DISSERTATION

Titel der Dissertation

Etablierung von Systemen zur biochemischen, zellbasierten und strukturellen Charakterisierung der NS2B/3 Protease des Frühsommer-Meningoenzephalitis Virus sowie anderer viraler Proteasen

Verfasserin

Dipl. Ing. Martina Kurz

angestrebter akademischer Grad

Doktorin der Naturwissenschaften (Dr. rer.nat.)

Wien, 2012

Für meine Eltern

Danke schön!!

Ganz besonders möchte ich mich bei Prof. Timothy Skern bedanken, der mich in seine Arbeitsgruppe aufgenommen, mich stets mit fachlichem Rat unterstützt und mir den nötigen Freiraum gegeben hat, meine Dissertation selbstständig durchzuführen. Weiters gilt mein Dank Prof. Franz X. Heinz bedanken, der mir die Möglichkeit gegeben hat, einen Teil meiner Dissertation am Institut für Virologie durchzuführen. Dass ich an zwei Instituten arbeiten durfte ist ein großes Geschenk, ich habe sehr viel gelernt.

Ein riesiges Dankeschön auch an all meine Kollegen: Carla, Chiara, Niki, Angelika, Junping, Melanie, Kathi und Ulli, ganz besonders aber an Jutta und David, mit denen ich von Anfang an zusammen gearbeitet habe und die ich persönlich wie auch fachlich ganz besonders schätze. Danke fürs Korrekturlesen, haufenweise Schoko, literweise Kaffee, lustige Festl und für „The poem of the day". Weiters herzlich bedanken möchte ich mich bei Sabrina, deren Projekt am Institut für Virologie ich beenden durfte.

Danke auch an all meine Freunde, ohne die alles nur halb so schön wäre, die mein Leben so interessant, lustig und abwechslungsreich machen und die mich durch alle Höhen und Tiefen der Diss begleitet haben. Und wo ich schon dabei bin, möchte ich auch gleich meinen Saunafreunden danken. Seit vielen Jahren begleitet ihr mich ins Wochenende und das genieße ich sehr.

Vielen Dank auch an meine Familie. Ich freue mich, dass ihr alle so stolz auf mich seid und so regen Anteil an meiner Arbeit nehmt.

Der größte Dank gilt allerdings meinen Eltern. Danke, dass ich mich immer auf euch verlassen kann, eure Liebe ist mein Fels in der Brandung. Ich verdanke euch viel und liebe euch so sehr.

Tja, und dann ist da noch mein Alex. Ohne dich hätte ich die schwierigen Zeit der Diss nicht so gut überstanden und die guten nicht so sehr genossen. Danke für dein Verständnis, deine Unterstützung und dafür, dass du immer genau weißt, was ich gerade brauche. Mit dir an meiner Seite ist alles einfach und alles ist gut. Ich liebe dich.

INDEX

Generation and genetic stability of tick-borne encephalitis virus mutants dependent on processing by FMDV 3C protease

NS2B/3 proteolysis at the C-prM junction of the tick-borne encephalitis virus polyprotein is highly membrane dependent

SUMMARY

Proteolysis by viral enzymes plays a fundamental role in the viral life cycle, a fact that makes these proteases an attractive target for the antiviral therapy. Infections for which specific antiviral therapeutics are still needed include those of the human immunodeficiency virus (HIV) (Piacenti et al., 2006), hepatitis C virus (Lamarre et al., 2003), West Nile virus (WNV), tick-borne encephalitis virus (TBEV) (Mansfield et al., 2009) and, as a prominent representative of the family of animal pathogens, the foot-and-mouth disease virus (FMDV) (Grubman and Baxt, 2004). However, the development of such an inhibitor is difficult, as viral proteolysis is a highly complex process. Cleavage sites are often insufficiently defined and the development of resistance to inhibitors due to the error prone nature of the viral polymerase is an enduring problem. To overcome these obstacles, we set out to develop TBEV based systems, both biochemical and cell-based, which allow us to examine the properties of viral proteases in general, and of the TBEV protease in particular. TBEV causes the tick-borne encephalitis (TBE), a potentially lethal neurological disease of humans (Gritsun et al., 2003; Mansfield et al., 2009). It belongs to the family of *Flaviviridae* and comprises a single positive-strand RNA genome, which is translated into one polyprotein. Subsequent processing by cellular proteases together with the viral NS2B/3 protease (NS2B/3^{pro}) yields the mature proteins.

In order to set up a cell based assay that allows us to investigate various viral proteases, we modified the genome of TBEV in such a way that the production of infectious virions became dependent on the activity of a heterologous protease. To this end, the NS2B/3^{pro} capsid cleavage site in the C-terminal region of TBEV was replaced by an optimised cleavage site for the FMDV 3C protease (3C^{pro}). Infectivity of this non-viable mutant could be regained by expression of the 3C^{pro} either from the same TBEV genome *in cis* or from a replicon *in trans*. To test the genetic stability of both systems, various TBEV mutants were passaged in the absence of an active 3C^{pro}. This led to the appearance of three revertants only in the mono-cistronic system, indicating that the *trans*-complementation system is more stable. In all three mutants, a NS2B/3^{pro} cleavage motif RR*C was introduced at an identical location, giving rise to protein C with a C-terminal extension of 14 amino acids. A comparison of the revertants reveals that the more hydrophobic the extension of protein C, the lower the rate of early RNA synthesis in general and the later the onset of early RNA replication and/or unpackaging in particular. Furthermore, our results show a need of positive charges at the C-terminus of protein C for efficient budding of the nucleocapsid, thereby highlighting the role of protein C as a multifunctional protein.

We established a system, in which a heterologous protease can recognise its corresponding cleavage motif in the C-terminal region of protein C and consequently allows the production of infectious TBEV particles. Addition of protease inhibitors to the trans-

complementation system should allow protease mutants to be detected that are resistant to the inhibitor. Thus, this system can be used to further illuminate the relation between inhibitors and protease resistance.

In order to investigate the biochemical and structural properties of the NS2B/3^{pro} we constructed a modular protein consisting of the protease domain of NS3 and the core-domain of cofactor NS2B. The protease was expressed in *E.coli* and purified to homogeneity via affinity and size exclusion chromatography.

The modular protease NS2B/3^{pro} was used in a TBEV based *trans*-cleavage assay for biochemical characterisation. After eliminating unwanted autocleavage by site directed mutagenesis, thereby improving the activity of the protease, reaction conditions were optimised to allow proteolytic cleavage of *in vitro* translated C-prM structural precursor protein. Efficiency and specificity of the NS2B/3^{pro} were enhanced by providing canine pancreatic microsomal membranes; unwanted signal peptidase I cleavage was prevented by lengthening the h-region of the signal peptide. In order to establish a system which allows us to find suitable substrates for the above described cell based system, we substituted the NS2B/3^{pro} cleavage site at the C-prM junction with a one for the HIV-1 protease (HIV-1^{pro}). This mutant was no longer processed by the NS2B/3^{pro}; instead it was recognised by the HIV-1^{pro} which was also expressed in *E.coli*. Again, cleavage was efficient and specific in the presence of microsomal membranes. Thus, with this system we are able to mimic *in vitro* the cleavage of NS2B/3^{pro} and HIV-1^{pro} *in vivo*. This assay will facilitate the search for suitable substrates for the above described cell based assay and should enable future studies of proteolysis requirements of various proteases, even those that require higher biosafety level handling, in a safe system.

Finally, an inactivated form of the NS2B/3^{pro} (NS2B/3^{pro} S135A) was used in crystallisation experiments for structural characterisation. Crystallisation proved to be difficult, most probably because of the flexibility of the linker between NS3 and NS2B. As a result, crystallisation was only successful when we fused the protease to the maltose binding protein (MBP) and crystallised it in the presence of maltose as a stabilising agent. We determined conditions which allow a reproducible crystallisation of the MBP-NS2B/3 S135A fusion protein. The obtained crystals were too small and fragile for X-ray analysis, but provide a good basis for further optimisation.

Taken together, the results of this thesis extend the understanding of basic molecular mechanisms of the TBEV and cleavage requirements of the proteases of TBEV, FMDV and HIV. The assays allow the study of various aspects of viral pathogenesis and development of resistance to inhibitors. Thus, they provide the basis for the analysis of structure function relationships of viral proteases and should therefore facilitate future drug design.

ZUSAMMENFASSUNG

Proteolyse durch virale Enzyme spielt eine fundamentale Rolle im viralen Lebenszyklus. Das macht virale Proteasen zu einem wichtigen Ansatzpunkt für die antivirale Therapie. Zu den Infektionen, für deren Behandlung immer noch Medikamente benötigt werden, gehören jene die durch das Humane Immundefizienz Virus (HIV, Piacenti et al., 2006), das Hepatitis C Virus (Lamarre et al., 2003), das West Nil Virus (WNV), sowie das Frühsommer-Meningoenzephalitis Virus (FSME) verursacht werden. Dasselbe gilt für einen bekannten Vertreter der Tierpathogene, das Maul- und Klauenseuchevirus (MKSV) (Grubman und Baxt, 2004).

Die Entwicklung solcher Inhibitoren gestaltet sich allerdings schwierig, da virale Proteolyse ein komplexer Vorgang ist. Spaltstellen sind oft ungenügend definiert; weiters ist das Entstehen von Resistenzen, ausgelöst durch das Fehlen eines Korrekturlesemechanismus der viralen Polymerase, ein großes Problem. Um diese Probleme zu beheben, haben wir biochemische und zellbasierte Systeme auf Grundlage des FSME Virus entwickelt, die es ermöglichen, virale Proteasen im allgemeinen, und die Protease des FSME Virus im speziellen, zu untersuchen. Das FSME Virus ist ein humanpathogenes Virus aus der Familie der *Flaviviridae*. Vom einzel- und (+)strängigen RNA Genom wird ein Vorläuferpolyprotein synthetisiert, das ko- und post-translational von der viralen NS2B/3 Protease (NS2B/3^{pro}) in seine funktionellen Einzelproteine zerlegt wird.

Um ein System zu etablieren, das uns erlaubt unterschiedliche Proteasen zu untersuchen, war es notwendig, das Genom des FSME-Virus so zu verändern, dass die Entstehung infektiöser Partikel von der Aktivität einer heterologen Protease abhängt. Zu diesem Zweck haben wir die ursprüngliche Erkennungssequenz der FSME Protease (NS2B/3^{pro}) durch eine Spaltstelle der MKSV Protease (3C^{pro}) ersetzt. Diese Mutante war erst wieder infektiös, wenn die korrespondierende Protease *in cis* oder in *trans* verfügbar war.

Um die genetische Stabilität der beiden Systeme zu testen, wurden unterschiedliche FSME Mutanten in der Abwesenheit einer aktiven 3C^{pro} passagiert. Das führte zum Entstehen von drei Revertanten im monocistronischen System, während sich das *trans* System als das stabilere erwies. In allen drei Mutanten entstand die NS2B/3^{pro} Spaltstelle RR*C an derselben Position des Genoms. Dies führte zu einer Verlängerung des Kapsidproteins um 14 Aminosäuren. Eine Analyse der Eigenschaften der Revertanten zeigte, dass eine hydrophobere Verlängerung des Kapsidproteins zu einer verspäteten und verringerten RNA Synthese führt. Effizientes Budding des Nukleokapsid ist nur möglich, wenn sich am C-terminalen Ende des Kapsidproteins positiv geladene Aminosäuren befinden. Diese Ergebnisse implizieren, dass das Kapsidprotein für eine Vielzahl viraler Prozesse von Bedeutung ist.

Wir konnten also ein System etablieren, in dem eine heterologe Protease ihre korrespondierende Spaltstelle in der C-terminalen Region des Kapsidproteins erkennt und so

die Produktion infektiöser Virionen erlaubt. Die Anwesenheit eines Inhibitors der heterologen Protease provoziert das Entstehen von Mutanten, die resistent gegen den Inhibitor sind. Eine Analyse der Revertanten kann dazu beitragen, die Wechselbeziehung zwischen Inhibitor und Resistenzentstehung aufzuklären.

Um die biochemischen und strukturellen Eigenschaften der NS2B/3[pro] untersuchen zu können, haben wir eine modulare Protease, bestehend aus der Gensequenz der Proteasedomaine des NS3 Proteins verknüpft mit jener der Kernregion des Kofaktors NS2B, hergestellt. Diese wurde in *E.coli* exprimiert und via Affinitätschromatographie und Gelfiltration aufgereinigt. Die aktive Protease wurde, nachdem wir eine unerwünschte Selbstspaltung eliminieren konnten, zur Etablierung eines *in vitro* Assays zur biochemischen Charakterisierung der Protease verwendet. Das Substrat, C-prM, wurde *in vitro* translatiert, die Reaktionsbedingungen für die proteolytische Prozessierung durch die Protease optimiert. Effizienz und Spezifität der NS2B/3 konnten durch pankreatische mikrosomale Membranen weiter erhöht werden. Die damit einhergehende, unerwünschte Prozessierung durch die Signalpeptidase I konnten wir durch eine Verlängerung der h-region des Signalpeptids hemmen. Eine C-prM Mutante mit einer Erkennungssequenz für die HIV-1 Protease anstelle jener für die NS2B/3[pro], wurde verwendet, um die biochemischen Eigenschaften der HIV-1 Protease zu untersuchen, die ebenfalls in *E.coli* exprimiert wurde. Dieser Assay bietet die Möglichkeit, geeignete Substrate für den zuvor beschriebenen zellbasierten Assay zu finden und sollte so dessen Etablierung vereinfachen. Weiters erlaubt er uns, Proteasen unterschiedlicher Viren, unabhängig von deren Biosicherheitsstufe, in einem sicheren System zu untersuchen.

Eine inaktivierte Variante der modularen Protease wurde für Kristallisationsexperimente zur strukturellen Charakterisierung verwendet. Die Kristallisation war schwierig, vermutlich wegen der durch die Verbindung von Protease und Kofaktor induzierten Flexibilität. Daher war die Kristallisation erst dann erfolgreich und reproduzierbar, als die Protease mit dem Maltose bindenden Protein (MBP) als Fusionsprotein exprimiert wurde. Die resultierenden Kristalle waren zu klein und fragil um mittels Röntgenkristallographie analysiert zu werden, sind allerdings eine gute Basis für eine weitere Optimierung der Kristallisationsbedingungen und stellen einen wichtigen Schritt in Richtung der Aufklärung der Struktur der NS2B/3[pro] dar.

Die Ergebnisse der Dissertation erweitern das Verständnis der grundlegenden molekularen Mechanismen des FSME Virus und jenes über die Proteasen des FSME, HI und MKS Virus. Die beschriebenen Systeme ermöglichen es, unterschiedliche Aspekte der Pathogenese und der Entstehung von Inhibitorresistenzen zu untersuchen. Sie sind die Basis für eine Untersuchung der Beziehung zwischen Struktur und Funktion der Proteasen und sollten daher eine zukünftige Entwicklung von Proteaseinhibitoren vereinfachen.

1. GENERAL INTRODUCTION

1.1. Classification, taxonomy and clinical importance of flaviviruses

1.1.1. Family *Flaviviridae* - Genus *Flavivirus*

The family of *Flaviviridae* contains three genera, the *Flavivirus*, *Pestivirus* and *Hepacivirus*. They are similar in terms of their replication strategy, morphology and composition of the viral genome. However, they lack serological cross-reactivity and differ in regard to their biological properties. The virions are small (50nm), spherical, enveloped particles with one or more membrane anchored glycoproteins. The envelope surrounds the nucleocapsid which is formed by multiple units of the capsid protein and the single-stranded positive-sense RNA genome. Upon infection, it is translated into a polyprotein which is subsequently processed into its individual proteins by virus and host cell proteases. Finally, the progeny viruses assemble by budding through the membranes of the endoplasmic reticulum (ER) (Lorenz et al., 2002).
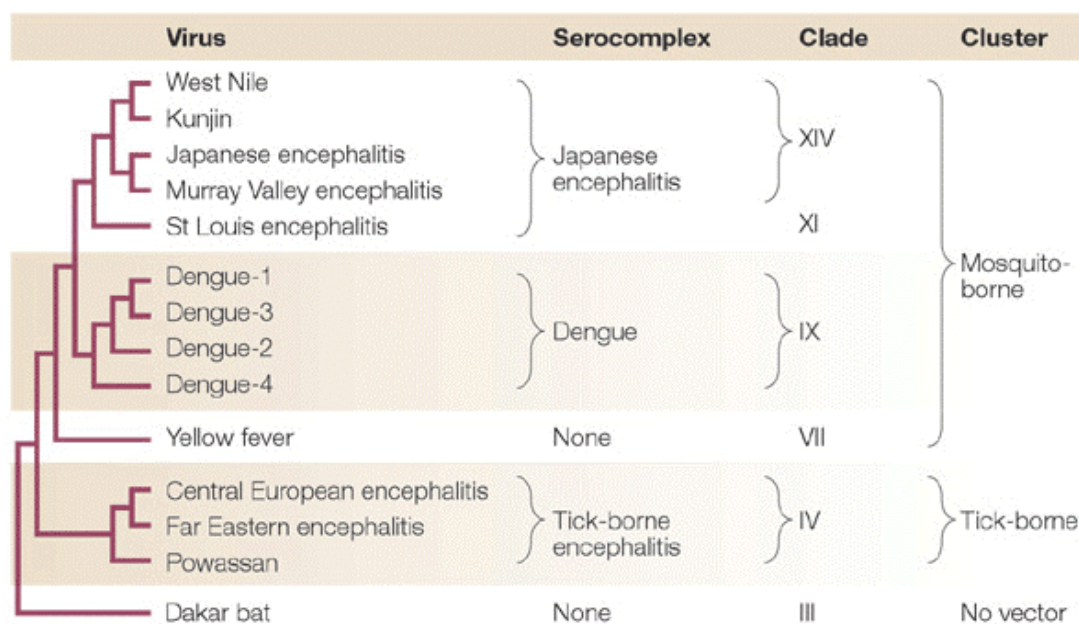
Figure 1. Classification of the genus *Flavivirus*. The serological (serocomplex) and phylogenetical (clade and cluster) classifications of these flaviviruses are shown on the right (adapted from Mukhopadhyay et al., Nature Reviews Microbiology, 2005).

The genus *Flavivirus* includes approximately 70 different viruses. Based on their route of transmission they can be divided into mosquito-borne, tick-borne and non-vectored viruses (Calisher et al, 1989). Furthermore, the viruses can be classified on the basis of serological criteria and sequence identity (Fig. 1) (Calisher et al., 1989; Kuno et al. 1998; Mukhopadhyay et al., 2005). The most important human pathogens are found among the mosquito-borne flaviviruses and include yellow fever virus (YFV), Japanese encephalitis virus (JEV), dengue

virus (DENV, subtypes 1, 2, 3 and 4) and West Nile virus (WNV) as well as one member of the tick-borne group, the tick-borne encephalitis virus (TBEV). Based on minor variations of their nucleotide sequence, TBEV is further classified into the European, the Far Eastern and Central Siberian subtype, each causing severe diseases with distinct clinical manifestations (Ecker et al., 1999).

1.1.2. Tick-borne encephalitis virus

TBEV is endemic in an area ranging from Northern China and Japan through the former Soviet countries to Europe. However, there are no reported cases for the British Isles, the Benelux countries, Portugal and Spain (Mansfield et al., 2009). Humans are occasionally infected by the consumption of non-pasteurised milk products from viremic livestock or by inhaling aerosolised infectious virions. However, most of the infections are caused by the bite of a chronically infected tick (Süss, 2003), mostly of the species *Ixodes ricinus* (European subtype), *Ixodes persulcatus* (Eurasia) and *Ixodes ovatus* (North-eastern China and Northern Japan) (Gubler et al., 2007). A majority of TBE infections remains asymptomatic but the infection can also lead to severe neurological diseases. After an incubation time of 3 to 7 days, a flu-like sickness occurs. In 70 % of cases, the patients recover within a week. In 20 to 30 % of patients a second stage of illness may occur which correlates with severe diseases of the central nervous system (CNS), including symptoms of meningitis, meningoencephalitis or meningoencephalomyelitis. 10 to 20 % of these patients experience long-term neuropsychiatric complications, such as spiral nerve paralysis; 1 to 2 % die (Charrel et al., 2004). This mortality rate is only valid in European countries and is much higher in Siberia (6 -8 %) and Far East (20 – 40 %). To prevent infection, a chemically inactivated whole virus vaccine derived from the prototypic European subtype, strain Neudoerfl, is available. The vaccine has proven to be efficient, safe and well-tolerated, and provides a 99 % protection rate after a basic immunisation of 3 doses. Booster doses are recommended every 3 to 5 years (Kunz et al., 1991). In Austria, about 90 % of the population is vaccinated. However, there does not exist any specific antiviral therapy for TBE. Thus, once manifested, the disease is not curable (Kaiser et al., 2008).

1.2. Molecular organisation of TBEV

1.2.1. Virion structure

TBEV are relatively small (50 nm), spherical particles, which are surrounded by a host cell derived lipid envelope (Fig. 2) (Lindenbach, Thiel and Rice, 2007). In this envelope, two of the three structural proteins, the membrane protein (M) and the envelope protein (E), are

anchored via transmembrane helices (Zhang et al., 2003). The dense core of the particles contains the nucleocapsid (NC). It is composed of multiple copies of capsid protein (C) complexed with one copy of the viral RNA genome and does not seem to have a symmetrical structure. Glycoprotein E mediates attachment and entry of the virus and is the major target of the host´s immune response. Functional protein M is produced from precursor M (prM) protein during maturation of the viral particle. Hereby, a structural reorganisation of the prM-E heterotrimers into antiparallel homodimers is induced which flattens the surface of the mature virus particle (Fig. 2) (Stiasny and Heinz, 2006).
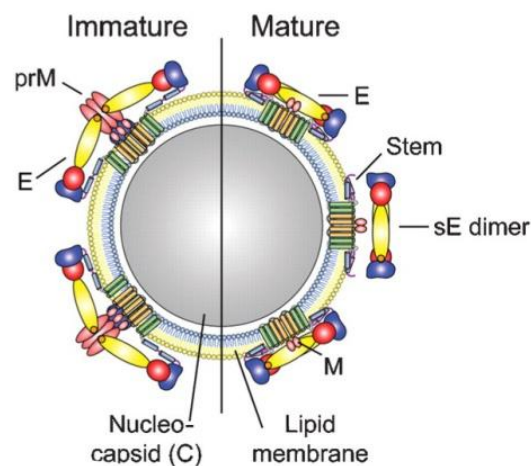
Figure 2. TBEV particle. Schematic representation of the immature and mature flavivirus particle. NC, structural glycoproteins prM/M, E, and the lipid membrane are indicated. (adapted from Stiasny and Heinz, 2006).

1.2.2. Genome organisation

The ~10.8-kb genome (Fig. 3) has only one open reading frame (ORF) and is infectious when introduced into a host cell. There, it is directly translated into a polyprotein that is subsequently cleaved into three structural proteins, C, prM and E and seven non-structural proteins, NS1 to NS5.

Figure 3. TBEV genome organisation. Scheme of the single-stranded positive-sense RNA genome (not drawn to scale). The structural proteins are indicated in dark grey, the non-structural ones in light grey. See text for details.

The 5' and 3' non-coding regions (NCR) of the genome are rich in RNA secondary structure elements. They are involved in a variety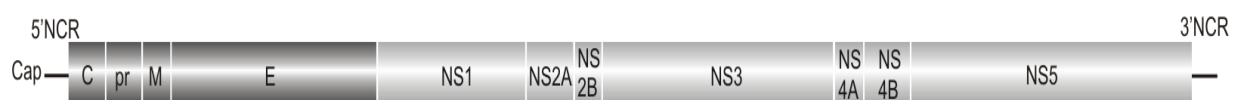 of important processes such as RNA replication and packaging as well as translation and binding of host factors (Alvarez et al., 2005; Gritsun et al., 1997; Khromykh et al., 2001; Kofler et al., 2006; Mandl et al., 1993). The 5´NCR is approximately 100 nucleotides long and carries a 5´type-I cap structure. The 3' NCR is 775 to 1125 nucleotides (nt) in length and can be further divided into a highly conserved core region (325 nt), involved in viral replication, and a variable region (450 to 800 nt) (Wallner et al., 1995; Markoff et al., 2003). In contrast to cellular messenger RNAs, it lacks a poly(A)tail (Lindenbach, Thiel and Rice, 2007).

1.2.3.   TBEV life cycle

Figure 4 shows a scheme of the TBEV life cycle. TBEV attaches to the host cell by interaction of the envelope protein E with one or more cellular receptors which have not been sufficiently identified yet. The ability to use multiple receptors might be responsible for the wide host range of flaviviruses, which replicate in arthropods and vertebrates. Several studies indicate an involvement of heparan sulphate (HS), a highly sulphated glycosaminoglycane that is found on many cell types, in TBEV attachment and entry (Kroschewski et al., 2003; Mandl et al., 2001; Germi et al., 2002). However, the availability of HS on the cell surface is not obligatory and its precise function remains unclear. After internalisation by receptor-mediated endocytosis, the virion is transported to prelysosomal endocytic vesicles (Chu et Ng, 2004), where the acidic environment induces a permanent trimerisation of the E protein. Thereby, the fusion peptide becomes exposed and inserts into the endosomal membrane; viral and cellular membranes (Allison et al., 1995; Heinz et al., 2003) fuse. As a result, the nucleocapsid is released into the host cell cytoplasm where capsid proteins and RNA dissociate. Translation is started immediately by the host cell translation machinery and results in one polyprotein that is cotranslationally integrated into the ER lumen (Mukhopadhyay et al., 2005). There, the polyprotein is processed by the host signal peptidase I (SPaseI) and the viral NS2B/3 protease (NS2B/3^{pro}). This processing generates the mature, functional proteins; three of which are structural ones (C, prM and E) that compose the viral particle, seven are non-structural proteins (NS1, NS2A, NS2B; NS3, NS4A, NS4B and NS5) that play a role in all stages of viral RNA synthesis. After generation of the viral polymerase and other necessary non-structural proteins, a replicase complex is assembled. First, negative-strand RNA copies are generated, which in turn serve as templates for the production of plus strand genomic RNAs. These are capped by the viral NS3 and NS5 proteins, which act as both triphosphatase and RNA helicase, or methyltransferase and RNA polymerase, respectively. The glycoproteins E and prM are cotranslationally translocated into the ER lumen, where they are both glycosylated while they

are still unfolded. This indicates that glycosylation might support correct folding of prM and E proteins (Helenius et al. 2004). After SPaseI cleavage at their N-termini, both proteins remain attached to the ER membrane and associate into prM/E heterodimers. This association is the driving power for the budding of the nucleocapsid (NC). The NC is formed on the cytoplasmic site of the ER membrane, when protein C interacts with the viral genomic RNA. By budding into the ER, the viral envelope is acquired; as a by-product, capsidless, non-infectious subviral particles (SVP) are produced. These immature virions are then transported to the late trans-Golgi Network (TGN), where prM is converted into its active form by the host protease furin (Stadler et al., 1997). This cleavage induces a structural reorganisation of protein E, resulting in the decomposition of the prM/E heterodimers. Instead, fusion-competent E protein homodimers are generated. Finally, Golgi vesicle and host cell plasma membrane fuse, thereby releasing mature virions and SVP (Mandl et al., 2005; Stiasny and Heinz, 2006).
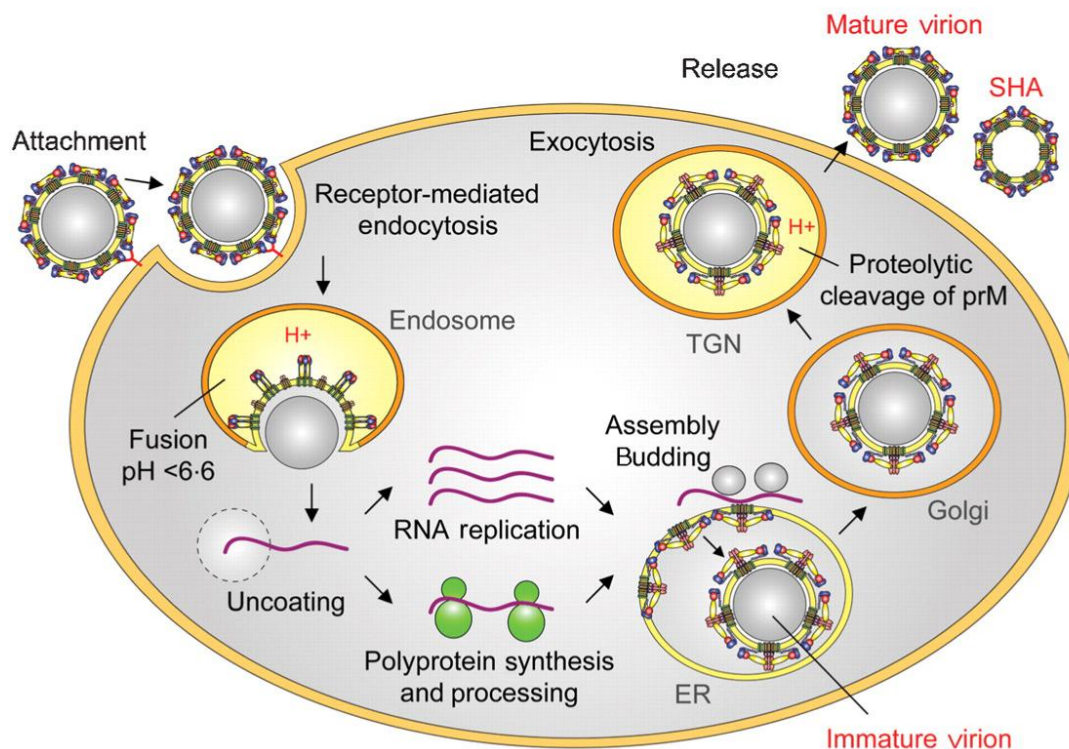
Figure 4. Flaviviral life cycle (adapted from Stiasny and Heinz, 2006). See text for details.

1.2.4.   Capsid protein C and pre-matrix protein prM

The capsid protein C is located at the N-terminus of the precursor polyprotein. It is anchored in the ER membrane via a C-terminal hydrophobic region, which is highly conserved among flaviviruses and is responsible for the translocation of protein prM into the lumen of the ER (Nowak et al., 1989) (Fig. 5). Mature protein C (Amberg et al., 1994; Lobigs et al. 1993), with a high content of basic residues, a mass of 11 kDa and an isoelectric point of 13, is

released from the polypeptide chain by cleavage of the viral NS2B/3 protease. The nucleocapsid is formed by multiple copies of protein C that surround a single copy of the genomic RNA. The first three-dimensional model of the YFV and DENV-2 capsid protein was proposed by Jones et al., 2003, based on nuclear magnetic resonance (NMR) spectroscopy analysis. Ma et al, 2004 improved this structure by using the same method. Again in 2004, Dokland and colleagues succeeded in resolving the structure of the C protein of the Kunjin subtype of the WNV by using high-resolution X-ray crystallography. In good agreement with biochemical evidence (Jones et al., 2003; Kiermayr et al., 2004), the structural data suggest that the elementary unit of the NC is a dimer of protein C, with each subunit composed of four alpha helices (α1 to α4) that are interspersed by short loops (Fig. 5A and B).

A

B
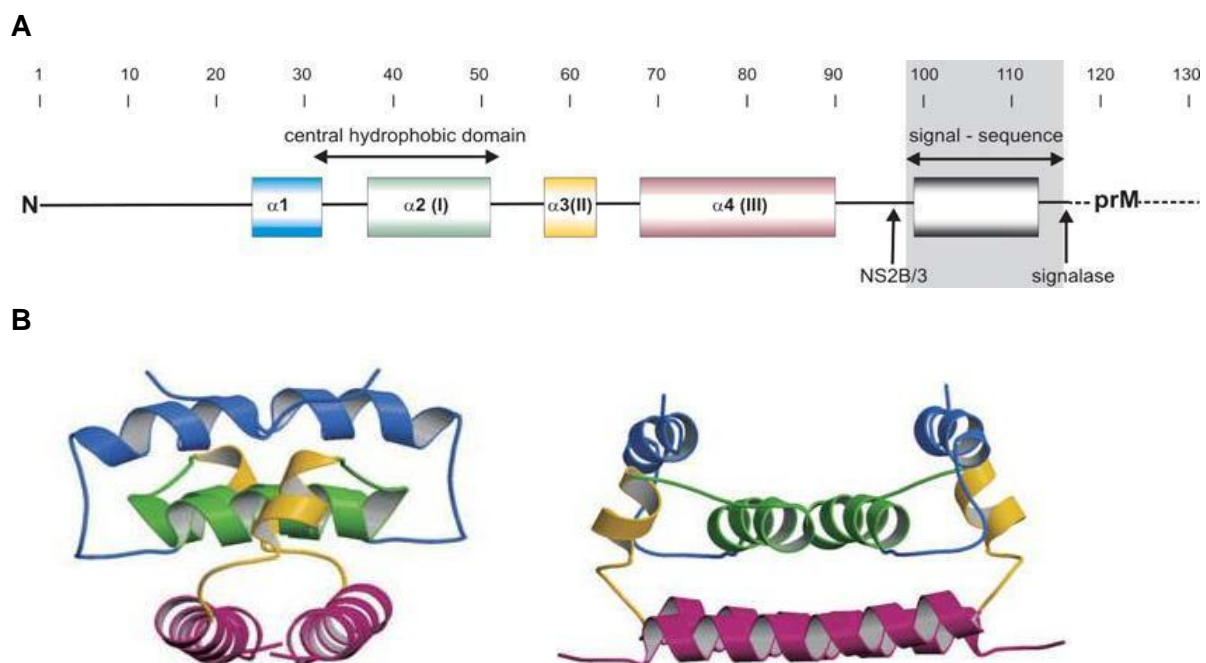
Figure 5. Capsid protein C. (A) Schematic diagram of TBEV protein C. The ER membrane is represented by a shaded field. The signal sequence for the translocation of prM into the ER as well as the cleavage sites of the viral NS2B/3 protease and the host signalase are indicated. (B) Views of the DV2 capsid protein dimer structure as determined by Ma et al., 2004. Alpha-helix α1 is shown in blue, α2 in green, α3 in yellow and α4 in magenta (adapted from Mandl 2005).

Helices α2 and α4 are primarily responsible for the intra-dimeric interactions. The α4 helices of both monomeric subunits interact with each other in a structural arrangement that has characteristics of a coiled-coil structure (Kofler et al., 2002; Ma et al., 2004). They form a surface containing a large number of positively charged side chains which presumably interacts with the negatively charged genomic RNA. This interaction is the basis for the formation of the NC. Hydrophobic helices α1 and α2 lie at the opposite site of the dimer and are responsible for

interactions with the membrane that surrounds the NC (Mandl et al. 2004) (Fig. 5B). Samsa et al., 2009 showed that protein C of DENV accumulates on the surface of ER-derived lipid droplets and that this amassment is important for the formation of viral particles. Substitution of only two amino acids in the hydrophobic core region of DENV protein C results in an accumulation of protein C in the cytoplasm rather than on lipid droplets. As a consequence, viral RNA synthesis is reduced.

1.2.5. Processing of flaviviral structural proteins

The TBEV polyprotein is integrated cotranslationally into the host cell´s ER membrane. This makes it impossible for the cytoplasmic viral protease to access the cleavage sites within the ER lumen. Therefore, viral and host cell proteases are required to process the polyprotein precursor into its mature proteins. Proteins prM and E are cotranslationally released from the polyprotein precursor by SPase I cleavage on the luminal side of the ER. In contrast, protein C and prM are separated by regulated cleavages by the viral NS2B/3^{pro} and the host cell SPase I on both sides of the ER (Fig. 6).
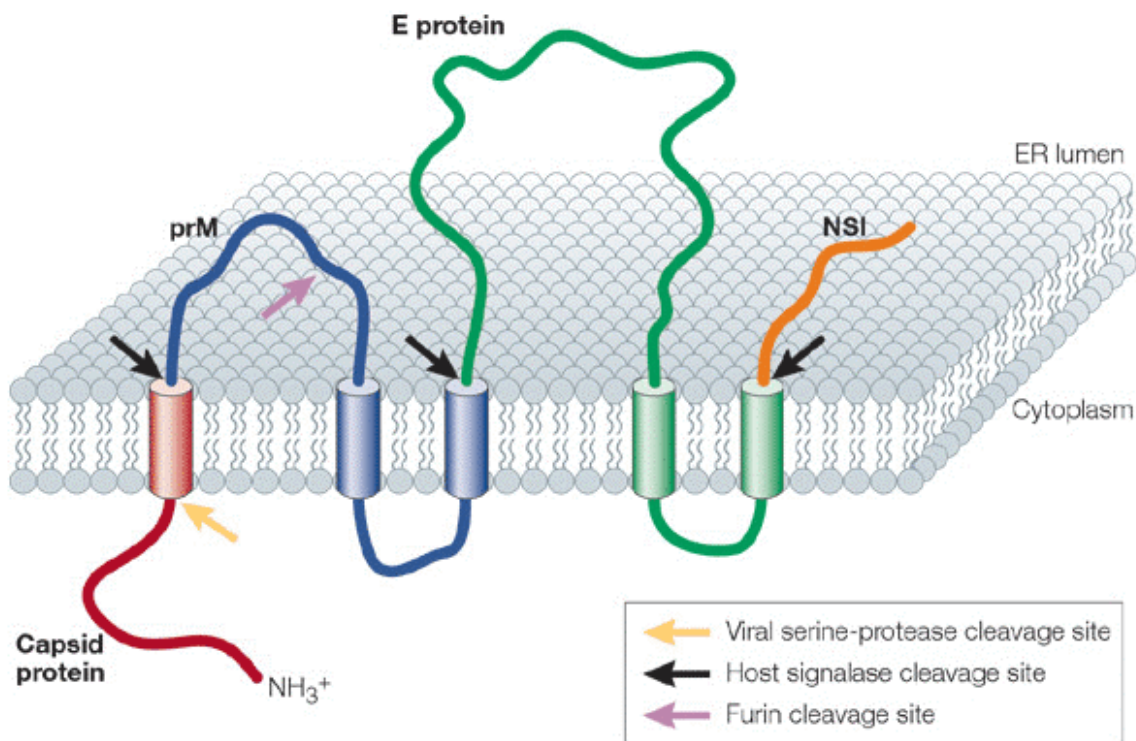

Figure 6. Topology of the flaviviral structural proteins across the ER membrane. Transmembrane helices are indicated by cylinders, cleavages sites for the various proteases are indicated by arrows (adapted from Mukhopadhyay et al., 2005). E, envelope; NS1, non-structural protein 1; prM, pre-membrane protein.

The C-terminus of protein C of TBEV contains two putative NS2B/3[pro] cleavage motifs which are separated by one single amino acid. Furthermore, it bears an internal hydrophobic signal sequence which directs the translocation of the prM protein into the ER lumen. This signal sequence is removed by cleavage of the NS2B/3[pro] on the cytosolic side of the ER membrane, thereby liberating mature protein C (Amberg et al.; 1994, Yamshchikov et al., 1994). Only after this cleavage has taken place, the signal peptide can move towards the ER lumen. As a consequence, the host cell SPaseI can assess its previously hidden cleavage site. Thereby, the N-terminus of protein prM is liberated, which in turn is important for the correct processing of protein E (Lorenz et al., 2002; Konishi et al., 1993). Mature protein C associates with the viral genome to form the nucleocapsid. By budding into the ER lumen, the viral envelope is acquired (Chambers et al., 1990). Budding is disturbed when the two cleavage events at the C-prM junction of the Murray Valley encephalitis virus polyprotein are uncoupled. As a result, the production of non-infectious SVP (Lobigs et al., 2004) is increased, emphasising the importance of this coordinated processing for efficient virion assembly (Amberg and Rice, 1999; Lobigs et al., 2004; Stocks et al., 1998).

1.3.   Viral proteases and their specificity

1.3.1.   TBEV NS2B/3 protease

The viral protease NS2B/3 is formed by a complex made of the protease domain of the NS3 protein and its cofactor, NS2B. The crystal structures of the protease domain of WNV and DENV-2 NS3 protein have already been determined. The same was accomplished for the full-length NS3 protein of MVEV and DENV in complex with the central hydrophilic domain of the NS2B cofactor. (Fig. 7) (Aleshin et al., 2007; Assenberg et al., 2009; Erbel et al., 2006; Luo et al., 2008). However, until now a structure of the TBEV NS2B/3[pro] has not been determined.

The NS3 protein comprises a helicase as well as a triphosphatase activity in its C-terminal domains. In addition, a protease activity is located at its N-terminus (Lobigs, 1992; Preugschat et al., 1990). The NS2B/3[pro] is a serine protease with a trypsin like fold, consisting of two ß-barrels with the catalytic triad (His51-Asp75-Ser135) being located in between. The hydrophilic core domain of the co-factor NS2B stabilises the protease and is also involved in substrate binding (Erbel et al., 2006). The N- and C-terminal flanking regions are rich in hydrophobic residues and anchor the NS3 protease complex to the ER membrane (Chambers et al., 1993; Clum et al., 1997). As a result, the protease is positioned near its cleavage sites, all of which are proximal to transmembrane regions (Bera et al., 2007). However, *in vitro* processing studies revealed that NS2B/3 precursors are capable of performing self-cleavage also in the absence of membranes (Shiryeav et al., 2007; Erbel et al., 2006; Chambers et al., 1995; Pugachev et al., 1993).

Without the NS2B domain, the NS3 protease is proteolytically inert. However, it has been shown by Leung et al. (2001) that the full length NS2B sequence can be substituted with its 35-48-residue core region without any loss of cofactor activity. Based on this knowledge, several groups have generated proteolytically active single chain NS2B-NS3 modular proteases by coupling the NS2B core region to the NS3 protease domain by a synthetic nona-peptide linker. Although this linker replaces the NS2B/3^{pro} cleavage site at the NS2B/NS3 junction, some of these proteases still exhibited auto-catalytic activity (Chappell et al., 2007, Shiryaev et al., 2007).

The TBEV NS3 protease usually cleaves after two basic amino acid residues at the P2 and P1 positions (nomenclature according to Schechter et al., 1967) followed by one with a short, uncharged side chain (Chambers et al., 1990; Rice et al., 1990). At the C-prM junction, the TBEV exhibits two potential NS2B/3 cleavage motifs which are separated by a single amino acid. Schrauf and co-workers (2009) showed that, in vivo, the downstream cleavage site is predominantly used and that the upstream cleavage motif cannot be accepted when the downstream cleavage motif had been removed by mutation.
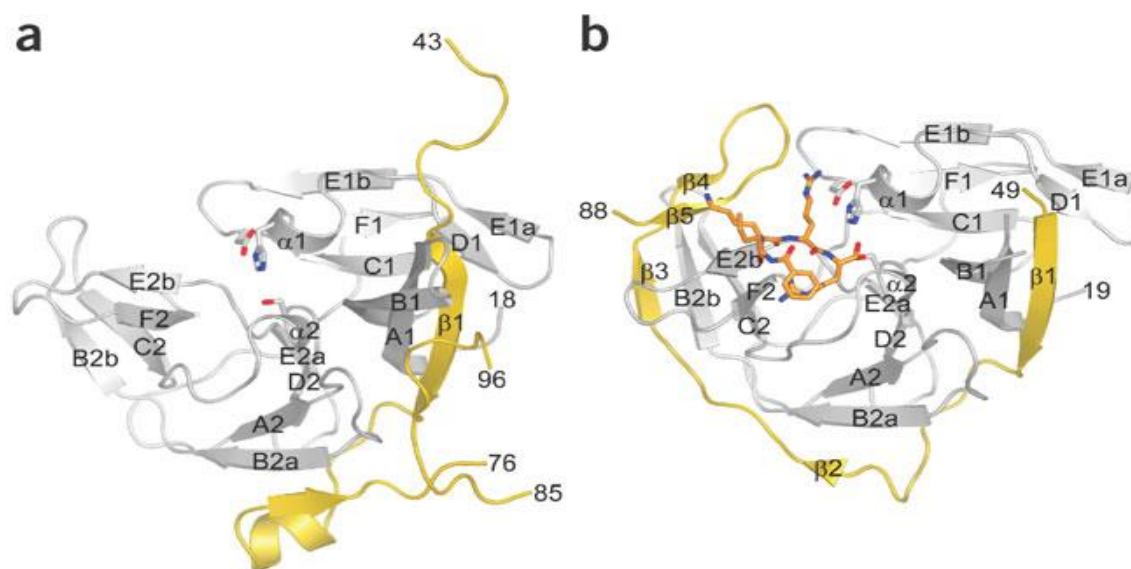
Figure 7. Structure of the NS2B-NS3 complex in the absence (A) or presence (B) of an inhibitor. (A) Open DEN NS2B-NS3 complex. (B) WNV NS2B-NS3 complex with bound inhibitor. NS2B is shown in yellow, NS3 in grey, the inhibitor in orange. Adapted from Erbel et al., 2006.

1.3.2. Foot and Mouth Disease Virus (FMDV) 3C protease

FMDV is a member of the *Aphtovirus* genus in the *Picornaviridae* family, an important group of mammalian single-strand, positive-sense RNA viruses that includes human rhinovirus, poliovirus and hepatitis A virus (Mason et al., 2003). The viral genome is translated as a single polyprotein that is cleaved into functional proteins by a virally encoded leader (L^{pro}) and 3C protease (3C^{pro}) (Fig. 8).
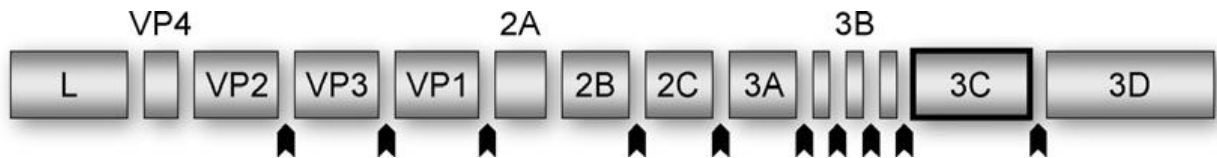
Figure 8. Processing of the FMDV polyprotein. 3C protease cleavages are indicated by arrows. Adapted from Curry *et al.*, 2006.

During translation, the L^{pro} is liberated from the N-terminus of the polyprotein by autolytic processing. A ribosomal skip mechanism inhibits correct peptide bonding at the C-terminus of protein 2A and is therefore responsible for the separation of structural and non-structural proteins (Mason et al., 2003). Ten of the 13 cleavages are performed by the highly conserved 3C^{pro} (Curry et al., 2006). Furthermore, the 3C^{pro} is responsible for several cleavages of transcription and translation-associated host cell proteins. In this way, the protease contributes to a pronounced inhibition of host-cell protein synthesis (Belsham et al., 2000; Li et al., 2001) while at the same time promoting viral replication.

Recent structural work on FMDV 3C^{pro} (Birtley et al, 2005; Sweeney et al, 2006; Yin et al, 2005) shows that its catalytic domain adopts a conformation similar to chymotrypsin-like serine proteases (Hedstrom et al., 2002). Nevertheless, in accordance with its catalytic triad being Cys163-His46-Asp84 (Birtley et al., 2005), the FMDV 3C^{pro} is a member of the rare group of chymotrypsin-like cysteine proteases (Barret & Rawlings, 2001). Presumably, Asp84 forms hydrogen bonds with His46 of the catalytic triad and is stabilised by Ser182. Even though these interactions have parallels with those of other, well studied, chymotrypsin-like serine proteases, the exact catalytic mechanism of FMDV 3C^{pro} remains unclear.
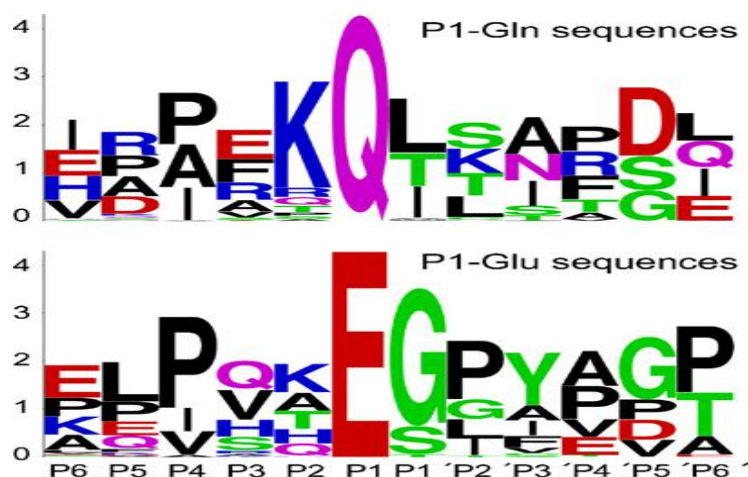

Figure 9. Sequence logos of the polyprotein junctions cleaved by FMDV 3C^{pro}. Adapted from Curry et al., 2006.

Like other picornaviral 3C proteases, the FMDV 3C^{pro} processes peptide sequences with a hydrophobic residue at P4, such as proline or isoleucine, a glutamine residue at P1 and

a small hydrophobic amino acid at P1´. However, unlike other picornaviral proteases, the FMDV 3C^{pro} tolerates larger residues at P1´ and cleaves sequence motifs with P1-glutamate at rates similar to those with P1-glutamine (Fig. 9) (Curry et al., 2006). The significance of P4, P1 and P1´positions was confirmed by Birtley et al., 2005.

1.3.3. Human Immunodeficiency Virus (HIV-1) protease

HIV is a member of the *Retroviridae* family, genus *lentivirus* (van Regenmortel et al., 2009). On the basis of serological properties and sequence analysis of the genome, the HIV isolates are grouped into two types, HIV-1 and HIV-2. HIV-1 is the most common and pathogenic one.

The positive sense single stranded RNA genome of HIV is 9.2 kb long. It is flanked by long terminal repeats (LTR) that regulate viral integration into the host genome, as well as several phases of viral gene expression. From the unspliced polycistronic mRNA, two large large precursor polyproteins, Gag and Gag-Pol, are translated. These are subsequently processed by the virally encoded, autocatalytically excised protease (HIV^{pro}) to yield the mature viral proteins (Frankel et al., 1998; Vogt et al., 1996). The protease, together with the viral reverse trancriptase and integrase, are products of the pol open reading frame. The structural proteins (matrix, capsid and nucleocapsid, and the peptides p1, p2 and p6) are generated by processing of the gag polyprotein.

Table 1. Processing sites in gag and gag-pol polyproteins recognized by HIV-1 protease. Adapted from Ozer et al., 2006.

Substrate sites in the Gag Polyprotein

matrix–capsid (ma–ca)	SQNY*PIVQ
capsid–p2 (ca–p2)	ARVL*AEAM
p2–nucleocapsid (p2–nc)	ATIM*MQRG
p1–p6	PGNF*LQSR
nucleocapsid–p1 (nc–p1)	RQAN*FLGK

Substrate sites in the Pol polyprotein

reverse transcriptase–RnaseH (rt–rh)	AETF*YVDG
RnaseH–integrase (rh–in)	RKIL*FLDG
protease–reverse transcriptase (pr–rt)	TLNF*PISP
transframe–protease (tf–pr)	SFNF*PQIT

The HIV-1 protease is an aspartic protease that is active as a homodimer. The two identical subunits are non-covalently associated (Lapatto et al., 1989; Navia et al., 1989;

Wlodawer et al., 1989). Each monomer is encoded by 99 amino acids and consists of nine β-sheets and one α-helix. The catalytic triad (Asp25, Thr26 and Gly27) lies in a loop at the end of the third ß-sheet; the active site is formed at the dimer interface, where each subunit contributes one catalytically relevant aspartic acid residue (Asp25). The N-termini of each subunit contribute to dimer stability (Ishima et al., 2000; Louis et al., 2003) whereas the C-termini are essential for dimerisation.

The HIV-1 polyprotein is processed by the HIV-1 protease at unique sites. Although a comparison of these sequences reveal a broad substrate specificity and the absence of a consensus sequence, these sites share several characteristics (Debouck et al., 1992). All HIV proteases prefer hydrophobic residues at the P1 and P1', small ones at the P2, aliphatic and glutamine/glutamiate residues at the P2' position; aromatic side chains are poorly accepted at P3' position. In contrast to cellular aspartic proteases, HIV-1 proteases recognise sequences which bear an aromatic residue such as tyrosine or phenylalanine at P1, followed by proline at P1´ (Pettit et al., 1991; Roberts et al., 1990). This property, together with its significance in the viral life cycle, makes the protease an important target for antiviral therapy (Sommadossi et al., 1999; Wlodawer et al., 1993). Although several inhibitory compounds have already been developed, their clinical utility is limited by the emergence of drug resistant viruses.

2. AIMS

Proteolysis by virally encoded enzymes is essential for the replication of many viruses (Krausslich and Wimmer 1988; Hellen et al., 1989). This makes protease inhibitors an important target for the antiviral therapy. However, the development of such inhibitors is difficult, as viral proteolysis is a highly complex process. Cleavage sites are often insufficiently defined and the development of resistance to inhibitors due to the high mutation rate of the viral RNA polymerase is an enduring problem.

To overcome these difficulties, detailed information on substrate specificity of the proteases and their resistance to inhibitors is needed. Therefore, novel systems are required that allow the generation of a large number of escape mutants in a random manner and to analyse them systematically. Thus, the overall objective of this work is to develop a sensitive system that allows us to examine the properties of various viral proteases. As an RNA virus, the TBEV itself could be used as such a tool. It produces a large number of mutants due to its error-prone polymerase, with approximately one error per synthesised genome. Thus, the presence of suboptimal substrates or protease inhibitors should provoke the production of infectious escape mutants. These can be easily selected over non-infectious virions by measuring their infectivity. Finally, analysis of the escape mutants should give important insights into the mechanisms leading to resistance to protease inhibitors.

Aim 1: Establishment of a cell based system to investigate various viral proteases

To establish a system that allows us to investigate several distinct proteases, we intended to modify the genome of TBEV in such a way that the production of infectious virions becomes dependent on the activity of a heterologous protease instead of the TBEV protease (NS2B/3^{pro}). To this end, the NS2B/3^{pro} capsid cleavage site in the C-terminal region of TBEV should be replaced by a heterologous cleavage site such as that of the FMDV 3C^{pro} or HIV-1^{pro}. The infectivity of these non-infectious mutants should then be regained by providing the corresponding heterologous protease either in *cis* or in *trans*. The establishment of such a system should allow us to study the substrate specificity of these heterologous proteases and to investigate its ability to become resistant to a corresponding inhibitor (manuscript 1).

Aim 2: Establishment of a biochemical system to investigate the NS2B/3^{pro} and to find suitable substrates needed to accomplish aim 1

To generate the above mentioned mutants of TBEV, with cleavage sites for a heterologous protease instead of one for the NS2B/3^{pro}, it is important to have a profound understanding of the cleavage mechanism and substrate specificity of the NS2B/3^{pro}. To this

end, we intend to establish a biochemical assay that allows us to investigate *in vitro* the cleavage of the NS2B/3^{pro}. Once established, this assay should be modified in a way to allow investigation of other viral proteases. This assay should then be used to test the suitability of diverse substrates to be processed by various heterologous proteases (e.g. HIV-1 protease) as described in aim 1 (manuscript 2).

Aim 3: Determination of the three dimensional structure of the TBEV NS2B/3^{pro}

Finally, we intended to determine the structure of the TBEV NS2B/3^{pro} by X-ray crystallography. To this end, expression and purification protocols should be optimised to yield NS2B/3 protein in a quantity and quality sufficient for structure determination. A 3D structure of the physiologically relevant NS2B/3^{pro} would represent a significant impulse for the development of antiviral substances against TBEV (manuscript 3).

3. MANUSCRIPT 1

Generation and genetic stability of tick-borne encephalitis virus mutants dependent on processing by FMDV 3C protease

Sabrina Schrauf*[1], Martina Kurz[§1], Christian Taucher*[†], Christian W. Mandl*[††], and Tim Skern[§¶]

Published in the *Journal of General Virology*

S. Schrauf developed the assay and performed the initial stability tests. All experiments were repeated by M. Kurz in order to verify the reproducibility of the results. Further stability test, as well as the generation and characterisation of the revertants was performed by M. Kurz. The paper was written by M. Kurz, S. Schrauf and T. Skern.

[1]These authors contributed equally to this work

*Institute of Virology, Medical University of Vienna, A-1095 Vienna, Austria

[†] Current address: Bristol-Myers Squibb GmbH, A-1101 Vienna, Austria

[††] Current address: Novartis Vaccines and Diagnostics, Inc., Cambridge, MA 02139, USA

[§] Max F. Perutz Laboratories, Medical University of Vienna, A-1030 Vienna, Austria

[¶]Corresponding Author: Tim Skern

Generation and genetic stability of tick-borne encephalitis virus mutants dependent on processing by the foot-and-mouth disease virus 3C protease

Sabrina Schrauf,[1]† Martina Kurz,[2]† Christian Taucher,[1]‡ Christian W. Mandl[1]§ and Tim Skern[2]

Correspondence
Tim Skern
timothy.skern@meduniwien.ac.at

[1]Institute of Virology, Medical University of Vienna, Kinderspitalgasse 15, A-1095 Vienna, Austria

[2]Max F. Perutz Laboratories, Medical University of Vienna, Dr. Bohr-Gasse 9/3, A-1030 Vienna, Austria

Mature protein C of tick-borne encephalitis virus (TBEV) is cleaved from the polyprotein precursor by the viral NS2B/3 protease (NS2B/3[pro]). We showed previously that replacement of the NS2B/3[pro] cleavage site at the C terminus of protein C by the foot-and-mouth disease virus (FMDV) 2A StopGo sequence leads to the production of infectious virions. Here, we show that infectious virions can also be produced from a TBEV mutant bearing an inactivated 2A sequence through the expression of the FMDV 3C protease (3C[pro]) either *in cis* or *in trans* (from a TBEV replicon). Cleavage at the C terminus of protein C depended on the catalytic activity of 3C[pro] as well as on the presence of an optimized 3C[pro] cleavage site. Passage of the TBEV mutants bearing a 3C[pro] cleavage site either in the absence of 3C[pro] or in the presence of a catalytically inactive 3C[pro] led to the appearance of revertants in which protein C cleavage by NS2B/3[pro] had been regained. In three different revertants, a cleavage site for NS2B/3[pro], namely RR*C, was now present, leading to an elongated protein C. Furthermore, two revertants acquired additional mutations in the C terminus of protein C, eliminating two basic residues. Although these latter mutants showed wild-type levels of early RNA synthesis, their foci were smaller and an accumulation of protein C in the cytoplasm was observed. These findings suggest a role of the positive charge of the C terminus of protein C for budding of the nucleocapsid and further support the notion that TBEV protein C is a multifunctional protein.

INTRODUCTION

Tick-borne encephalitis virus (TBEV), from the family *Flaviviridae*, is an enveloped, positive-stranded RNA virus. Its genome is translated into one large polyprotein that traverses the endoplasmic reticulum (ER) several times. This polyprotein is co- and post-translationally cleaved by viral and cellular proteases (Lindenbach *et al.*, 2007). For productive infection, proteolytic processing and assembly of the virus particle must be closely co-ordinated (Amberg & Rice, 1999; Lindenbach *et al.*, 2007). Structural protein processing leads ultimately to the formation of protein C and the two glycoproteins prM (membrane) and E

(envelope) (Lobigs, 1993). The latter remain anchored in the ER membrane, whereas mature protein C participates in RNA packaging and particle assembly with the newly synthesized mRNA to form the nucleocapsid (Mandl, 2005). Preformed nucleocapsid has not yet been detected in infected cells. Given that budding of flavivirus membranes is driven by the interactions of prM and E independently of protein C or the assembled nucleocapsid (Lobigs & Lee, 2004), the co-ordinated cleavage events at the C–prM junction appear to be a key to the efficient incorporation of nucleocapsid during flavivirus assembly.

Cleavage by the viral protease NS2B3 (NS2B/3[pro]) in the C-terminal region of protein C is a prerequisite for the subsequent signallase cleavage inside the lumen of the ER that generates the N terminus of prM (Amberg & Rice, 1999). The signallase cleavage site of prM is thought to be maintained in a predominantly cryptic conformation until cleavage of C by NS2B/3[pro]. This prevents cotranslational signallase cleavage of prM and should result in transient

---

†These authors contributed equally to this work.

‡Present address: Bristol-Myers Squibb GmbH, Columbusgasse 4, A-1101 Vienna, Austria.

§Present address: Novartis Vaccines and Diagnostics, Inc., 350 Massachusetts Ave, Cambridge, MA 02139, USA.

A supplementary table is available with the online version of this paper.

expression of a C–prM intermediate at the putative flavivirus assembly site on the ER membrane until cytosolic processing by the viral protease has taken place (Lobigs & Lee, 2004). The importance of the concerted progression of the cleavage events at the C–prM junction has been documented in yellow fever virus, Murray Valley encephalitis virus and TBEV. Inhibition of protein C cleavage or disturbance of the cleavage order interfered with or even abrogated viral growth and infectivity (Amberg & Rice, 1999; Chappell *et al.*, 2005; Lee *et al.*, 2000; Lobigs & Lee, 2004; Schrauf *et al.*, 2008; Stocks & Lobigs, 1998).

The exact position of the TBEV NS2B/3^{pro} protein C cleavage site and the cleavage sequence itself are highly flexible (Schrauf *et al.*, 2008). Furthermore, this NS2B/3^{pro} cleavage site can be replaced by the FMDV (foot-and-mouth disease virus) 2A protein sequence in both TBEV and West Nile Virus (WNV) (Schrauf *et al.*, 2009). This 2A sequence enables the release of mature protein C from the polyprotein in the absence of proteolysis through a 'StopGo' mechanism that promotes reinitiation on the translating ribosome (Atkins *et al.*, 2007). The release of protein C thus occurs co-translationally and prematurely compared with that in the wild-type. In addition, the mechanism of action of FMDV 2A results in a protein C with a 19 aa C-terminal extension. Examination of the resulting TBEV mutants showed an altered buoyant density and a reduction of early RNA synthesis upon viral infection but not with electroporation of naked RNA, suggesting previously unrecognized roles of protein C in uncoating and/or early RNA replication (Schrauf *et al.*, 2009).

Here, we further characterize the function(s) of protein C by making the C/prM cleavage dependent on the synthesis of a heterologous protease, the FMDV 3C^{pro} (Curry *et al.*, 2007).

RESULTS

Constructing a system allowing protein C cleavage to be performed by a heterologous protease

TBEV variants in which the NS2B/3^{pro} cleavage site at the C terminus of the protein C had been replaced by the 20 aa FMDV 2A protein still produce infectious virions (Schrauf *et al.*, 2009). Generation of the mature protein C is accomplished by a 'StopGo' mechanism in which a PGP motif located at the C-terminal end of the FMDV 2A protein causes the ribosomes to skip peptide bond formation (Donnelly *et al.*, 2001).

Mutants lacking the PGP motif cannot perform the StopGo process, preventing separation of protein C from prM and consequently virion production, even in the presence of viral RNA and protein synthesis. Could the expression of the 3C^{pro} of FMDV rescue mutants lacking PGP by cleaving between the C terminus of protein C and the N terminus the FMDV 2A sequence? To address this question, we took the mutant TBEV 2AΔ3 (Fig. 1a) (Schrauf *et al.*, 2009) lacking

the PGP motif at the C terminus of 2A. At the N terminus of the 2A sequence, there is a suboptimal cleavage motif for the FMDV 3C^{pro}, RGKQ*TLN (the asterisk indicates the cleavage site). To improve this cleavage site, we inserted the amino acids PV and QV between the glycine and lysine residues, optimizing the P4 and P3 positions (Fig. 1a). It has been previously demonstrated (Birtley *et al.*, 2005) that the P4 position is also important for 3C^{pro} cleavage. To further enhance the system, we also introduced the mutation E122G into protein E (PV-2AΔ3-H and QV-2AΔ3-H, Fig. 1a). This substitution enables the virus to use heparan sulphate to bind to BHK-21 cells, thus increasing the specific infectivity and binding affinity by up to 10-fold (Kroschewski *et al.*, 2003).

To express FMDV 3C^{pro}, we used a TBEV replicon lacking the coding region for the structural proteins prM and E (Fig. 1b). This replicon cannot produce infectious viral particles on its own but can replicate RNA and express a heterologous protein from a second ORF at the 3′ end (Gehrke *et al.*, 2005). We placed the FMDV 3C^{pro} coding sequence after the EMCV IRES element (ΔME-I-3C, Fig. 1b). We also prepared a replicon expressing an inactive 3C^{pro} (ΔME-I-3Ci, Fig. 1b).

The rationale was that when one of the TBEV 2AΔ3 mutants and the protease-encoding TBEV replicon are present within the same cell, the FMDV 3C^{pro} expressed from the replicon should recognize its cleavage motif at the C–prM junction of the TBEV-2AΔ3 mutant. This will liberate mature protein C from the signal sequence for prM and allow the signallase to access its cleavage site at the N terminus of prM. These cleavages then enable the generation of all structural proteins leading to the production of single-round infectious particles.

Heterologous expression of the FMDV 3C protease results in virion production

To ensure that replication and protein production still function correctly, we first singly transfected BHK-21 cells with RNAs of ΔME-I-3C, PV-2AΔ3-H and QV-2AΔ3-H. Viral protein expression was analysed after transfection and after the first passage of the supernatant by immuno-fluorescence (IF) using an antiserum that predominantly recognizes the structural protein E as well as the non-structural protein NS1. Cells transfected with all RNAs showed viral protein expression (Fig. 2a, upper panel). Thus, all RNAs were competent for RNA replication, as we have previously shown that only replication-competent mutants generate and process sufficient proteins to be visible by IF (Kofler *et al.*, 2002). The efficiencies of transfection were approximately 30–40 % for ΔME-I-3C and 80 % for PV-2AΔ3-H and QV-2AΔ3H (Fig. 2a, upper panel). As the replicons do not express protein E, it is possible that the efficiency of their transfection is underestimated compared with those of the modified viruses that express both protein E and NS1 that can be recognized by the antiserum used. Twenty-four hours later,
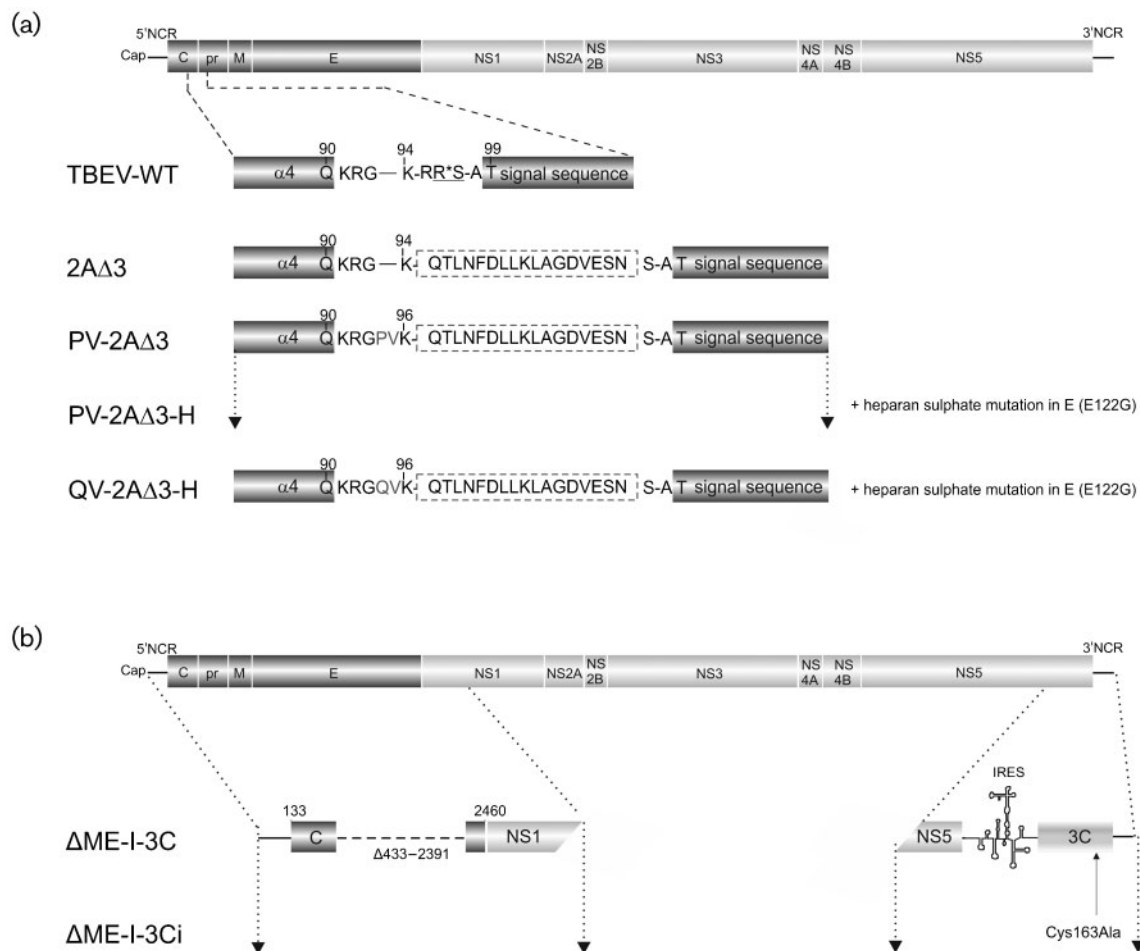
Fig. 1. Structural organization of flavivirus mutants. (a) Drawing of the TBEV full-length genome and expanded views of the protein C of TBEV WT and each TBEV mutant (not drawn to scale). The NS2B/3 cleavage site is underlined and the cleavage position is marked by an asterisk. The inserted FMDV 2AΔ3 protein sequence is surrounded by a dashed line. The amino acids PV and QV preceding residue K96 were inserted to optimize the FMDV 3C^{pro} cleavage sequence. The numbers at the top refer to the amino acid positions within protein C. (b) Diagram of the TBEV genome and an expanded view of the structural protein region and the 3′NCR of each constructed TBEV replicon. The 3C^{pro} sequence was amplified by PCR from the FMDV strain O/ SAR/19/00 and placed behind an EMCV IRES as described in methods. Engineered mutations are shown with the corresponding designation.

phase-contrast microscopy revealed that cells transfected with these showed a strong cytopathic effect (CPE) with the concomitant death of the majority of cells. Passaging of supernatants from the transfected cells did not, however, lead to the production of viable virus (Fig. 2a, lower panel).

We then transfected PV-2AΔ3-H or QV-2AΔ3-H RNAs together with RNA from the replicon ΔME-I-3C. After 24 h, cells stained almost 100 % positive for viral protein expression (Fig. 2b, upper panel), indicating a high electroporation efficiency. As the two replicating constructs both produce the NS1 protein, the antiserum (predominantly recognizing proteins E and NS1) cannot distinguish between them; therefore, we were not able

to determine how many cells were double-positive. Nevertheless, IF analysis of singly transfected cells (80–90 % for the TBEV-mutant RNAs and 30–40 % for the ΔME-I-3C replicon RNA) indicated that at least 20–30 % of the co-transfected cells expressed both constructs. We then used the supernatants of the electroporated cells to infect fresh BHK cells. Using supernatants transfected with the RNAs PV-2AΔ3-H and ΔME-I-3C, about 40–50 % of the fresh BHK cells were positive for the production of viral proteins as detected by IF using the antiserum staining for the proteins E and NS1, indicating that virions had been produced by doubly transfected cells (Fig. 2b, lower panel). In contrast, no positive cells were observed with supernatants from cells transfected with QV-2AΔ3-H or ΔME-I-3C RNA (Fig. 2b), suggesting that
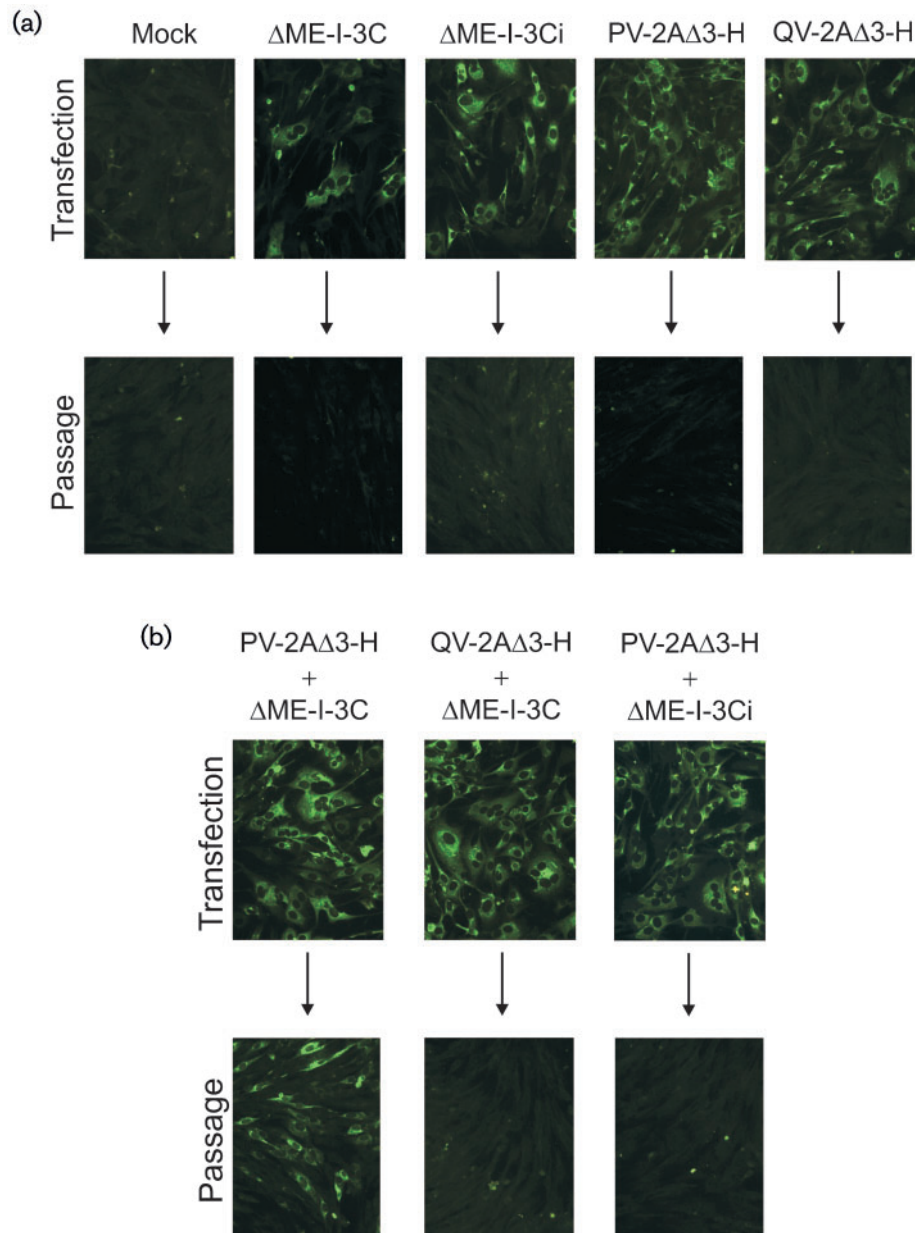
Fig. 2. Cotransfection of TBEV PV-2AΔ3-H and QV-2AΔ3H with TBEV replicons expressing active or inactive 3C^{pro}. (a and b) BHK-21 cells were transfected with mutant RNAs or were mock transfected and viral protein expression was detected with a polyclonal serum 1 day post-transfection. Supernatants harvested 1 day post-electroporation were used to inoculate fresh BHK-21 cells (arrows), and viral protein expression monitored by IF 1 day post-infection (bottom of panels a and b).

the only cleavage site recognized was that with the optimal PV sequence at P3 and P4.

Virion production depends on expression of an active 3C protease

The above results implied that the FMDV 3C^{pro} could rescue the defect in PV-2AΔ3-H. To provide further evidence and to exclude any involvement of NS2B/3^{pro}, we constructed a ΔME replicon containing an inactive 3C^{pro}

(ΔME-I-3Ci; Fig. 1b) in which the active site cysteine (Cys163) had been mutated to alanine. Cells transfected with ΔME-I-3Ci RNA alone showed viral protein expression and a CPE (Fig. 2a and data not shown). However, the supernatant of cells co-transfected with ΔME-I-3Ci and PV-2AΔ3-H RNAs did not contain any infectious particles (Fig. 2b). This strengthens the notion that an active 3C^{pro} is required for virion production from PV-2AΔ3-H and that the 3C^{pro} can replace the NS2B/3^{pro} in protein C cleavage.

Protein prM is only formed in the presence of an active 3C protease

To assess protein processing at the C–prM junction in cells transfected with full-length TBEV mutants and the ΔME replicons, we examined protein prM production. All constructs used [i.e. the full-length TBEV mutants (Fig. 1a) and the ΔME replicon (Fig. 1b)] code for protein C; therefore, its analysis provides no information on 3C^{pro} cleavage. However, protein prM is only expressed in cells containing the TBEV mutants when 3C^{pro} cleaves at the cytoplasmic side of the signal sequence for prM, allowing the signallase access to its cleavage site at the N terminus of protein prM. We therefore lysed cells transfected with various RNAs 22 h after transfection and detected the structural proteins E and prM with polyclonal antisera raised against all structural proteins of TBEV and against prM itself, respectively (Fig. 3). A prominent band corresponding to protein E, indicating viral protein synthesis and processing, was detected in lysates of cells transfected with a full-length TBEV mutant (Fig. 3, upper panel). In contrast, the band was absent from the lysate transfected with the ΔME replicon alone. With the anti-prM serum (Fig. 3, lower panel), a prominent band with a size corresponding to that of protein prM was only detected in lysates of cells transfected with combinations of PV-2AΔ3-H + ΔME-I-3C or PV-2AΔ3 + ΔME-I-3C (Fig. 3, lower panel, lanes 4 and 7). Cell lysates of the full-length TBEV mutants with the PV-2AΔ3 motif either alone or in combination with ΔME-I-3Ci failed to produce any protein prM, indicating no protein processing at the C–prM junction in the absence of an active 3C^{pro}. Furthermore, no protein prM was visible in lysates from cells transfected with RNAs containing the QV-Δ3-H motif alone or in combination with the 3C^{pro}, indicating that the 3C^{pro} cleavage requires a suitable P4 residue. We did not, however, observe any uncleaved C-2AΔ3-prM protein in the Western blot with the anti-prM antiserum (Fig. 3, lower panel). We believe that this is due to the high background at the predicted position (about 35 kDa) as well as a previously observed inability of this antiserum to recognize the C–prM protein. No differences were visible between lysates containing the PV-2AΔ3 mutant with or without the heparan sulphate mutation; hence, this additional mutation has no impact on processing of the structural proteins (Fig. 3, lanes 4 and 7).

Supplying the 3C protease *in cis* also results in TBEV mutant virion production

Can 3C^{pro} also be provided *in cis*? Accordingly, we constructed full-length TBEV mutants containing the 3C^{pro} cleavage motif P-V-K-Q*T at the C terminus of protein C as well as the coding region for an active or inactive 3C^{pro} as a second cistron under IRES control (Fig. 4a). The protein E heparan sulphate mutation was present in all constructs.

RNAs of PV-2AΔ3-H-I-3C and PV-2AΔ3-H-I-3Ci were transfected into BHK-21 cells and viral protein expression was examined by IF analysis (Fig. 4b, upper panel). About 30–40 % of cells transfected with both mutants exhibited a positive IF staining, indicating that both mutants were competent for replication and protein expression. Next, supernatants of the transfected cells were applied to fresh cells and infection of these cells was again monitored by IF (Fig. 4b, lower panel). RNA from PV-2AΔ3-H-I-3C produced infectious virions and could be propagated further as indicated by the detection of positive cells after the first passage. No infectious virions were obtained with PV-2AΔ3-H-I-3Ci containing an inactive 3C^{pro}.

Thus, the presence of the 3C^{pro} cleavage site and the 3C^{pro} sequence on the same genome also allows virion production. However, the efficiency was lower than when 3C^{pro} was supplied *in trans*.

Passage of 3C protein C cleavage site mutants in the absence of an active 3C protease provokes resuscitating mutations

Now, we wished to examine how quickly mutants could arise in the absence of an active 3C^{pro} that would allow the formation of a new NS2B/3^{pro} protein C cleavage site and thus lead to the production of infectious virions. Accordingly, passaging experiments on BHK-21 cells were performed. First, *in vitro*-synthesized RNAs of PV-2AΔ3-H, QV-2AΔ3-H, PV-2AΔ3-H-I-3Ci as well as PV-2AΔ3-H together with RNA of the replicon ΔME-I-3Ci were electroporated into BHK-21 cells. Then, supernatants from the transfected cells
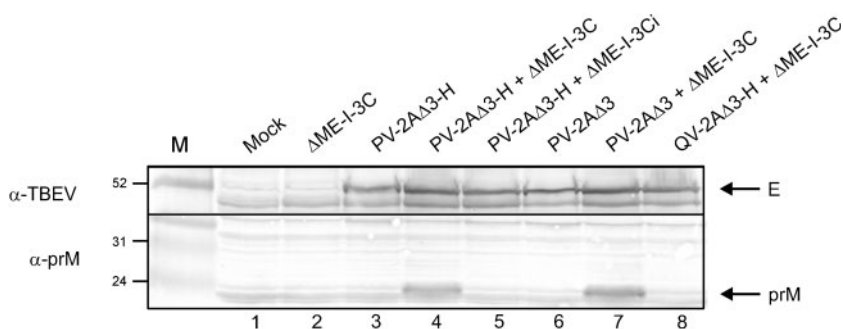

Fig. 3. Western blot analysis of structural protein processing. Cells transfected with mutant RNAs were lysed 22 h post-transfection, and structural proteins prM and E were detected using polyclonal sera. Positions of marker proteins in kDa are marked
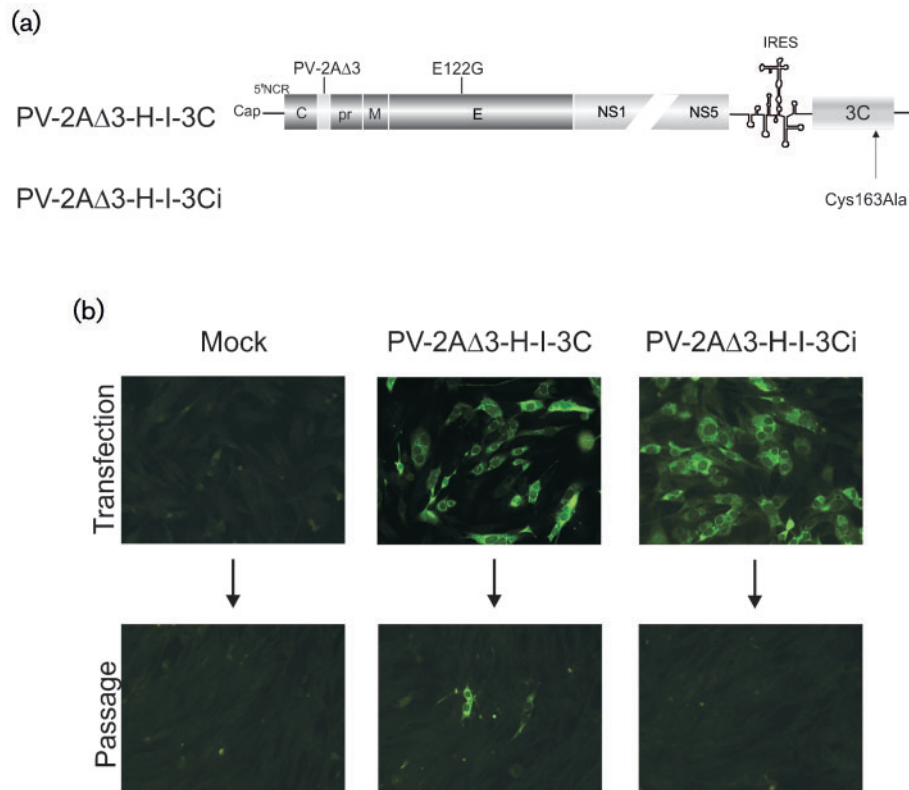
Fig. 4. Analysis of TBEV mutants expressing the 3C^{pro} cleavage site and the 3C^{pro} on the same construct. (a) Schematic of TBEV full-length mutants with the PV-2AΔ3 insertion within the C-terminal region of protein C, the heparan sulphate mutation in E and the 3C^{pro}. (b) IF analysis of viral protein expression and viral infectivity. BHK-21 cells were transfected with mutant RNAs or were mock transfected, and viral protein expression was detected by IF with the polyclonal serum 1 day post-transfection (top). Supernatants harvested 1 day post-electroporation were also transferred onto fresh BHK-21 cells (arrows), and viral protein expression monitored by IF 1 day post-infection (bottom).

were harvested 1 or 2 days post-transfection and transferred onto fresh BHK-21 cells; infection of cells was determined by IF. Cells inoculated with supernatants from day 1 were all negative for viral protein expression (Fig. 5a). Surprisingly, however, in one of three attempts, 100 % of the cells inoculated with 2 day supernatants derived from PV-2AΔ3-H transfected cells and about 50 % of cells inoculated with those from QV-2A3Δ-H and PV-2AΔ3-H-I-3Ci exhibited a bright IF staining (Fig. 5a), indicating the production of infectious virions. In contrast, no positive IF was observed with supernatants from cells transfected with the two RNAs PV-2AΔ3-H + ΔME-I-3Ci (data not shown).

To identify the responsible mutations, viral RNA was isolated from supernatants 3 days post-infection of the second passage and the entire structural protein-coding region (C–prM–E), as well as the domain comprising the heterologous inactive 3C protease was sequenced. No mutations were found within the protease domain; however, all three mutants had acquired a one nucleotide insertion and a one nucleotide deletion in the structural protein coding region that led to a short change of the reading frame. In PV-2AΔ3-H, 6 aa within the 2AΔ3

protein sequence were provided from another reading frame (Fig. 5b). In contrast, QV-2AΔ3-H and PV-2AΔ3-H-I-3Ci, 22 aa in the C-terminal region of protein C and the 2AΔ3 protein sequence were provided from the same reading frame as in PV-2AΔ3-H (Fig. 5b). These alterations not only changed 2AΔ3 hydrophobicity but also introduced a potential NS2B/3^{pro} cleavage motif of R-R*C at identical locations in all three revertants.

To ensure that these frame-shifts were responsible for the restoration of infectivity, we introduced the mutations observed in Fig. 5 into the full-length TBEV-2AΔ3 clone and transfected the RNAs PV-2AΔ3-mut1, PV-2AΔ3-mut2 and QV-2AΔ3-mut into BHK-21 cells. Infectious virions were indeed obtained after passage of the transfection supernatants (data not shown). These passaging experiments also revealed that the virions of PV-2AΔ3-mut2 and QV-2AΔ3-mut had a lower infectivity than PV-2AΔ3-mut1, possibly due to the changes in protein C.

Was the sequence R-R*C being recognized by the TBEV NS2B/3^{pro}? Accordingly, we transfected BHK-21 cells with WT-TBEV and mutant RNAs and lysed cells 48 h
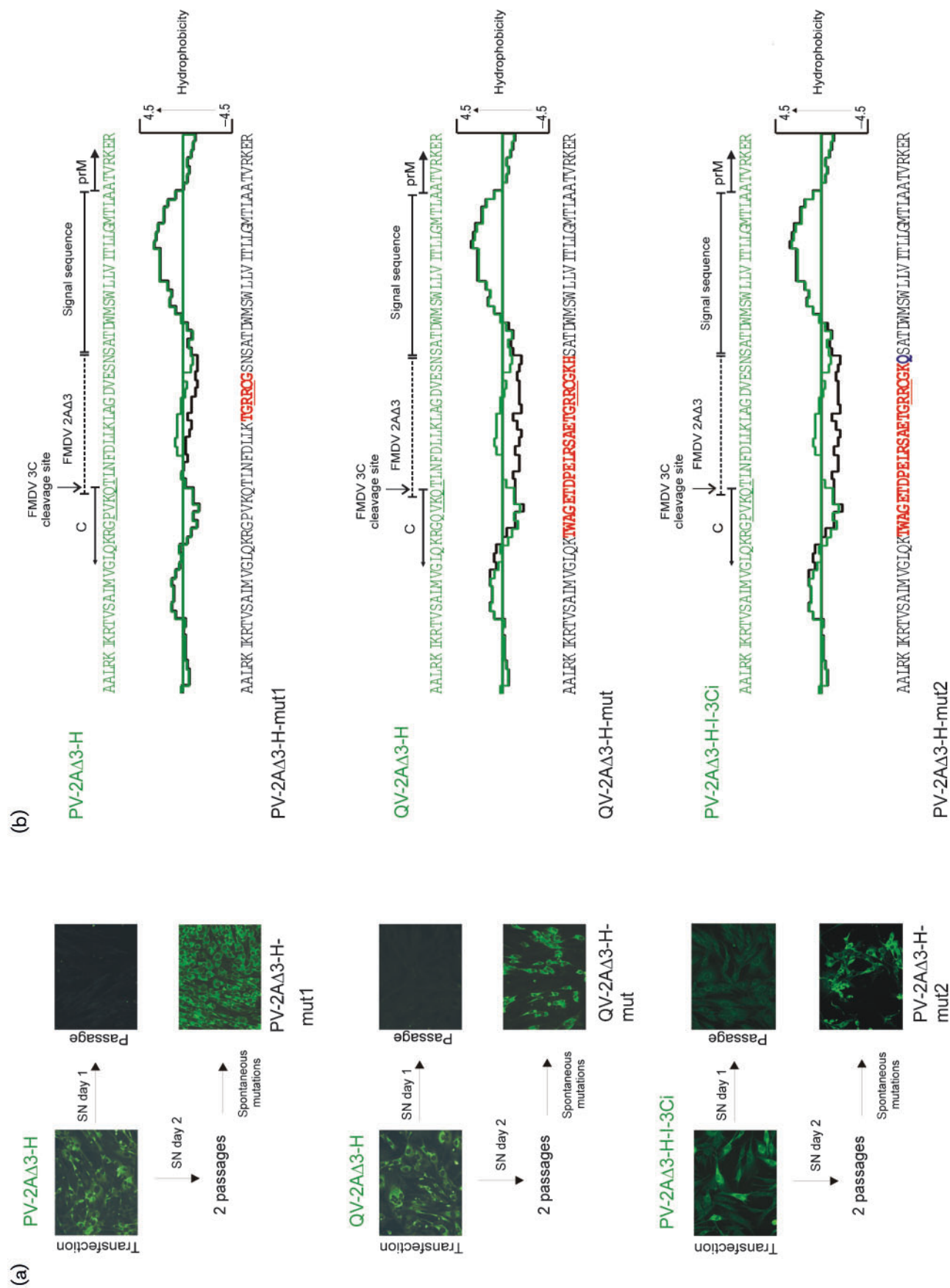
post-transfection. As expected, WT-TBEV lysates showed prominent bands for the structural proteins E, prM and C (Fig. 6, lane 2), whereas the mutants showed only prM and E but no C band. Instead a new band was visible, termed C* (Fig. 6, lane 3, 4 and 5), running slightly slower than the wild-type. The size would correspond to a protein C extended by 14 aa, further indicating that the R-R*C motif is processed by the NS2B/3^{pro}. Moreover, the ratio of intracellular protein C to the surface proteins prM and E differed in the mutants. Thus, PV-2AΔ3-mut1 has a C/E ratio similar to that of the WT-TBEV, whereas PV-2AΔ3-mut2 and QV-2AΔ3-mut (Fig. 6, lane 4 and 5) show a much higher ratio, suggesting that protein C accumulates in the cytoplasm.

QV-2AΔ3-mut1 and PV-2AΔ3-mut are impaired in viral spread

To determine differences in infectious virus progeny production, we examined the ability of the revertants to form foci. WT-TBEV was used as a positive control. As a further control, TBEV-2A was included into the experiments, as its protein C comprises four amino acids more than the second-site revertants and has already been characterized. TBEV-2A foci are 10 times smaller than those produced by the WT-TBEV, and can only be visualized 70 h post-infection, clearly indicating a defect in viral spread (Schrauf et al., 2009). Interestingly, despite the 14 aa extension of its protein C, foci formed by PV-2AΔ3-mut1 were only marginally smaller than those formed by the WT-TBEV and could be detected 50 h post-infection (Fig. 7a, bottom). In contrast, PV-2AΔ3-mut2 and QV-2A3Δ-mut formed foci that could be visualized only 70 h post-infection and were considerably smaller than those of the WT-TBEV. However, they were still about three to four times larger than those of TBEV-2A (Fig. 7a, top).

To quantify the observed differences in infectivity, we compared the growth properties of WT-TBEV, TBEV-2A and the three second site mutants. We performed multistep growth curves in BHK-21 cells at a low m.o.i. of 0.01 as shown in Fig. 7(b). As previously observed (Schrauf et al., 2009), TBEV-2A showed a delayed and reduced particle release. In contrast, PV-2AΔ3-mut1, which forms large foci, produced infectious progeny close to WT level, whereas QV-2AΔ3-mut1 and PV-2AΔ3-mut2 showed infectivity titres 1 and 1.5 log units below those of the WT but still between 2.5 and 1.5 log units higher than TBEV-2A (Fig. 7b).

PV-2AΔ3-mut1 exhibits a delay in the onset of early RNA replication

The observed defects in infectivity and viral spread of TBEV-2A were proposed to result from an impairment of unpackaging and early RNA synthesis (Schrauf et al., 2009). To determine the effect of protein C extensions of the revertants on viral replication, we examined early RNA

synthesis after infection. As expected, TBEV-2A exhibited a delayed and reduced RNA replication (Fig. 8). Interestingly, mutant PV-2AΔ3-mut1 also showed a delay in the onset of RNA replication during the first 6 h post-infection. Nevertheless, after the onset of RNA replication, no further impairment in RNA synthesis was detectable. Indeed, already 9 h post-infection, this mutant produced as many RNA molecules as the WT-TBEV, which is in good accordance with the results shown in Fig. 7. In contrast, PV-2AΔ3-mut2 and QV-2AΔ3-mut produced similar amounts of RNA copies (Fig. 8). Thus, the reduction of foci size and infectivity of PV-2AΔ3-mut2 and QV-2AΔ3-mut does not appear to be the result of an impaired RNA replication.

DISCUSSION

We show here that the production of infectious flaviviral virions can be dependent on protein C cleavage by a heterologous protease, the FMDV 3C^{pro}. Virion production was enabled by providing the enzyme either *in cis* or *in trans* from a TBEV replicon. Virion assembly was more efficient when the 3C protease was provided *in trans*. Most probably, in the *cis* construction, protein synthesis from the 5′ end of the TBEV genome was lowered by the presence of the IRES element in the 3′ non-coding region. This has also been observed in bicistronic TBEV mutants in which an IRES element controls the expression of the structural proteins prM and E (Orlinger *et al.*, 2006).

Cell-based antiviral assays in which the production of infectious virions depends on cleavage by HCV NS3 protease have been reported (Filocamo *et al.*, 1997; Hahm *et al.*, 1996; Lai *et al.*, 2000). Our system is, however, the first to enable the recovery of a non-infectious virus through a second replicon expressing the required protease. To assess the genetic stability of this system, we investigated how rapidly monocistronic full-length TBEV mutants PV-2AΔ3-H and QV-2AΔ3-H would acquire spontaneous mutations during passage on BHK-21 cells in the absence of the 3C^{pro}. Similarly, we also investigated the bicistronic PV-2AΔ3-H-I-3Ci construction that contains an inactive 3C^{pro} as well as the monocistronic full-length TBEV mutant PV-2AΔ3-H together with the TBEV replicon expressing an inactive 3C^{pro}. Multiple amino acid changes at the C–prM junction were observed in the PV-2AΔ3-H and QV-2AΔ3-H mutants as well as in PV-2AΔ3-H-I-3Ci but not in the PV-2AΔ3-H transfected with the TBEV replicon expressing the inactive 3C^{pro}. This indicates that the *trans*-complementation system is more stable than the *cis* system.

The specific changes in the revertants all resulted from the insertion of one nucleotide and the deletion of a second. Throughout, an NS2B/3^{pro} cleavage motif R-R*C was introduced at an identical location in the 2AΔ3 region so that cleavage at this sequence by NS2B/3^{pro} should give rise to a C-terminal extension of 14 aa in protein C in all three
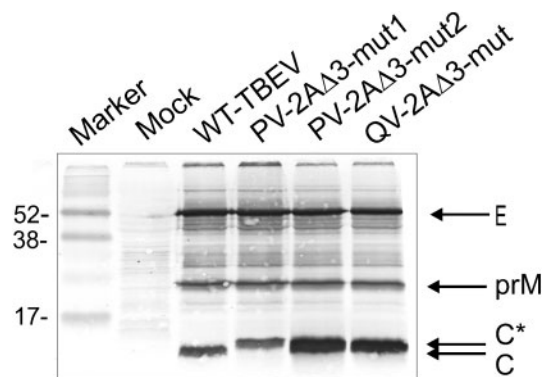
Fig. 6. Analysis of structural-protein processing. Cells transfected with WT or mutant RNAs were lysed 48 h post-transfection, and proteins were detected by immunoblotting using a rabbit polyclonal serum against the TBEV structural proteins. Positions of proteins E, prM, C and C* as well as marker proteins in kDa are marked.

revertants. Furthermore, in PV-2AΔ3-H, six amino acids within the 2AΔ3 protein sequence were mutated, whereas in QV-2AΔ3-H and PV-2AΔ3-H-3Ci, the insertion and deletion led to the mutation of the same 22 aa, located in both the C-terminal region of protein C and the 2AΔ3 protein sequence. The new amino acid sequences differed only in the most C-terminal amino acid of the mutated region (Fig. 5b and c).

Introduction of the mutations into TBEV-2AΔ3 confirmed that they restored the production of infectious virus progeny. In addition, protein C of all three mutants showed a slower mobility on SDS-PAGE than the WT (Fig. 6), strongly suggesting that cleavage was performed by NS2B/3^{pro} at the putative cleavage site.

Differences in the properties of the revertants were evident. PV-2AΔ3-H-mut1 reached wild-type levels both in infectivity and RNA synthesis and showed only a minor effect on viral spread. However, a delay in the onset of early RNA replication was detectable. The other mutants PV-2AΔ3-H-mut2 and QV-2AΔ3-H-mut that bear the almost identical 22 aa change also grew to high titres, just 1 and 1.5 log units lower than WT-TBEV, respectively. Both exhibited a smaller foci size, even though their RNA replication was similar to the WT-TBEV level.

The origin of these differences might be the more hydrophilic character of the protein C extension of all three revertants (compared with that of TBEV-2A), which would favour both the infectivity and early RNA synthesis. This agrees with our previous work on TBEV-2A, indicating that a hydrophobic extension to protein C can interfere with particle formation, viral spread and early RNA replication (Schrauf *et al.*, 2009). The different behaviour of the revertants in the onset of early RNA replication described here can be also attributed to the
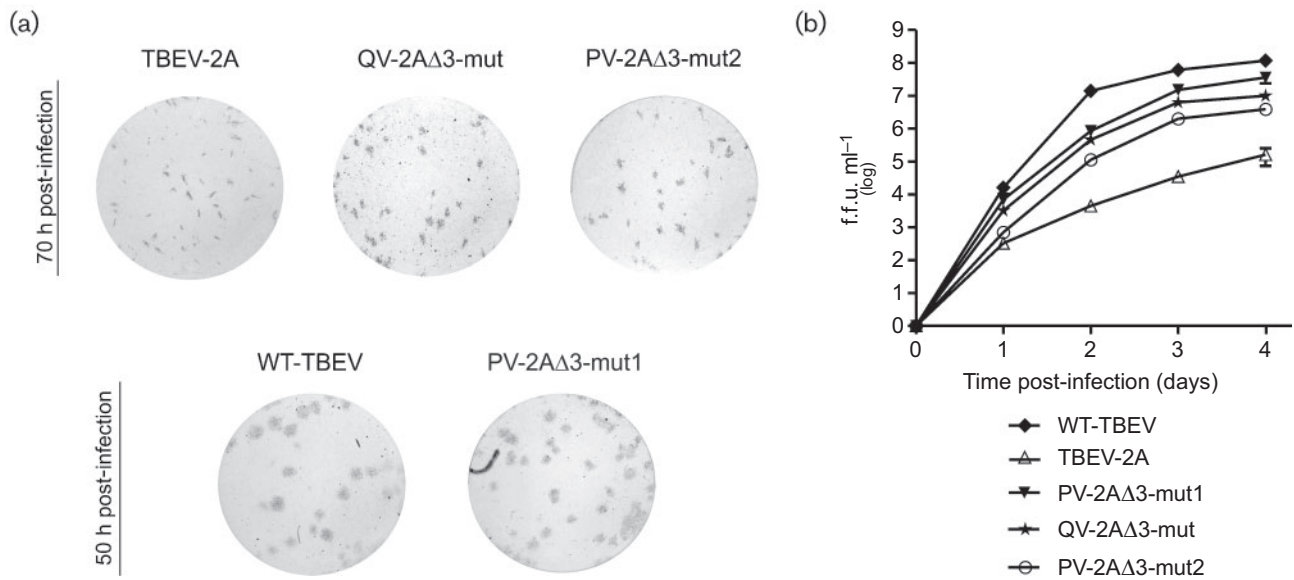
Fig. 7. Spread and growth properties of second site revertants in BHK-21 cells. Foci morphology (a) and infectivity titres (b) were determined using a focus forming assay. BHK-21 cells were infected (m.o.i. of 0.01) and samples were taken at different times. All data points represent mean values from two independent experiments. Error bars indicate SD.

composition of the protein C extension. A comparison of all three revertants and TBEV-2A reveals that the more hydrophobic the extension of protein C, the lower the rate of early RNA synthesis in general and the later the onset of early RNA replication and/or unpackaging in particular.
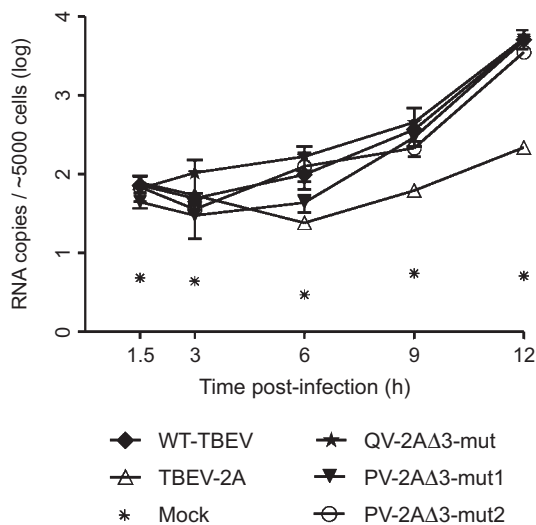
Fig. 8. RNA replication of second site revertants, TBEV-2A and WT-TBEV after infection of BHK-21 cells. Cells were infected (m.o.i. of 1), samples were taken at different time points, and intracellular RNA levels were measured. Data points represent geometric mean values from two independent experiments. Error bars indicate standard deviations.

The amino acid changes of mutants PV-2AΔ3-H-mut2 and QV-2AΔ3-H-mut resulted in a more hydrophilic 2AΔ3 region than in the case of PV-2AΔ3-H-mut1. Consequently, these two mutants have no impairment in the onset of early RNA synthesis, in marked contrast to the 6 h delay of PV-2AΔ3-H-mut1. TBEV-2A, however, which possesses an extension of mainly hydrophobic amino acids showed both a delay as well as a reduction in early RNA synthesis. Finally, the observed difference between the revertants and TBEV-2A also suggests that the re-establishment of a NS2B/3^{pro} cleavage in the C-terminal region of C is advantageous for viral spread.

Despite these beneficial contributions, the more hydrophilic character of the extension did not aid viral spread. PV-2AΔ3-H-mut2 and QV-2AΔ3-H-mut, which carry more hydrophilic amino acids in the 2AΔ3 region than PV-2AΔ3-H-mut1, formed considerably smaller foci that could be visualized only 3 days post-infection. In contrast, the foci of PV-2AΔ3-H-mut1 were only marginally smaller than those of the WT. The difference appears to lie in the C-terminal part of protein C, where only PV-2AΔ3-H-mut2 and QV-2AΔ3-H-mut have acquired additional mutations. These mutations led to the loss of two basic amino acids in the C-terminal part of protein C. Ma, Dokland and their co-workers (Dokland et al., 2004; Ma et al., 2004) have shown for dengue virus and WNV protein C that protein C occurs as a dimer with the two C-terminal helices forming a broad interface for interaction with RNA. Therefore, the loss of positive charges close to the helix may hinder correct dimer formation and/or RNA binding. Given that the RNA production upon infection was at

wild-type levels (Fig. 8), we can exclude a defect in unpackaging or synthesis of RNA. In contrast, we propose that the accumulation of protein C in the cytoplasm is crucial (Fig. 6), indicating that the reduced infectivity of PV-2AΔ3-H-mut2 or QV-2AΔ3-H-mut is caused by a defect in formation and/or budding of the NC, thus limiting viral spread. In contrast, the protein C of the PV-2AΔ3-H-mut1 was not affected by the amino acid changes and no accumulation of protein C in the cytoplasm occurred (Fig. 6), indicating that budding was undisturbed. Our results highlight therefore a previously undetected role of the C terminus of protein C for budding of the NC and the necessity of positive charges.

METHODS

Plasmids and cells. All TBEV constructs derive from the cDNA of TBEV strain Neudoerfl (GenBank accession no. U27495) (Mandl *et al.*, 1997). BHK-21 (ATCC CCL10) cells were cultured as described previously (Schrauf *et al.*, 2008). Oligonucleotides are listed in Supplementary Table S1 (available in JGV Online).

Full-length TBEV mutants. pTNd/5′ (Schrauf *et al.*, 2008) and all derivatives of pTNd/5′ contain the 5′ one-third of the cDNA of TBEV strain Neudoerfl genome. Infectious full-length RNA was generated after *in vitro* ligation of the 5′ cDNA clones with the 3′ cDNA clone pTNd/3′ (Schrauf *et al.*, 2008), which contains the 3′ two-thirds of the TBEV genome. Plasmid p2AΔ3/5′ was constructed by deleting the PGP motif at the C-terminal end of the FMDV-2A protein sequence of TBEV-2A/5′. For plasmid pPV-2AΔ3/5′, the P4 (proline) and the P3 positions (valine) of the 3C^{pro} cleavage site were introduced upstream of lysine 94 of the TBEV protein C using the forward primer P43-F, the reverse primer P43-R and p2A3Δ/5′ as template. Plasmid pPV-2A3Δ-H/5′, containing the P4 and P3 positions of the 3C^{pro} cleavage site as well as the heparan sulphate mutation E122G in protein E, was generated using the forward primer E122G-F, the reverse primer E122G-R and pPV-2A3Δ/5′ as template. Plasmid pQV-2A3Δ-H/5′, containing valine instead of proline at the P4 position of the 3C^{pro} cleavage site, was created using the forward primer PQ-F, the reverse primer PQ-R and pPV-2A3Δ-H/5′ as template.

Plasmid pTNd/c contains a full-length genomic cDNA of the WT TBEV strain Neudoerfl. TBEV-2A contains a functional FMDV-2A sequence that replaces the NS2B/3 cleavage site between protein C and prM. RNA of TBEV mutants with resuscitating mutations in the 2AΔ3 region was isolated, reverse transcribed and cloned into the pTNd/c cDNA using restriction sites for *Sal*I and *Sna*BI; the cDNAs have no heparan sulphate mutation.

TBEV replicons. Replicons ΔME-I-3C and ΔME-I-3Ci are derivatives of clone ΔME–eGFP (Gehrke *et al.*, 2005), expressing the eGFP under the control of the ECMV IRES. The replicon ΔME-I-3C was created by a three-step PCR strategy. PCR-I was performed using the forward primer *Avr*II-F, the reverse primer IRES-3C-R and plasmid ME–ΔeGFP as template. PCR-II amplified the 3C^{pro} with the forward primer IRES-3C-F, the reverse primer 3C-*Age*I-R and plasmid pET-3C (provided by Francois Maree, Onderstepoort, South Africa, containing the FMDV 3C sequence of strain O/SAR/19/00) as template. PCR-III was then performed with primers *Avr*II-F and 3C-*Age*I-R using PCR products I and II as template. Then, the *Avr*II–*Age*I fragment of ΔME–eGFP was replaced by the *Avr*II and *Age*I restricted PCR product III to generate ΔME-I-3C. Replicon ΔME-I-3Ci, containing a cysteine-to-alanine mutation at position 163 of the

3C protease, was constructed in a similar way as described above, with the primer combinations *Avr*II-F and C163A-R for PCR-I and C163A-F and *Age*I-R for PCR-II. As template, ΔME-I-3C was used.

Plasmid ΔME-I-3C was amplified in *Escherichia coli* Copy Cutter EPI400 (Epicentre) *E. coli* HB101 was used for all other plasmids.

RNA transcription and transfection. Transcription of RNA *in vitro* and transfection into BHK-21 cells by electroporation were as described previously (Mandl *et al.*, 1997; Schrauf *et al.*, 2008).

Detection of viral proteins. Viral protein expression was examined by IF staining and Western blotting. For IF, transfected BHK-21 cells were disseminated into 24-well tissue culture plates containing microscope coverslips and permeabilized by acetone-methanol fixation (1 : 1) 1 day post-transfection. Expression of the structural proteins C, prM and E and the non-structural protein NS1 was visualized by successive incubation with a rabbit polyclonal anti-TBEV serum predominantly recognizing protein E and protein NS1 and fluorescein-isothiocyanate-conjugated anti-rabbit antibody (Jackson Immune Research Laboratory). For passaging experiments, supernatants from transfected BHK-21 cells seeded in 25 cm^2 tissue culture flasks were transferred onto fresh BHK-21 cells 1 day post-transfection and viral protein expression was analysed as described above 1 day post-infection. For the selection of resuscitating mutations, supernatants from transfected cell cultures were applied to fresh cells 2 days post-electroporation and, after a second round of passage, viral protein expression determined.

Formation of the structural proteins prM and E was monitored by immunoblotting (Schrauf *et al.*, 2008).

Virus stock preparation. To generate virus stocks, BHK-21 cells were electroporated with RNAs of WT-TBEV, TBEV-2A, PV-2AΔ3-mut1, PV-2AΔ3-mut2 and QV-2AΔ3-mut, respectively. Supernatants were harvested 48 h post-electroporation. RNA was isolated from these stocks, followed by reverse transcription. Constructs were confirmed by sequencing. Titre and focus morphology were determined by focus assay.

Multistep growth curves. BHK-21 monolayers grown in six-well plates were infected with stock preparations of respective viruses at a m.o.i of 0.01. Aliquots of supernatants were collected at various time points and infectious particles were quantified by focus assay.

Focus assay. To quantify the production of infectious virus particles, an immunochemical focus assay was employed (Schrauf *et al.*, 2009).

RNA replication. Intracellular RNA replication after infection was determined as described previously (Schrauf *et al.*, 2009).

Protein sequence analysis. Hydrophobicity plots were done using the DNASTAR programme PROTEAN (Kyle-doolitle scale, window size 11).

ACKNOWLEDGEMENTS

REFERENCES

Amberg, S. M. & Rice, C. M. (1999). Mutagenesis of the NS2B-NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. *J Virol* **73**, 8083–8094.

Atkins, J. F., Wills, N. M., Loughran, G., Wu, C. Y., Parsawar, K., Ryan, M. D., Wang, C. H. & Nelson, C. C. (2007). A case for "StopGo": reprogramming translation to augment codon meaning of GGN by promoting unconventional termination (Stop) after addition of glycine and then allowing continued translation (Go). *RNA* **13**, 803–810.

Birtley, J. R., Knox, S. R., Jaulent, A. M., Brick, P., Leatherbarrow, R. J. & Curry, S. (2005). Crystal structure of foot-and-mouth disease virus 3C protease. New insights into catalytic mechanism and cleavage specificity. *J Biol Chem* **280**, 11520–11527.

Chappell, K. J., Nall, T. A., Stoermer, M. J., Fang, N. X., Tyndall, J. D., Fairlie, D. P. & Young, P. R. (2005). Site-directed mutagenesis and kinetic studies of the West Nile Virus NS3 protease identify key enzyme-substrate interactions. *J Biol Chem* **280**, 2896–2903.

Curry, S., Roqué-Rosell, N., Zunszain, P. A. & Leatherbarrow, R. J. (2007). Foot-and-mouth disease virus 3C protease: recent structural and functional insights into an antiviral target. *Int J Biochem Cell Biol* **39**, 1–6.

Dokland, T., Walsh, M., Mackenzie, J. M., Khromykh, A. A., Ee, K. H. & Wang, S. (2004). West Nile virus core protein; tetramer structure and ribbon formation. *Structure* **12**, 1157–1163.

Donnelly, M. L., Luke, G., Mehrotra, A., Li, X., Hughes, L. E., Gani, D. & Ryan, M. D. (2001). Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. *J Gen Virol* **82**, 1013–1025.

Filocamo, G., Pacini, L. & Migliaccio, G. (1997). Chimeric Sindbis viruses dependent on the NS3 protease of hepatitis C virus. *J Virol* **71**, 1417–1427.

Gehrke, R., Heinz, F. X., Davis, N. L. & Mandl, C. W. (2005). Heterologous gene expression by infectious and replicon vectors derived from tick-borne encephalitis virus and direct comparison of this flavivirus system with an alphavirus replicon. *J Gen Virol* **86**, 1045–1053.

Hahm, B., Back, S. H., Lee, T. G., Wimmer, E. & Jang, S. K. (1996). Generation of a novel poliovirus with a requirement of hepatitis C virus protease NS3 activity. *Virology* **226**, 318–326.

Kofler, R. M., Heinz, F. X. & Mandl, C. W. (2002). Capsid protein C of tick-borne encephalitis virus tolerates large internal deletions and is a favorable target for attenuation of virulence. *J Virol* **76**, 3534–3543.

Kroschewski, H., Allison, S. L., Heinz, F. X. & Mandl, C. W. (2003). Role of heparan sulfate for attachment and entry of tick-borne encephalitis virus. *Virology* **308**, 92–100.

Lai, V. C., Zhong, W., Skelton, A., Ingravallo, P., Vassilev, V., Donis, R. O., Hong, Z. & Lau, J. Y. (2000). Generation and characterization of a hepatitis C virus NS3 protease-dependent bovine viral diarrhea virus. *J Virol* **74**, 6339–6347.

Lee, E., Stocks, C. E., Amberg, S. M., Rice, C. M. & Lobigs, M. (2000). Mutagenesis of the signal sequence of yellow fever virus prM protein: enhancement of signalase cleavage In vitro is lethal for virus production. *J Virol* **74**, 24–32.

Lindenbach, B. D., Thiel, H. J. & Rice, C. M. (2007). Flaviviridae: the viruses and their replication. In *Fields Virology*, pp. 1101–1152. Edited by D. M. Knipe & P. M. Howley. Philadelphia: Lippincott Williams & Wilkins.

Lobigs, M. (1993). Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3. *Proc Natl Acad Sci U S A* **90**, 6218–6222.

Lobigs, M. & Lee, E. (2004). Inefficient signalase cleavage promotes efficient nucleocapsid incorporation into budding flavivirus membranes. *J Virol* **78**, 178–186.

Ma, L., Jones, C. T., Groesch, T. D., Kuhn, R. J. & Post, C. B. (2004). Solution structure of dengue virus capsid protein reveals another fold. *Proc Natl Acad Sci U S A* **101**, 3414–3419.

Mandl, C. W. (2005). Steps of the tick-borne encephalitis virus replication cycle that affect neuropathogenesis. *Virus Res* **111**, 161–174.

Mandl, C. W., Ecker, M., Holzmann, H., Kunz, C. & Heinz, F. X. (1997). Infectious cDNA clones of tick-borne encephalitis virus European subtype prototypic strain Neudoerfl and high virulence strain Hypr. *J Gen Virol* **78**, 1049–1057.

Orlinger, K. K., Hoenninger, V. M., Kofler, R. M. & Mandl, C. W. (2006). Construction and mutagenesis of an artificial bicistronic tick-borne encephalitis virus genome reveals an essential function of the second transmembrane region of protein e in flavivirus assembly. *J Virol* **80**, 12197–12208.

Schrauf, S., Schlick, P., Skern, T. & Mandl, C. W. (2008). Functional analysis of potential carboxy-terminal cleavage sites of tick-borne encephalitis virus capsid protein. *J Virol* **82**, 2218–2229.

Schrauf, S., Mandl, C. W., Bell-Sakyi, L. & Skern, T. (2009). Extension of flavivirus protein C differentially affects early RNA synthesis and growth in mammalian and arthropod host cells. *J Virol* **83**, 11201–11210.

Stocks, C. E. & Lobigs, M. (1998). Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. *J Virol* **72**, 2141–2149.

4. MANUSCRIPT 2

NS2B/3 proteolysis at the C-prM junction of the tick-borne encephalitis virus polyprotein is highly membrane dependent

Martina Kurz[§], Nikolas Stefan [§†], Junping Zhu [§*] and Tim Skern[§¶]

Initial cloning, expression and purification of the NS2B/3 protease as well as the initial setup of the cleavage assay have been done by N. Stefan and J. Zhu. All further cloning, expression, purification and optimisation steps that were necessary to make the protease cleavage specific and efficient, as well as all experiments including the HIV-1 protease were performed by M. Kurz. The manuscript was written by M. Kurz and T. Skern.

[§] Max F. Perutz Laboratories, Medical University of Vienna, A-1030 Vienna, Austria

[†] Current address: Dept. of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland

[*] Current address: Dept. of Pathogenic Biology, You An Men Beijing, 100069, P. R. China

Corresponding author: Tim Skern

Abstract

The replication of tick-borne encephalitis virus (TBEV), like that of all flaviviruses, is absolutely dependent on proteolytic processing. Production of the mature proteins C and prM from their common precursor requires the activity of the viral NS2B/3 protease (NS2B/3^{pro}) at the C-terminus of protein C and the host signal peptidase I (SPaseI) at the N-terminus of protein prM. Recently, we have shown in cell culture that the cleavage of protein C and the subsequent production of TBEV particles can be made dependent on the activity of the foot-and-mouth disease virus 3C protease, but not on the activity of the HIV-1 protease (HIV-1^{pro}) (S. Schrauf et al., 2012 and personal communication). To investigate this failure, we developed an *in vitro* cleavage assay to assess the two cleavage reactions performed on the C-prM precursor. Accordingly, a recombinant modular NS2B/3^{pro}, consisting of the protease domain of NS3 linked to the core-domain of cofactor NS2B, was expressed in *E. coli* and purified to homogeneity. This enzyme could cleave a C-prM protein synthesised in rabbit reticulocyte lysates. However, cleavage was only specific when protein synthesis was performed in the presence of canine pancreatic microsomal membranes and required the prevention of signal peptidase I (SPaseI) activity by lengthening the h-region of the signal peptide. Substitution of the NS2B/3^{pro} cleavage motif in C-prM by a HIV-1^{pro} motif inhibited NS2B/3^{pro} processing in the presence of microsomal membranes but allowed cleavage by HIV-1^{pro} at the C-prM junction. However, a second cleavage site for HIV-1^{pro} was observed in the prM part in both the presence and absence of microsomal membranes. This indicates that the eight amino acid HIV-1^{pro} cleavage site disturbed the correct membrane insertion of the C-prM precursor and provides an explanation why C-prM cleavage and TBEV particle production could not be made dependent on HIV-1^{pro} activity in cell culture.

Introduction

Tick-borne encephalitis virus (TBEV), a member of the family of *Flaviviridae* in the genus *Flavivirus* (Lindenbach et al., 2007) is a small (~50 nm) enveloped virus with a single stranded, positive sense RNA genome. In addition to the lipid envelope and the RNA, three structural proteins (capsid (C), membrane (M, derived from a precursor prM) and envelope (E)) are present in the virion. The genetic information in the viral RNA is expressed as a single polyprotein that meanders in and out of the endoplasmic reticulum (ER) membrane. The structural proteins are located in the amino-terminal part of the polyprotein; the rest of the polyprotein comprises the non-structural proteins NS1, NS2A, NS2B, NS3, NS4A, NS4B, and NS5 (Lindenbach et al., 2007). NS2B and NS3 (designated here NS2B/3^{pro}) comprise the viral protease that is required for polyprotein processing, along with furin and signal peptidase I (SPaseI), two host cell proteases. These cleavages occur co- and post-translationally. In

addition, during translation and processing, the surface proteins prM and E are glycosylated in the lumen of the endoplasmic reticulum (ER). The production of these structural proteins is vital as protein C is the initiator molecule for viral assembly; multiple copies of this protein encapsidate a newly synthesised RNA molecule to produce the nucleocapsid (NC). Furthermore, prM is required to ensure that protein E is manufactured in a viable conformation. Correctly processed prM and E are added to the growing particle by budding of the NC through the ER. The resulting immature particles cross the trans-Golgi network (TGN), permitting furin cleavage of prM; this reaction transforms the immature particles into infectious virions (Stadler et al., 1997).

Processing is thus a prerequisite for the assembly of the TBEV particle; not surprisingly, the two processes are closely co-ordinated (Lobigs, 1993). For instance, the C-terminus of the protein C contains an internal hydrophobic signal sequence responsible for translocating the prM protein into the lumen of the ER. On the cytosolic side of the ER membrane, protein C is cleaved off the signal sequence by NS2B/3[pro] (Amberg et al., 1994; Yamshchikov and Compans, 1994). This cleavage is a prerequisite to allow the host cell SPaseI cleavage at the signal sequence on the luminal side of the ER membrane to generate the N-terminus of protein prM. Initiation of virion assembly has been proposed to depend upon the timing of these cleavage events at the termini of the signal sequence separating proteins C and prM (Amberg and Rice, 1999; Lee et al., 2000; Lobigs and Lee, 2004; Stocks and Lobigs, 1998). Protein prM is in its turn a prerequisite to ensure correct synthesis and transport of protein E (Konishi and Mason, 1993; Lorenz et al., 2002).

In TBEV, investigation of the cleavage by NS2B/3[pro] in the C-terminal region of protein C is complicated by the presence of two potential cleavage sites (KR*G and RR*G, underlined in Fig. 2 (Mandl et al., 1991)). Schrauf et al. (Schrauf et al., 2009) showed that, in vivo, the downstream cleavage site is used and that the upstream cleavage motif cannot be accepted when the downstream cleavage motif had been removed by mutation.

Schrauf et al. (Schrauf et al., 2009) showed further that, in the absence of the downstream NS2B/3[pro] cleavage site, the C-terminus of protein C could be released from the polypeptide chain by the introduction of the foot-and-mouth disease virus (FMDV) 2A sequence. The presence of the sequence Asn-Pro-Gly-Pro at the C-terminus of this 20 amino acid protein causes the ribosome to pause, release the first protein and then continue translation of the mRNA. In contrast, upon removal of the residues Pro-Gly-Pro from the 20 amino acid sequence, the separation of C and prM was inhibited, leading to a replication deficient virus. Virion production could be rescued by the introduction of an FMDV 3C[pro] cleavage site at the N-terminus of the 2A sequence and the expression of the 3C[pro] either *in cis* from a second ORF or *in trans* from a TBEV replicon (Schrauf et al., 2012). However, we were unable to achieve success in such a system with the HIV-1[pro] (Schrauf et al., unpublished). We

reasoned that this failure might be due to the presence of additional cleavage site(s) for the HIV-1[pro] in the C-prM region. To investigate this idea, we set out to establish an *in vitro* assay system for proteolytic cleavage in the C-terminal region of protein C.

Material and methods

Oligonucleotides and plasmids

Oligonucleotide sequences can be found in supplementary table 1. Constructs encoding TBEV European subtype strain Neudörfl (EMBL: U27495) were derived from pBR-TBEV-ΔME-EGFP (kind gift of Dr. P. Schlick). Plasmids coding for the co-factor and the protease were constructed in two steps. First, the hydrophilic cofactor region of NS2B (Fig. 1A, corresponding to amino acids 1404 to 1453 of the TBEV polyprotein (UniProt: P14336) was amplified by primers NS2B-f and NS2B-r-3′linker. These primers, introducing an *Nco*I restriction site at the 5′ end and the first half of the synthetic linker containing an *Xma*I recognition sequence at the 3′ end, were used to clone the fragment into pCite-A1, designated as pCite-A1-NS2B. Then, a fragment encoding the 190 N-terminal residues of NS3 (Fig. 1B), the protease domain (UniProt: P14336, residues 1490 to 1679) was amplified with primers 5′-Linker-NS3-F and NS3-R-6xHis. Thereby, the second half of the synthetic linker containing an *Xma*I recognition site was introduced at the 5′end of the protease region. Furthermore, six histidine codons and a *Bam*HI site were added to the 3′end. These enzymes were used to ligate the protease domain to pCite-A1-NS2B, finally named pCite-A1-NS2B/3, coding for the tethered NS2B/3 protease (Fig. 2A).

NS2B/3 S135A with alanine substitutions at S135 in the catalytic triad of NS3 were made by site-directed mutagenesis with primers NS3 S135A-f and NS3 S135A-r.

To generate the stabilised proteases with the new linker GGGGSGGGG, a cleavage site for *Bgl*II was introduced downstream of the linker with primers NS3-BglI-f and NS3-BglI-r. Cassettes coding for the desired mutations K90A and R93A of NS2B as well as the new linker were designed and introduced by use of the BglII and XhoI. The mutation G to A in the linker region of construction NS2B/3 R93A occurred serendipitously during mutagenesis. As the ensuing protease proved to be stable and active, it was used in all subsequent experiments.

For expression in *E. coli*, sequences coding for the desired tethered proteases were cloned from pCite-A1-NS2B/3 into the expression vector pET11d by use of *Nco*I and *Bam*HI.

To generate the plasmids coding for the substrates (Fig 2B and 2C), plasmid pTNd/5 which contains the 5′part of the genome of TBEV European subtype strain Neudörfl, was used as template (kind gift of Dr. R. Kofler). Using primers C-prM-f and C-prM-r, we produced PCR fragments encoding the capsid protein, transmembrane region and pre-membrane protein prM.

Thus, C-prM corresponds to residue 1 to 280 of the TBEV polyprotein. The plasmid was named pCITE-C-prM.

To generate pCITE-C-2AΔ3-prM and pCite-C-HIV-prM, we used plasmid pBR-C-2AΔ3-prM and pBR-C-HIV-prM, both kind gifts of Dr. S. Schrauf. By use of *Age*I and *Mlu*I, we isolated the desired fragment and cloned it into pCite-C-prM using the same cleavage sites.

In order to inhibit signal peptidase cleavage, nine leucine residues were introduced into the transmembrane regions of substrates by use of primers TM-9L-f and TM-9L-r.

In vitro transcription and translation

All pCITE plasmids were linearised with *Bam*HI. *In vitro* transcription with T7 RNA polymerase and *in vitro* translation were as described in Schlick et al. (Schlick and Skern, 2002). *In vitro* translation reactions (typically 20 µl) contained 70% RRL (Promega), 20 µCi of ^{35}S-methionine (1000 Ci/mmol, American Research Company) and amino acids (except methionine) at 20 µM. After pre-incubation for 1 min at 30°C, translation was started by addition of RNA. The reaction was stopped at designated time points by immediate transfer to ice and addition of unlabelled methionine and cysteine to a final concentration of 2mM. Where indicated, translation reactions were supplemented with 4 to 8% pancreatic microsomal membranes (Promega).

Protease assay

For self-cleavage reactions, translation products with or without microsomal membranes were incubated in assay buffer (20mM Tris/HCl, pH 7, 50 mM NaCl) at 30°C and quenched by addition of an SDS-PAGE loading buffer to a final concentration of 2% SDS.

Trans-cleavage reactions were performed at 30°C in a final volume of 10µl containing substrate and purified NS2B/3^{pro} at the indicated final concentrations and assay buffer. After the indicated times, the reaction was stopped by addition of SDS sample buffer.

Electrophoresis and immunoblotting

The polyacrylamide gel electrophoresis system of Dasso und Jackson (Dasso and Jackson, 1989), containing 15% polyacrylamide, was used to separate proteins. ^{35}S-containing proteins were detected by fluorography.

Expression and purification

The plasmids pET 11d-NS2B/3, pET 11d-NS2B/3 S135A, pEt11d-NS2B/3 R50A G59A were used for high-level, inducible expression of hexa-histidine-tagged recombinant proteins.

Cultures of *E. coli* strain BL21(DE3)pLysS transformed with the expression plasmids were grown in 10 ml of LB medium containing 34 µg/ml chloramphenicol and ampicillin at 37°C overnight. This overnight culture was diluted 1:100 in fresh LB medium containing chloramphenicol and ampicillin at the concentrations stated above and incubated at 37°C, 225 rpm until an OD_{600} of 0.5 to 0.6 had been reached. The expression of the recombinant proteins was induced by addition of isopropyl-β-D-thiogalactopyranose (IPTG) to a final concentration of 0.1 mM. Cultures were incubated for additional 4 hours, and cells were harvested by centrifugation. Cell pellets were resuspended in 30 ml resuspension buffer containing 50mM Tris-HCl, 50mM NaCl, pH 8 and lysed with a Bandelin Sonoplus sonicator (4 cycles, 40%, 30 seconds and once continuous for 30 seconds) and centrifuged at 15,000 rpm for 30 min at 4°C. The supernatant was loaded onto a 5 ml HiTrap chelating column (GE Healthcare) loaded with Nickel and pre-equilibrated with lysis buffer. The protein was eluted with a linear gradient of 0 to 500 mM imidazole in lysis buffer. Fractions containing NS2B-NS3 proteins, determined by 12.5% SDS-PAGE, were pooled, concentrated using Amicon Ultra-4 Centrifugal filter Device 10kDa cut-off (Millipore). Subsequently, it was loaded on to a HiLoad TM 26/60 Superdex 75 column (GE Healthcare) equilibrated with resuspension buffer and run according to the manufacturer´s instructions. Fractions were again analysed by SDS-PAGE. Aliquots were stored until use at -80° C.

Expression and purification of the HIV-1[pro]

Penta-stabilised HIV-1[pro] Q7K, L33I, L63I C67 C95A was a kind gift of Dr. J. Tözsér (Louis et al., 1999). Briefly, the protease was expressed in *E. coli* and refolded by dialysis against 25mM formic acid for 2 hours and then against 100mM Na-acetate-trihydrate, 1mM DTT, 1mM EDTA, 10% glycerol and 0,03% Triton X-100, pH 5 five times for 2 hours. This buffer was also used in HIV-1[pro] trans-cleavage experiments.

Results

Generation of an active TBEV NS2B/3[pro]

The proteolytic domain of the multifunctional flaviviral NS3 protein serine protease resides in the amino-terminal domain. The protein folds into two ß-barrels with the hydrophilic core being stabilised by the co-factor NS2B (Erbel et al., 2006). Without the NS2B domain, the NS3 protease is proteolytically inert. It has been shown by Leung et al. (Leung et al., 2001) that the cofactor activity of the NS2B core region is comparable to that of the entire NS2B sequence. This knowledge allowed single chain NS2B/3[pro] to be generated by genetically fusing the NS2B core region to the NS3 protease domain by a short linker region. Indeed,

NS2B/3^{pro} of several flaviviruses have been expressed in this way (Bessaud et al., 2005; Leung et al., 2001; Nall et al., 2004; Shiryaev et al., 2006; Yusof et al., 2000). To generate a similar construction for TBEV, the pertinent regions of NS2B and NS3 in the TBEV polyprotein were analysed by sequence alignment with WNV (Fig. 1) and related flaviviruses (data not shown). Residues 46 to 94 of TBEV NS2B (in green in Fig. 1A) and residues 1 to 190 (in red in Fig. 1B) of NS3 were identified as the central hydrophilic co-factor region of NS2B and the protease domain of NS3, respectively. These regions were amplified by PCR as indicated in Methods and linked by a DNA sequence encoding the flexible linker indicated in Fig. 2A. To prevent the loss of the NS2B activating peptide through autocatalytic cleavage, we attempted to ensure that no cleavage site for the NS3 protease was present in the flexible linker. A C-terminal hexa-histidine tag was introduced at the C-terminus for purification. A similar construct containing the inactivating mutation S135A was also prepared. Both wild-type and the inactivated mutant could be produced either by *in vitro* transcription and translation assay in rabbit reticulocyte lysates (RRL) or by expression in *E. coli*.

To test whether the NS2B/3 hybrid protein was indeed stable, the NS2B/3^{pro} and NS2B/3^{pro} S135A RNAs were translated in RRL in the presence of ^{35}S methionine for 30 min at 30°C followed by addition of unlabelled methionine and incubation at 30°C for the indicated times. Synthesised proteins were separated by SDS-PAGE and detected by fluorography. Both NS2B/3^{pro} and NS2B/3^{pro} S135A RNAs were translated into a protein of 33 kDa, slightly higher than the calculated Mr of 28 kDa. During the incubation of the active protease, the 33 kDa band slowly disappeared overnight and was replaced by bands of apparent molecular weight 28 kDa and 12 kDa. However, as the smaller band was actually running with the buffer front of the gel, its molecular weight cannot be given with any confidence. In contrast, the 33 kDa band of the inactive NS2B/3^{pro} S135A remained stable (Fig. 3), suggesting that autolysis of the active protease was responsible for the conversion of the 33 kDa species into the two smaller ones.

Similar results were obtained using the bacterially expressed NS2B/3^{pro} (data not shown). Mass spectroscopy of the purified 28 kDa band from SDS-PAGE revealed that it lacked the N-terminal 47 amino acids of the NS2B sequence. Thus, despite the lack of a corresponding cleavage site, the active NS2B/3^{pro} was still able to perform autolysis between residues K90 and E91 (Fig. 2 and 3), giving rise to fragments of a calculated Mr of 22 and 6 kDa, respectively. To prevent such autolysis, we constructed two modified NS2B/3 proteins in which either one (R93A) or two (K90A R93A) basic residues at the C-terminus of 2B were replaced with alanine (Fig. 2A). In addition, the linker sequences were also modified to more closely resemble that used in other flaviviral NS2B/3 constructs (Fig. 2A). RNA from the resulting constructs NS2B/3^{pro} R93A and NS2B/3^{pro} K90A R93A was translated in RRLs; both synthesised proteins were stable over the time-course of the experiment (Fig. 3). NS2B/3^{pro} R93A was selected for use in all subsequent experiments.

In Vitro Processing of the Structural Protein Precursor C-prM

The precursor (C-prM) of C and prM is connected by a short hydrophobic domain that spans the membrane of the ER. C is cleaved into its mature form by the NS2B/3[pro] at the cytosolic side whereas the SPaseI cleaves prM to its mature form in the ER (Fig. 2B and 2C). SPaseI only gains access to its cleavage site once NS2B/3[pro] has cleaved at the C-terminus of protein C. To investigate this cleavage reaction *in vitro*, we incubated purified recombinant NS2B/3[pro] R93A with C-prM synthesised in RRL. Control experiments were performed with buffer or with recombinant inactive NS2B/3pro S135A (final concentration 1 µg/µl). Proteins were separated by SDS-PAGE and detected by fluorography.

A single specific band corresponding to unprocessed C-prM running at 35 kDa was detected when the substrate was incubated with buffer or inactive NS2B/3[pro] (Fig. 4A, lanes 2 to 4 and data not shown). When incubated with active NS2B/3[pro] at a final concentration of 1µg/µl (Fig. 4A, lane 5 to 8), the substrate was converted into two bands running at 27 kDa, corresponding to prM* (prM plus amino-terminal transmembrane region), and about 10 kDa, corresponding to protein C. We tested several assay conditions to optimise the cleavage efficiency but were not able to find conditions in which amounts of protein less than 1 µg/µl led to cleavage. Thus, the cleavage efficiency of the purified protease remained rather low in comparison with recombinant proteases from related viruses (Bera et al., 2007; Chappell et al., 2008; Shiryaev et al., 2007).

NS2B/3[pro] cleavage of prM *in vitro* does not depend on the presence of its cleavage site

The canonical NS2B/3[pro] recognition site in the C-terminal region of protein C of TBEV was shown previously to be at R-R*S ((Schrauf et al., 2008); Fig. 2A). To test whether the cleavage of C-prM by NS2B/3[pro] was dependent on the presence of its cleavage site, we generated a C-prM mutant (C-2AΔ3-prM) in which the downstream NS2B/3[pro] cleavage site was replaced by a version of the 2A sequence of foot and mouth disease virus (FMDV) lacking the characteristic Pro-Gly-Pro motif at its C-terminus (see Fig. 2). This form is unable to allow the ribosomes to skip the formation of a peptide bond so that protein C cannot be released from the polyprotein (Donnelly et al., 2001). A TBEV mutant bearing this truncated 2A (C-2AΔ3-prM) was replication deficient, as protein C could not be processed by the NS2B/3[pro] (Schrauf et al., 2009). This provides further evidence that the upstream cleavage site is not used. Nevertheless, in this *in vitro* system described here, we observed, in the absence of the downstream cleavage site at the C-terminus of protein C, cleavage of the C-2AΔ3-prM with purified NS2B/3[pro] (Fig. 4B, lane 3). The most likely interpretation is that the upstream cleavage was substituting for the downstream site.

It should be noted that, despite repeated attempts, we were not able to detect the cleavage product for protein C. We interpret this failure to mean that the modified protein C containing all or part of the 2A protein of FMDV is unstable in RRLs. We have observed this phenomena previously when expressing other viral proteins in this system (Sousa et al., 2006).

Signal peptidase cleavage of C-prM in the presence of microsomal membranes occurs in the absence of NS2B/3^{pro} cleavage

We postulated that the specificity of the NS2B/3^{pro} reaction could be increased by the addition of canine pancreatic microsomal membranes (subsequently referred to as membranes) to allow the signal peptide that separates the two proteins to insert into a membrane. Upon translation of C-prM in the presence of the above-mentioned membranes for 30 min, a decrease in its electrophoretic mobility could clearly be seen after a further 60 min of incubation at 30°C (Fig 5A, compare lane 2 with lane 3), suggesting that N-glycosylation had occurred at residue Asp 144 of prM (Fig. 5A, lane 3). After 120 min, however, the amount of the C-prM species had decreased; instead, a new band had appeared at 28 kDa, corresponding to glycosylated prM lacking the trans-membrane region (Fig. 5A, lane 4).

The most likely explanation for the above result is that the membrane bound signal peptidase (SPaseI) has processed its luminal cleavage site in the absence of the viral protease. According to previous studies, however, the SPaseI cleavage should not occur until the NS2B/3^{pro} has performed processing at the C-terminus of protein C (Lobigs, 1993; Lobigs and Lee, 2004; Stocks and Lobigs, 1995). As we wished primarily to investigate NS2B/3^{pro} cleavage, we looked for an approach to inhibit the SPaseI reaction whilst permitting the cleavage by NS2B/3^{pro}. To this end, we first used a known inhibitor of SPaseI, N-methoxysuccinyl-Ala-Ala-Pro-Val-chloromethyl ketone (Lundin et al., 2008; Nilsson et al., 2002). Indeed, this compound, at concentrations of 750 µM or higher, did inhibit the proposed SPaseI cleavage reaction (data not shown). However, this concentration also inhibited the activity of NS2B/3^{pro} (data not shown), leading to the abandonment of this approach.

We therefore took an alternative approach to eliminate SPaseI cleavage. Signal peptides have been shown to possess a common structure, namely a short, positively charged amino-terminal n-region, a central hydrophobic h-region (blue in Fig. 2B and 2C) and a polar, mostly uncharged c-region containing the signal peptidase cleavage site. In addition to certain other differences, one appreciable difference between cleaved signal peptides and non-cleaved signal anchors is the length of the h-region. Consequently, it is possible to convert a signal peptide into a signal anchor by lengthening its h-region (Nilsson 1994). To this end, we introduced a stretch of nine leucine residues into the h-region of the signal peptide between C and prM (Fig. 2C). RNA from this construct, termed C-9L-prM, was then incubated in RRLs in

the presence of microsomal membranes and the stability of the translated protein examined. Figure 5B shows that the protein remained intact in both the absence (lanes 2 and 3) and presence of membranes (lanes 7 and 8), indicating that extension of the h-region conveyed resistance to SPaseI cleavage.

It is worth noting that, in this experiment, the glycosylation was complete during the translation phase of the reaction (compare Fig 5A lane 2 with Fig 5B lane 7) and that there was hardly any detectable difference between the protein in the glycosylated and non-glycosylated form (compare left and right panels of Fig. 5B). We postulate that the presence of the nine leucine residues is the reason for these observations.

NS2B/3[pro] shows higher activity and specificity in the presence of microsomal membranes

Having generated a form of C-prM that was resistant to cleavage by SPaseI, we could now examine its ability to act as substrate for NS2B/3[pro]. Initial control experiments in the absence of microsomal membranes revealed that C-prM and C-9L-prM were processed equally with respect to extent and position of cleavage (data not shown). We therefore incubated C-9L-prM that had been translated in the absence or presence of microsomal membranes with different concentrations of NS2B/3[pro]. Interestingly, the concentration of NS2B/3[pro] could be reduced by 10 to 20-fold (compare Fig. 5B lanes 4 and 5 with lanes 9 and 10) when microsomal membranes were used. NS2B/3[pro] at a final concentration of 500 ng/µl (Fig. 5B, lane 4) was required to cleave C-9L-prM in the absence of membranes, whereas 25 ng/µl (Fig. 5B, lane 10) were sufficient to obtain cleavage of this substrate when membranes were used.

To confirm that the NS2B/3[pro] enzyme was indeed recognising the correct cleavage site, we designed a C-9L-prM mutant in which the downstream NS2B/3[pro] cleavage site was replaced by a cleavage site of the HIV-1[pro] (C-HIV-9L-prM). Fig. 6 (lanes 4 and 5) shows that the protein is not cleaved by NS2B/3[pro] in the presence of membranes, even though the upstream cleavage site was still available. However, once again cleavage was observed in the absence of membranes (Fig. 6, lane 8).

Thus, we concluded from the above set of experiments that the NS2B/3[pro] was indeed recognising its true cleavage site but only in the presence of membranes.

The cleavage between C and prM can be performed by HIV-1[pro]

The presence of the HIV-1[pro] cleavage site in C-HIV-9L-M allowed us to investigate whether C-HIV-9L-prM was cleaved by HIV-1[pro] in the presence or absence of membranes. As a control, we used the construction C-9L-prM lacking the HIV-1[pro] cleavage site. In the absence of microsomal membranes, both constructs were processed by the HIV-1[pro] (Fig. 7A, lane 2 to

5), showing a similar pattern of two cleavage products of about 28 kDa and 20 kDa. The 28 kDa band presumably arises through cleavage at the C-prM junction to give prM*, whereas the 20 kDa band indicates cleavage at a position within protein C or prM. To verify that HIV-1^{pro} is responsible for the cleavages, we added increasing concentrations of pepstatin, a known HIV-1^{pro} inhibitor (Wondrak et al., 1991). Figure 7B shows a concentration dependent inhibition of cleavage, indicating that the cleavage is indeed due to HIV-1^{pro} processing.

Next we investigated HIV-1^{pro} cleavage in the presence of microsomal membranes (Fig 7A, lanes 6 to 8). Cleavage of the C-9L-prM construct was no longer observed (upper panel); however, with the C-HIV-9L-prM construct, the same cleavage products were observed as in the absence of membranes (lower panel). Thus, the presence of the membranes eliminated the cleavage of the C-prM construct lacking the HIV-1^{pro} site, indicating that the protein was correctly inserted in the membrane. In addition, the cleavage products obtained with the C-HIV-9L-prM indicate that correct processing at the HIV-1^{pro} sequence was occurring, but suggested that a second cleavage site was present within the prM protein that was still accessible even in the presence of membranes. Nevertheless, this experiment shows clearly that it is possible, at least *in vitro*, to make the cleavage between proteins C and prM dependent on HIV-1^{pro}.

Discussion

We present here a system in which the cleavage in the C-terminal region of protein C of TBEV can be specifically performed by purified recombinant viral NS2B/3 protease. Such cleavage required the presence of microsomal membranes to prevent non-specific proteolysis by the NS2B/3^{pro} as well as introduction of nine leucine residues at the C-terminus of the transmembrane segment of protein C to prevent incorrect SPaseI cleavage. Cleavage of the protein C-9L-prM in the presence of membranes was efficient and specific, requiring only 25 ng/µl to initiate cleavage. This NS2B/3^{pro} cleavage could be prevented by the replacement of the NS2B/3^{pro} cleavage site by a sequence from the FMDV 2A protein and thus demonstrated the specificity of cleavage, even though the putative upstream cleavage site in the C-terminal region of protein C was still present. The replacement of the NS2B/3^{pro} cleavage site by that of the HIV-1^{pro} cleavage site also prevented NS2B/3^{pro} cleavage, further demonstrating the specificity of the reaction. The cleavage between C and prM could now be carried out by the HIV-1^{pro}. However, a second cleavage was observed with the HIV-1^{pro}. This may explain why it was not possible to make a TBEV that was dependent on cleavage with HIV-1^{pro} as was the case for FMDV 3C^{pro} (Schrauf et al., 2012).

To establish this system, a number of difficulties had to be overcome. Firstly, the modular NS2B/3 protein that we constructed was capable of undergoing autocatalytic cleavage, even in the absence of a dibasic sequence in the proximity of the linker. This reaction could be

prevented by the substitution of an arginine residue (R93) two amino acids upstream of the synthetic linker. This suggests that the presence of this basic residue, coupled with the proximity of the linker to the active site, allowed slow autoproteolytic cleavage to take place on a non-canonical cleavage sequence.

This recombinant NS2B/3^{pro} was purified to homogeneity and was able to cleave specifically in the C-terminal region of protein C, provided that translation took place in the presence of microsomal membranes and that SPaseI cleavage had been prevented by extension of the trans-membrane helix. This is in contrast to experiments that examined the cleavage of West Nile virus C-prM using infected cells as a source of the active WNV NS2B/3^{pro} (Yamshchikov and Compans, 1994) and suggests a fundamental difference in the cleavage of the two viral polyproteins of WNV and TBEV.

Substitution of the NS2B/3^{pro} cleavage motif in C-prM by of HIV-1^{pro} allowed us to examine cleavage of C-HIV-9L-prM by the HIV-1^{pro} in the presence of membranes. In addition to a band corresponding to prM*, indicating cleavage at the HIV-1^{pro} sequence, an additional cleavage product with a size of about 20 kDa was observed. Where might the additional HIV-1^{pro} cleavage sequence be located? Sequence analysis via an online HIV-1^{pro} cleavage site prediction program (Chou, 1996; Chou et al., 1993; Shen and Chou, 2008) revealed a TVIR/AEGK motif spanning amino acids 134 to 143 with a high probability of cleavage by HIV-1^{pro}. Cleavage at this site in prM would result in the formation of two bands of 20 and 6 kDa, respectively, in good agreement with the observed 20 kDa band; the 6 kDa product would not be retained on the polyacrylamide gel. However, this site of the prM protein should be inside the membrane vesicles and not available for the protease. In contrast, the construction C-9L-prM without an HIV-1^{pro} site was not cleaved by HIV-1^{pro} in the presence of membranes. This implies that there is a difference in the orientation of the two proteins in the membrane that can only derive from the presence of the HIV-1^{pro} cleavage site in the C-terminal region of protein C. This mis-orientation of the polyprotein containing the HIV-1^{pro} cleavage site may explain why it was not possible to obtain virions when we tried to make the C-prM cleavage dependent on the HIV-1^{pro} in cell culture using a bi-cistronic virus (Schrauf et al, unpublished). As the cleavage at the C-terminus was possible with the FMDV 3C^{pro} (Schrauf et al., 2012), the failure with the HIV-1^{pro} must lie with this enzyme and/or with its cleavage site and not with the system *per se*.

This system will allow us to adjust the cleavage site of HIV-1^{pro} to prevent the cryptic cleavage and mis-orientation and thus to obtain a higher chance of constructing a TBEV that is dependent on the activity of the HIV-1^{pro}. In addition, this system may also serve as a model to examine the orientation of the transmembrane anchor of the TBEV protein C and the requirements for correct SPaseI processing.

Acknowledgements

References

Amberg, S.M., Nestorowicz, A., McCourt, D.W., Rice, C.M., 1994. NS2B-3 proteinase-mediated processing in the yellow fever virus structural region: in vitro and in vivo studies. J. Virol. 68, 3794-3802.

Amberg, S.M., Rice, C.M., 1999. Mutagenesis of the NS2B-NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. J. Virol. 73, 8083-8094.

Bera, A.K., Kuhn, R.J., Smith, J.L., 2007. Functional characterization of cis and trans activity of the Flavivirus NS2B-NS3 protease. The Journal of biological chemistry 282, 12883-12892.

Bessaud, M., Grard, G., Peyrefitte, C.N., Pastorino, B., Rolland, D., Charrel, R.N., de Lamballerie, X., Tolou, H.J., 2005. Identification and enzymatic characterization of NS2B-NS3 protease of Alkhurma virus, a class-4 flavivirus. Virus research 107, 57-62.

Chappell, K.J., Stoermer, M.J., Fairlie, D.P., Young, P.R., 2008. West Nile Virus NS2B/NS3 protease as an antiviral target. Current medicinal chemistry 15, 2771-2784.

Chou, K.C., 1996. Prediction of human immunodeficiency virus protease cleavage sites in proteins. Analytical biochemistry 233, 1-14.

Chou, K.C., Zhang, C.T., Kezdy, F.J., 1993. A vector projection approach to predicting HIV protease cleavage sites in proteins. Proteins 16, 195-204.

Dasso, M.C., Jackson, R.J., 1989. Efficient initiation of mammalian mRNA translation at a CUG codon. Nucleic Acids Res. 17, 6485-6497.

Donnelly, M.L., Luke, G., Mehrotra, A., Li, X., Hughes, L.E., Gani, D., Ryan, M.D., 2001. Analysis of the aphthovirus 2A/2B polyprotein 'cleavage' mechanism indicates not a proteolytic reaction, but a novel translational effect: a putative ribosomal 'skip'. J. Gen. Virol. 82, 1013-1025.

Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S.P., Yin, Z., Keller, T.H., Vasudevan, S.G., Hommel, U., 2006. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nature structural & molecular biology 13, 372-373.

Konishi, E., Mason, P.W., 1993. Proper maturation of the Japanese encephalitis virus envelope glycoprotein requires cosynthesis with the premembrane protein. J. Virol. 67, 1672-1675.

Lee, E., Stocks, C.E., Amberg, S.M., Rice, C.M., Lobigs, M., 2000. Mutagenesis of the signal sequence of yellow fever virus prM protein: enhancement of signalase cleavage In vitro is lethal for virus production. J. Virol. 74, 24-32.

Leung, D., Schroder, K., White, H., Fang, N.X., Stoermer, M.J., Abbenante, G., Martin, J.L., Young, P.R., Fairlie, D.P., 2001. Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. The Journal of biological chemistry 276, 45762-45771.

Lindenbach, B.D., Thiel, H.J., Rice, C.M., 2007. Flaviviridae: the viruses and their replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology Lippincott Williams & Wilkins, Philadelphia, pp. 1101-1152.

Lobigs, M., 1993. Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral proteinase NS3. Proc. Natl. Acad. Sci. U.S.A. 90, 6218-6222.

Lobigs, M., Lee, E., 2004. Inefficient signalase cleavage promotes efficient nucleocapsid incorporation into budding flavivirus membranes. J. Virol. 78, 178-186.

Lorenz, I.C., Allison, S.L., Heinz, F.X., Helenius, A., 2002. Folding and dimerization of tick-borne encephalitis virus envelope proteins prM and E in the endoplasmic reticulum. J. Virol. 76, 5480-5491.

Louis, J.M., Clore, G.M., Gronenborn, A.M., 1999. Autoprocessing of HIV-1 protease is tightly coupled to protein folding. Nature structural biology 6, 868-875.

Lundin, C., Kim, H., Nilsson, I., White, S.H., von Heijne, G., 2008. Molecular code for protein insertion in the endoplasmic reticulum membrane is similar for N(in)-C(out) and N(out)-C(in) transmembrane helices. Proc. Natl. Acad. Sci. U.S.A. 105, 15702-15707.

Mandl, C.W., Iacono-Connors, L., Wallner, G., Holzmann, H., Kunz, C., Heinz, F.X., 1991. Sequence of the genes encoding the structural proteins of the low-virulence tick-borne flaviviruses Langat TP21 and Yelantsev. Virology 185, 891-895.

Nall, T.A., Chappell, K.J., Stoermer, M.J., Fang, N.X., Tyndall, J.D., Young, P.R., Fairlie, D.P., 2004. Enzymatic characterization and homology model of a catalytically active recombinant West Nile virus NS3 protease. J. Biol. Chem. 279, 48535-48542.

Nilsson, I., Johnson, A.E., von Heijne, G., 2002. Cleavage of a tail-anchored protein by signal peptidase. FEBS Lett. 516, 106-108.

Schlick, P., Skern, T., 2002. Eukaryotic initiation factor 4GI is a poor substrate for HIV-1 proteinase. FEBS Lett 529, 337.

Schrauf, S., Kurz, M., Taucher, C., Mandl, C., Skern, T., 2012. Generation and genetic stability of tick-borne encephalitis virus mutants dependent on processing by FMDV 3C protease. The Journal of General Virology 93, 504-515.

Schrauf, S., Mandl, C.W., Bell-Sakyi, L., Skern, T., 2009. Extension of flavivirus protein C differentially affects early RNA synthesis and growth in mammalian and arthropod host cells. J. Virol. 83, 11201-11210.

Schrauf, S., Schlick, P., Skern, T., Mandl, C.W., 2008. Functional analysis of potential carboxy-terminal cleavage sites of tick-borne encephalitis virus capsid protein. J. Virol. 82, 2218-2229.

Shen, H.B., Chou, K.C., 2008. HIVcleave: a web-server for predicting human immunodeficiency virus protease cleavage sites in proteins. Analytical biochemistry 375, 388-390.

Shiryaev, S.A., Aleshin, A.E., Ratnikov, B.I., Smith, J.W., Liddington, R.C., Strongin, A.Y., 2007. Expression and purification of a two-component flaviviral proteinase resistant to autocleavage at the NS2B-NS3 junction region. Protein Expr Purif 52, 334-339.

Shiryaev, S.A., Ratnikov, B.I., Chekanov, A.V., Sikora, S., Rozanov, D.V., Godzik, A., Wang, J., Smith, J.W., Huang, Z., Lindberg, I., Samuel, M.A., Diamond, M.S., Strongin, A.Y., 2006. Cleavage targets and the D-arginine-based inhibitors of the West Nile virus NS3 processing proteinase. The Biochemical journal 393, 503-511.

Sousa, C., Schmid, E.M., Skern, T., 2006. Defining residues involved in human rhinovirus 2A proteinase substrate recognition. FEBS Lett. 580, 5713-5717.

Stadler, K., Allison, S.L., Schalich, J., Heinz, F.X., 1997. Proteolytic activation of tick-borne encephalitis virus by furin. J. Virol. 71, 8475-8481.

Stocks, C.E., Lobigs, M., 1995. Posttranslational signal peptidase cleavage at the flavivirus C-prM junction in vitro. J. Virol. 69, 8123-8126.

Stocks, C.E., Lobigs, M., 1998. Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. J. Virol. 72, 2141-2149.

Wondrak, E.M., Louis, J.M., Mora, P.T., Oroszlan, S., 1991. Purification of HIV-1 Wild-Type Protease and Characterization of Proteolytically Inactive HIV-1 Protease Mutants by Pepstatin-A Affinity Chromatography. FEBS Lett. 280, 347-350.

Yamshchikov, V.F., Compans, R.W., 1994. Processing of the intracellular form of the west Nile virus capsid protein by the viral NS2B-NS3 protease: an in vitro study. J. Virol. 68, 5765-5771.

Yusof, R., Clum, S., Wetzel, M., Murthy, H.M., Padmanabhan, R., 2000. Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. The Journal of biological chemistry 275, 9963-9969.

Figures

Figure 1

A   <u>NS2B</u>

```
TBEV    --SFSEPLTVVGVMLTLASGMMRHTSQEALCALAVASFLLLMLVLGTRKMQLVAEWSGCV 58
WNV     GWPATEVMTAVGLMFAIVGGLAELDIDSMAIPMTIAGLMFAAFVISGKSTDMWIERTADI 60
         . :* :*.**:*::..*: .    :.    .:::*.::: :*:. :. ::   * :. :

TBEV    EWYPELVNEGGEVSLRVRQDAMGNFHLTELEKEERMMAFWLIAGLAASAIH-WSGILGVM 117
WNV     TWESDAEITGSSERVDVRLDDDGNFQLMNDPGAPWKIWMLRMACLAISAYTPWAILPSVI 120
         * .:    *..  : ** *  ***:* :        : :  :* ** **   *: : .*:

TBEV    GLWTLTEMLRSSRR 131
WNV     GFWIT---LQYTKR 131
        *:*     *: ::*
```

B   <u>NS3</u> (N-terminal parts)

```
TBEV    SDLVFSGQGGRERGDRPFEVKDGVYRIFSPGLFWGQNQVGVGYGSKGVLHTMWHVTRGAA 60
WNV     GGVLWDTPSPKEYKKG--DTTTGVYRIMTRGLLG-SYQAGAGVMVEGVFHTLWHTTKGAA 57
        ..::::.  . :*   .   :.. *****:: **:  . *.*.*   :**:**:**.*:***

TBEV    LSIDDAVAGPYWADVREDVVCYGGAWSLEEKWKG-ETVQVHAFPPGRAHEVHQCQPGELI 119
WNV     LMSGEGRLDPYWGSVKEDRLCYGGPWKLQHKWNGHDEVQMIVVEPGKNVKNVQTKPGVFK 117
        *  .:.  .***..*:** :****.*.*:.**:* : **: .. **:  :   * :** :

TBEV    LDTGRKLGAIPIDLVKGTSGSPILNAQGVVVGLYGNGLKT-NETYVSSIAQGEAEKSRPN 178
WNV     TPEG-EIGAVTLDYPTGTSGSPIVDKNGDVIGLYGNGVIMPNGSYISAIVQGERMEEP-- 174
         * ::**:.:*  .*******:: :* *:******:   * :*:*:*.*** :.

TBEV    LPQAVVGTGWTSKGQITVLDMHPGSGKTHRVLPELIRQCIDRRLRTLVLAPTRVVLKEME 238
WNV     -APAGFEPEMLRKKQITVLDLHPGAGKTRKILPQIIKEAINKRLRTAVLAPTRVVAAEMS 233
         . * . .    * ******:***:***::*:*::*::.*::**** *******  **.
```

Fig. 1. Alignment of tick-borne encephalitis virus (TBEV) and West Nile virus (WNV) amino acid sequences. (A) Complete sequences of NS2B. A conserved region involved in the formation of the active site in WNV is underlined. (B) N-terminal portions of NS3. Catalytic triad residues (His 54, Asp 78 and Ser 138) are indicated in blue. The parts of NS2B and NS3 used for constructing the recombinant TBEV NS2B/3 are green and red, respectively.
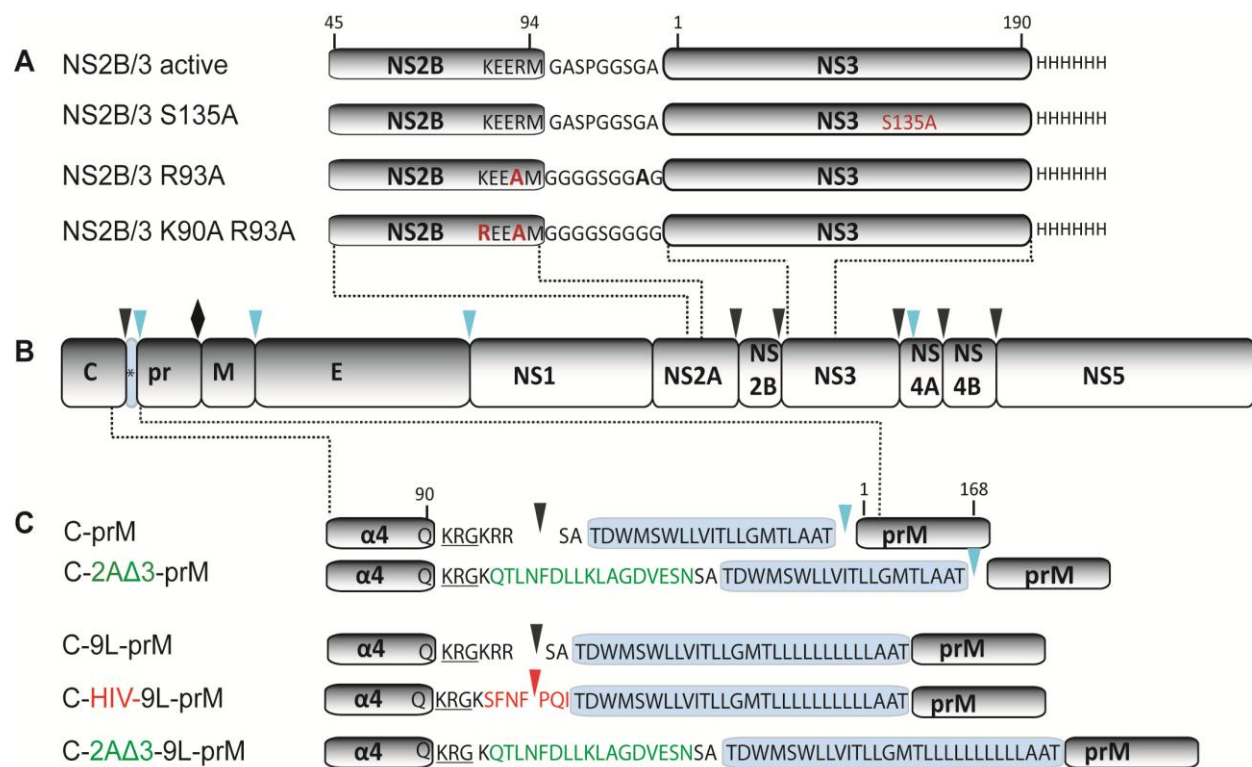
Figure 2

Fig. 2. (A) Variants of recombinant proteases. The 48 amino acid residue-long, central portion of NS2B was linked with the NS3^{pro} sequence via a GASPGGSGA or GGGGSGGGAG linker, respectively. All protease constructs were C-terminally tagged with a (his)x6 tag. (B) Schematic diagram of the TBEV polyprotein with cleavage sites for the viral protease (black arrowhead), host signal peptidase I (blue arrowhead) and furin (diamond). (C) Variants of substrates used in this study. Schematic drawing of the C-terminal region of the capsid protein and the signal sequence for prM (not drawn to scale). Engineered mutations are shown together with the corresponding designations. The potential but unused upstream NS2B/3^{pro} cleavage site is underlined. The downstream cleavage site of NS2B/3 (black), and cleavage sites of SPaseI (blue) and HIV-1 protease (red) are indicated by arrowheads. Numbers at the top refer to the amino acid positions within protein C or prM. The transmembrane region is highlighted in blue.

Figure 3
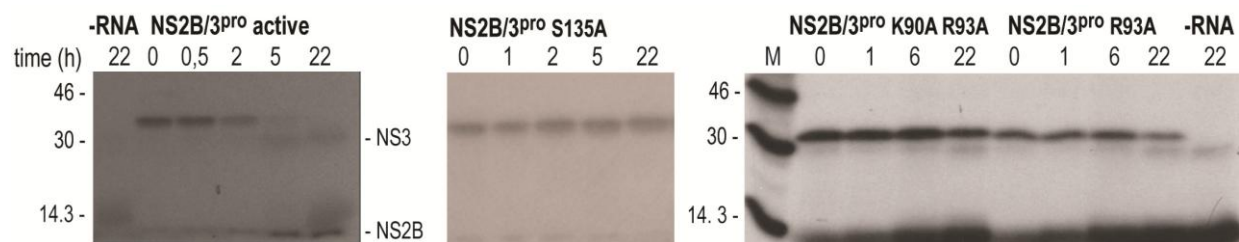
Fig. 3. Analysis of autocatalytic activity of TBEV NS2B/3^{pro}. *In vitro* transcribed RNAs encoding the indicated proteins were translated in RRLs in the presence of ^{35}S-methionine for 30 min at 30°C. Following the addition of unlabelled methionine, samples were incubated for the indicated times in buffer containing 50mM NaCl, 50mM Tris/HCl, pH7. Proteins were separated by SDS-PAGE and ^{35}S-containing proteins detected by fluorography.

Figure 4

Fig.4. Processing of TEBV polyprotein subfragments containing the C and prM proteins by purified recombinant NS2B/3^{pro}. (A) C-prM. (B) C-2AΔ3-prM. *In vitro* translated substrates were incubated with buffer or NS2B/3^{pro} for the indicated times. Proteins were separated by SDS-PAGE and ^{35}S-containing proteins detected by fluorography.

Figure 5

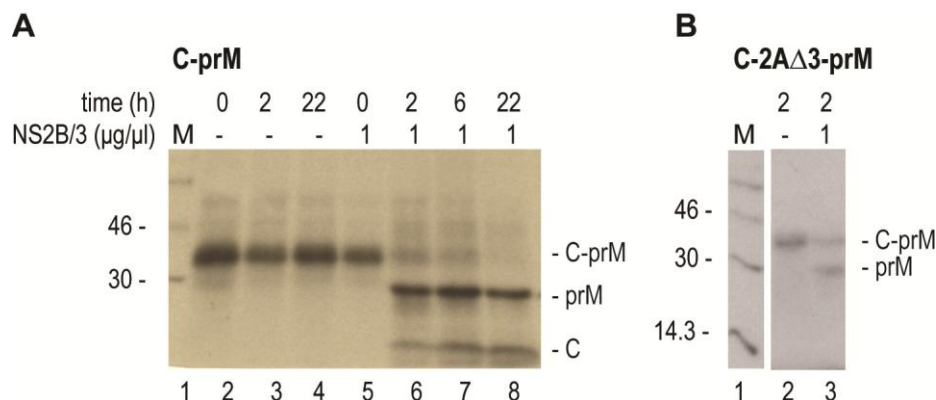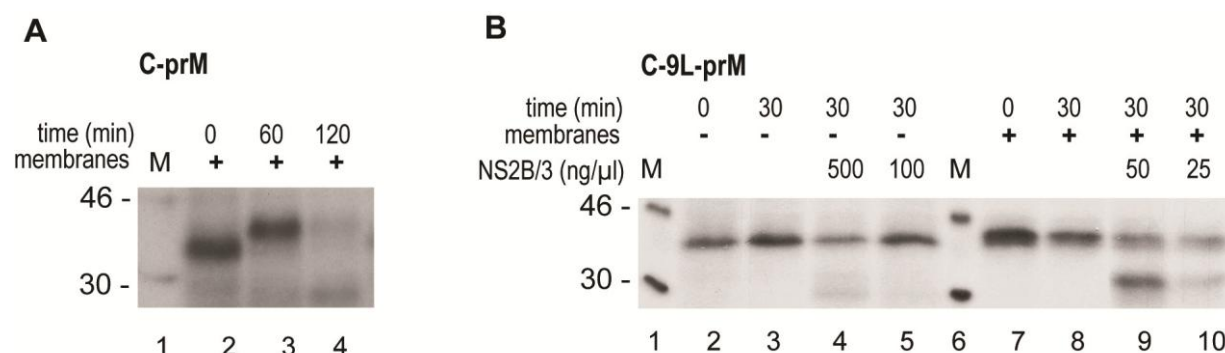**Fig.5.** Effect on proteolytic processing of the synthesis of C-prM subfragments in the presence of microsomal membranes (A) C-prM. Translation products were incubated at 30°C for the indicated times. (B) C-9L-prM was synthesised in the presence or absence of microsomal membranes and then incubated in buffer or with NS2B/3^{pro} at the indicated concentrations and times at 30°C. Proteins were separated by SDS-PAGE and ^{35}S-containing proteins detected by fluorography.

Figure 6
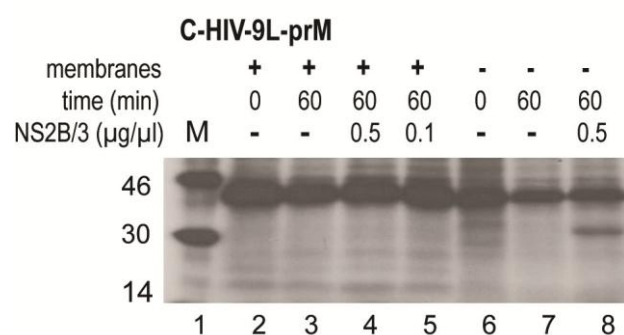
Fig. 6. Effect of the replacement of the NS2B/3 cleavage site in C-9L-prM with that of HIV-1^{pro} on NS2B/3^{pro} processing. C-HIV-9L-prM was translated in the presence (lanes 2 – 5) or absence (lanes 6 to 8) of microsomal membranes and incubated with buffer or NS2B/3^{pro} at the indicated final concentrations and for the indicated times at 30°C. Proteins were separated by SDS-PAGE and ^{35}S-containing proteins detected by fluorography.

Figure 7
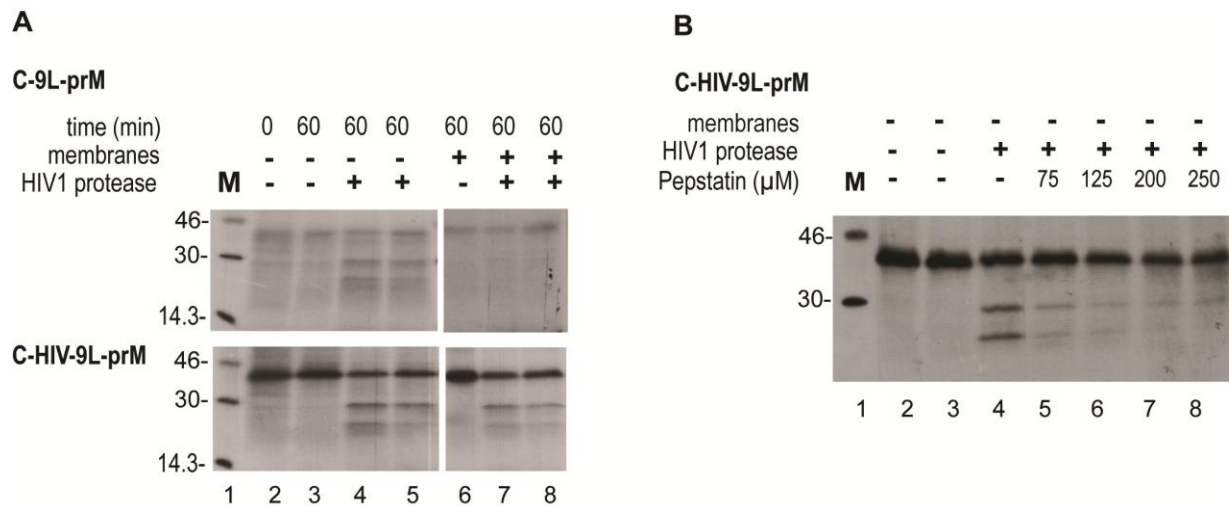
Fig. 7. Cleavage at the TBEV C-prM junction by HIV-1[pro]. (A) HIV-1[pro] cleavage in the absence (lanes 1 to 3) and presence (lanes 5 to 6) of microsomal membranes. C-9L-prM (upper panel) and C-HIV-9L-prM (lower panel) were translated with or without microsomal membranes and incubated with purified HIV-1[pro] for 30 min at 30 °C. (B) C-HIV-9L-prM was translated in the absence of microsomal membranes and incubated with HIV-1[pro] in the absence or presence of pepstatin at the indicated concentrations. Proteins were separated by SDS-PAGE and ^{35}S-containing proteins detected by fluorography.

## Supplementary Table 1: Sequences of employed oligonucleotides

C-prM-f   5´CCTTACCATGGTAAAGAAGGCCATCCTGAAA

C-prM-r   5´TTAGGATCCTTATCAAGCGTAAACCGGTGCCAA 3´

NS2B-f   5´CCTACCATGGCTAGAAAGATGCAGCTGGTT 3´

NS2B-r-3´linker   3´GGATGCGTCAGTCCCGGGAGAAGCACCCATTCTCTTCTTTCTC 3´

5´-Linker-NS3-F   5´AGCGGTACTGCACCCGGGGGTAGCGGCGCTTCTGACCTGGTTTTCTCT 3´

3´NS3-R-6xHis   5´TTAGGATCCTTATCAGTGATGATGGTGGTGATGTGATGTCCAGCCCGTACC 3´

NS3-S135A-f   5´GTGAAAGGAACAGCAGGCAGCCCCATT 3´

NS3-S135A-r   5´AATGGGGCTGCCTGCTGTTCCTTTCAC 3´

NS3-BglI-f   5´GGCGCTTCAGATCTGGTTTTC 3´

NS3-BglI-r   5´GAAAACCAGATCTGAAGCGCC 3'

K47A, R50A-f   5´TCGAGGCGGAAGAGGCGATGGGTGGCGGTGGCTCTGGTGGCGGTGGCTCA 3´

K47A R50A-r   5´GATCTGAGCCACCGCCACCACAGCCACCGCCACCCATCGCCTCCTTCCGCC 3´

R50A-f   5´ TCGAGAAAGAAGAGGCGATGGGTGGCGGTGGCTCTGGTGGCGGTGGCTCA 3`

R50A-r   5´ GATCTGAGCCACCGCCACCAGAGCCACCGCCACCCATCGCCTCCTTCTTC 3`

TM-9L-f: 5´CTGTTGCTGTTGCTAGCTGCAACGGTGAGGAAAGAAAGG 3´

TM-9L-r.   5´CAACAGCAACAGAAGCGTCATCCCCAACAGAGTGATGACCAG 3´

prM-Δglyc-f   5´GGTGCGTGTGGAGGCCGGCACCTGTGTGATCC 3´

prM-Δglyc-r   5´GGATCACACAGGTGCCGGCCTCCACACCACC 3´

5. MANUSCRIPT 3

Towards structure determination of the TBEV NS2B/3 protease – Expression, purification and crystallisation of various NS2B/3[pro] constructs

Martina Kurz[§], Nikolas Stefan [§†] and Tim Skern[§ ¶]

Manuscript in preparation

Initial cloning, expression and purification of the NS2B/3 S135A protease was executed by N. Stefan. All further cloning, expression, purification, optimisation, crystallisation as well as writing the manuscript was performed by M. Kurz.

[§] Max F. Perutz Laboratories, Medical University of Vienna, A-1030 Vienna, Austria

[†] Current address: Dept. of Biochemistry, University of Zurich, CH-8057 Zurich, Switzerland

5.1.   Introduction

The tick-borne encephalitis (TBE) serocomplex of viruses belongs to the family of *Flaviviridae*, genus *Flavivirus*. It comprises a number of human pathogens that cause serious neurological diseases. The most important member is the tick borne encephalitis virus (TBEV), the causative agent of a potentially lethal neuroinfection, TBE (Robertson et al., 2009). In most cases, the virus is transmitted to humans through the bite of an infected tick, but it can also be acquired by consuming unpasteurised dairy products from viremic livestock (Süss, 2003).

In Europe, two vaccines are available: FSME-IMMUN (Baxter, Austria), and Encepur (Novartis, Germany), both of which are based on the European subtype (Heinz and Holzmann 2007). At the moment, there exists no specific treatment for TBE; therapy consists of symptomatic and supportive care and management of complications. Thus, a therapeutic agent against TBEV would be highly valuable as it could potentially save many lives throughout the world.

The TBEV is a small, spherical, enveloped virus with a single positive strand RNA genome of 10.8 kb. Upon infection, the genome is translated into one single polyprotein which is subsequently cleaved into its mature, functional proteins by host cell proteases together with the viral protease NS3. The viral cofactor NS2B is essential for the catalytic activity of the NS3 protease. As the TBEV protease NS2B/3 (NS2B/3^{pro}) is indispensible for the processing of its polyprotein and subsequently for virus replication, it is an attractive target for the antiviral therapy.

The crystal structures of the West Nile virus or the dengue virus NS2B/3^{pro} complex, with and without an inhibitor, respectively, have already been solved (Erbel et al., 2006) and provided important insight into the mechanism of substrate recognition and activation of the NS3^{pro}. The same can be expected from the 3D structure of the physiologically relevant TBEV dNS2B/3^{pro} (D´Arcy et al., 2006), which should therefore provide an important impetus to the development of a TBEV protease inhibitor.

Thus, we set out to optimise expression and purification of the NS2B/3^{pro} complex in order to obtain protein of sufficient quality and quantity to allow crystallisation. To obtain well diffracting, well-ordered protein single-crystals is the vital aim of protein crystallography and an important step towards solving the 3D structure of the NS2B/3^{pro}.

5.2.   Material and methods

5.2.1.   Cloning, expression, purification and chemical modification of NS2B/3[pro] S135A and NS2B/3[pro] S135A AAA

5.2.1.1. Cloning

Constructs encoding tick-borne encephalitis virus European subtype strain Neudörfl (EMBL: U27495) were derived from pBR-TBEV-ΔME-E-EGFP (kind gift of Dr. Petra Schlick). Plasmids coding for the inactive NS2B/3[pro] S135A were constructed in two steps as described in Kurz et al., 2012 (manuscript 2) and named pCite-A1-NS2B/3 S135A.

For expression in *E. Coli,* sequences coding for the tethered proteases were cloned from pCite-A1-NS2B/3 into the expression vector pET11d by use of *Nco*I and *Bam*HI.

Plasmid pET11d-NS2B/3 S135A K47A E48A E49A (for simplicity subsequently referred to as NS2B/3 S135A AAA) was made by site-directed mutagenesis with primers NS2B/3-AAA-f (5´GACTGAGCTCGAGGCTGCAGCGAGAATGGGTGC 3´) and NS2B/3-AAA-r (5´GCACCC ATTCTC GCTGCAGCCTCGAGCTCAGTC 3´).

5.2.1.2. Expression and purification

The plasmids pET 11d-NS2B/3 S135A and pEt11d-NS2B/3 S135A AAA were used for high-level, inducible expression of hexa-histidine-tagged (His$_6$) recombinant proteins as described in Kurz et al., 2012 (manuscript 2).

5.2.1.3. Lysine methylation (according to Walter et al., 2006)

NS2B/3[pro] S135A was expressed in *E.coli* BL21 DE3 pLysS and purified using HiTrap chelating column and HiLoad TM 26/60 Superdex 75 as stated above. The protein was eluted from the latter with buffer containing 50mM HEPES and 250mM NaCl, pH 7.5. Fractions were analysed by SDS-PAGE, those containing the protein were pooled. For each ml of protein solution 20µl 1M dimethyl-boran complex (freshly prepared in H$_2$O) and 40µl 1M formaldehyde were added and left under stirring at 4° C for 2 hours. Then, the same amount of dimethyl-boran complex and formaldehyde was added again, and left stirring at 4°C for another 2 hours. Finally, 10 µl of the 1M dimethyl-boran complex were added per ml protein solution and left stirring overnight at 4°C, covered with parafilm.

The sample was concentrated to 3ml using the Amicon Ultra-4-Centrifugal filter Device and applied to the HiLoad TM 26/60 Superdex 75 column preequilibrated with buffer containing 50mM Tris/HCl, pH 7.5, 100mM NaCl and 5% glycerol. Fractions were analysed via SDS-

PAGE, pooled and concentrated to 5.25 and 10.5 mg/ml, followed by centrifugation (14000 rpm, 4°C, 45 min) to remove precipitate. The supernatant was used for crystallisation trials.

5.2.2. Cloning, expression and purification of fusion protein MBP-NS2B/3 S135A

5.2.2.1. Cloning

Plasmid pet15b-MBP-NS2B/3 was generated as follows: The coding sequence of NS2B/3 S135A was cut out of pET NS2B/3 S135A by using primers NS2B/3 *Nde*I (5′CAGTCATATGGCTAGAAAGATGCAGCTG-3′) and NS2B/3-*Bam*HI (5′TGCATGGATCC TCACTATGATGTCCAGCCCGTACCCAC 3′). Thereby, a cleavage site for *Nde*I and *Bam*HI was inserted and the C-terminal His_6-tag was removed. These enzymes were then used to introduce the NS2B/3 S135A coding sequence into vector pET15b-MBP (kind gift of Dr. Gang Dong) downstream of the maltose binding protein. The His_6-tag at the N-terminus of the MPB coding sequence is removable by thrombin cleavage.

5.2.2.2. Expression and purification

MBP-NS2B/3^{pro} S135A fusion protein was expressed in *E.coli* strain BL21 DE3 grown in LB medium containing 100mg/ml ampicillin and 35µmol/ml Chloramphenicol at 37°C overnight. This culture was diluted 1:100 in fresh LB medium containing ampicillin and incubated at 30°C, 225 rpm until an OD_{600} of 0.5 to 0.6 was reached. The expression of the recombinant proteins was induced by addition of isopropyl-β-D-thiogalactopyranose (IPTG) to a final concentration of 0.5 mM. Cultures were incubated overnight at 16°C, and cells were harvested by centrifugation. Cell pellets were resuspended in 30 ml lysis buffer containing 50mM Tris-HCl, 200mm NaCl, pH 7.5, 5 % glycerol and lysed with a Bandelin Sonoplus sonicator (4 cycles, 40 %, 30 seconds, 1 cycle, 100 %, 30 seconds) followed by centrifugation at 17,000 rpm for 30 minutes at 4°C. The supernatant was loaded onto a 5-ml HiTrap chelating column (GE Healthcare) loaded with Ni^{++} and pre-equilibrated with lysis buffer. The protein was eluted with a linear gradient of 10 to 500mM imidazole in lysis buffer. Fractions containing MBP-NS2B/3 protein were determined by 12.5 % SDS-PAGE, pooled and diluted to a concentration of 1 mg/ml with lysis buffer to reduce the imidazole and NaCl to a final concentration of 60 and 160 mM, respectively. The N-terminal His_6 tag was removed by incubation overnight with 50 Units/ml thrombin. Then, the protein was diluted with lysis buffer to further reduce the imidazole concentration, then concentrated to a volume smaller than 13 ml using Amicon Ultra-4 Centrifugal filter Device 10kDa cut-off (Millipore). To remove the His_6 tag, the protein was loaded on to a HiLoad TM 26/60 Superdex 75 column (GE Healthcare) equilibrated with buffer containing 20mM Tris/HCl, pH 7.4, 100mM NaCl and 5 % glycerol. Fractions were again

analysed by SDS-PAGE, protein containing fractions pooled and concentrated to 25, 50 and 66 mg/ml, respectively. The protein samples were incubated with an equimolar amount of maltose for 1 hour at 4°C, followed by centrifugation (60 minutes, 14000 rpm, 4°C). This solution was used for crystallisation trials.

5.2.3. Crystallisation techniques

5.2.3.1. Pre crystallisation test

To optimise the protein concentration prior to initial screens a pre-crystallisation test (PCT, Hampton research) was performed according to the manufacturers' instructions. When necessary, the protein concentration was adjusted by further concentrating or diluting with elution buffer.

5.2.3.2. Crystallisation trials

Vapour-diffusion crystallisation experiments were performed using a nano-robot and the sitting drop method. In a typical experiment, 0.1 or 0.2 µl protein solution were added to 0.1 µl screening solution on 96-well Intelliplates (Robbins Instruments, Sunnyvale, USA), the reservoir well contained 70 µl screening solution.

The screening solutions used for the experiments were Index, Crystal Screen, Crystal Screen Cryo I and II and SaltRX (Hampton research), the PEGs Suite (NExtal Biotechnologies), Wizard I and II (Emerald BioStructures), JCSG+ Suite, PEG Suites (Qiagen) and PACT HT96 (Molecular Dimensions Limited). Monitoring of the crystal growth started immediately; photos were taken routinely after 12 and 24 hours, and at day 2, 3, 5, 9, 17 and 31 after setting the drops by using the RIGAKU Crystal Web Site.

5.2.3.3. Streak seeding

Crystals were pulverised into crystalline particles by vortexing. The nuclei were streaked with a seeding wand onto the surface of a new drop which had been equilibrated overnight at 196K.

5.3. Results and Discussion

In order to be able to determine a molecules´ structure via X-crystallography, well-diffracting crystals are needed. To generate them, a sufficient amount of pure, soluble and stable protein is necessary. To this end, the NSB/3[pro] S135A was overexpressed in *E.coli* and

purified with affinity and size exclusion chromatography (SEC). To maximise the amount of soluble protein, several parameters of the expression procedure, including induction and cell lysis, had to be optimised

5.3.1. Expression

5.3.1.1. Optimisation of induction

Induction parameters including inducer concentration, period of induction and the cell concentration at which the inducer is to be added to the fermentation broth were optimised in order to increase the yield of NS2B/3^{pro} S135A.

5.3.1.1.1. Optimisation of IPTG concentration

In order to determine the effect of the inducer concentration on induction efficiency, six parallel shake-flask fermentations were performed. The expression was done as described in Material and Methods. Plasmid pET 11d NS2B/3 S135A was transformed into competent *E.coli* BL21 DE3 LysS cells and grown in LB medium containing ampicillin and chloramphenicol over night at 37°C. The starter culture was 1:10 diluted in LB medium containing selective antibiotics and grown at 30°C for around 1 ½ hours until an OD_{600} of 0.6 was reached. Expression was induced with IPTG at a final concentration of 0.0 mM, 0.05 mM, 0.10 mM, 0.25 mM, 0.50 mM and 1.0 mM, respectively. After incubation for 4 hours at 30°C, cells were harvested and sonicated as described in Materials and Methods. The total protein fraction was analysed via SDS-PAGE (Fig. 1).
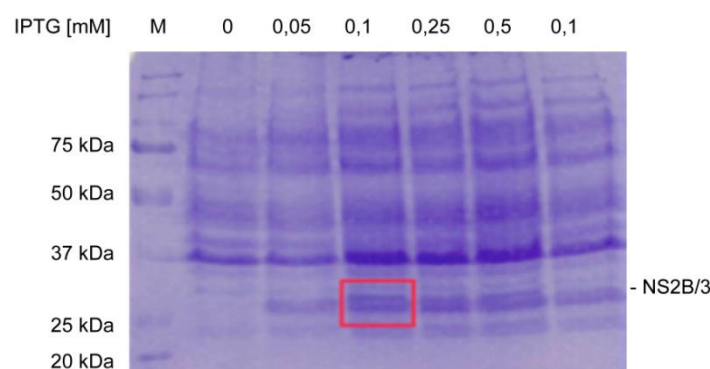

Figure 1. Optimisation of IPTG concentration. Protein expression was performed as described in detail in Material and Methods. Protein expression was induced with various IPTG concentrations between 0.05 and 1 mM. Aliquots were taken 4 hours post induction, sonicated, centrifuged and analysed via SDS-PAGE. M indicates the Precision Plus Protein All Blue Standards (BioRad) Marker.

The protein expression reached a maximum at an IPTG concentration of 0.1 mM. A further increase in the IPTG concentration did not improve protein production. Thus, an IPTG concentration of 0.1 mM was used in all further experiments.

5.3.1.1.2. Optimisation of starting time and period of induction

In order to determine the optimal starting time of induction, six shake-flask fermentations were performed in parallel. Each culture was induced at different phases of growth, namely at absorbance values of 0.6 and 0.9 at 600 nm. Following the addition of 0.1 mM IPTG to each of the culture media, the incubation was continued over a 4, 5 or 6-hour period, and protease expression was determined in cell extracts after sonication.

As shown in figure 2, a maximal yield was obtained with an induction at an OD_{600} of 0.6 and harvest of cells after 4 hours.
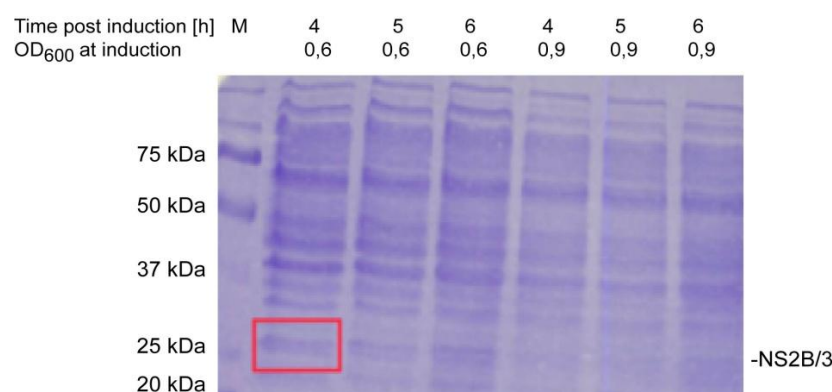

Figure 2. Optimisation of optimal starting time and period of induction. The protein expression was performed as described in detail in Material and Methods. The starter culture was 1:10 diluted in LB medium with selective antibiotics and grown at 30°C until an OD_{600} of 0.6 and 0.9 was reached. Protein expression in both cultures was induced with the optimised IPTG concentration of 0.1 mM. Aliquots were taken 4, 5 and 6 hours post induction. Then, cells were sonicated, centrifuged and analysed via SDS-PAGE.

5.3.1.2. Optimisation of cell lysis: Sonication vs. French Press

For the extraction of soluble protein from bacteria various methods are available; many of these need specialised, expenisve equipment. We compared two simple methods for cell disruption, namely French pressing and sonication.

NS2B/3[pro] S135A was expressed under optimised conditions as described in Material and Methods. 2 ml of induced and uninduced cells were sonicated on ice (4 times, 30 seconds, 20 %). In a second setup, induced cells only were applied to one or two cycles of French Press.

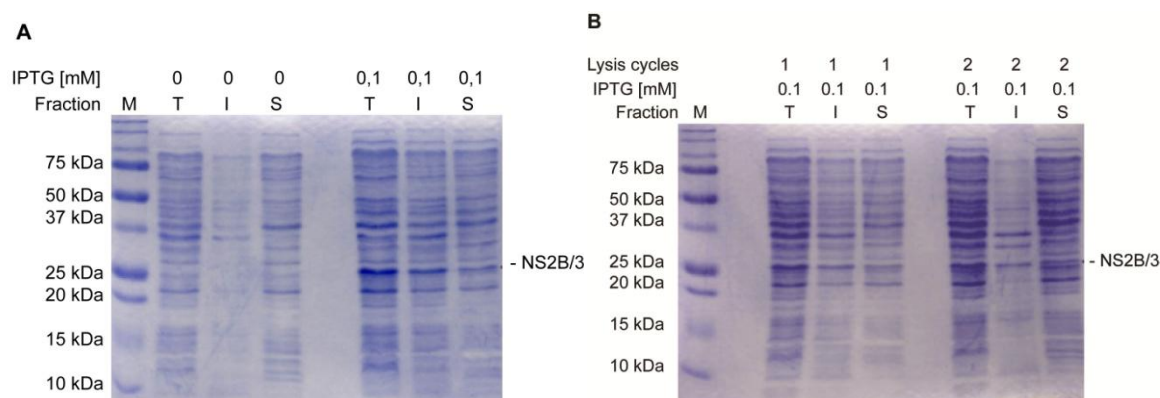Soluble and insoluble fractions were separated by centrifugation. The SDS-PAGE results are shown in figure 3.

Figure 3. Efficiency of cell lysis. Cells were lysed by (A) sonication or (B) French press. Total (T), insoluble (I) and soluble (S) fractions were analysed via SDS PAGE; 5 µl of the samples were loaded on a 15% SDS PAGE; M = Marker

The method of cell lysis did not have a major influence on the amount of soluble protein produced. Thus, sonication was used in all further protein preparation procedures, as it is less labour and time intensive.

5.3.2. Purification

5.3.2.1. Affinity chromatography

The purification of the NS2B3/3^{pro} S135A is described in detail in the Material and Methods section. The NS2B3^{pro} S135A carries a His$_6$-tag at the C-terminus, which was used for purification via affinity chromatography. The soluble fraction of the cell lysate was loaded onto the Hi-Trap column pre-equilibrated with Nickel. The protein was eluted with a linear gradient of imidazole in lysis buffer.

5 µl of each fraction of the second peak A8 to A12 were analysed via SDS-PAGE (Fig. 4, picture inlay). The most prominent bands run with a molecular weight slightly above 25 kDa. This is in good agreement with the calculated molecular weight of the NS2B/3^{pro} S135A of 28 kDa, indicating that the protease was already highly pure and concentrated after the first purification step.
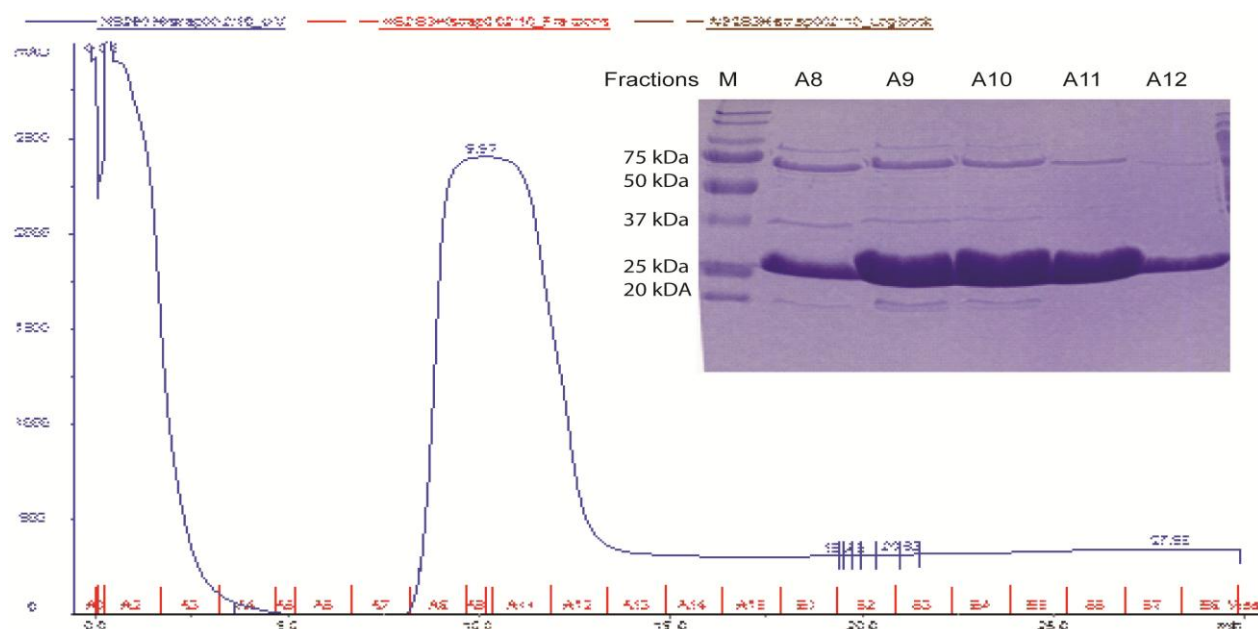
Figure 4. Example of an affinity chromatography profile. The gel picture shows the eluted fractions A8 to A12 corresponding to the second peak in affinity chromatogram).

5.3.2.2 Size exclusion chromatography (SEC)

Fractions A8 to A12 from the affinity chromatography were pooled, concentrated by centrifugation and applied to SEC.

Peaks A and C were analysed via SDS-PAGE. Interestingly, the molecular weight of the protein in both fractions was the same. Western blot analysis showed that both peaks held NS2B/3[pro] (data not shown). We assume that the protein eluted with the first peak is aggregated protein. Thus, eluat from peak A was not used for further experiments.
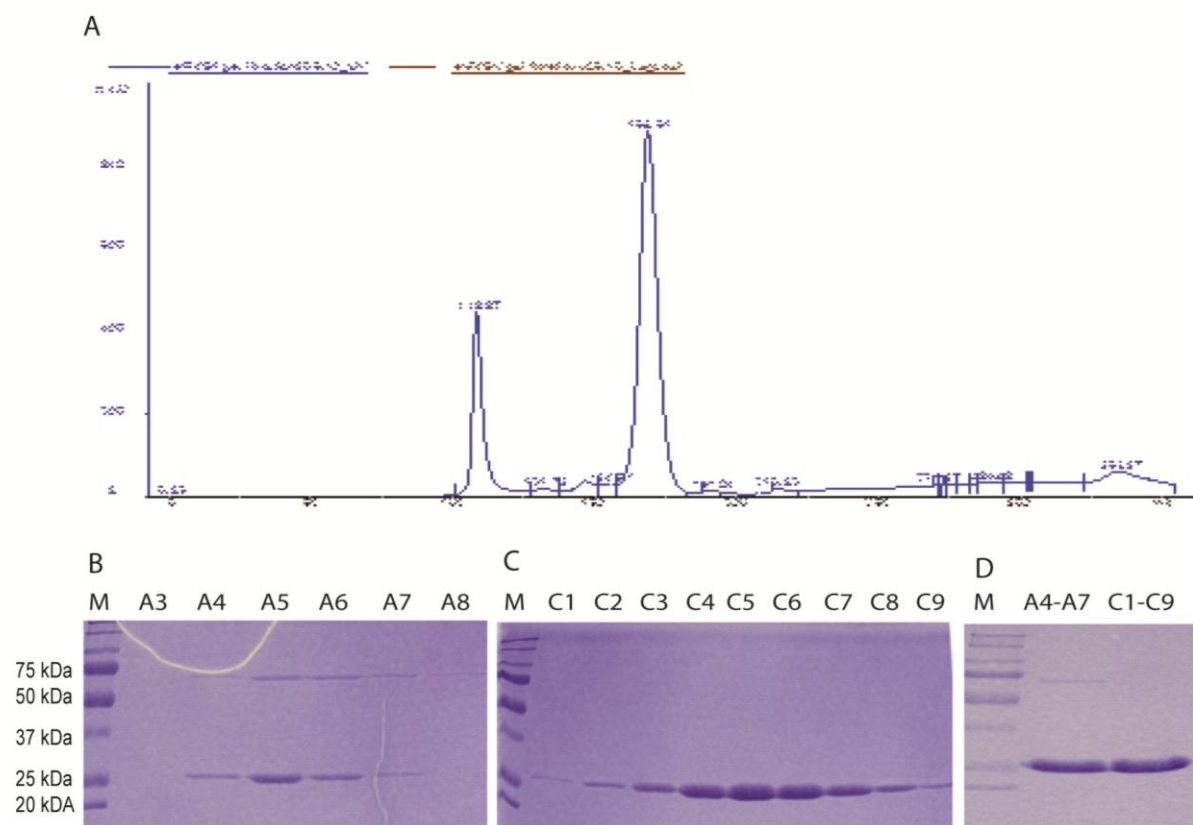
Figure 5. Example of a size exclusion chromatography profile and SDS-PAGE after gel filtration. (A) SEC profile, (B) and (C) SDS-PAGE of fractions of peak A and C, respectively (2,5 µl/lane). (D) The fraction of both peaks A and C were concentrated separately and analysed by SDS-PAGE (0.5 µl, 11 mg/ml)

5.3.3. Quality analysis

Homogeneity, both chemical and conformational, stability and solubility of a biological macromolecule are key factors that have a strong effect on the probability of obtaining crystals. Thus, these parameters have to be tested.

5.3.3.1. Dynamic light scattering (DLS)

Dynamic light scattering is applied in protein crystallography to obtain information about a protein in solution. Hereby, the proteins hydrodynamic radius is determined, as well as information about the oligomeric state of the protein sample. The extent of size distribution is a parameter for the polydispersity of the sample and correlates with protein crystallisation success. Thus, DLS represents a simple and non-invasive screening technique for evaluating the crystallisability of protein solutions.

NS2B/3^{pro} S135A from the initial purification at a concentration of 11 mg/ml in Tris/HCl, pH7, 50mM NaCl and 5 % was analysed using DLS to determine its suitability for crystallisation

as described by Ferre-D'Amare Burley (1994) and Zulauf & D'Arcy (1992). The protein was filtered through 0.22 µm centrifugal units prior to analysis, all measurements were made at 296K.

The protein preparation was found to be monodisperse to a very high degree when examined by DLS.
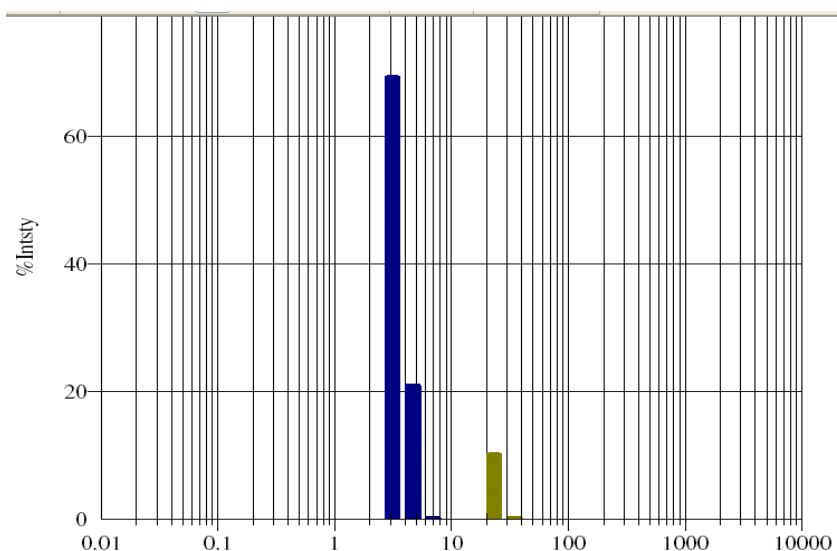
Figure 6. Size distribution profiles from the DLS. The results indicate a monodisperse preparation for the concentrated sample peak C of the SEC profile.

5.3.3.2. Thermofluor

In order to determine whether the NS2B/3^{pro} S135A is folded and which buffer could help to improve stability and folding of the protein, a Thermofluor measurement was performed.

The method exploits the fact that it is possible to distinguish between folded and unfolded proteins through exposure to a hydrophobic fluorophore which is quenched in aqueous solution. Thus, as long as the protein is properly folded and the hydrophobic areas are buried in the interior of the protein, no fluorescence signal is observed. However, with increasing temperature, the protein starts to unfold, thereby exposing its internal hydrophobic domains. These bind the fluorophore, as a consequence, the quenching effect of the water is abolished. Thus, this interaction gives rise to a fluorescence signal which is a function of temperature. In this way, the melting temperature Tm of the protein can be determined. Conditions which cause an increase in the melting temperature of the protein stabilise it. Furthermore, it is assumed that a more stable protein is more prone to crystallise (Ericsson et al., 2006)

The Thermofluor experiment was performed on a 96-well real-time PCR machine (BioRad) as done by Nettleship et al., 2008 and Lo et al., 2006. This way, 96 different buffer conditions could be tested simultaneously.
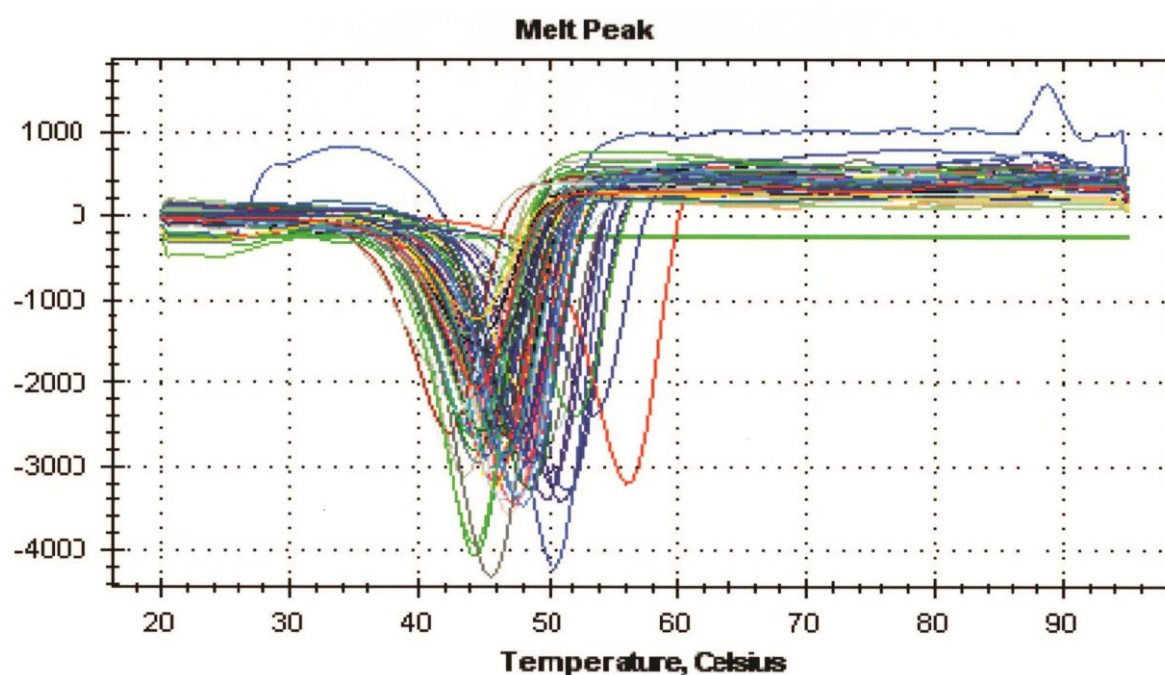
Figure 7. Melting curves of NS2B/3pro S135A in different buffer systems. A temperature gradient was applied, starting from 4 C and increasing to 95 C, measuring the fluorescence signal every 0.5 min. SYPRO® Orange (Invitrogen) was used as fluorophore, the absorption was measured at a wavelength of 600 nm. A sigmoidal curve (thermogram) was obtained by plotting the fluorescence intensity against the temperature.

The Thermofluor experiments showed that the protein was correctly folded in almost any buffer system tested. Exceptionally high melting temperatures were determined in buffer containing 2-(N-morpholino)ethanesulfonic acid (MES, red curve, peak at 56°C), pH 5.8 and in Tris/HCl buffer (violet curve, peak at 54°C), pH 7.5, respectively.

The latter one is the buffer system with which the protein was initially eluted from the SEC column. To test whether the MES buffer could further stabilise the NS2B/3^{pro} S135A and thus favour its crystallisation, we expressed and purified it again as described in Materials and Methods. This time, the sample was eluted in buffer containing 50 mM MES, 50 mM NaCl and 5% glycerol and concentrated to 10, 20 and 30 mg/ml. As before, the samples were analysed by SDS-PAGE and DLS and proven to be of sufficient purity and homogeneity for crystallisation (data not shown).

5.3.4. Crystallisation

At the beginning of the crystallisation process, purified, highly concentrated protein is diluted in an aqueous solution with a precipitant concentration lower than that needed for crystallisation. In the curse of slow evaporation of the solvent, both protein and precipitate concentration are enhanced; thereby, precipitating conditions are generated which allow crystallisation to occur. These conditions remain stable until crystal growth terminates. Once suitable protein crystals become available, X-ray diffraction data can be collected.

The crystallisation experiment starts by testing a wide range of screening conditions to find the precipitant that favours crystallisation. Commercially available crystallisation suites allow a combination of various chemical conditions to be tested in a systematic manner.

5.3.4.1. NS2B3^{pro} S135A

5.3.4.1.1. Buffer systems

The protein solutions in Tris and MES buffer were both set up for crystallisation trials as described in Material and Methods. For the protein solutions either three (Tris) or two (MES) different protein concentrations were used in 8 different commercially available screening kits. Each of these kits tests 96 different buffer solutions. They were used at two different protein to precipitant ratios (1:1 and 2:1) and at 277K and 296K, respectively. This makes a total of more than 7500 conditions tested.

Unfortunately, no crystal or crystalline material was found in any of the tested conditions.

5.3.4.1.2 Surface engineering

Molecular flexibility has a negative effect on the formation of a highly ordered crystal lattice. Such flexibility arises from motions of different domains of the polypeptide chain or from small movement due to solvent exposed, mobile amino acid side chains (Walter et al, 2006).

One possibility to eliminate this latter problem is to selectively substitute these residues with smaller, hydrophobic ones (Lawson et al., 1991). Such surface entropy reduction (SER) mutagenesis is an established technique. Still, as it requires the modification of expression plasmids, it remains time-intensive and laborious.

As an alternative, it is possible to reduce the surface entropy by chemically modifying the proteins surface (Rayment et al., 1993). The most frequently used technique is the reductive methylation of free amino groups. Hereby, primary amines, like the N-terminus and lysine residues, are converted into tertiary amines. In this way, repulsive charges of exposed side chains are reduced.

Lysine Methylation

Lysine and glutamic acid residues on the surface of proteins have long side chains that are mobile. Addition of methyl groups immobilises them and might therefore improve protein crystallisability.

NS2B/3[pro] S135A was expressed, purified and its lysines methylated as described in Materials and Methods except that the final elution step was done with buffer containing 50 mM Tris/HCl, pH 7.5, 100 mM NaCl and 5% glycerol. Fractions were analysed via SDS-PAGE, protease containing fractions were pooled and concentrated to 5.25 and 10.5 mg/ml. Finally, the sample was centrifuged at 14000 rpm, 4°C for 45 minutes, to remove precipitate. The supernatant was highly pure as confirmed by SDS-PAGE (data not shown) and was thus used for crystallisation trials.

Even though more than 1000 screening conditions were tested with above mentioned screening kits, again no crystal growth could be observed.

Surface Entropy Reduction mutagenesis

A study done by Cooper et al., 2007 strongly supports the notion that substitution of high-entropy surface residues with ones that favour intermolecular contacts might facilitate the production of high quality crystals for structural analysis.

To find appropriate residues at the surface of the NS2B/3 protein, the online program UCLA MBI — SERp Server was used (http://nihserver.mbi.ucla.edu/SER/). Good candidates for mutation typically have clusters scores of 3.0 and above.

Table 1: Results of the SERp online tool that identifies sites that are most suitable for mutation designed to enhance crystallisability by a Surface Entropy Reduction approach.

Cluster #1:   **Residues 47 - 49:** EKEE

☐ K 47 => A

☐ E 48 => A   SERp Score: 7.39

☐ E 49 => A

Cluster #2:   **Residues 148 - 150:** EEKWKGETVQVHA

☐ E 145 => A

☐ E 146 => A   SERp Score: 7.12

☐ K 147 => A

Cluster #3:   **Residues 230 - 233:** AQGEAEK

☐ E 227 => A

☐ E 229 => A   SERp Score: 6.87

☐ K 230 => A

As suggested by the program, residues K47, E48 and E 49 were mutated to alanine. Expression and purification were done as described previously and yielded again high amounts of pure, stable protein which was used for crystallisation assays.

Although more than thousand conditions for the protein with reduced surface entropy were tested, again no crystal or crystalline material was found.

5.3.4.2. NS2B/3^{pro} S135A with a large-affinity tag (maltose binding protein, MBP)

As already more than 10.000 conditions had been tested, we changed the crystallisation strategy and made some major changes to the protein.

In order to improve expression, solubility and folding of a protein, or to facilitate protein purification via affinity chromatography, it is a well-known technique to bind the protein of interest to a large-affinity tag, such as maltose-binding protein (MBP) or glutathione-S-transferase (GST) (Beckwith et al., 2000). As a side effect, these fusion tags often generate a conformational diversity, which might be detrimental to crystal growth. Therefore, large fusion tags are often removed prior to crystallisation by a potentially problematic cleavage step (Smyth et al., 2003).

However, it is possible to avoid such cleavages and keep the tag for the crystallisation trials. This might have, beside the above mentioned disadvantages, several advantages: Such fusion proteins can generate scaffolds which stabilise the protein of interest, thereby favouring crystallisation. In this way, the 3D structure of the human T cell leukaemia virus I gp21 ectodomain (Kobe et al., 1999), the MATa2 homeodomain (Ke et al., 2003) and the SarR protein (Liu et al., 2001) could be solved. In these structures, MBP is responsible for most of the crystal contacts and consequently for the formation of the crystal (Derewenda, 2004).

5.3.4.2.1. Sample preparation and crystallisation

Plasmid pet15b-MBP-NS2B/3 was generated as described in detail in Material and Methods. The initial expression protocol had to be changed, as the MPB-NS2B/3^{pro} S135A was predominately found in the insoluble fraction when expressed in *E.coli* BL21 DE3 pLysS (data not shown). The protein was soluble when we expressed it in *E.coli* BL21 DE3 and in the absence of chloramphenicol. Thus, this bacterial strain was used for further expression procedures. The purification was done as described before. In an additional step, the N-terminal His$_6$ tag was removed by thrombin cleavage, followed by another SEC. Finally, the protein samples were concentrated to 25, 50 and 66 mg/ml, respectively. All proteins solutions were then supplemented with an equimolar amount of maltose to stabilise the MBP. After centrifugation to remove precipitation, all samples were analysed via SDS-PAGE (Fig. 8). The

fusion protein has a calculated molecular weight of 68 kDa. Although very minor degradation of the fusion protein can be detected when 0.5 µl were applied, all samples were used for crystallisation trials at 277K and 296K. Again, several commercially available precipitant kits were used, including the Crystal suite, Wizard I and II.
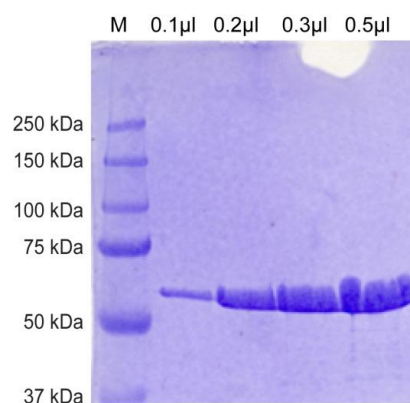
Figure 8. Analysis of concentrated samples of MBP-NS2B/3^{pro} S135A. 0.1, 0.2, 0.3 and 0.5 µl of the final sample (66 mg/ml) were analysed via SDS-PAGE.

Here, potentially crystalline material was found between day 17 and 31 in the Crystal Screen (Hampton research) in condition number 32 (2 M ammonium sulphate), at a precipitate to protein ratio of 1:2 (i.e. 0.666 M ammonium sulphate). In contrast, at a precipitate to protein ratio of 1:1 (i.e. 1 M ammonium sulphate), oils were found already 12 hours after the drops were set (Figure 9).
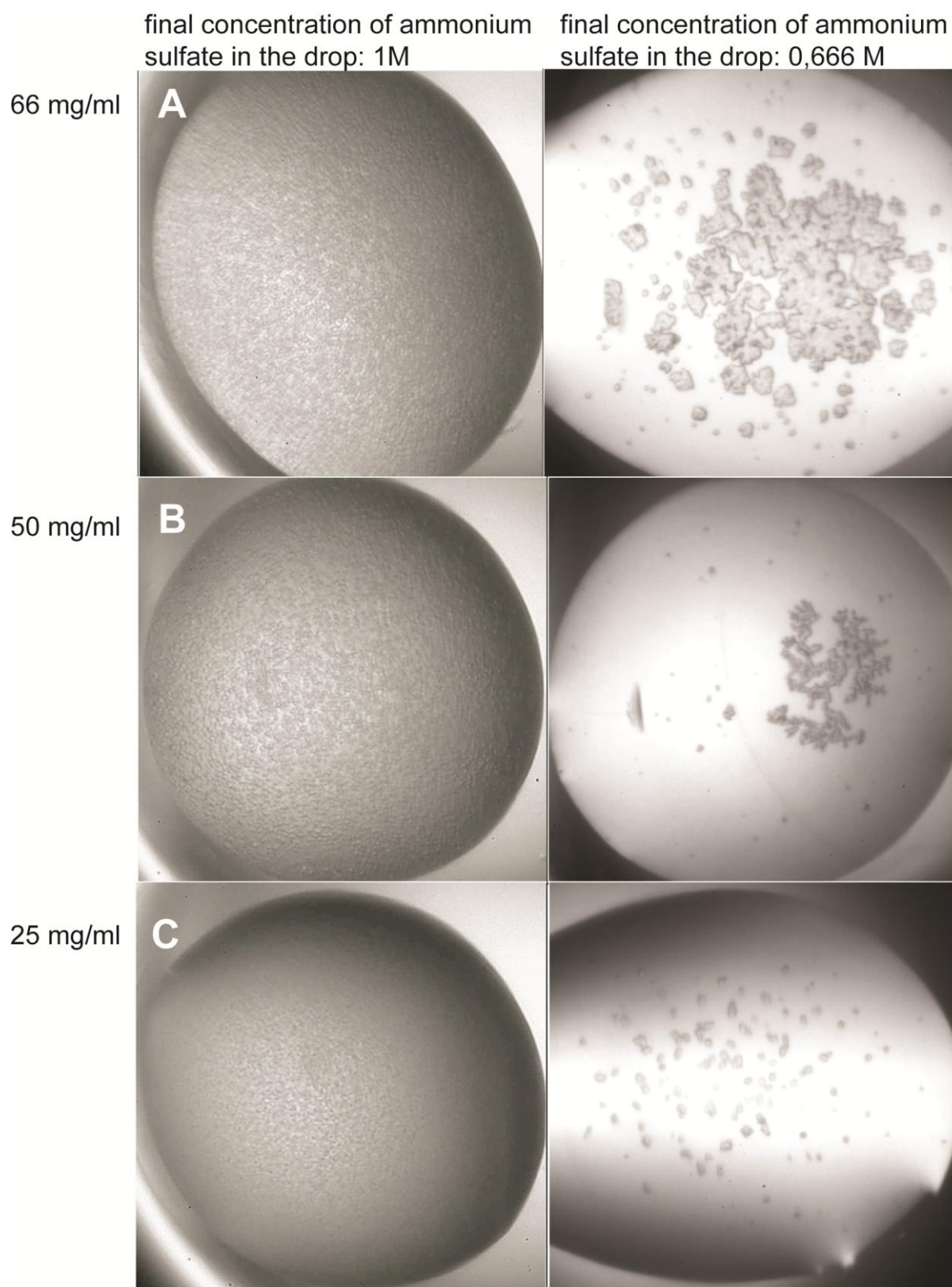
Figure 9. **Oils (left panel) and crystals (right panel) of the MBP-NS2B/3pro complex.** Protein solutions at a concentration (stated on the left) of 66 mg/ml (A), 50 mg/ml (B) and 25 mg/ml (C) were mixed with precipitant solution (2 M ammonium sulphate) either 1:1 (i.e. final conc. ammonium sulphate: 1M; left panel) or 2:1 (i.e. final conc. ammonium sulphate: 666 mM, right panel). Crystals were formed to an extent proportional to the protein concentration (compare right panels A with B and C).

5.3.4.2.2. Characterisation of the crystals – is it salt or protein?

At a concentration of 666 mM ammonium sulphate (Fig. 9, right panel), the extent of crystalline material varied in a relation directly proportional to the amount of protein used. This indicates that these are indeed protein crystals. However, as salt crystals form very easily in protein drops, further characterisation of the crystals was needed.

Proteins show epifluorescence at 340 nm when exposed to UV light. Thus, we investigated the crystals with a microscope equipped with a UV light source. The crystal showed epifluorescence (data not shown), further confirming the protein character of the crystals.

As the crystals were too fragile to be mounted, a further characterisation via SDS-PAGE or western blot could not be performed. Thus, it is possible that the crystals are composed of degraded protein instead of full-length MBP-NS2B/3^{pro} S135A.

However, several facts speak against this possibility. First, the MBP as a single protein crystallised in the presence of polyethylenglycol but not in 2M ammonium sulphate (Souza et al., 2008), the condition in which we observed crystals. Second, the MBP was stable during prior crystallisation experiments (Quiocho et al., 1997; Souza et al., 2008). The same is true for the NS2B/3^{pro} S135A, which did not show any degradation during prior stability experiments (data not shown) and did also not crystallise in 2 M ammonium sulphate. Consequently, we can exclude that what we found are crystals of the single proteins of MBP or NS2B/3^{pro} S135A, nor of any of their degradation products. Thus, we conclude that we finally succeeded in crystallising the full length MBP-NS2B/3 S135A.

5.3.4.2.3. Optimisation of crystal growth

Unfortunately, the optimal conditions for nucleation of crystals do not favour their subsequent growth (Bergfors et al., 2003). Spontaneous nucleation needs an elevated degree of supersaturation whereas the growth of diffraction quality crystals occurs at lower levels. Thus, the ideal experiment separates nucleation from growth, thereby fulfilling the different requirements of both processes. This can be achieved by seeding previously nucleated crystals into new drops, preequilibrated at lower levels of supersaturation.

Thus, to improve the size and quality of the crystals, streak seeding was performed as described in Materials and Methods. Again, crystals could be observed at day 19, indicating that crystal growth is reproducible (data not shown). Nevertheless, with this approach, the size and quality of the crystals could not be improved. Thus, further optimisation is needed.

5.4. Conclusion

Crystallisation properties of proteins differ greatly and are difficult to predict. Some proteins crystallise in numerous conditions, even though they are only rudimentarily purified. Crystallisability of others is very limited and restricted to a few conditions even after intensive purification. For the latter, it may prove difficult to identify initial hits. This was the case with the NS2B/3^{pro}. Although SDS-PAGE, DLS and Thermofluor results confirmed the purity and correct folding of the protein, no crystalline material could be detected in more than 10.000 conditions tested. Neither surface entropy reduction by lysine methylation nor mutagenesis nor the use of several different buffer systems, temperatures and protein concentrations led to any success. One possible reason might be that NS2B and NS3 are connected by a flexible linker, which could be detrimental to crystallisation. This theory is supported by the fact that fusing the protease to the large MBP finally led to the growth of crystals. MBP is known to be responsible for most of the crystal contacts in many of the successfully crystallised fusion proteins (Derewenda et al., 2004), and might therefore minimise the inhibitory effect of the linker.

As a result, after more than 20,000 conditions tested, we succeeded in reproducibly growing protein crystals of the fusion protein MBP-NS2B/3^{pro} S135A. Although the crystals are of insufficient quality and size, these data are the basis for further optimisation and an important step toward structure determination of the functional NS2B/3^{pro} S135A complex.

Once suitable protein crystals are available, X-ray diffraction data can be collected. Each structure determination by X-ray crystallography includes the calculation of the phases of the diffraction data. As the structure of the MBP has already been determined, molecular replacement can be directly applied to the native dataset (Smyth et al., 2003). Therefore, the crystallographic phase problem can be solved by using the MBP as a search model. This should facilitate and accelerate future structure determination.

Another possibility to determine the 3D structure of a protein at atomic resolution is nuclear magnetic resonance (NMR) spectroscopy where the data is recorded while the protein is in solution. Considering that many proteins perform their physiological functions in their soluble form, knowledge of the molecular structures in solution is highly relevant. Nevertheless, NMR spectroscopy is restricted to proteins of relatively low molecular weight (< 30 kDa). Furthermore, results are less clear than those obtained with X-ray crystallography.

As the NS2B/3^{pro} S135A has a molecular weight of 28 kDa, it is a possible candidate for structure determination via NMR spectroscopy. First, we had to examine whether the protease can be expressed in bacteria that are grown in M9 minimal medium, containing all the salts and trace elements needed by the bacteria, but with glucose and ammonium chloride as the only carbon and nitrogen source, respectively. These elements can be isotopically labelled; they become incorporated into the protein during expression. We tried to express the NS2B/3^{pro}

S135A in minimal medium and succeeded (data not shown), thus fulfilling the first prerequisite for future NMR spectroscopy analysis.

As NMR data could further increase the knowledge generated from X-ray crystallography data, determining the structure of the NS2B/3^{pro} S135A by NMR spectroscopy might be another goal worth pursuing.

6. GLOBAL DISCUSSION

Proteolysis by viral enzymes plays a fundamental role in the life cycle of many viruses (Krausslich and Wimmer 1988; Hellen et al., 1989); this makes them an interesting target for the antiviral therapy. Several protease inhibitors have already been developed; for instance, inhibitors that target the HIV protease are the basis of the anti-retroviral therapy (Piacenti et al., 2006). Infections against which specific antiviral therapeutics are needed include those of the hepatitis C virus (Lamarre et al., 2003), West Nile virus as well as tick-borne encephalitis virus (TBEV) (Mansfield et al., 2009). The same is true for animal pathogens, such as the foot and mouth disease virus (FMDV), which infects cloven-hoofed animals and causes considerable economic losses in livestock production (Grubman and Baxt, 2004).

Although extensive knowledge on proteolysis requirements has been generated throughout the last decades, viral proteolysis is still a cryptic and highly complex process. Cleavage sites are often insufficiently defined, influenced by neighbouring residues or by their structural context. This makes cleavage efficiencies unpredictable. Furthermore, the development of resistance to inhibitors due to the high mutation rate of the viral RNA polymerase is a constant problem. To overcome these obstacles, detailed information on substrate specificity and resistance to inhibitors is needed. Therefore, novel systems are required that allow the generation of a large number of escape mutants in a random manner and to analyse them systematically.

Thus, the overall objective of this work was to develop sensitive systems that allow us to examine the properties of various viral proteases in general and of the TBEV protease in particular. We used the TBEV itself as such a tool for several reasons. Firstly, due to its error prone RNA polymerase, TBEV produces a large number of mutants with approximately one error per synthesised genome. From this pool of mutants, infectious TBEV particles or single round infectious VLPs can be selected over the non-infectious ones by simply measuring their infectivity. Thirdly, by investigating the infectivity, we determine a parameter which is directly relevant for viral pathogenesis. The presence of suboptimal substrates or protease inhibitors should provoke the generation of TBEV escape mutants. Examination of the nature of the mutations will give important information about the rate at which resistance occurs and shed light onto the responsible mechanisms. In this way, information about the qualification of an inhibitor in clinical applications is generated.

To make this system usable for other viral proteases, it is necessary to inactivate NS2B/3^{pro} cleavage at the C-prM junction by introducing a cleavage site from another protease, thereby preventing the production of viral particles. This then allows the reactivation of the TBEV through the expression of a corresponding heterologous protease. In this way, the described above system can be used to examine the properties of the heterologous protease.

Furthermore, it allows the analysis of the consequences of the altered processing for C and prM properties and their relevance for the TBEV life cycle.

As described in detail in manuscript 1 (Schrauf and Kurz et al., 2012), we modified the genome of TBEV in such a way that the production of infectious virions became dependent on the activity of the FMDV 3C protease (3C^{pro}) instead of the TBEV protease (NS2B/3^{pro}). Virion production was achieved by providing an optimised FMDV 3C^{pro} cleavage site at the C-prM junction and the protease itself either *in cis* or *in trans* from a TBEV replicon, with the *trans* complementation system being more efficient and stable. In the monocystronic system, viruses regained a viable NS2B/3 cleavage site in the absence of an active 3C^{pro}. Analysis of the revertants revealed a previously undetected role of the C-terminus of protein C and the necessity of positive charges for efficient budding of the NC. Finally, the results suggest that the re-establishment of a NS2B/3^{pro} cleavage in the same region is beneficial for viral spread over a co-translational separation of C and prM, as described by Schrauf et al., 2009. This underscores the importance of the concerted progression of NS2B/3^{pro} and SPase1 cleavage, presumably because in this way, protein C is kept in close proximity to the ER membrane, thereby promoting efficient budding of the nucleocapsid.

Taken together, this assay was used to investigate principal molecular mechanisms of the TBEV itself as well as characteristics of the NS2B/3 and FMDV 3C protease cleavage. Furthermore, it provides the basis for future inhibitor studies. The conceptional basis of this system, i.e. the use the inherent replicative infidelity of the viral polymerase to introduce random mutations into a heterologous gene of interest, is, to our knowledge, unique, and should find broad application in various fields of virology. Finally, by using the replicon system, problems usually associated with infectious material, are avoided. This simplifies handling and manipulation of all pathogens; however, it is of special interest for BSL-3 and BSL-4 viruses.

To generate mutants of TBEV with cleavage sites for heterologous proteases but not the NS2B/3^{pro}, it is important to have a profound understanding of the cleavage mechanism and substrate specificity of the NS2B/3^{pro}. Therefore, we established a biochemical assay to investigate *in vitro* the cleavage of the NS2B/3^{pro}. Furthermore, this assay allows us to test the suitability of diverse cleavage sites to be processed by various heterologous proteases.

We expressed both active and inactive variants of NS2B/3^{pro} of TBEV as a His-tagged protein in *E.coli*. The NS2B/3 fusion proteins were expressed in a soluble form in bacteria and could be purified to homogeneity. As substrate, we expressed a fragment of TBEV coding for the C and prM region in rabbit reticulocyte lysates. Specific cleavage at the C-prM junction required the presence of microsomal membranes, which in turn required a lengthening of the transmembrane segment at the C-prM junction to inhibit incorrect processing by the SPaseI. Substitution of the NS2B/3^{pro} cleavage motif at the C-prM junction by one of the HIV-1^{pro} allowed us to examine cleavage by the HIV-1^{pro} and revealed the presence of a potential

second HIV-1^{pro} cleavage site in the prM coding sequence. This second cleavage was observed in the absence and presence of membranes, indicating that the protein was inserted into the membrane inside-out. We could pin this misorientation down to the presence of the HIV-1^{pro} cleavage site in the C-terminal region of protein C. Furthermore, we conclude that this misorientation is the cause for our failure to generate a TBEV dependent on cleavage of the HIV-1^{pro} as we report in manuscript 1 for the FMDV 3C^{pro} (Schrauf and Kurz et al., 2012). However, the biochemical assay will allow us to modify the cleavage site of HIV-1^{pro} in order to prevent the cryptic cleavage as well as incorrect membrane insertion of the protein. As a result, the assay should promote the construction of a TBEV that is dependent on the activity of the HIV-1^{pro}.

As proteolytic cleavages are highly dependent on the structural context of their substrate, especially when membrane proteins are involved, it is important to examine their characteristics in a natural environment. However, biochemical assays provide some advantages over cell based ones. First, they present a fast and safe method to investigate characteristics of proteases without the need of special laboratory equipment or infectious material. Secondly, in contrast to cell based assays, they offer the opportunity to study isolated mechanisms without being influenced by associated processes; i.e. the system described in detail in manuscript 2 (Kurz et al., 2012, submitted), may provide a model for examining the orientation of the transmembrane anchor of the TBEV protein C as well as the requirements for correct SPaseI processing. Furthermore, due to the inherent glycosylation activity of the ribosomal membranes, the influence of glycosylation of protein M on NS2B/3^{pro} processing can be studied. In this way, this assay is a valuable tool for itself. However, it is also an essential supplement for the above described cell based assay.

Finally, in order to promote the design of an inhibitor for the TBEV NS2B/3^{pro}, we attempted to solve its 3D structure by X-ray crystallography. Structures of the proteases of other flaviviruses (e.g. WNV and DENV) are already available. However, our results (discussed in manuscript 2) indicate a fundamental difference in the processing of TBEV and WNV polyprotein, as Yamshchikov and Compans, 1994, who examined the cleavage of WNV C-prM by WNV NS2B/3^{pro}, did not observe SPaseI cleavage when C-prM was translated in the presence of microsomal membranes and when NS2B/3^{pro} was excluded from the reaction. This is in contrast to our results and underlines, together with a sequence identity of less than 50 % of the two viral proteases, the need to solve the structure of the TBEV NS2B/3^{pro}.

Therefore, we attempted to obtain crystals with various preparations of the inactive NS2B/3^{pro} by using the sitting drop method. After intensive screening at 4°C and 20°C of a huge variety of conditions, we did not succeed in obtaining crystals of the protease. In order to minimise repulsive charges of the protein, thereby enhancing the chance of getting crystals, lysine methylation as well as surface entropy reduction mutagenesis was performed. Both

attempts were not successful. Only when we fused the NS2B/3 coding sequence to that of the maltose binding protein (MBP), thereby minimising a potential inhibitory impact of the flexible linker, did we obtain crystals in one of the conditions tested. Crystallisation occurred in a protein concentration dependent manner. Crystals were of suboptimal quality and size; however, they could be grown reproducibly. Thus, this work is the basis for further crystal optimisation and an important step toward structure determination.

Taken together, the results of this thesis extend the knowledge and understanding of basic molecular mechanisms of the TBEV and further clarify cleavage requirements of the proteases of TBEV, FMDV and HIV. The concept of the established assays allows the study of various aspects of viral pathogenesis and development of resistance to inhibitors. Therefore, this work provides the basis for analysing structure function relationships of viral proteases and should therefore facilitate future drug design.

7. REFERENCES

Aleshin, A. E., S. A. Shiryaev, A. Y. Strongin, and R. C. Liddington. 2007. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. Protein Sci 16:795-806.

Allison, S. L., J. Schalich, K. Stiasny, C. W. Mandl, C. Kunz, and F. X. Heinz. 1995. Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. J Virol 69:695-700.

Alvarez, D. E., M. F. Lodeiro, S. J. Luduena, L. I. Pietrasanta, and A. V. Gamarnik. 2005. Long-range RNA-RNA interactions circularize the dengue virus genome. J Virol 79:6631-43.

Amberg, S. M., A. Nestorowicz, D. W. McCourt, and C. M. Rice. 1994. NS2B-3 protease-mediated processing in the yellow fever virus structural region: in vitro and in vivo studies. J Virol 68:3794-802.

Amberg, S. M., and C. M. Rice. 1999. Mutagenesis of the NS2B-NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. J Virol 73:8083-94.

Assenberg, R., E. Mastrangelo, T. S. Walter, A. Verma, M. Milani, R. J. Owens, D. I. Stuart, J. M. Grimes, and E. J. Mancini. 2009. Crystal structure of a novel conformational state of the Flavivirus NS3 protein: Implications for Polyprotein Processing and Viral Replication. J Virol.

Bedard, K. M., and B. L. Semler. 2004. Regulation of picornavirus gene expression. Microbes Infect 6:702-13.

Belsham, G. J., G. M. McInerney, and N. Ross-Smith. 2000. Foot-and mouth disease virus 3C protease induces cleavage of translation initiation factors eIF4A and eIF4G within infected cells. J Virol 74:272-80.

Birtley, J. R., and S. Curry. 2005. Crystallization of foot-and-mouth disease virus 3C protease: surface mutagenesis and a novel crystal-optimization strategy. Acta Crystallogr D Biol Crystallogr 61:646-50.

Birtley, J. R., S. R. Knox, A. M. Jaulent, P. Brick, R. J. Leatherbarrow, and S. Curry. 2005. Crystal structure of foot-and-mouth disease virus 3C protease. New insights into catalytic mechanism and cleavage specificity. J Biol Chem280:11520-7.

Calisher, C. H., N. Karabatsos, J. M. Dalrymple, R. E. Shope, J. S. Porterfield, E. G. Westaway, and W. E. Brandt. 1989. Antigenic relationships between flaviviruses as determined by cross-neutralization tests with polyclonal antisera. J Gen Virol 70 (Pt 1):37-43.

Chambers, T. J., C. S. Hahn, R. Galler, and C. M. Rice. 1990. Flavivirus genome organization, expression, and replication. Annu Rev Microbiol 44:649-88.

Chambers, T. J., A. Nestorowicz, S. M. Amberg, and C. M. Rice. 1993. Mutagenesis of the yellow fever virus NS2B protein: effects on proteolytic processing, NS2B-NS3 complex formation, and viral replication. J Virol 67:6797-807.

Chambers, T. J., A. Nestorowicz, and C. M. Rice. 1995. Mutagenesis of the yellow fever virus NS2B/3 cleavage site: determinants of cleavage site specificity and effects on polyprotein processing and viral replication. J. Virol. 69:1600–1605.

Chappell, K. J., M. J. Stoermer, D. P. Fairlie, and P. R. Young. 2007. Generation and characterization of proteolytically active and highly stable truncated and full-length recombinant West Nile virus NS3. Protein Expr Purif 53:87-96.

Charrel RN, Attoui H, Butenko AM et al. (2004): Tick-borne virus diseases of human interest in Europe. Clin. Microbiol. Infect. 10(12), 1040-1055.

Chu, J. J., and M. L. Ng. 2004. Infectious entry of West Nile virus occurs through a clathrin-mediated endocytic pathway. J Virol 78:10543-55.

Clum, S., K. E. Ebner, and R. Padmanabhan. 1997. Cotranslational membrane insertion of the serine protease precursor NS2B-NS3(Pro) of dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. J Biol Chem 272:30715.

Cooper D.R., T. Boczek, .K. Grelewska et al. (2007) Protein crystallization by surface entropy reduction: optimization of the SER strategy. Acta Crystallographica Section D Biological Crystallography ISSN 0907-4449.

Curry, S., N. Roque-Rosell, P. A. Zunszain, and R. J. Leatherbarrow. 2007. Foot-and-mouth disease virus 3C protease: recent structural and functional insights into an antiviral target. Int J Biochem Cell Biol 39:1-6.

D'Arcy, A., Chaillet, M., Schiering, N., Villard, F., Lim, S. P., Lefeuvre, P. and Erbel, P. (2006). Purification and crystallization of dengue and West Nile virus NS2B-NS3 complexes. Acta Cryst. Sec. F 62, 157-162.

Debouck, C. 1992. The HIV-1 protease as a therapeutic target for AIDS. AIDS Res Hum Retroviruses 8:153-64.

Derewenda ZS. 2004. The use of recombinant methods and molecular engineering in protein crystallization. Methods. 2004 Nov; 34(3):354-63.

Dokland, T., M. Walsh, J. M. Mackenzie, A. A. Khromykh, K. H. Ee, and S. Wang. 2004. West Nile virus core protein; tetramer structure and ribbon formation. Structure 12:1157-63.

Ecker, M., Allison, S.L., Meixner, T. and Heinz, F.X. (1999) Sequence analysis and genetic classification of tick-borne encephalitis viruses from Europe and Asia. J Gen Virol, 80 (Pt 1), 179-185.

Erbel, P., N. Schiering, A. D'Arcy, M. Renatus, M. Kroemer, S. P. Lim, Z. Yin, T. H. Keller, S. G. Vasudevan, and U. Hommel. 2006. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nat Struct Mol Biol 13:372-3.

Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P 2006.. Thermofluor-based high- throughput stability optimization of proteins for structural studies. Anal Biochem. 357:289-98.

Frankel, A. D., and J. A. Yong. 1998. HIV-1: fifteen proteins and a RNA. Annu. Rev. Biochem. 67:1-25.

Germi, R.; Crance, J.M.; Garin, D.; Guimet, J.; Lortat-Jacob, H.; Ruigrok, R.W; Zarski, J.P. & Drouet, E. (2002). Heparan sulfate-mediated binding of infectious dengue virus type 2 and yellow fever virus. Virology, Vol.292, pp. 162-168.

Gritsun, T. S., K. Venugopal, P. M. Zanotto et al. 1997. Complete sequence of two tick-borne flaviviruses isolated from Siberia and the UK: analysis and significance of the 5' and 3'-UTRs. Virus Res 49:27-39.

Gritsun, T. S., Lashkevich, V. A. & Gould, E. A. (2003). Tick-borne encephalitis. Antiviral Res 57, 129–146.

Grubman MJ and Baxt B, 2004. Foot-and-mouth disease. Clin. Microbiol. Rev. April 2004 vol. 17 no. 2 465-493

Gubler, D. J., G. Kuno, and R. Markoff. 2007. Flaviviruses, p. 1154-1252. In D. M. Knipe and P. M. Howley (ed.), Fields Virology, 5th ed. Lippincott Williams & Wilkins, Philadelphia.

Hedstrom, L. 2002. Serine protease mechanism and specificity. Chem Rev 102:4501-24.

Heinz, F. X., and S. L. Allison. 2003. Flavivirus structure and membrane fusion. Adv Virus Res 59:63-97.

Heinz, F.X., Holzmann, H., Essl, A. and Kundi, M. (2007) Field effectiveness of vaccination against tick-borne encephalitis. Vaccine, 25, 7559-7567.

Hellen, C.U. Krausslich H. G., and E. Wimmer. 1989. Proteolytic processing of polyproteins in the replication of RNA viruses. Biochemistry 28, 9881-9890

Ishima, R., Ghirlando, R., Tözsér, J., Gronenborn, A. M., Torchia, D. A. & Louis, J. M. (2001). Folded Monomer of HIV-1 protease. J. Biol. Chem. 276, 49110-49116.

Jones, C. T., L. Ma, J. W. Burgner, T. D. Groesch, C. B. Post, and R. J. Kuhn. 2003. Flavivirus capsid is a dimeric alpha-helical protein. J. Virol. 77:7143–7149.

Ke A. and C. Wolberger (2003), "Insights into binding cooperativity of ATa1/MATalpha2 from the crystal structure of a MATa1 homeodomain-maltose binding protein chimera," protein Science 12, 306-312

Khromykh, A. A., H. Meka, K. J. Guyatt, and E. G. Westaway. 2001. Essential role of cyclization sequences in flavivirus RNA replication. J Virol 75:6719-28.

Kobe B., R.J. Center, B.E. Kemp, P. Poumbourios. (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. Proc. Natl.Acad. Sci. USA, 4319–4324.

Kofler, R. M., F. X. Heinz, and C. W. Mandl. 2002. Capsid protein C of tick-borne encephalitis virus tolerates large internal deletions and is a favourable target for attenuation of virulence. J Virol 76:3534-43.

Kofler, R. M., V. M. Hoenninger, C. Thurner, and C. W. Mandl. 2006. Functional analysis of the tick-borne encephalitis virus cyclization elements indicates major differences between mosquito-borne and tick-borne flaviviruses. J Virol 80**:**4099-113.

Krausslich H.G. and E. Wimmer. 1988. Viral proteinases. Annu Rev Biochem 57, 701 ´754

Kroschewski, H., S. L. Allison, F. X. Heinz, and C. W. Mandl. 2003. Role of heparan sulfate for attachment and entry of tick-borne encephalitis virus. Virology 308:92-100.

Kuno, G., G. J. Chang, K. R. Tsuchiya, N. Karabatsos, and C. B. Cropp. 1998. Phylogeny of the genus Flavivirus. J Virol 72:73-83.

Kunz, C., H. Hoffmann, and H. Dippe. 1991. Early summer meningoencephalitis vaccination, a preventive medicine measure with high acceptance in Austria. Wien. Med Wochenschr 141**:**273-276.

Lamarre D, Paul C. Anderson, Murray Bailey et al. 2003. An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. Nature 426, 186-189.

Lapatto, R., T. Blundell, A. Hemmings, et al. 1989. X-ray analysis of HIV-1 protease at 2.7 A resolution confirms structural homology among retroviral enzymes. Nature 342**:**299-302.

Lawson DM, Artymiuk PK, Yewdall SJ, Smith JMA, Livingstone JC, Treffy A, Luzzago A, Levi S, Arosio P, Cesareri G, Thomas CD, Shaw WV, Harrison PM. 1991. Solving the structure of human H ferritin by genetically engineered intermolecular crystal contacts. Nature 349541-544.

Lee, E., C. E. Stocks, S. M. Amberg, C. M. Rice, and M. Lobigs. 2000. Mutagenesis of the signal sequence of yellow fever virus prM protein: enhancement of signalase cleavage In vitro is lethal for virus production. J Virol 74:24-32.

Leung, D.; Schroder, K.; White, H.; Fang, N.X.; Stoermer, M.J.; Abbenante, G.; Martin, J.L.;Young, P.R. & Fairlie, D.P. (2001). Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. Journal of Biology Chemistry, Vol.276, pp. 45762–45771

Li, W., N. Ross-Smith, C. G. Proud, and G. J. Belsham. 2001. Cleavage of translation initiation factor 4AI (eIF4AI) but not eIF4AII by foot-and-mouth disease virus 3C protease: identification of the eIF4AI cleavage site. FEBS Lett 507**:**1-5.

Lin, C., S. M. Amberg, T. J. Chambers, and C. M. Rice. 1993. Cleavage at a novel site in the NS4A region by the yellow fever virus NS2B-3 protease is a prerequisite for processing at the downstream 4A/4B signalase site. J Virol 67**:**2327-35.

Lin, C., T. J. Chambers, and C. M. Rice. 1993. Mutagenesis of conserved residues at the yellow fever virus 3/4A and 4B/5 dibasic cleavage sites: effects on cleavage efficiency and polyprotein processing. Virology 192**:**596-604.

Lin, C. W., H. D. Huang, S. Y. Shiu, W. J. Chen, M. H. Tsai, S. H. Huang, L. Wan, and Y. J. Lin. 2007. Functional determinants of NS2B for activation of Japanese encephalitis virus NS3 protease. Virus Res 127**:**88-94.

Lindenbach, B. D., H. J. Thiel, and C. M. Rice. 2007. Flaviviridae: the viruses and their replication, p. 1101-1152. In D. M. Knipe and P. M. Howley (ed.), Fields Virology 5th ed. Lippincott Williams & Wilkins, Philadelphia.

Liu, Y., A. Manna, R. Li, W. E. Martin, R. C. Murphy, A. L. Cheung, and G. Zhang. 2001. Crystal structure of the SarR protein from Staphylococcus aureus. Proc. Natl. Acad. Sci. USA 98:6877-6882.

Lo MC, Aulabaugh A, Jin G, Cowling R, Bard J, Malamas M & Ellestad G. (2004) Evaluation of fluorescence-based thermal shift assays for hit identification in drug discovery. Anal Biochem 332, 153–159.

Lobigs, M. 1992. Proteolytic processing of a Murray Valley encephalitis virus non-structural polyprotein segment containing the viral protease: accumulation of a NS3-4A precursor which requires mature NS3 for efficient processing. J. Gen. Virol. 73 (Pt. 9), 2305– 2312.

Lobigs, M. 1993. Flavivirus premembrane protein cleavage and spike heterodimer secretion require the function of the viral protease NS3. Proc Natl Acad Sci U S A 90:6218-22.

Lobigs, M., and E. Lee. 2004. Inefficient signalase cleavage promotes efficient nucleocapsid incorporation into budding flavivirus membranes. J Virol 78:178-86.

Luo, D., T. Xu, C. Hunke, G. Gruber, S. G. Vasudevan, and J. Lescar. 2008. Crystal structure of the NS3 protease-helicase from dengue virus. J Virol 82:173-83.

Ma, L., C. T. Jones, T. D. Groesch, R. J. Kuhn, and C. B. Post. 2004. Solution structure of dengue virus capsid protein reveals another fold. Proc Natl Acad Sci U S A 101:3414-9.

Mandl, C. W., L. Iacono-Connors, G. Wallner, H. Holzmann, C. Kunz, and F. X. Heinz. 1991. Sequence of the genes encoding the structural proteins of the low-virulence tick-borne flaviviruses Langat TP21 and Yelantsev. Virology 185:891–895.

Mandl, C. W., H. Holzmann, C. Kunz, and F. X. Heinz. 1993. Complete genomic sequence of Powassan virus: evaluation of genetic elements in tick-borne versus mosquito-borne flaviviruses. Virology 194:173-84.

Mandl, C. W., H. Kroschewski, S. L. Allison, R. Kofler, H. Holzmann, T. Meixner, and F. X. Heinz. 2001. Adaptation of tick-borne encephalitis virus to BHK-21 cells results in the formation of multiple heparan sulfate binding sites in the envelope protein and attenuation in vivo. J Virol 75:5627-37.

Mandl, C. W. 2005. Steps of the tick-borne encephalitis virus replication cycle that affect neuropathogenesis. Virus Res 111:161-74.

Mansfield KL, N. Johnson, L. P. Phipps, J. R. Stephenson, A. R. Fooks and T. Solomon. 2009. Tick-borne encephalitis virus – a review of an emerging zoonosis. J Gen Virol vol. 90 no. 8 1781-1794

Marquet 1991

Markoff, L. 2003. 5'- and 3' non-coding regions in flavivirus RNA. Adv Virus Res 59:177-228.

Mason, P. W., M. J. Grubman, and B. Baxt. 2003. Molecular basis of pathogenesis of FMDV. Virus Res **91:**9-32.

Mukhopadhyay, S., R. J. Kuhn, and M. G. Rossmann. 2005. A structural perspective of the flavivirus life cycle. Nat Rev Microbiol 3:13-22.

Navia, M. A., P. M. Fitzgerald, B. M. McKeever, C. T. Leu, J. C. Heimbach, W. K. Herber, I. S. Sigal, P. L. Darke, and J. P. Springer. 1989. Threedimensional structure of aspartyl protease from human immunodeficiency virus HIV-1. Nature 337**:**615-20.

Nettleship JE, Brown J, Groves MR & Geerlof A (2008) Methods for protein characterization by mass spectrometry, thermal shift (ThermoFluor) assay, and multiangle or static light scattering. Methods Mol Biol 426, 299–318.

Nowak, T., P. M. Farber, G. Wengler, and G. Wengler. 1989. Analyses of the terminal sequences of West Nile virus structural proteins and of the in vitro translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. Virology 169:365-76.

N. Ozer, T. Haliloglu, C. A. Schiffer. 2006. "Substrate Specificity of HIV-1 Protease by a Biased Sequence Search Method", Proteins: Structure, Function and Bioinformatics, 64(2):444-456.

Pettit, S. C., J. Simsic, D. D. Loeb, L. Everitt, C. A. Hutchison, 3rd, and R. Swanstrom. 1991. Analysis of retroviral protease cleavage sites reveals two types of cleavage sites and the structural requirements of the P1 amino acid. J Biol Chem **266:**14539-47.

Piacenti FJ. (2006) An update and review of antiretroviral therapy. Pharmacotherapy 26:1111–1133.

Preugschat, F., Yao, C.W., Strauss, J.H. 1990. In vitro processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. J. Virol. 64 (9), 4364–4374.

Pugachev, K.V., Nomokonova, N.Y., Dobrikova, E., Wolf, Y.I., 1993. Site directed mutagenesis of the tick-borne encephalitis virus NS3 gene reveals the putative serine protease domain of the NS3 protein. FEBS Lett. 328 (1–2), 115– 118.

Quiocho FA, Spurlino JC, Rodseth LE. 1997) Extensive features of tight oligosaccharide binding revealed in high-resolution structures of the maltodextrin transport/chemosensory receptor. Structure. 5(8):997-1015.

Rice, C. M., and J. H. Strauss. 1990. Production of flavivirus polypeptides by proteolytic processing. Semin Virol 1, 357-367.

Rieger, M., M. Nübling, R. Kaiser et al. 1998. FSME Infektionen durch Rohmilch - welche Rolle spielt dieser in südwestdeutschen FSME Endemiegebieten. Gesundheitswesen **60**:348-356.

Roberts NA, Joseph A. Martin, Derek Kinchington et al. 1990. Rational Design of Peptide-Based HIV Proteinase Inhibitors. Science. New Series, Vol. 248, No. 4953 (Apr. 20, 1990), pp. 358-361

Robertson SJ, Mitzel DN, Taylor RT Best SM, Bloom ME. 2009. Tick-borne flaviviruses: dissecting host immune responses and virus countermeasures. Immunol Res; 43(1-3) :172-86.

Samsa, M. M., J. A. Mondotte, N. G. Iglesias, I. Assunção-Miranda, G. Barbosa-Lima, A. T. Da Poian, P. T. Bozza, and A. V. Gamarnik. 2009. Dengue Virus Capsid Protein Usurps Lipid Droplets for Viral Particle Formation. PLoS Pathog **5:**e1000632.

Schechter, I., and A. Berger. 1967. On the size of the active site in proteases. I. Papain. Biochem Biophys Res Commun **27:**157-62.

Schrauf, S., C. W. Mandl, L. Bell-Sakyi, and T. Skern. 2009. Extension of flavivirus protein C differentially affects early RNA synthesis and growth in mammalian and arthropod host cells. J Virol **83:**11201-10.

Schrauf, S., P. Schlick, T. Skern, and C. W. Mandl. 2008. Functional analysis of potential carboxy-terminal cleavage sites of tick-borne encephalitis virus capsid protein. J Virol **82:**2218-29.

Shiryaev, S.A., Aleshin, A.E., Ratnikov, B.I., Smith, J.W., Liddington, R.C., Strongin, A.Y. 2007. Expression and purification of a two-component flaviviral protease resistant to autocleavage at the NS2B-NS3 junction region. Protein Expr Purif 52(2), 334-339.

Smyth, D. R., Mrozkiewicz, M. K., McGrath, W. J., Listwan, P., and Kobe, B. (2003). Crystal structures of fusion proteins with large affinity tags. Protein Sci. 12: 1313-1322.

Sommadossi, J. P. 1999. HIV protease inhibitors: pharmacologic and metabolic distinctions. AIDS

Souza, C. S., Ferreira, L. C. S., Thomas, L., Barbosa, J. A. R. G.,Balan, A. Crystallization, data collection and data processing of maltose-binding protein (MalE) from the phytopathogen Xanthomonas axonopodis pv. citri (2009) Acta Crystallogr Sect F Struct Biol Cryst Commun. 2009 February 1; 65(Pt 2): 105–107. 2008

Spurolino JC, SgII, Guang-Ying LuSI, and Florante A. Quiocho S. 1991 The 2.3-A Resolution Structure of the Maltose- or Maltodextrinbinding Protein, A Primary Receptor of Bacterial Active Transport and Chemotaxis*J Biol Chem. 1991 Mar 15;266(8):5202-19.

Stadler, K., S. L. Allison, J. Schalich, and F. X. Heinz. 1997. Proteolytic activation of tick-borne encephalitis virus by furin. J Virol 71:8475-81.

Stiasny, K., and F. X. Heinz. 2006. Flavivirus membrane fusion. J Gen Virol 87**:**2755-66.

Stocks, C. E., and M. Lobigs. 1998. Signal peptidase cleavage at the flavivirus C-prM junction: dependence on the viral NS2B-3 protease for efficient processing requires determinants in C, the signal peptide, and prM. J Virol **72:**2141-9.

Süss J. 2003. Epidemiology and ecology of TBE relevant to the production of effective vaccines. Vaccine; 21(Suppl 1):S19–35.

Sweeney, T. R., N. Roque-Rosell, J. R. Birtley, R. J. Leatherbarrow, and S.Curry. 2007. Structural and mutagenic analysis of foot-and-mouth disease virus 3C protease reveals the role of the beta-ribbon in proteolysis. J Virol 81:115-24.

Vogt, V. M. 1996. Proteolytic processing and particle maturation. Curr. Top. Microbiol. Immunol. 214**:**95-131.

Wallner, G., C. W. Mandl, C. Kunz, and F. X. Heinz. 1995. The flavivirus 3'- noncoding region: extensive size heterogeneity independent of evolutionary relationships among strains of tick-borne encephalitis virus. Virology 213:169-78.

Walter TS, **Meier C**, **Assenberg F et al.,** 2006. Lysine methylation as a routine rescue strategy for protein crystallization. Structure 14/11, 1617-1622.

Wlodawer, A., M. Miller, M. Jaskolski, B. K. Sathyanarayana, E. Baldwin, I.T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. Kent. 1989.Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1protease. Science 245:616-21.

Wlodawer, A., and J. W. Erickson. 1993. Structure-based inhibitors of HIV-1 protease. Annu. Rev. Biochem. 62**:**543-585.

Yamshchikov VF, Compans RW. 1994. Processing of the intracellular form of the west Nile virus capsid protein by the viral NS2B-NS3 protease: an in vitro study. J Virol. 1994 Sep;68(9)

Zhan, Y., Song, X., and Zhou, G.W. 2001. Structural analysis of regulatory protein domains using GST-fusion proteins. Gene 281: 1–9.

Zhang, W., P. R. Chipman, J. Corver, P. R. Johnson, Y. Zhang, S.Mukhopadhyay, T. S. Baker, J. H. Strauss, M. G. Rossmann, and R. J.Kuhn. 2003. Visualization of membrane protein domains by cryo-electronmicroscopy of dengue virus. Nat Struct Biol 10**:**907-12.

8. CURRICULUM VITAE

Personal data

Name	Martina Kurz
Date of birth	March 31, 1975
Place of birth	Vienna, Austria
Citizenship	Austria

Education

Since 04/2007	**Doctorial thesis:** University of Vienna, Supervisor: Prof. T. Skern Department of Medical Biochemistry and Institute of Virology (Prof. F. X. Heinz) within the FWF project: *"Using viral replicative infidelity to test viral proteases"*
10/1997 - 11/2006	**Diploma study:** Food- and Biotechnology, focus on biotechnology University of Natural Resources and Applied Life Sciences (BOKU), Vienna. Graduation date: November 2006
	Diploma thesis: BOKU, Supervisor: Dr. A. Egorov Polymun Scientific Immunbiologische GmbH and Institute of Applied Microbiology (Prof. H. Katinger), Vienna *"Construction and evaluation of influenza A vectors expressing GFP and IL2 proteins"*
09/1995 - 10/1997	**Kolleg** Bau- und Umwelttechnik, HTL Mödling Graduation date/Matura: October 1997
09/1989 - 06/1994	**Handelsakademie**, Maygasse, 1130 Vienna Graduation Date/Matura: June 1994

Employment history

04/2007 – 08/2010	**Medical University,** Vienna (MUV) Doctorial thesis; focus on virology, biochemistry
11/1997 – 07/2006	**Siemens AG Austria**, (half time employment) Measurement and control technology, strategy, organisation

9. Poem of the day

<u>Mutations</u>
Mutations,
Mutations,
Mutations.   Carla

<u>Thinking</u>
I stink... when I think…   Carla

<u>Breadcrumbs</u>
They are nice.   Carla

<u>Breadcrumbs</u>
They are.   Tina

<u>Yoghurt</u>
It is gone (like the protein).   Tina

<u>Nobel Sorrow</u>
One week too soon.
One hour too late.   David

<u>34</u>
34… very young
Nevertheless, applying for a leader position…
Enviousness.   David

<u>Exhaustion</u>
Inoculation of
two colonies…   David and Carla

Communication…
It's difficult.   Carla