYIEVLHLTQGTCW  240  Segment_1_RKI_PB2
181  SQLTITKEKKEELQDCKISPLMVAYMLERELVRKTRFLPVAGGTSSVYIEVLHLTQGTCW  240  Segment_1_NIBSC_PB2
     ************************************************************

241  EQMYTPGGEVRNDDVDQSVIIAARNIVRRAAVSADPLASLLEMCHSTQIGGIRMVDILRQ  300  Segment_1_RKI_PB2
241  EQMYTPGGEVRNDDVDQSLIIAARNIVRRAAVSADPLASLLEMCHSTQIGGIRMVDILRQ  300  Segment_1_NIBSC_PB2
     ******************:*****************************************

301  NPTEEQAVDICKAAMGLRISSSFSFGGFTFKRTSGSSVKREEEVLTGNLQTLKIRVHEGY  360  Segment_1_RKI_PB2
301  NPTEEQAVDICKAAMGLRISSSFSFGGFTFKRTSGSSVKREEEVLTGNLQTLKIRVHEGY  360  Segment_1_NIBSC_PB2
     ************************************************************

361  EEFTMVGRRATAILRKATRRLIQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF  420  Segment_1_RKI_PB2
361  EEFTMVGRRATAILRKATRRLIQLIVSGRDEQSIAEAIIVAMVFSQEDCMIKAVRGDLNF  420  Segment_1_NIBSC_PB2
     ************************************************************

421  VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEPIDNVMGMIGILPDMTPSIEMSMRGVRI  480  Segment_1_RKI_PB2
421  VNRANQRLNPMHQLLRHFQKDAKVLFQNWGVEPIDNVMGMIGILPDMTPSIEMSMRGVRI  480  Segment_1_NIBSC_PB2
     ************************************************************

481  SKMGVDEYSSTERVVVSIDRFLRVRDQRGNVLLSPEEVSETQGTEKLTITYSSSMMWEIN  540  Segment_1_RKI_PB2
481  SKMGVDEYSSTERVVVSIDRFLRIRDQRGNVLLSPEEVSETQGTEKLTITYSSSMMWEIN  540  Segment_1_NIBSC_PB2
     **********************:*************************************

541  GPESVLVNTYQWIIRNWETVKIQWSQNPTMLYNKMEFEPFQSLVPKAIRGQYSGFVRTLF  600  Segment_1_RKI_PB2
541  GPESVLVNTYQWIIRNWETVKIQWSQNPTMLYNKMEFEPFQSLVPKAIRGQYSGFVRTLF  600  Segment_1_NIBSC_PB2
     ************************************************************

601  QQMRDVLGTFDTAQIIKLLPFAAAPPKQSRMQFSSFTVNVRGSGMRILVRGNSPVFNYNK  660  Segment_1_RKI_PB2
601  QQMRDVLGTFDTAQIIKLLPFAAAPPKQSRMQFSSFTVNVRGSGMRILVRGNSPVFNYNK  660  Segment_1_NIBSC_PB2
     ************************************************************

661  ATKRLTVLGKDAGTLTEDPDEGTAGVESAVLRGFLILGKEDRRYGPALSINELSNLAKGE  720  Segment_1_RKI_PB2
661  ATKRLTVLGKDAGTLTEDPDEGTAGVESAVLRGFLILGKEDKRYGPALSINELSNLAKGE  720  Segment_1_NIBSC_PB2
     ******************************************:*****************

721  KANVLIGQGDVVLVMKRKRDSSILTDSQTATKRIRMAIN  759  Segment_1_RKI_PB2
721  KANVLIGQGDVVLVMKRKRDSSILTDSQTATKRIRMAIN  759  Segment_1_NIBSC_PB2
     ***************************************
```

figure S 5: Alignment of AA consensus sequences of segment 1 coding for PB2 of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1    MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKGRWTTNTE   60  Segment_2_RKI_PB1
1    MDVNPTLLFLKVPAQNAISTTFPYTGDPPYSHGTGTGYTMDTVNRTHQYSEKGRWTTNTE   60  Segment_2_NIBSC_PB1
     ************************************************************

61   TGAPQLNPIDGPLPEDNEPSGYAQTDCVLEAMAFLEESHPGIFENSCIETMEVVQQTRVD  120  Segment_2_RKI_PB1
61   TGAPQLNPIDGPLPEDNEPSGYAQTDCVLEAMAFLEESHPGIFENSCIETMEVVQQTRVD  120  Segment_2_NIBSC_PB1
     ************************************************************

121  KLTQGRQTYDWTLNRNQPAATALANTIEVFRSNGLAAANESGRLIDFLKDVMESMNKEEMG  180  Segment_2_RKI_PB1
121  KLTQGRQTYDWTLNRNQPAATALANTIEVFRSNGLTANESGRLIDFLKDVMESMNKEEMG  180  Segment_2_NIBSC_PB1
     ***********************************:************************

181  ITTHFQRKRRVRDNVTKKMITQRTIGKRKQRLNKRSYLIRALTLNTMTKDAERGKLKRRA  240  Segment_2_RKI_PB1
181  ITTHFQRKRRVRDNMTKKMITQRTMGKKKQRLNKRSYLIRALTLNTMTKDAERGKLKRRA  240  Segment_2_NIBSC_PB1
     **************:*********:**:*********************************

241  IATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKAKLANVVRKMMTNSQDTELSF  300  Segment_2_RKI_PB1
241  IATPGMQIRGFVYFVETLARSICEKLEQSGLPVGGNEKKAKLANVVRKMMTNSQDTELSF  300  Segment_2_NIBSC_PB1
     ************************************************************

301  TITGDNTKWNENQNPRMFLAMITYMTRNQPEWFRNILSIAPIMFSNKMARLGKGYMFESK  360  Segment_2_RKI_PB1
301  TITGDNTKWNENQNPRMFLAMITYMTRNQPEWFRNVLSIAPIMFSNKMARLGKGYMFESK  360  Segment_2_NIBSC_PB1
     ***********************************:************************

361  SMKLRTQIPAEMLASIDLKYFNDSTRKKIEKIRPLLIEGTASLSPGMMMGMFNMLSTVLG  420  Segment_2_RKI_PB1
361  SMKLRTQIPAEMLASIDLKYFNDSTRKKIEKIRSLLIEGTASLSPGMMMGMFNMLSTVLG  420  Segment_2_NIBSC_PB1
     ********************************* ***************************

421  VSILNLGQKRYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVDRFYRTCKLLGINMSKK  480  Segment_2_RKI_PB1
421  VSILNLGQKRYTKTTYWWDGLQSSDDFALIVNAPNHEGIQAGVDRFYRTCKLLGINMSKK  480  Segment_2_NIBSC_PB1
     ************************************************************

481  KSYINRTGTFEFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKNNMINNDLG  540  Segment_2_RKI_PB1
481  KSYINRTGTFEFTSFFYRYGFVANFSMELPSFGVSGINESADMSIGVTVIKNNMINNDLG  540  Segment_2_NIBSC_PB1
     ************************************************************

541  PATAQMALQLFIKDYRYTYRCHRGDTQIQTRRSFEIKKLWEQTRSKAGLLVSDGGPNLYN  600  Segment_2_RKI_PB1
541  PATAQMALQLFIKDYRYTYRCHRGDTQIQTRRSFEIKKLWEQTRSKAGLLVSDGGPNLYN  600  Segment_2_NIBSC_PB1
     ************************************************************

601  IRNLHIPEVCLKWELMDEDYQGRLCNPLNPFVSHKEIESMNNAVMMPAHGPAKNMEYDAV  660  Segment_2_RKI_PB1
601  IRNLHIPEVCLKWELMDEDYQGRLCNPLNPFVSHKEIESMNNAVMMPAHGPAKNMEYDAV  660  Segment_2_NIBSC_PB1
     ************************************************************

661  ATTHSWIPKRNRSILNTSQRGVLEDEQMYQRCCNLFEKFFPSSSYRRPVGISSMVEAMVS  720  Segment_2_RKI_PB1
661  ATTHSWIPKRNRSILNTSQRGVLEDEQMYQRCCNLFEKFFPSSSYRRPVGISSMVEAMVS  720  Segment_2_NIBSC_PB1
     ************************************************************

721  RARIDARIDFESGRIKKEEFTEIMKICSTIEELRRQK  757  Segment_2_RKI_PB1
721  RARIDARIDFESGRIKKEEFTEIMKICSTIEELRRQK  757  Segment_2_NIBSC_PB1
     ************************************
```

figure S 6: Alignment of AA consensus sequences of segment 2 coding for PB1 of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1    MGQEQDTPWILSTGHISTQKREDGQQTPKLEHRNSTRLMGHCPKTMNQVVMPKQIVYWKQ   60  Segment_2_RKI_PB1-F2
1    MGQEQDTPWILSTGHISTQKREDGQQTPKLEHRNSTRLMGHCQKTMNQVVMPKQIVYWRR   60  Segment_2_NIBSC_PB1-F2
     ****************************************** ***************::

61   WLSLRNPILVFLKTRVLKRWRLFSKHE   87  Segment_2_RKI_PB1-F2
61   WLSLRNPILVFLKTRVLKRWRLFSKHE   87  Segment_2_NIBSC_PB1-F2
     ***************************
```

figure S 7: Alignment of AA consensus sequences of segment 2 coding for PB1-F2 of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1     MEDFVRQCFNPMIVELAEKTMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGES   60  Segment_3_RKI_PA
1     MEDFVRQCFNPMIVELAEKTMKEYGEDLKIETNKFAAICTHLEVCFMYSDFHFINEQGES   60  Segment_3_NIBSC_PA
      ************************************************************

61    IIVELGDPNALLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEI  120  Segment_3_RKI_PA
61    IIVELGDPNALLKHRFEIIEGRDRTMAWTVVNSICNTTGAEKPKFLPDLYDYKENRFIEI  120  Segment_3_NIBSC_PA
      ************************************************************

121   GVTRREVHIYYLEKANKIKSEKTHIHIFSFTGEEMATKADYTLDEESRARIKTRLFTIRQ  180  Segment_3_RKI_PA
121   GVTRREVHIYYLEKANKIKSEKTHIHIFSFTGEEMATKADYTLDEESRARIKTRLFTIRQ  180  Segment_3_NIBSC_PA
      ************************************************************

181   EMASRGLWDSFRQSERGEETIEERFEITGTMRKLADQSLPPNFSSLENFRAYVDGFKPNG  240  Segment_3_RKI_PA
181   EMASRGLWDSFRQSERGEETIEERFEITGTMRKLADQSLPPNFSSLENFRAYVDGFEPNG  240  Segment_3_NIBSC_PA
      ************************************************************:***

241   YIEGKLSQMSKEVNARIEPFLKTTPRPLRLPNGPPCSQRSKFLLMDALKLSIEDPSHEGE  300  Segment_3_RKI_PA
241   YIEGKLSQMSKEVNARIEPFLKTTPRPLRLPNGPPCSQRSKFLLMDALKLSIEDPSHEGE  300  Segment_3_NIBSC_PA
      ************************************************************

301   GIPLYDAIKCMRTFFGWKEPNVVKPHEKGINPNYLLSWKQVLAELQDIENEEKVPKTKNM  360  Segment_3_RKI_PA
301   GIPLYDAIKCMRTFFGWKEPNVVKPHEKGINPNYLLSWKQVLAELQDIENEEKIPKTKNM  360  Segment_3_NIBSC_PA
      ************************************************************:******

361   KKTSQLKWALGENMAPEKVDFDDCKDVGDLKQYDSDEPELRSLASWIQNEFNKACELTDS  420  Segment_3_RKI_PA
361   KKTSQLKWALGENMAPEKVDFDDCKDVGDLKQYDSDEPELRSLASWIQNEFNKACELTDS  420  Segment_3_NIBSC_PA
      ************************************************************

421   SWIELDEIGEDVAPIEHIASMRRNYFTSEVSHCRATEYIMKGVYINTALLNASCAAMDDF  480  Segment_3_RKI_PA
421   SWIELDEIGEDVAPIEHIASMRRNYFTSEVSHCRATEYIMKGVYINTALLNASCAAMDDF  480  Segment_3_NIBSC_PA
      ************************************************************

481   QLIPMISKCRTKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKY  540  Segment_3_RKI_PA
481   QLIPMISKCRTKEGRRKTNLYGFIIKGRSHLRNDTDVVNFVSMEFSLTDPRLEPHKWEKY  540  Segment_3_NIBSC_PA
      ************************************************************

541   CVLEIGDMLLRSAIGQVSRPMFLYVRTNGTSKIKMKWGMEMRRCLLQSLQQIESMIEAES  600  Segment_3_RKI_PA
541   CVLEIGDMLIRSAIGQVSRPMFLYVRTNGTSKIKMKWGMEMRRCLLQSLQQIESMIEAES  600  Segment_3_NIBSC_PA
      *********:**************************************************

601   SVKEKDMTKEFFENKSETWPIGESPKGVEESSIGKVCRTLLAKSVFNSLYASPQLEGFSA  660  Segment_3_RKI_PA
601   SVKEKDMTKEFFENKSETWPIGESPKGVEESSIGKVCRTLLAKSVFNSLYASPQLEGFSA  660  Segment_3_NIBSC_PA
      ************************************************************

661   ESRKLLLIVQALRDNLEPGTFDLGGLYEAIEECLINDPWVLLNASWFNSFLTHALS  716  Segment_3_RKI_PA
661   ESRKLLLIVQALRDNLEPGTFDLGGLYEAIEECLINDPWVLLNASWFNSFLTHALS  716  Segment_3_NIBSC_PA
      ********************************************************
```

figure S 8: Alignment of AA consensus sequences of segment 3 coding for PA of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

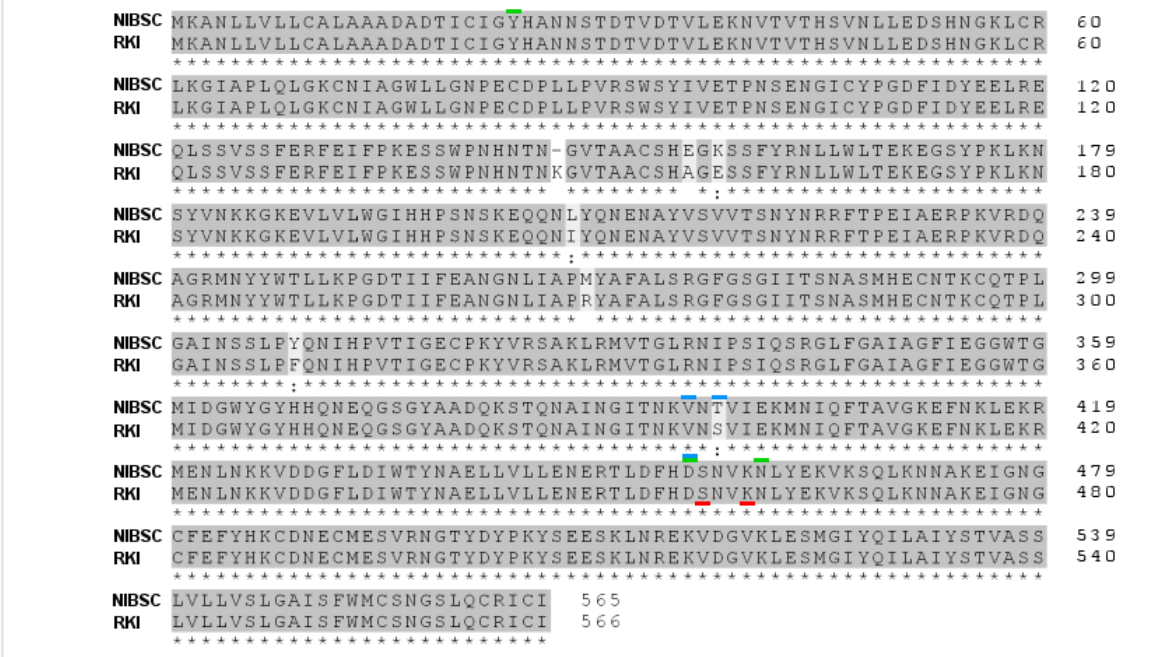


figure S 9: Alignment of AA consensus sequences of segment 4 coding for HA of the RKI- and the NIBSC-strain.

The virus seed of RKI (Amp. 3138) corresponds to a homogeneous population. Substitutions in the sequence during the virus adaptation processes are indicated in red. In contrast, the virus seed from NIBSC (#06/114) comprises various virus variants; substitutions in the sequence are indicated in green. The positions of substitutions, acquired during the adaptation processes are indicated in blue. The amino acid assembly was performed at http://services.uniprot.org/clustalw. Reprinted with permission [4].

```
1    MASQGTKRSYEQMETDGERQNATEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNS   60  Segment_5_RKI_NP
1    -ASQGTKRSYEQMETDGERQNATEIRASVGKMIGGIGRFYIQMCTELKLSDYEGRLIQNS   59  Segment_5_NIBSC_NP
      ************************************************************

61   LTIERMVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEEIRRIW  120  Segment_5_RKI_NP
60   LTIERMVLSAFDERRNKYLEEHPSAGKDPKKTGGPIYRRVNGKWMRELILYDKEEIRRIW  119  Segment_5_NIBSC_NP
     ************************************************************

121  RQANNGDDATAGLTHMMIWHSNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAG  180  Segment_5_RKI_NP
120  RQANNGDDATAGLTHMMIWHSNLNDATYQRTRALVRTGMDPRMCSLMQGSTLPRRSGAAG  179  Segment_5_NIBSC_NP
     ************************************************************

181  AAVKGVGTMVMELVRMIKRGINDRNFWRGENGRKTRIAYERMCNILKGKFQTAAQKAMMD  240  Segment_5_RKI_NP
180  AAVKGVGTMVMELVRMIKRGINDRNFWRGENGRKTRIAYERMCNILKGKFQTAAQKAMMD  239  Segment_5_NIBSC_NP
     ************************************************************

241  QVRESRNPGNAEFEDLTFLARSALILRGSVAHKSCLPACVYGPAVASGYDFEREGYSLVG  300  Segment_5_RKI_NP
240  QVRESRNPGNAEFEDLTFLARSALILRGSVAHKSCLPACVYGPAVASGYDFEREGYSLVG  299  Segment_5_NIBSC_NP
     ************************************************************

301  IDPFRLLQNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGTKVVPRGKLST  360  Segment_5_RKI_NP
300  IDPFRLLQNSQVYSLIRPNENPAHKSQLVWMACHSAAFEDLRVLSFIKGTKVLPRGKLST  359  Segment_5_NIBSC_NP
     ************************************************:*******

361  RGVQIASNENMETMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISIQPTFSVQRNLPF  420  Segment_5_RKI_NP
360  RGVQIASNENMETMESSTLELRSRYWAIRTRSGGNTNQQRASAGQISIQPTFSVQRNLPF  419  Segment_5_NIBSC_NP
     ************************************************************

421  DRTTVMAAFTGNTEGRTSDMRAEIIRMMESARPEDVSFQGRGVFELSDEKAASPIVPSFD  480  Segment_5_RKI_NP
420  DRTTIMAAFNGNTEGRTSDMRTEIIRMMESARPEDVSFQGRGVFELSDEKAASPIVPSFD  479  Segment_5_NIBSC_NP
     ****:****.**********:*************************************

481  MSNEGSYFFGDNAE  494  Segment_5_RKI_NP
480  MSNEGSYFFGDNAE  493  Segment_5_NIBSC_NP
     **************
```

figure S 10: Alignment of AA consensus sequences of segment 5 coding for NP of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1    MNPNQKIITIGSICMVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNST   60  Segment_6_RKI_NA
1    MNPNQKIITIGSICLVVGLISLILQIGNIISIWISHSIQTGSQNHTGICNQNIITYKNST   60  Segment_6_NIBSC_NA
     **************:*********************************************

61   WVKDTTSVILTGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLT  120  Segment_6_RKI_NA
61   WVKDTTSVILTGNSSLCPIRGWAIYSKDNSIRIGSKGDVFVIREPFISCSHLECRTFFLT  120  Segment_6_NIBSC_NA
     ************************************************************

121  QGALLNDRHSNGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGI  180  Segment_6_RKI_NA
121  QGALLNDKHSSGTVKDRSPYRALMSCPVGEAPSPYNSRFESVAWSASACHDGMGWLTIGI  180  Segment_6_NIBSC_NA
     *******:**.*************************************************

181  SGPDNGAVAVLKYNGIITETIKSWRKKILRTQESECACVNGSCFTIMTDGPSDGLASYKI  240  Segment_6_RKI_NA
181  SGPDNGAVAVLKYNGIITETIKSWRKKILRTQESECACVNGSCFTIMTDGPSDGLASYKI  240  Segment_6_NIBSC_NA
     ************************************************************

241  FKIEKGKVTKSIELNAPNSHYEECSCYPDTGKVMCVCRDNWHGSNRPWVSFDQNLDYQIG  300  Segment_6_RKI_NA
241  FKIEKGKVTKSIELNAPNSHYEECSCYPDTGKVMCVCRDNWHGSNRPWVSFDQNLDYQIG  300  Segment_6_NIBSC_NA
     ************************************************************

301  YVCSGVFGDNPRPKDGTGSCGPVYVDGANGVKGFSYRYGNGVWIGRTKSHSSRHGFEMIW  360  Segment_6_RKI_NA
301  YICSGVFGDNPRPEDGTGSCGPVYVDGANGVKGFSYRYGNGVWIGRTKSHSSRHGFEMIW  360  Segment_6_NIBSC_NA
     *:***********:*************************************************

361  DPNGWTETDSEFSVRQDVVAMTDWSGYSGSFVQHPELTGLDCIRPCFWVELIRGRPKEKT  420  Segment_6_RKI_NA
361  DPNGWTETDSKFSVRQDVVAMTDWSGYSGSFVQHPELTGLDCMRPCFWVELIRGRPKEKT  420  Segment_6_NIBSC_NA
     **********:*******************************:*****************

421  IWTSASSISFCGVDSDTVDWSWPDGAELPFTIDK  454  Segment_6_RKI_NA
421  IWTSASSISFCGVNSDTVDWSWPDGAELPFSIDK  454  Segment_6_NIBSC_NA
     *************:****************:***
```

figure S 11: Alignment of AA consensus sequences of segment 6 coding for NA of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1     MSLLTEVETYVLSIIPSGPLKAEIAQRLEDVFAGKNTDLEVLMEWLKTRPILSPLTKGIL   60  Segment_7_RKI_M1
1     MSLLTEVETYVLSIIPSGPLKAEIAQRLENVFAGKNTDLEVLMEWLKTRPILSPLTKGIL   60  Segment_7_NIBSC_M1
      ******************************:******************************

61    GFVFTLTVPSERGLQRRRFVQNALNGNGDPNNMDKAVKLYRKLKREITFHGAKEISLSYS  120  Segment_7_RKI_M1
61    GFVFTLTVPSEQGLQRRRFVQNALNGNGDPNNMDKAVKLYRKLKREITFYGAKEIALSYS  120  Segment_7_NIBSC_M1
      ***********:************************************:*****:****

121   AGALASCMGLIYNRMGAVTTEVAFGLVCATCEQIADSQHRSHRQMVTTTNPLIRHENRMV  180  Segment_7_RKI_M1
121   AGALACCMGLIYNRMGTVTAEVVFGLVCATCEQIADSQHRSHRQMVATTNPLIRHENRMV  180  Segment_7_NIBSC_M1
      *****.**********:**:**.**********************:**************

181   LASTTAKAMEQMAGSSEQAAEAMEVASQARQMVQAMRTIGTHPSSSAGLKNDLLENLQAY  240  Segment_7_RKI_M1
181   LASTTAKAMEQMAGSSEQAAEAMEVAIRARQMVQAMRTIGTHPSSSAGLKDDLLENLQAY  240  Segment_7_NIBSC_M1
      *************************** :**********************:*********

241   QKRMGVQMQRFK  252  Segment_7_RKI_M1
241   QKRMGVQMQRFK  252  Segment_7_NIBSC_M1
      ************
```

figure S 12: Alignment of AA consensus sequences of segment 7 coding for M1 of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1     MSLLTEVETPIRNEWGCRCNGSSDPLAIAANIIGILHLILWILDRLFFKCIYRRFKYGLK   60  Segment_7_RKI_M2
1     MSLLTEVETPIRNEWGCRCNGSSDPLVIASSIIGILHLILWILDRLFFKCIYRLFKYGLK   60  Segment_7_NIBSC_M2
      **************************.**:.********************** ******

61    GGPSTEGVPKSMREEYRKEQQSAVDADDGHFVSIEL-   96  Segment_7_RKI_M2
61    RGPSTEGVPKSMREEYQKEQQSAVDADDGHFVSIELE   97  Segment_7_NIBSC_M2
       ***************:******************
```

figure S 13: Alignment of AA consensus sequences of segment 7 coding for M2 of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1     MDPNTVSSFQVDCFLWHVRKRVADQELGDAPFLDRLRRDQKSLRGRGSTLGLDIETATRA   60  Segment_8_RKI_NS1
1     MDPNTVSSFQVDCFLWHVRKRVADQELGDAPFLDRLRRDQKSLRGRGSTLGLDIKTATRA   60  Segment_8_NIBSC_NS1
      ************************************************************:*****

61    GKQIVERILKEESDEALKMTMASVPASRYLTDMTLEEMSREWSMLIPKQKVAGPLCIRMD  120  Segment_8_RKI_NS1
61    GKQIVERILKEESDEALKMTMASVPASRYLTDMTLEEMSRDWSMLIPKQKVAGPLCIRMD  120  Segment_8_NIBSC_NS1
      ****************************************:*******************

121   QAIMDKNIILKANFSVIFDRLETLILLRAFTEEGAIVGEISPLPSLPGHTAEDVKNAVGV  180  Segment_8_RKI_NS1
121   QAIMDKNIILKANFSVIFDRLETLILLRAFTEEGAIVGEISPLPSLPGHTAEDVKNAVGV  180  Segment_8_NIBSC_NS1
      ************************************************************

181   LIGGLEWNDNTVRVSETLQRFAWRSSNENGRPPLTPKQKREMAGTIRSEV  230  Segment_8_RKI_NS1
181   LIGGLEWNDNTVRVSETLQRFAWRSSNENGRPPLTPKQKREMAGTIRSEV  230  Segment_8_NIBSC_NS1
      ************************************************
```

figure S 14: Alignment of AA consensus sequences of segment 8 coding for NS1 of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

```
1     MDPNTVSSFQDILLRMSKMQLESSSEDLNGMITQFESLKLYRDSLGEAVMRMGDLHSLQN   60  Segment_8_RKI_NS2
1     MDPNTVSSFQDILLRMSKMQLESSSEDLNGMITQFESLKLYRDSLGEAVMRMGDLHSLQN   60  Segment_8_NIBSC_NS2
      ************************************************************

61    RNEKWREQLGQKFEEIRWLIEEVRHKLRVTENSFEQITFMQALHLLLEVEQEIRTFSFQL  120  Segment_8_RKI_NS2
61    RNEKWREQLGQKFEEIRWLIEEVRHKLKITENSFEQITFMQALHLLLEVEQEIRTFSFQL  120  Segment_8_NIBSC_NS2
      ***************************::******************************

121   I  121  Segment_8_RKI_NS2
121   I  121  Segment_8_NIBSC_NS2
      *
```

figure S 15: Alignment of AA consensus sequences of segment 8 coding for NS2 (NEP) of the RKI- and the NIBSC-strain.
The amino acid assembly was performed at http://services.uniprot.org/clustalw.

12.9 SOPs and protocols

12.9.1 Thawing of cells

Internal SOP (file Z_02_Auftauenzellen_250603_IB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.2 Passaging MDCK cells

12.9.2.1 Serum-containing

Internal SOP (Z_04_Passagieren_MDCK_200606_SK.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.2.2 Serum-free

Internal SOP (Z_05_Passagieren_MDCK_serumfrei_211105_SK.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.3 Preparation of isotonic phosphate-buffered saline (PBS)

Internal SOP (file M_01_Herstellung_PBS_260907_CB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.4 Preparation of caso-bouillon for sterility testing

Internal SOP (file M_05_Herstellung_Von_CASO_Bouillon_290606_CB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.5 Preparation of cell culture and virus production media

12.9.5.1 Glasgow-MEM-medium from powder

Internal SOP (M_03_Herst_von_GMEM_aus_Pulvermedium_231105_SK.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.5.2 Glasgow-MEM-medium from prepared solutions

Internal SOP (file M_04_Herst.von Glasgow-MEM-Vollmed._Z-Med._200606_SK.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.5.3 Smif 8 PGd-medium from powder

H_2O_{MQ}	fill up to 10 L
Smif 8-PGd powdermedium (FH Emden)	30.6 g
NaCl (Roth)	62.4 g
$NaHCO_3$ (Roth)	20.0 g
Ethanolamine (98%, Sigma)	1.6 g
L-glutaminic acid (Merck)	10 µL
D- (+)-glucose (H_2O-free, Roth)	36.5 g
Pluronic F68 10% (GIBCO Invitrogen)	100 mL
pH	7.3
osmolarity	300 mOs/kg

12.9.6 Virus propagation

Internal SOP (file V_03_Virenvermehrung_Kulturgefaesse_260907_CB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.7 Preparation of trypsin-EDTA-stock solution (10x) for cell detaching

Internal SOP (file M_07_Trypsinherstellung_200607_CB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.8 Preparation of trypsin for virus propagation

Internal SOP (file V_02_Trypsin_fuer_Virus_170107_CB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.9 Preparation of peptone solution (20%)

Herstellungsmenge	LAB-M-Peptone	Milli-Q-Wasser
200ml	40g	160g
500ml	100g	400g

Peptonpulver in 80°C warmen dest. H_2O lösen, in kleine Schottflaschen verteilen und bei 120°C 20 min autoklavieren

12.9.10 Preparation of Alsevers solution

Internal SOP (file M_13_Alseversloesung_040107_CB.doc).Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.
Solution should be sterile for stabilizing chicken erythrocytes

12.9.11 Preparation of chicken erythrocytes

Internal SOP (file V_07_Erythrocytenloesung_07.06.07_CB.doc). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.12 Hemagglutination-assay

Internal SOP, Version: 2.1 and 2.2 (04.12.2006 and 20.01.2011, respectively). Please ask Dr. habil. Y Genzel or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.13 β-propiolactone inactivation

Internal SOP, Version: 1.1 (26.10.06). Please ask Dr. M. Wolff or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.14 N-glycosylation pattern analysis ($NaBH_3CN$-based, V1.2)

Internal SOP, please ask Dr. E. Rapp or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.15 N-glycosylation pattern analysis (picoline borane-based, V1.5)

Internal SOP, please ask Dr. E. Rapp or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.16 Purification of labeled N-glycans by HILIC

Internal SOP, please ask Dr. E. Rapp or Prof. Dr.-Ing. U. Reichl (MPI for Dynamics of Complex Technical Systems, Magdeburg, Germany) for access.

12.9.17 Native influenza virus deglycosylation

12.9.17.1 Virus concentration

- isolate virus from virus-containing, cleared supernatant by ultra-centrifugation (90 min, 4 °C, 31.000 rpm, 70Ti rotor, UZ-tubes) => from approx. 240 mL virus-containing supernatant
- resuspend virus pellets in 20 – 25 µL V-medium each making a final volume of 160 -200 µL

12.9.17.2 Deglycosylation procedure

add 190 µL of the concentrated, recovered sample from step (1)

add 6:7 µL protease inhibitor (40x in sterile H_2O_{MQ}; this corresponds to 1 tablet in 250 µL; #11777700, Roche)

add 50 µL reaction buffer (R9150, Sigma-Aldrich)

add 10 µL endoglycosidase F2

add 10 μL endoglycosidase F3

add 10 μL α-galactosidase (G7163, Sigma-Aldrich)

incubate at 37 °C, 450 rpm over night

add 10 μL reaction buffer (R9025, Sigma-Aldrich)

add 10 μL endoglycosidase F1

incubate at 37 °C, 450 rpm over night

add 10 μL reaction buffer (R0266, Sigma-Aldrich)

add 10 μL α-mannosidase (M7257, Sigma-Aldrich)

add 10 μL α-neuramidase (N8271, Sigma-Aldrich)

add 10 μL β-*N*-*a*cetylglucosaminidase (A6805, Sigma-Aldrich)

add 10 μL β-galactosidase (G0413, Sigma-Aldrich)

add 2 μL α-galactosidase (G7163, Sigma-Aldrich)

add 2 μL endoglycosidase F3

incubate at 37 °C, 450 rpm over night

Isolate virus by ultra-centrifugation at 31.000 rpm (Beckman Coulter, Rotor 70Ti, 31000 rpm, $Ø_{65.7\ mm}$ 70714 g) for 90 min at 4 °C in 100 mM Tris (pH7)

transfer the supernatant into a labeled falcon tube for later APTS-labelling

resuspend the virus pellet in 20 - 30 μL 100 mM Tris-HCl (pH7) and store 2 aliquots 1x 7 – 10 μL: for later *N*-glycoanalysis (could be diluted by a factor of 5) and 1x 13 – 20 μL for immunogenicity studies, pellet should later be washed once again in 100 mM Tris pH7)

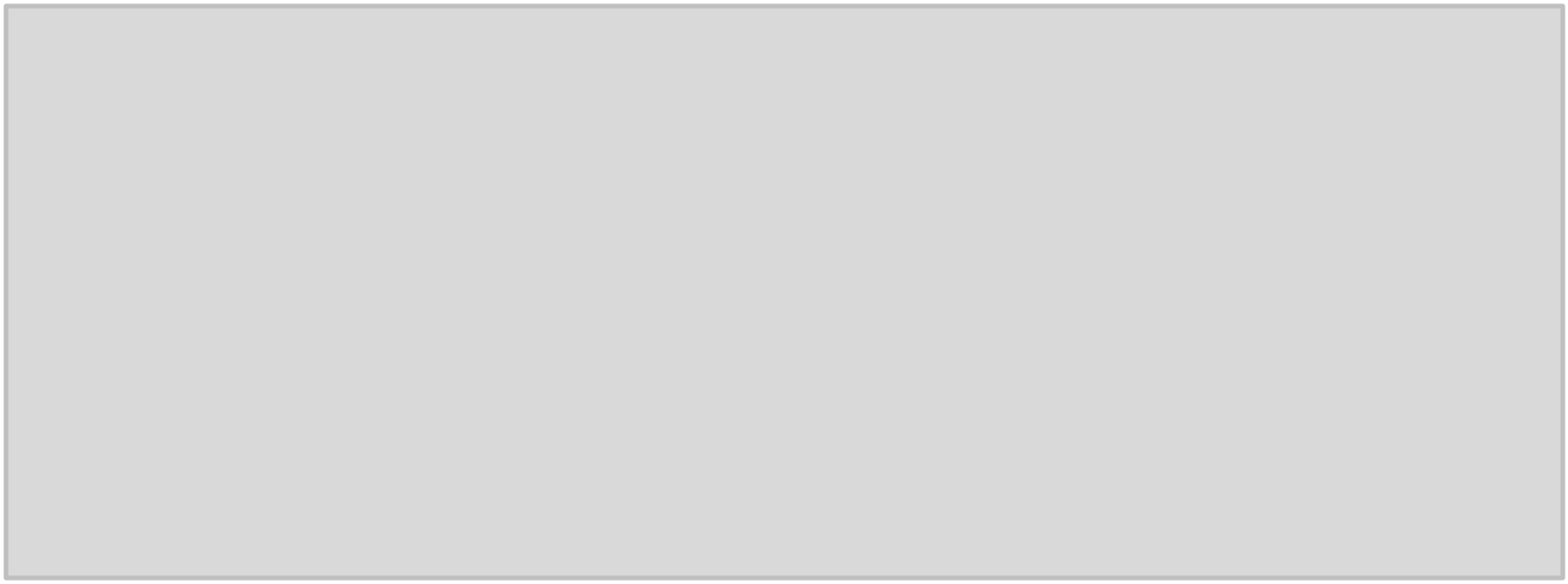
Store samples at -80 °C

12.10 Principles of next-generation pyrosequencing

12.10.1 DNA library preparation

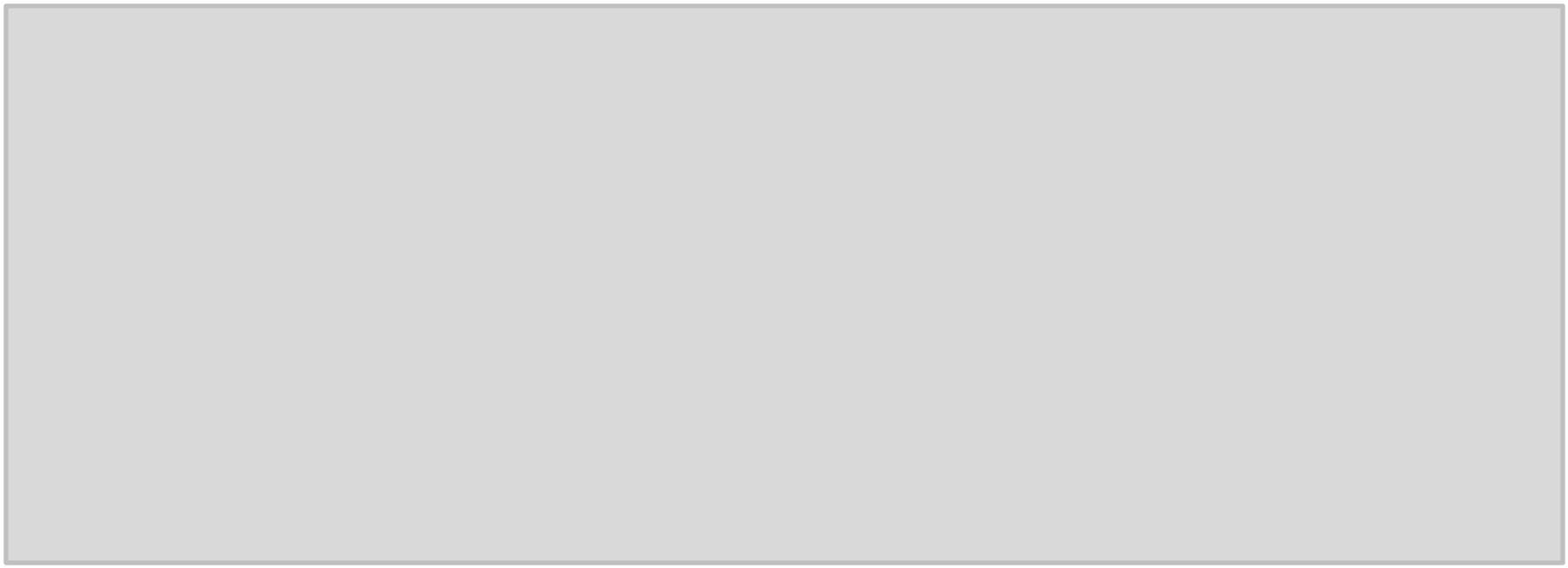
See / from http://www.genomicsdisorders.nl/GSFLX_Poster1.pdf

12.10.2 emPCR – emulsion PCR set up



See / from http://www.genomicsdisorders.nl/GSFLX_Poster2.pdf

12.10.3 emPCR – emulsion PCR breaking and enrichment



See / from http://www.genomicsdisorders.nl/GSFLX_Poster3.pdf

12.10.4 Sequencing

See / from http://www.genomicsdisorders.nl/GSFLX_Poster3.pdf

13 Own work

13.1 Reviewed journal articles, book sections and statement on authorship

This dissertation contains material that has previously been published by the author elsewhere. In the following a detailed overview of articles and book sections is given in the order of their publication date using numbers according to the reference list (section 7).

[8] Genzel Y, Rödig JV, Rapp E, Reichl U: **Vaccine production - upstream processing with adherent or suspension cell lines**. In: *Animal cell biotechnology - methods and protocols.* Edited by Pörtner R, vol. 25: Humana Press; 3rd ed. 2013, in press.

Copyright 2013/2014. Springer.

The author contributed to study conception/design, data acquisition/analysis/interpretation, manuscript writing and final approval of published version with respect to the *N*-glycosylation part of the manuscript.

[6] Rödig JV, Rapp E, J. B, Kampe M, Kaffka H, Bock A, Genzel Y, Reichl U: **Impact of cultivation conditions on *N*-glycosylation of influenza A virus hemagglutinin produced in MDCK cell culture**. *Biotechnology and Bioengineering* 2013, 110(6):1691-1703.

Copyright 2013 Wiley Periodicals, Inc.

The author contributed to study conception/design, data acquisition/analysis/interpretation, wrote the manuscript and contributed to final approval of published version.

[1] Hütter J*, Rödig JV*, Höper D, Reichl U, Seehofer PH, Rapp E, Lepenies B: **Toward animal cell culture-based influenza vaccine design: viral hemagglutinin *N*-glycosylation markedly impacts immunogenicity.** *J Immunol* 2013, 190(1):220-230.

*authors contributed equally

The author contributed to study conception/design, data acquisition/analysis/interpretation, manuscript writing and final approval of published version. For a detailed description of each author's contribution see section 4.9, 3rd and 4th paragraph..

[9] Genzel Y, Behrendt I, Rodig J, Rapp E, Kueppers C, Kochanek S, Schiedner G, Reichl U: **CAP, a new human suspension cell line for influenza virus production**. *Appl Microbiol Biotechnol* 2013, 97(1):111-122.

The author contributed to study conception/design, data acquisition/analysis/interpretation, manuscript writing and final approval of published version with respect to the *N*-glycosylation part of the manuscript.

[4] Roedig JV, Rapp E, Höper D, Genzel Y, Reichl U: **Impact of Host Cell Line Adaptation on Quasispecies Composition and Glycosylation of Influenza A Virus Hemagglutinin.** *PLos ONE* 2011, 6(12).

The author contributed to study conception/design, data acquisition/analysis/interpretation, wrote the manuscript and contributed to final approval of published version.

[7] Rödig J, Rapp E, Djeljadini S, Lohr V, Genzel Y, Jordan I, Sandig V, Reichl U: **Impact of Influenza Virus Adaptation status on HA *N*-Glycosylation Patterns in Cell Culture-Based Vaccine Production**. *Journal of Carbohydrate Chemistry* 2011, 30:281-290.

The author contributed to study conception/design, performed data acquisition/analysis/interpretation, wrote the manuscript and contributed to final approval of published version.

13.2 Conference proceedings and statement on authorship

This dissertation contains material that has previously been published by the author elsewhere. In the following a detailed overview of proceedings is given in the order of their publication date using numbers according to the reference list (section 7).

[5] Roedig JV, Rapp E, Genzel Y, Reichl U: **Impact of different influenza cultivation conditions on HA *N*-Glycosylation.** *BMC proceedings* 2011, 5 Suppl 8:P113.

The author contributed to study conception/design, performed data acquisition/analysis/interpretation, wrote the manuscript and contributed to final approval of published version.

[147] Rödig J, Rapp E, Hennig R, Schwarzer J, Reichl U: **Optimized CGE-LIF-Based Glycan Analysis for High-Throughput Applications**. In: *Proceedings of the 21st Annual Meeting of the European Society for Animal Cell Technology (ESACT), June 7–10, 2009.* Dublin, Ireland: Springer Science+Business Media B.V.; 2011.

The author contributed to study conception/design, data acquisition/analysis/interpretation, wrote the manuscript and contributed to final approval of published version.

13.3 Conference contributions

In the following conference conrtibutions are summarized and numbered chronologically from recent to longer ago.

13.3.1 Oral presentations

(1) 12/2012 ACTIP Meeting, Marburg

J. Rödig, E. Rapp, J. Hütter, B. Lepenies, J. Bohne, Y. Genzel, D. Höper, U. Reichl. Process Conditions Can Markedly Affect the *N*-Glycosylation Pattern and Immunogenicity of Influenza A Virus.

(2) 09/2012 30th Annual Meeting of Biotechnologists, DECHEMA (Society for Chemical- and Biotechnology e.V.), Karlsruhe (Germany).

J. Rödig, E. Rapp, J. Hütter, B. Lepenies, J. Schwarzer, Y. Genzel, D. Höper, U. Reichl. **Impact of process conditions on influenza A virus HA *N*-glycosylation and immunogenicity.**

13.3.2 Poster presentations

(1) 05/2011 European Society for Animal Cell Culture Technology (ESACT), Vienna (Austria).

J. Rödig, E. Rapp, Y. Genzel, U. Reichl. Impact of Different Influenza Cultivation Conditions on HA *N*-Glycosylation.

(2) 03/2011 21st Annual Meeting of the Society for Virology, Heidelberg (Germany).

J. Rödig, E. Rapp, D. Höper, Y. Genzel, U. Reichl. Does Altered *N*-Glycosylation of Influenza A Virus Hemagglutinin in Vero Cells Prevent Efficient Virus Replication?

(3) 03/2011 5th Glycan Forum, Berlin (Germany).

J. Rödig, E. Rapp, Y. Genzel, U. Reichl. Impact of Influenza Virus Adaptation Status on HA *N*-Glycosylation Patterns in Cell Culture-Based Vaccine Production.

(4) 03/2011 5th Glycan Forum, Berlin (Germany).
J. Hütter, J. Rödig, U. Reichl, P. H. Seeberger, E. Rapp, B. Lepenies. Impact of Influenza A Virus Hemagglutinin *N*-Glycosylation on Immunogenicity.

(5) 09/2009 27th Annual Meeting of Biotechnologists, DECHEMA (Society for Chemical- and Biotechnology e.V.), Mannheim (Germany).
J. Rödig, E. Rapp, J. Schwarzer, Y. Genzel, U. Reichl. Impact of viral Adaptation on the N-Glycosylation Pattern of Influenza A Hemagglutinin.

(6) 06/2009 European Society for Animal Cell Culture Technology (ESACT), Dublin (Great Britain).
J. Rödig, E. Rapp, R. Hennig, J. Schwarzer, U. Reichl. Optimized CGE-LIF-Based Glycan Analysis for High Throughput Applications.

13.4 Awards

ACTIP* fellowship 2012
*animal cell technology industrial partners

13.5 Supervised students

Michael Kampe (Studienarbeit; 03/2010 – 10/2010)
Helene Kaffka (studentische Hilfswissenschaftlerin; 01/2011 – 10/2012)