DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
212

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS
212

ALEKSEI LULLA

Alphaviral nonstructural protease and its polyprotein substrate: arrangements for the perfect marriage

TARTU UNIVERSITY PRESS

Institute of Technology, Faculty of Science and Technology, University of Tartu, Estonia.

Dissertation was accepted for the commencement of the degree of Doctor of Philosophy in virology on April 4, 2012 by the Scientific Council of the Institute of Molecular and Cell Biology, University of Tartu, Estonia.

Supervisor: Andres Merits, PhD, Professor of Applied Virology,
Institute of Technology, University of Tartu, Estonia

Opponent: Ilya Frolov, PhD, Professor,
Department of Microbiology and Immunology,
University of Alabama at Birmingham, USA

Commencement: Room No 217, Riia 23, Tartu, on May 18, 2012, at 12.00.

Publication of this thesis is granted by the Institute of Technology, University of Tartu and by the Graduate School of Biomedicine and Biotechnology created under auspices of European Social Fund.

ISSN 1024–6479
ISBN 978–9949–19–981–5 (trükis)
ISBN 978–9949–19–982–2 (PDF)

Tartu Ülikooli Kirjastus
www.tyk.ee
Tellimus nr 188

All truths are easy to understand once they are discovered;
the point is to discover them.
Galileo Galilei

TABLE OF CONTENTS

LIST OF ORIGINAL PUBLICATIONS

The current thesis is based on the following original publications, referred to in the text by their Roman numerals:

I. **Lulla, A.**, V. Lulla, K. Tints, T. Ahola, and A. Merits. 2006. Molecular determinants of substrate specificity for Semliki Forest virus nonstructural protease. Journal of Virology, 80:5413–22.

II. Tamberg, N., V. Lulla, R. Fragkoudis, **A. Lulla**, J.K. Fazakerley, and A. Merits. (2007). Insertion of EGFP into the replicase gene of Semliki Forest virus results in a novel, genetically stable marker virus. The Journal of General Virology, 88: 1225–30.

III. **Lulla, A.**, V. Lulla, and A. Merits. 2012. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. Journal of Virology, 86:553–65.

The articles listed above have been printed with the permission of the copyright owners.
Some unpublished data is also presented.

My personal contribution to the articles referred to in this thesis is as follows:

I. participated in experimental design, performed the experiments, analyzed the data and wrote manuscript.

II. participated in experimental design, performed the experiments and analyzed the data.

III. participated in experimental design, performed the experiments, analyzed the data and wrote manuscript.

LIST OF ABBREVIATIONS

aa	amino acid
BHK cells	baby hamster kidney cells
BMV	Brome mosaic virus
CHIKV	Chikungunya virus
CPV	cytopathic vacuole
CSE	conserved sequence element
dsRNA	double-stranded RNA
EEEV	Eastern equine encephalitis
EGFP	enhanced green fluorescent protein
ER	endoplasmic reticulum
GMP	Guanosine-5'-monophosphate
GTP	Guanosine-5'-triphosphate
GTPase	guanosine triphosphatase
HCV	Hepatitis C virus
HEV	Hepatitis E virus
HIV	Human immunodeficiency virus
IRES	internal ribosome entry site
mRNA	messenger RNA
ns	nonstructural
nsP	nonstructural protein
nt	nucleotide
NTPase	nucleotide triphosphatase
p.i.	post-infection
PAR	poly-ADP-ribose
PARP	poly-ADP-ribose polymerase
RdRp	RNA-dependent RNA polymerase
RI	replicative intermediate
RNA	ribonucleic acid
RRV	Ross River virus
RTPase	RNA triphosphatase
SARS	Severe acute respiratory syndrome
SF1 / SF2	Superfamily 1 / Superfamily 2
SFV	Semliki Forest virus
SINV	Sindbis virus
SLiM	short linear motif
ssRNA	single-stranded RNA
SUD	SARS unique domain
ts	temperature-sensisitive
UTR	untranslated region
VEEV	Venezuelan equine encephalitis virus
WEEV	Western equine encephalitis
wt	wild-type

ALPHAVIRUSES: REVIEW OF LITERATURE

Introduction

Alphaviruses from the *Togaviridae* family are plus-stranded, enveloped RNA viruses. The almost 30 catalogued alphaviruses are antigenically classified into 6 serocomplexes and have historically been divided into geographic groups that consist of Old World and New World alphaviruses (95, 307). Chikungunya virus (CHIKV), O'Nyong'nyong virus, Semliki Forest virus (SFV), Ross River virus (RRV), Getah virus, Sindbis (SINV) and multitudes of their derivatives are representatives of Old World alphaviruses, whereas Venezuelan equine encephalitis virus (VEEV), Eastern equine encephalitis (EEEV) and Western equine encephalitis (WEEV) are examples from the group of New World alphaviruses. Recently, phylogenetic analysis of complete genomic sequences of all known members of *Alphavirus* genus suggested its aquatic origin with subsequent emergence into terrestrial hosts (96). Symptomatically, infections with Old World viruses often result in rashes, fevers and arthritic syndromes in humans (390). New World alphaviruses may in turn endanger their hosts with encephalitis, whereas Highlands J virus (a cousin of EEEV) may cause mortality in domestic birds (457). The symptoms are also behind the peculiar names of certain viruses. For example, O'Nyong'nyong means "weakening of the joints" in the Nilotic language of Uganda, while Chikungunya stands for "that which bends up" in the Makonde language of Tanzania. However, the geographic location of strain isolation or outbreak is more frequently used for naming purposes: examples include the Semliki River region of Uganda and the city of Sindbis in Egypt. Alphaviruses are reported to be arthropod-borne (arboviruses) and are maintained in natural reservoirs, which are typically rodents and birds, by transmission between blood-sucking insect vectors and susceptible vertebrate hosts, with possible transoceanic exchanges performed by birds. Most of the vectors are represented by mosquito species, typically from *Aedes* and *Culex* genera, for which alphaviral infection is usually lifelong, persistent and asymptomatic (390). The exceptions include Southern elephant seal virus from the SFV complex, for which lice were identified as vectors capable of transmitting alphaviruses to marine mammals (201). Vectorology of the deviant alphaviral relatives that infect fish species, including the economically important rainbow trout that can be affected by Salmonid alphavirus, is also unclear. The transmission of fish alphaviruses is unlikely arthropod-mediated and can even occur horizontally via direct water-borne transmission (254). The involvement of mosquitoes in the transmission of alphaviruses is responsible for the geographic prevalence of these viruses in the tropics and subtropics; nevertheless, significant morbidity is caused by the Sindbis-related Ockelbo, Pogosta or Karelian fever viruses in Fennoscandia (206). During years 2005–2008 massive outbreaks caused by CHIKV strains in Southeast Asia, including India and islands in the Indian Ocean, severely affected millions of

people and drew great public attention. In Europe the medical status of that otherwise third-world problem has particularly raised after CHIKV managed to establish itself in a local mosquito population in Italy, albeit for short period (365). Additionally, VEEV and CHIKV have been added to the list of potent bioterrorism tools, and several nations, including the USA and the former USSR, reportedly developed weapons-grade VEEV, strategic stockpiles of which most likely still exist (35). Nevertheless, for a long time, alphavirus infections were not considered to be medically significant; as a consequence, neither vaccines for general human vaccination nor antiviral agents exist, making healthcare resources poorly prepared to fight against future outbreaks that involve Old World or New World alphaviruses.

Two members of an *Alphavirus* genus, SFV and SINV, are generally recognized as "workhorses" of virology due to their decades of service as convenient models for the study of the fundamentals of virology. For historical reasons, SFV was a favorite pet virus in the European virology labs, whereas the US virological community was in favor of using SINV as the primary model in alphaviral research. Regardless of this division of labor, alphavirus studies have significantly enriched textbooks with information regarding viral replication strategies and virus-host relationships. Notable topics include virus entry mechanisms using clathrin-mediated endocytosis and membrane fusion processes, structural studies of virion organization, regulation of the progression of the viral replication through switching the template specificity of the viral replicase complex as a consequence of proteolytic maturation of its component proteins and also the mechanism of the production of structural proteins by subgenomic RNA expression under conditions of viral infection-induced host cell metabolism and antiviral defense suppression. Extensive studies resulting in a relatively deep understanding of alphaviral life principles have culminated in the construction of gene delivery, viral vaccine and expression vectors, earning a distinguished biotechnological reputation for alphaviruses (15, 240, 315, 362).

Virion of alphavirus

Alphaviruses are particularly famous for being excellent model systems for the structural studies of enveloped viruses, which have been carried out over many years using a variety of biophysical methods (reviewed in (167)). Cryo-electron microscopy and image reconstruction analyses of whole viruses yielded a resolution of 9 Å for both SFV (248) and SINV virions (268), while additional X-ray crystallographic studies of the component structural proteins and their complexes further advanced the field (227, 431). Evidently, alphavirus virions are spherical particles that are 65–70 nm in diameter with a molecular mass of 5.2×10^6 Da. A single copy of an approximately 11.5-kb, plus-strand genomic RNA is contained within a virion, protected by several layers of macromolecular assemblies that ensure its structural integrity (Figure 1). The RNA

molecule is encapsidated by 240 copies of the capsid protein to form an icosahedral nucleocapsid core that is approximately 45 nm in diameter. The first 110 N-terminal amino acids (aa) of the capsid protein could not be defined in the X-ray structures but are highly basic and presumed to bind to the genomic RNA, which does not appear to assume a regular symmetry within the nucleocapsid core (47, 167). The nucleocapsid is enveloped into a host-derived lipid bilayer, which in turn is covered by an almost continuous surface layer of the envelope glycoproteins E1, E2 and E3 arranged in an icosahedral lattice with a T=4 arrangement (47). The carboxy-terminal domain of the capsid protein contains a hydrophobic pocket that interacts with the carboxy-terminal domain (33 aa) of the E2 protein (166). The lipid membrane is enriched with cholesterol and sphingolipids, which are crucial for both virus entry and budding (382, 448). The envelope proteins are post-translationally glycosylated, although the glycosylation positions are not conserved among alphaviruses (183). In the virion E1 and E2 interact to form 240 rigid heterodimers that are organized into 80 trimers, which resemble spikes protruding from the membrane of the envelope (398, 459).

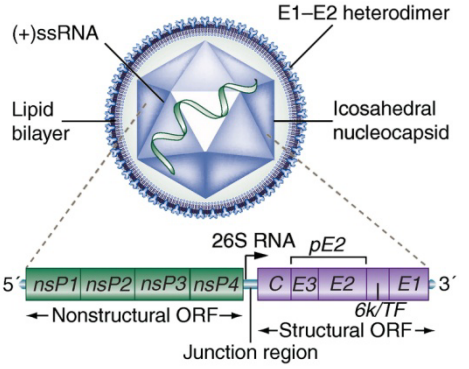


Figure 1. The structural and genomic features of alphavirus virion (adapted from (95) with permission of Future Medicine Ltd).

Alphavirus entry

Historically, the process of receptor-mediated endocytosis was primarily studied using alphavirus entry models, thus, this step of virus infection has been described in great details (167, 219, 250, 257, 298). Nevertheless, the alternative pathway of alphavirus entry without involvement of endocytosis was also proposed (187). The E2 protein is responsible for binding to the cellular surface receptor(s), which are, in general, not clearly identified, although they have been proposed to be of a proteinaceous nature (391). Heparan sulfate may also assist in the initial binding steps (37, 341). The diversity of vertebrate and

invertebrate host types for alphaviruses suggests that either a considerably conserved receptor is utilized for attachment or/and multiple cellular receptors are employed. The existing evidence suggests that a highly abundant laminin receptor can be a cellular receptor for SINV entry (435). Recent study also identified NRAMP2 (Natural Resistance-Associated Macrophage Protein) (331) as an alternative receptor for SINV entry, although the generalizability of these findings for other alphaviruses is questionable. Eventually, receptor-bound virions are internalized by clathrin-dependent endocytosis (178). As the virion-containing endosomal vesicles mature, the pH value is decreased, which leads to the dissociation of the E1-E2 heterodimers (433). At a neutral pH, the E2 glycoprotein hides the underlying E1 fusion peptide, whereas at low pH, the destabilization of the E1-E2 complex leads to the exposure of the fusion loop at the distal tip of an E1 protein (395). This fusion peptide is inserted into the endosomal membrane, where it leads to trimerization of the E1 proteins, which is associated with the melting of the endosomal and viral membranes, thus creating a pore in the membrane that is used to deliver the nucleocapsids into the cytoplasm (179). The induced structural changes in the envelope surface are translated into a concomitant alteration of the nucleocapsid core structure, thus exposing the RNA molecule (118, 441). The ability of the capsid protein to bind ribosomes led to the development of a model in which the interaction of capsid proteins in incoming nucleocapsids with ribosomes facilitates nucleocapsid disassembly (375, 443). The putative ion-channel properties of the E1 and 6K proteins may also contribute to the process of RNA release into the cytoplasm (256, 444).

Structural features of the alphaviral RNA

The genomic RNA of the typical alphavirus is approximately 11,500–11,800 nucleotides (nt) long and has 5' and 3' untranslated regions (UTRs) that are of varied length in different alphaviruses. There can be up to 90 nt for the 5'UTR, while the 3'UTR ranges from 77 nt for Pixuna virus to 609 nt for Bebaru virus and often contains repeats (120, 289, 348). Two protein coding regions within the genomic RNA are separated by an intergenic region, which is approximately 50-nt in the case of SFV (Figure 1). Only the first open reading frame, of approximately 7500 nt, is employed for translation from the genomic RNA to produce the nonstructural (ns) polyprotein P1234, components of which constitute the virus-specific portion of the replicase complex (Figure 1, 2). Viral RNA replication starts with the synthesis of an antigenomic minus strand of RNA. To perform efficient minus-strand synthesis, the polymerase or its *trans*-acting enhancer factor(s) must recognize a polyA fragment of at least 11 residues and approximately 19 highly conserved residues of the 3' conserved sequence element (3'CSE) immediately upstream of the polyA, which apparently lacks any significant secondary structure because of its high proportion of

A/U residues (147, 148). The individual nsP4 polymerase favors binding of UUUUUAA in the 3'CSE and is capable of the synthesis of the full-length minus strand in the absence of other viral proteins *in vitro* (403). Nevertheless, as for most plus-strand RNA viruses, the 5' region of the RNA is indispensable for minus-strand synthesis, and the complement of the 3'CSE on the minus strand is a co-promoter that is important for the synthesis of the plus strand of the RNA (103). Interaction between these two RNA regions has been demonstrated *in vitro* (98). Additionally, cellular factors, e.g., from the translation machinery (115), or viral replicase components may assist in bringing the RNA ends into proximity *in vivo* (103). The RNA 5' region exhibits significant length and sequence variation among alphaviruses and is highly structured, comprising 4 stem-loop regions that cover almost 230 residues (288). As evidenced by mutagenic analyses and template competition assays, the first stem-loop and the AU-rich region at the very beginning of the viral genomic RNA are engaged in the minus-strand synthesis through the binding of a component of the minus-strand ternary synthetic complex to this region (128, 280). The RNA minus strand is non-capped and non-polyadenylated; it also lacks polyU at the 5' end because the residues form polyA tract employed for synthesis initiation are not copied, but it contains unpaired guanosine at its 3' end (442). The minus strand or, more probably, the partly double-stranded RNA intermediate is a template for both genomic and subgenomic RNA syntheses. In the case of genome replication, the complements of the stem-loop regions on the 3' end of the minus strand are used as promoter elements (103). Stem-loop 1 is critical for efficient genomic RNA synthesis, and the other stem-loop structures appear to act as enhancers (281). Curiously, the 51-nt sequence element that forms stem-loops 3 and 4 is indeed highly conserved among diverse alphaviruses but was shown to be required only in insect cells and not in vertebrate cells, suggesting its possible recognition by host-cell-type-specific factors (85). Subgenomic 26S RNA synthesis is initiated through the recognition of an internal conserved sequence within the minus-strand template. Approximately 24 nucleotides, –19/+5, constitute the minimal promoter, although the –98/+14 region is required for maximal efficiency of subgenomic RNA synthesis in SINV (447); the role of this sequence has been confirmed by mutagenesis and analysis of duplication of the subgenomic promoter regions (222, 321). The subgenomic RNA is identical to the 3' third of genome; accordingly, it lacks the 5' end of the genomic RNA and therefore can only be used as mRNA for the production of structural proteins and not as a template for RNA replication. Both the genomic and subgenomic RNAs are capped by the viral machinery and polyadenylated, likely by the terminal transferase activity of nsP4 (409). Approximately 10% of the newly synthesized genomic RNA strands become packaged into virions (415), and, although capsid selectivity for RNA packaging is not exceptional, the encapsidation is significantly assisted by the presence of *cis*-active RNA-packaging signals in the genomic but not in the subgenomic RNAs (107, 233). Deletion analyses and

studies of defective interfering RNA species led to the identification of RNA elements that greatly enhance the selectivity of capsid binding and that are located in different parts of the ns-polyprotein coding region. Evidently, nucleotides 945–1076 of nsP1 region of SINV genome and nucleotides 2767–2990 in nsP2 region of SFV genome serve as packaging signals, whereas several putative packaging signals were mapped for RRV (180, 445). Additionally, the subgenomic RNAs of several viruses, including SFV and SINV, contain remarkably stable stem-loop structures within their 5' ends, allowing for the cap-independent translation of the structural polyprotein to overcome the limitations of the virus-induced shutdown of cellular macromolecular synthesis (104, 349, 378, 425).

RNA replication cycle of alphavirus

The genomic RNA molecule released from the virion is directly decoded to produce the ns-polyprotein, which contains essential viral components for the assembly of the replicase complex (355). For most alphaviruses, the translation of the genomic RNA predominantly yields the P123 polyprotein product, due to the opal termination codon (UGA) in the end of the nsP3 region (224, 225). In 10–20% of cases, translational read-through of this termination codon results in an extended version of the polyprotein, P1234 (88), which also contains, in contrast to the prematurely terminated product, an ns-protease recognition site of approximately 6 residues (237) followed by the nsP4 region. Presence of an opal termination codon between nsP3 and nsP4 may have influence on the host-specific infectivity (271) and contribute to virulence (414). In several alphaviruses from both the Old World and the New World groups, including SFV and several CHIKV strains, the opal terminator is absent, and the only translational product is P1234. The embedded ns-protease is then exclusively responsible for the three cleavage events within P1234 that yield several cleavage intermediate products and, finally, the mature nsP1, nsP2, nsP3 and nsP4 proteins. All of the processing intermediates and individual proteins are believed to play their own specific roles in the viral lifecycle (reviewed in (200)), which will be described and discussed in detail in the following sections. The viral polymerase activity provided by the nsP4 protein is accompanied by multiple biochemical activities from the P123 polyprotein, and in the early stages of infection (up to 6 h post-infection (p.i.)) is directed first towards the production of the antigenomic minus-strand templates which, in their turn, serve for the multiplication of genome-polarity RNAs (182). As evidenced by experiments with translation inhibitors, minus-strand synthesis requires the constant production of newly synthesized replicase polyproteins and ceases within the first 4–6 h p.i. (357, 359), whereas the synthesis of the plus strands continues, with an emphasis on the production of subgenomic RNAs during the late stages of infection due to the far more efficient initiation at the subgenomic

promoter compared with the genomic promoter (103). Recent measurements suggested that the minus-strand copy number reaches maximal values at 6 h p.i. and remains at this level stably until 12 h p.i., when the cell starts to die (246). The absolute numbers appear to be proportional to the cell size, with ~27,000 minus strand copies in BHK-21 cells in contrast to approximately 5,000–7,000 copies in chicken embryo fibroblasts (CEF), which have one-fifth the amount of RNA/cell compared with BHK-21 cells. This difference in copy numbers apparently indicates that the exhaustion of host factors may be a constraint on the total number of replication-transcription complexes (246, 437). As with all plus-strand RNA viruses, the RNA synthesis is asymmetric, whereas calculation suggests that almost 200,000 copies of each of the genomic and 26S-subgenomic plus strands are synthesized by 8 h p.i. in vertebrate cells (415).

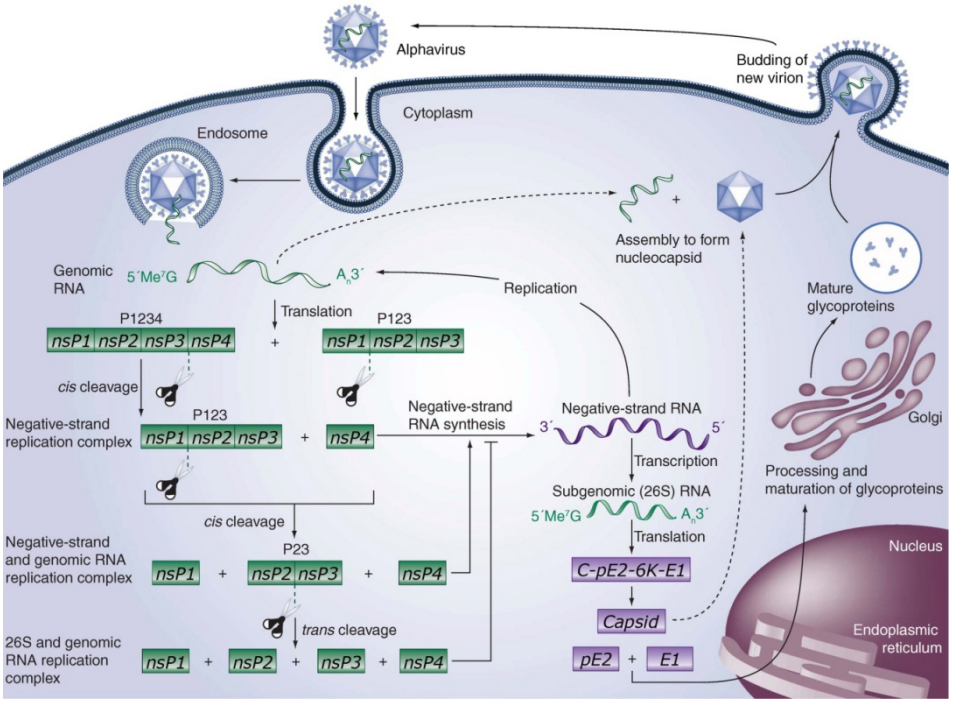


Figure 2. Replication cycle of alphaviruses (adapted from (95) with permission of Future Medicine Ltd).

Sites of alphavirus replication

The viral genomic RNA must participate in several competing processes, including translation to produce viral proteins, replication and encapsidation (Figure 2). Obviously, these processes must be highly regulated and therefore separated to avoid unwanted interference. To achieve this goal, many viruses induce rearrangements of the cellular membranes to create novel, virus-induced vesicles as the most efficient way to sequester, protect and concentrate viral replication-related factors and templates and to coordinate the replication steps (23, 62, 153, 184, 193, 231, 242, 262, 275, 282, 344, 352, 387, 440, 465, 466). Recent studies suggest that alphaviral replication proteins are targeted first to the plasma membrane, where they form spherular invaginations (109). These spherules appear to contain partly double-stranded replicative intermediate (RI) RNAs inside the cavities (200), whereas ns-proteins localise at the cytoplasmic side of the vesicle necks, which obviously play a role of an open channels for nucleotide import and product RNA export (196). It is not clear, whether the nature of the spherule framework is mostly membranous or whether it contains a proteinaceous skeleton that is covered by membrane (200). For certain viruses it was proposed that multiple copies of self-interacting, membrane-bound viral replication proteins may form capsid-like shells that resemble retrovirus and double-stranded RNA (dsRNA) virus virions (364). Alternative findings revealed that replication proteins at least from certain enteroviruses and flaviviruses through the recruitment of the host effector protein Arf1 GTPase and its guanidine exchange factor GBF1 may engage phosphatidylinositol-4-kinase III beta to the membranes, where it mediates the synthesis of phosphatidylinositol 4-phosphate (PI4P) lipids, thus leading to the generation of organelles with unusual lipid composition and microenvironment (153). Analogously, recruitment of phosphatidylinositol-4 kinase III alpha (PI4KIIIα) by Hepatitis C virus (HCV) NS5A protein was demonstrated to contribute to the structural integrity of membraneous web that is formed by the vesicles of heterogeneous size that contain replication complexes of HCV (323). Evidently, both unprocessed alphaviral P123 and replicating viral RNA are needed to induce spherule formation (109, 345). Later, as alphavirus infection proceeds, the plasma-membrane-bound spherules become detached and internalized in an active endocytic process, requiring a functional actin-myosin network and phosphatidylinositol 3-kinase activity (385), although this relocalization of the replication complexes appears to be dispensable for RNA synthesis (109). Eventually, the escalation of alphaviral replication leads to the disruption of secretory pathway integrity and the reorganization of the internal membrane compartments. Therefore, at later stages of infection, the relocalized spherules become integrated into more complex assemblies. Finely distinguishable by electron microscopy, the spherules appear in large vesicular structures (0.6 to 2 μm in diameter) designated as “type I cytopathic vacuoles”, hereafter referred to as CPVs (132, 303). The CPVs contain a number of double-membrane-bound

spherules of approximately 50nm in diameter that line the vacuole membrane at regular intervals and are connected to the cytoplasm via necks, where the active replication proceeds (93, 94, 133). These cytopathic vacuoles are rich in late endosomal and lysosomal marker proteins (196). Nevertheless, it has not yet been definitively determined, whether the existing endosomes and/or lysosomes become internally redesigned or whether the alphaviral proteins can divert the new endosomes' biogenesis pathway to establish this viral progeny-producing infrastructure. The intrinsic affinity of the viral replicase complex to the endo/lysosomal membranes appears to be unique to the togaviruses but apparently is evolutionarily justified (245). Ns-proteins were identified to be localized to the cytoplasmic face of the limiting membranes of the CPVs (112, 196). It appears that in alphaviruses, similarly to many other RNA viruses, packaging of the nascent genomic RNA molecules is coupled to replication, so that the nucleocapsid assembly sites are adjacent to the exit pores of the CPVs (466). Importantly, the CPVs seem to have direct connections to the ER through thread-like ribonucleoprotein structures that extend from the base of the spherules (112). The proximity to the ER membranes is certainly not coincidental but provides a clear advantage in allowing *in situ* synthesized structural glycoproteins to pass into the ER and, subsequently, the Golgi (131), to form preassembled spikes at the outer membrane surface awaiting assembled nucleocapsids (119, 166, 380).

The replicative machinery of alphaviruses

Despite the small size of their genome, alphaviruses, as well as many other RNA viruses, in general demonstrate versatility and self-sufficiency in the functioning of self-encoded parts in the replication process (reviewed in (390)). Although direct, cell-free experiments testing the self-completeness of the viral-encoded replicase entirely recreated using highly purified ns-polyprotein and its fragments have not yet been completed, it appears that, in terms of performing the actions most vital for viral RNA replication and preservation, alphaviruses rely mostly on themselves and not on the recruitment of host accessories. Obviously, being obligate parasites by definition, viruses, and alphaviruses are hardly an exception, divide and conquer the host resources, including the precisely tuned cellular nanomachinery that normally ensures the well-established internal life of the host cell that the viruses happen to invade. Nevertheless, it is the viral-encoded self that secures the very existence and procreation of viruses, whereas the capacity to be open and communicate, strengthened by the evolutionarily shaped ability to balance robustness with genetic flexibility, determines the quality of viral life.

As already mentioned, the alphavirus genome contains two open reading frames that are decoded to produce the ns- and structural polyproteins. As the name implies, the structural proteins derived from the structural polyprotein

form the enveloped virions containing encapsidated genomic RNA. In turn, the ns-polyprotein components are used to assemble the RNA replication complexes (reviewed in (200)). The viral replication machinery has a highly modular organization, in which the ns-polyprotein becomes divided into four mature component proteins during the replication process, and each of these mature viral proteins has a multi-domain structural organization itself (Figure 3). For reasons of simplicity, the functionality of the replicase components is usually explained by describing the properties of the individual ns-proteins (130). However, it should be stressed that the viral replicase is evolutionarily designed to be formed *in cis* as a functionally self-contained complex system and that it becomes operational only as a whole. Accordingly, vital steps of replication, such as RNA copying, strand separating, RNA capping, protection of replicating RNA and others are obligatorily cooperative tasks that are shared by functionally interacting replicase domains, which nevertheless may physically belong to different mature nsP-s. Therefore, studies of the replicase components in isolation can hardly uncover their true functionality within the sophisticated structure of the replication complex.

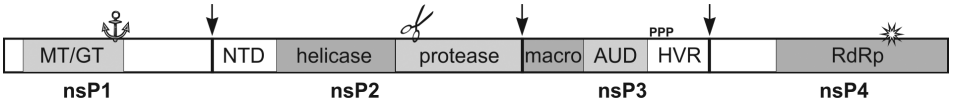


Figure 3. Modular organization of the ns-polyprotein of alphaviruses. Arrows indicate positions of proteolytic cleavages within P1234 polyprotein that eventually yield matured proteins nsP1, nsP2, nsP3 and nsP4. NsP1 features include methyltransferase and guanylyltransferase (MT/GT) activities in addition to membrane-binding properties (anchor sign). NsP2 posseses domains related to helicase and protease functions that are assisted by N-terminal domain (NTD); scissors indicate location of catalytic residues of protease. The body of nsP3 consists of macro domain and centrally located alphavirus-unique domain (AUD) followed by hypervariable region (HVR) that apparently lacks regular structure and contains phosphorylation sites (^{PPP}). N-terminal region of nsP4 also predicted to be unstructured is presumably involved in numerous intra-replicase interactions and assists in RNA-dependent RNA polymerase (RdRp) activity of nsP4, location of catalytic residues of which is indicated by multi-pointed star.

Notably, individual viral proteins likely retain their enzymatic properies also outside the context of viral replication complex. As a result, such activities may potentially lead to rather unpredictable consequences inside the cell. In can be inferred that viruses have taken measures to use some of these functionalities in their favor and minimize the potentially hazardous effects of others. Therefore, the mature viral proteins most likely have their own pathogenesis-related properties that are needed for the subversion of cellular homeostatic mechanisms and for the remodeling of the internal organization of the cell.

Nonstructural components of alphavirus replicase

NsP4 protein (614 aa in SFV) is a catalytic subunit of viral replicase that possesses RNA-dependent RNA polymerase (RdRp) activity (335), which is a vital function for all RNA viruses; the single known exception is Hepatitis delta virus. Although the crystal structure of the alphaviral polymerase is not yet known, its domain organization can, to some extent, be deduced bioinformatically. These analyses suggest that approximately 100 N-terminal aa residues are rather intrinsically disordered and specific for the genus *Alphavirus* only; the genetic evidence suggests that this domain participates in distinct interactions with other nsPs in the context of differentially functioning RNA synthetic complexes (337). The polymerase catalytic activity resides within the remainder of the nsP4, which contains signatures characteristic for RdRps, including a conserved GDD motif (171, 409). The alphaviral RNA polymerase belongs to the same group as the polymerases from *Bromoviridae*, *Tobamoviridae*, *Tobraviridae* and *Hepeviridae* (191, 192). RdRps are generally highly conserved enzymes and, because many of the solved crystal structures of the viral RNA polymerases follow the principally similar "fingers-palm-thumb" organization of sub-domains (50, 87, 278), it would not be surprising to find that alphaviral nsP4 follows this trend.

Thorough studies of alphaviral RNA polymerase activity resulted in many details about the requirements for RNA recognition and synthesis. The initial experiments demonstrated that the deletion of 97 N-terminal aa residues from the SINV nsP4 released a soluble core segment of the polymerase that possessed primitive terminal nucleotidyl transferase activity, mediating the non-templated addition of nucleotides, with a preference for adenosine, instead of RNA-copying activity (409). It was later shown more rigorously, using purified components that the N-terminal portion of nsP4 is of extreme importance. Its presence significantly enhanced the nucleotide transfer activity of full-length nsP4 compared to the Δ97nsP4 core domain alone, which suggested that the N-terminal region may assist in RNA substrate recognition (335). The terminal nucleotide transfer activity of alphaviral polymerase appears to be of extreme importance for viral fitness. It was demonstrated that, to restore replication competence, SINV is capable of repairing genomes after truncations of the termini, regenerating the poly(A) tail and even generating new AU-rich sequences at the 3' ends of defective genomes or in response to introduced mutations within the polymerase (320).

It was also demonstrated that the presence of unprocessed P123 is mandatory for the purified full-length nsP4 to synthesize minus-strand RNA *de novo*, whereas Δ97nsP4 core polymerase failed to do so (335). Interestingly, nsP4 isolated from mammalian cells, where it was initially co-produced with uncleavable P123 using vaccinia virus vectors, retained the ability to synthesize minus strands even without P123 (403). This data favored the conclusion that the N-terminal domain of nsP4 acts as a molecular switch, activated by protein-

protein interactions with P123, leading to corresponding conformational changes within nsP4 itself that render alphaviral polymerase capable of recognizing the promoter for minus-strand synthesis (335, 337). This promoter preference is nevertheless altered during the course of viral replication. Cross-linking studies identified two distinct regions (aa residues 329–334 and 531–538) within nsP4 of SINV which preferentially bound oligonucleotides representing the genomic and subgenomic promoters found on the minus-strand copy of RNA, respectively, suggesting that nsP4 itself has advanced specificity towards different promoters (228, 229). The change in promoter specificity coincides with changes in the polyprotein composition that are caused by sequential proteolytic processing. It was shown that the presence of unprocessed P123 switches nsP4 into minus-strand synthesizing mode, whereas the formation of the mature nsP1, nsP2 and nsP3 proteins evidently triggers further conformational changes within the polymerase (214–217, 370). These changes prevent further recognition of the promoter for minus-strand synthesis and instead provide for the recognition of the genomic or subgenomic promoters by the polymerase complex (182, 355, 359).

To become functional, nsP4 must be released from the P1234 polyprotein precursor immediately after it has been synthesized. Unprocessed P1234 is nonfunctional as a replicase, although it contains all of the necessary components (214, 259, 370; Paper I). The reason for this lack of function is that the polymerase activity of nsP4 is dependent on its exact N-terminus, which is highly conserved among alphaviruses (371). Importantly, the very first aa of nsP4 is tyrosine, which both defines its fate as an individual protein inside the cell, serving as a signal according to the N-end rule in the pathway of protein degradation by the proteasome (464), and plays a vital role in the polymerase activity (371). In mutagenesis experiments, most aromatic aa residues were accepted as substitutions for this tyrosine, although they still led to defects in replication, which were compensated for by changing the putative promoter site for genomic RNA synthesis, mostly through the addition of AU sequences to the 5' end of the genomic RNA (368, 372). These genetic interactions between the first aromatic aa of the nsP4 and the 5' end of the viral genome may suggest the speculative conclusion that in the initiation of RNA synthesis the conserved N-terminal tyrosine may be involved in the aa templating mechanism, in which the functionally important tyrosine side chain can mimic a pyrimidine nucleobase (284, 285).

Such proteolytic activation of polymerase activity is part of a general strategy for replication of several viruses and was carefully studied, for example, in the case of picornaviruses, which are known to employ 3CD, a functional intermediate product of the proteolytic maturation of the polyprotein (21). In the case of poliovirus, the 3C component of this precursor is functional as a protease, whereas the 3D polymerase (3Dpol) part is inactive until polyprotein processing is complete (39, 151). The initial explanation for this phenomenon came from structural analysis, which revealed that the N-terminal

glycine of 3Dpol becomes buried in a pocket at the base of the finger region and, through a network of hydrogen bonds, helps to correctly position the conserved aspartate that is involved in nucleotide selection at the catalytic site of the polymerase (146, 404). Therefore, the absence of the free N-terminus of 3Dpol in its binding pocket in the case of 3CD was expected to disrupt the polymerase activity. Surprisingly, the structure of the poliovirus 3CD revealed an arrangement of the active-site residues that was very similar to that in 3Dpol, despite the disruption of network of interactions that was expected to position the key residues in the active site (249). Analysis of this data led to the conclusion that changes in molecular flexibility, rather than significant structural rearrangements, are the true determinants of polymerase activity. This molecular flexibility is, then, in part regulated through the insertion of the N-terminus of 3Dpol into its binding pocket (123). The aa composition of the N-terminus of alphaviral nsP4 suggests that it is more likely buried than exposed. However, whether a direct analogy can be drawn from the case of picornavirus and applied to alphaviral nsP4 remains to be explored.

NsP1 protein (535 aa in SFV) is the main determinant of the subcellular localization of the replication complexes due to its intrinsic ability for membrane binding (304), which is mediated by the internally located membrane-binding peptide, aa 245–264 of SFV nsP1 (7). This region within nsP1 was predicted to fold into an amphipathic alpha-helix that ostensibly becomes inserted laterally into the membrane in a monotopic fashion so that hydrophobic aa residues present on one side of the helix could interact with the lipid acyl chains of one leaflet of the lipid bilayer and the positively charged residues located on the other side of the helix stabilize this construction through ionic interactions with the negatively charged groups of the phospholipids (208, 386). During the course of the viral replication, the string of consecutive cysteine residues within nsP1, aa 418–420 of SFV nsP1, become palmitoylated, which determines the final anchoring of the nsP1-containing complexes (4, 203). A distant relative of alphaviruses, Brome mosaic virus (BMV), was recently shown to cause membrane curvature upon the recruitment of membrane-shaping reticulon-homology proteins (63, 64). It is, then, an appealing hypothesis worthy of investigation whether multiple interactions of nsP1 molecules with membranes within a short distance can modulate bilayer curvature with possible help provided by the recruitment of specialized cellular proteins (26, 73, 255, 461). The special affinity for anionic phospholipids of the individual nsP1 protein along with its specific targeting to the plasma membrane results in the protrusion of filopodia by an as-yet-unexplained mechanism (4, 204). The induction of filopodia, leading to the formation of so-called virological synapses (165) or tunneling nanotubes (140), is characteristic feature of the membrane-binding proteins of many viruses and for retroviruses in particular (9, 212). However, to which extent filopodia-like structures may contribute to the pathogenesis and serve as an additional intercellular route for

spreading infection between connected cells in case of alphaviruses remains to be further clarified.

The nsP1 protein performs extensively collaborative work with RdRp on the maturation of newly synthesized plus-strand RNAs. For these actions, it possesses capping enzymatic activities leading to the addition of a cap-0 structure to the 5' ends of the RNAs (5, 205, 261). Because negative strands of RNA are not capped, a mechanism should exist for the discrimination of the RNA molecules, or alternatively, distinct conformational states of the protein could exist, such that the capping function of the enzyme becomes activated at the specific point of replication. For a number of viruses, including flaviviruses (234), the RNA-capping domain is integrated into the RdRp protein (316), whereas in alphaviruses, its separation obviously provides additional options for control. In alphaviruses, the N-terminal region of nsP1 (aa residues 30–265) is conserved and is related to the methyltransferase domains of other viral and cellular proteins, whereas the other regions are alphavirus-specific (334). It has been confirmed biochemically and genetically that nsP1 has a guanine-7-methyltransferase activity that is responsible for the transfer of a methyl group from an S-adenosyl-methionine donor molecule to the guanine moiety of guanosine-5'-triphosphate (GTP), resulting in a transient complex, whereas m^7-GMP remains covalently attached to the His-residue that is absolutely conserved among alphavirus-like viruses, which is His38 in the case of SFV nsP1 (6). This intermediate is then used to transfer the methylated GMP, in a guanylyltransferase reaction, onto the 5' end of the RNA with a previously removed γ-phosphate to finalize the capping reaction (5). The sequence of reactions is unique to the alpha-like superfamily of viruses, and similar findings were demonstrated for Bamboo mosaic virus (154), Brome mosaic virus (3), Tobacco mosaic virus (258) and Hepatitis E virus (244). Conventional capping reaction that is mediated by cellular and other viral enzymes normally proceeds in a way that guanylyl is first transferred to the γ-phosphate-cleaved 5' end of the RNA and only then is methylated at guanine position N-7, and in some cases also at the 2'-O position of the ribose ring to create cap-0 or cap-1, respectively, the latter of which is predominant in higher eukaryotes (59, 121). The peculiarity of the alphaviral capping mechanism thus lies in the reversed sequence of events, making this mechanism quite unique. Site-directed mutagenesis of the conserved residues implicated in capping activity results in loss of viral infectivity, suggesting that both of the enzymatic activities that participate in the capping reaction by nsP1 are vital for alphaviruses (434). This requirement seems obviously explicable, taking into account the important role that the cap structure holds for RNA stability and the efficiency of protein translation, as well as the fact that cellular N7-methyltransferases reside in the nuclear compartment and are incapable of supporting the cytoplasmic production of alphaviral RNA. However, it should not be overlooked that alphaviruses are known to cause a general transcriptional and translational shutdown (116, 124) and turn this shutdown into an advantage by directing the

cellular protein-synthesizing capacities for the exclusive production of viral proteins (253). Therefore, despite having their own capping apparatus, several alphaviruses (for example, SFV and SINV) have devised a strategy of cap-independent translation by employing stable IRES-like stem-loop structures in the 5' regions of their subgenomic RNAs (104, 378) to withstand the general translation-initiation shutdown mediated by the impotency of eukaryotic initiation factor 2 (eIF2) phosphorylated by alarmed protein kinase R (PKR) to support cap-dependent translation (349, 425).

Because the membrane-binding peptide is located somewhat centrally inside of the methyltransferase domain, it is not surprising that a feedback mechanism exists to allosterically regulate the capping activity of nsP1, mediated through its interaction with the membranes. In support of this idea, it was found that the depletion of phospholipids by detergents abolished the enzymatic activities of nsP1 in SFV, and supplementing with phosphatidylserine restored the activity (7); nevertheless, nsP1 of SINV did not demonstrate a similar dependence and remained biochemically active without any requirement for membrane association (410). NsP1 itself, through direct protein-protein interactions with nsP4, appears to be a major regulator of the promoter selectivity of the latter, therefore defects in nsP4 are most frequently compensated by complementary changes in the nsP1 coding region (84, 368). Additionally, a number of mutations in nsP1 leading to a temperature-sensitive (ts) phenotype were selected in SFV and SINV that specifically switched off the minus-strand-synthesizing mode of RdRp. These mutations, ts11 (A348T) in SINV nP1 and ts14 (D119N) in SFV, lie in the methyltransferase domain, highlighting the tight interconnection between the functions of the polymerase and the capping enzyme (238, 358, 438). Interestingly, the 80 C-terminal-most aa of nsP1, being less conserved than the remainder of the protein are not part of methyl-transferase domain and are predicted to be rather unstructured by the secondary structure prediction servers (330, 333). The aa composition of this region is somewhat spectacular, including a cluster of negatively charged residues that precedes the protease cleavage site, in contrast to the rather positive charge of the rest of the protein. Although the aa composition of this stretch is different in different alphaviruses, its presence is a common feature for most of them. Importantly, one of the ts-mutations of SFV, ts10 (E529D of SFV nsP1), which leads to a defect in viral RNA synthesis, is found in this area, suggesting that this non-conserved stretch is hardly random because even equivalent (acidic residue to another acidic) mutation is capable of disrupting this well-tuned macromolecular machine (239). It is tempting to speculate that by the analogy with NS4A of HCV (232) this negatively charged stretch in nsP1 may function as an electrostatic switch designed to dynamically regulate replicase activity.

NsP2 protein is the largest (799 aa in SFV) of all of the alphaviral proteins and therefore possesses many functional elements. It is considered to be a master regulator of viral replication activities, expressed through a number of its

enzymatic activities. The domain organization of nsP2, revealed by biochemical and bioinformatic studies, suggests that approximately 450 N-terminal aa residues are distributed between three domains: an N-terminal domain of approximately 180 aa and two RecA-like domains, which constitute the core of the alphaviral helicase (99, 200). The alphavirus nsP2 sequence contains signatures universally characteristic for the Superfamily 1 (SF1)-related helicases (81, 377), including motif I (Walker A motif) and motif II (Walker B motif). Motif I, sequence GVPGSG$\underline{K}^{192}$S in SFV nsP2, is known to be involved in NTP binding via the interaction of a conserved lysine residue with the β- and γ-phosphate groups of the NTP-Mg^{2+} complex, while a conserved catalytic aspartate of motif II, sequence YV$\underline{D}^{252}$EAFA in SFV nsP2, is known to interact with Mg^{2+} ions. Nucleotide triphosphatase (NTPase) (172, 324) and subtle RNA helicase (duplex-unwinding) activities (122) were demonstrated biochemically for the SFV helicase domain. SF1 helicases possess 5'-to-3' unwinding activity and seem to prefer stretches of single-stranded RNA adjacent 5' to the duplex region to initiate strand separation. For viruses bearing their own capping apparatus, the same active site used for NTPase activity is also used for RNA triphosphatase (RTPase) activity, and in accord with this shared function, the mutagenesis of conserved residues in the NTPase domain abolishes any RTPase activity as well (172). The RNA triphosphatase reaction is absolutely required to precede other capping reactions as it removes γ-phosphate from the 5' end of the RNA to which the cap structure will be transferred (421). The phenotypes of a number of alphaviral ts-mutants linked to the helicase domain demonstrate its importance for the viral replication cycle. Interesting examples include, for example, ts21 (C304Y in SINV nsP2), which fails to switch to subgenomic RNA synthesis (143).

However, there are a number of open questions related to the actual helicase function; thus, it is not quite clear whether alphaviral helicase acts at the separation of double-stranded RNA intermediates or for the melting of secondary structure elements (313); nor is it clear whether it is constantly active or needs special triggers for activation. RecA-like domains are thought to play a rather simple role of the acquisition of a closed conformation upon the binding of NTP at the interface of domains and the relaxation of this structure upon NTP hydrolysis (99, 162). The physical movement of the RecA domains in helicases is normally translated to movement along the nucleic acid via a conserved motif containing an arginine residue (“arginine finger”) that directly interacts with the RNA, usually found in the second RecA domain (40, 207). For the same reason, most crystal structures of functional helicases from the SF1 and SF2 groups analyzed thus far contain a third helicase domain, the role of which is to accommodate the RNA molecule in the special cleft created between this domain and the tandem RecA domains and to help in the transformation of the NTP hydrolysis energy into RNA movement alongside this cleft (296, 377). This third helicase domain, which sometimes appears inserted between the RecA domains or adjacent to these domains in the primary

sequence, occupies a similar spatial position in three-dimensional space, so that the overall helicase structure resembles a Y-shape (100). Nevertheless, RNA-accomodating helicase domain is usually very helicase-specific and therefore little or no homology is evident even between the structurally well-studied helicases of different flaviviruses, making the prediction of such domains difficult (450). Whether such a domain is present in the alphavirus ns-polyprotein is not clear. A number of candidate domains are available for the putative role of the third helicase domain. The N-terminal domain of nsP2 appears to be the most logical candidate due to its position adjacent to the RecA tandem domains. However, it is rather suspicious that no ts mutations were identified within this domain, which would then be functionally coupled to helicase function; whereas within the other helicase subdomains, a number of mutations are found in both SINV and SFV (143, 238). Additionally, in case of the SFV nsP2, single-stranded RNA (ssRNA) stimulated the NTPase activity <2-fold (122), a similar extent of stimulation was observed for NTPases from the Rubella virus (135) and Turnip yellow mosaic virus (168), which helicase domains also belong to SF1 group of helicases. At the same time, for the HCV NS3 helicase (SF2), which has all three domains in place, polyU is able to potentiate NTPase activity up to 27-fold (310), also 15- to 20-fold stimulation was reported in case of SARS coronavirus (399), Dengue virus (24) and Bovine diarrhea virus helicases (138), all of which belong to SF2 group. This observation provides another portion of indirect evidence suggesting that the helicase region of nsP2 may lack the ability to accommodate an RNA template. If so, then this putative third domain may be separated in the primary structure of ns-polyprotein and may physically reside in yet another matured ns-protein (such as nsP1 or nsP3). It remains to be experimentally demonstrated whether this assumption answers the question about the existence of possible triggers of helicase functionality. Certainly, another line of evidence argues against nsP2 being a helicase whose activities are needed to displace nascent RNA chains. These arguments take into account the conclusion that none of the many ts mutations in the helicase region produces a phenotype comparable with the ones arising from polymerase defects, exemplified in the blocking of the elongation reaction (143, 238), which would be the case for a true strand-separating helicase. Nevertheless, the interpretation of the experimental results obtained with ts-mutants leads to the conclusion that the helicase domain may be especially important during the early stages of replication and likely shares the responsibility for the activation of the plus-strand-synthesizing mode of the alphavirus replicase complex (77). Similar conclusions were also made for BMV (8).

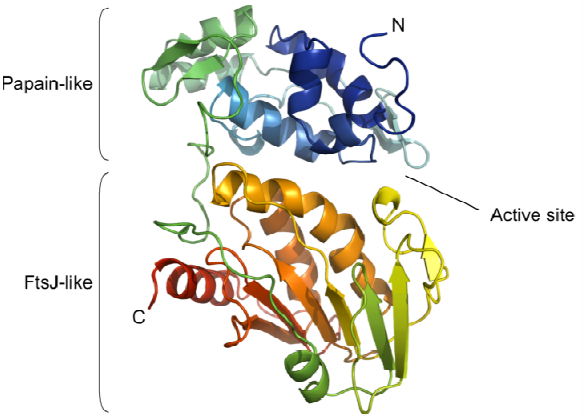


Figure 4. Model of the SFV protease domain built on the basis of the crystal structure of the VEEV protease domain, Protein Data Bank entry: 2hwk (339). Papain-like protease and FtsJ-like methyltransferase subdomains are marked; location of the active site is indicated.

The C-terminal part of nsP2 possesses a protease activity that is responsible for the proteolytic maturation of the ns-polyprotein (66, 150, 259, 389). The tertiary structure of the corresponding domain of the VEEV has been solved at 2.45-Å resolution, which has provided much insight into its functionality (339). Structurally, the C-terminal domain of nsP2 consists of two subdomains (Figure 4), the first of which contains the residues that constitute the catalytically active site, viz. Cys 478 and His 548 in case of SFV nsP2, embedded into the papain-like fold, and the second of which is structurally related to the FtsJ-like fold that is characteristic of RNA methyltransferases (38). The overall fold of this viral papain-like protease is sufficiently unique to deserve a separate C9 family of alphaviral proteases within the MEROPS peptidase database (16, 322). The FtsJ-methyltransferase-like domain found in the protease region appears to be incomplete in most alphaviruses, except those operating in fish species, therefore its potentially redundant activity with respect to nsP1 methyltransferase activity could be excluded (252). Nevertheless, its scaffold definitely plays an important role in replication, as mutations of the conserved exposed residues hamper replication efficiency. Even more strikingly, the specificity determinants of protease activity are shared between the subdomains, and the substrate-accommodating cleft appears in the interdomain region (338). Additionally, the analysis of a number of nsP2 ts-mutants, namely SINV ts24 (G736S) and ts113 (N700K), and alanine scanning of the conserved residues suggest that the FtsJ-like region is directly involved in subgenomic promoter recognition (143, 252). Curiously, the long loop region connecting the subdomains is also conserved among diverse alphaviruses, suggesting that it has a more important role, possibly in the allosteric regulation of proteolytic activity, than being a mere

linker. Most importantly for this dissertation, proteolytically active alphaviral full-length nsP2 and the protease domain Pro (39 kDa) can be efficiently produced in recombinant form; thus, their properties were extensively studied (422, 458; Paper I and Paper III). The interpretation of the experimental results and their contribution to the understanding of substrate specificity and the regulation of alphaviral proteolytic processing are discussed in the following sections of this thesis.

The processing of P1234 is found to be strictly regulated during the course of replication (150, 423, 463). As previously mentioned, the 3/4 site, residing between nsP3 and nsP4, is processed immediately after or during polyprotein translation. This cleavage leads to the activation of the polymerizing activity of nsP4, the recruitment of the viral template and the functionality of the minus-strand-synthesizing mode of the early replicase complex P123+nsP4 (182, 215–217, 370, Figure 2). It is not clear how many rounds (or a single one) of minus-strand synthesis are in fact performed on a single input of a positive strand, but molecular genetics and replication reconstitution studies suggest that, to activate recognition of the genomic promoter and switch replicase efforts over to plus-strand production, the 1/2 cleavage site located between nsP1 and nsP2 must be processed (182, 217). Presumably, internal replicase reconfiguration due to conformational changes followed by release from self-inhibition or the activation of additional viral transcription factors leads to the preferential utilization of minus-strand templates for genomic and subgenomic RNA synthesis. The analysis of processing in infected cells identifies P23 as an existing form of the polyprotein (149, 182), although *in vitro* experiments suggest that cleavage of the 2/3 site found between nsP2 and nsP3 immediately follows 1/2 site processing (423). *In vitro* experiments demonstrated that the 1/2 site is preferentially or exclusively cleaved *in cis* (259, 423), while from the structural analysis of the protease domain it was concluded that the 2/3 site can become accessible exclusively *in trans* (339). It is generally believed that the 2/3-site processing signifies the point of no return for the viral replication because mature individual proteins are incapable of supporting the initiation of minus-strand synthesis (216). Accumulated pool of liberated nsP2 consequently accounts for the swift processing of incoming P1234 polyproteins preferentially at the 2/3 site, thus precluding them from forming new replication foci (463). The existence of certain ts-mutants that are capable of reverting the activities of late replicase complexes to minus-strand synthesis in the presence of protein synthesis inhibitors is probably the best evidence that we have to conclude that protein-protein interactions-induced changes are indeed the main factors for the altered promoter-recognition preferences of the polymerase (77, 354). There remains a great deal of uncertainty surrounding our actual understanding of the sequence or extent of events; however, biochemical and structural data continue to accumulate, promising to eventually reveal some snapshots of reality.

The multi-talented nsP2 protein of Old World alphaviruses has also been confirmed as a negative regulator of general cellular transcription and trans-

lation to favour viral macromolecular syntheses over host ones. Transcription and translation inhibiting functions appear to be distinctly independent because certain mutations, namely of proline residues P726 in SINV nsP2 and its equivalent P718 in SFV nsP2, can selectively relieve the inhibition of cellular transcription (74, 102, 116, 126, 396). The effect of the wild-type (wt) SFV or SINV infection on cellular translation is so profound, however, that after just few hours of progressing viral replication, the cellular translation machinery becomes dedicated to the processing of viral mRNAs, hence viral gene products, especially structural proteins, become predominant in the proteome (253, 349, 396, 412, 425). Translation machinery subversion is evidenced from the recruitment of ribosomes to the sites of viral replication, potentially mediated by nsP2 interacting directly with the ribosomal proteins (264). Additionally, the accumulation of unprocessed pre-rRNAs is observed, which apparently indicates that the genesis of the new ribosomal units also becomes disturbed (116). A range of mutations affecting the nsP2 coding sequence with a bias to its C-terminal region were selected for the property of making viral RNA replication non-cytotoxic for the cell and therefore persistent in many studied cases (1, 41, 101, 102, 221, 301). The general conclusion that can be drawn from numerous studies is that mature nsP2 protein is self-sufficiently cytotoxic (116). It should not be left unnoticed, however, that like most biological entities, cytotoxicity as a property studied in the biological context (virus infection) is not a discrete parameter ascribed to a protein, but depending on the cell type, comprises a complex function of the proper localization and effective concentration of effector protein in the cell, which in turn is dependent on the intrinsic capacity of viral protein to support replication. This complexity makes the elucidation of the cytotoxicity determinants of nsP2 using replicating virus a difficult task because mutations that are expected to change a single property usually have more broad and unpredictable effects on replication (396; our unpublished data). Therefore, most mutations that hamper the release of nsP2 from the polyprotein precursor or those that reduce its overall quantity through lessened viral genomic mRNA production during replication, including any defects in minus-strand synthesis or failure to properly switch to plus strand production, may result in a reduction of the toxic effect (246, 359). After the completion of P1234 processing or replication complex breakdown, the mature nsP2 protein of most Old World viruses becomes translocated into the nucleus due to the presence of the nuclear localization signal in its C-terminal region (PRR^{649}RV in SFV) (265, 325, 326, 396, 436). This nuclear transfer is crucial for exhibiting the inhibitory properties of nsP2, whereas the proteolytic properties associated with the C-terminal part of this protein do not seem to be as relevant (106, 116, 117). Identified toxicity-reducing mutations, for example P726L or N779K in SINV, appear at the solvent-exposed loops in the structure and are therefore assumed to be important for as-yet-unidentified interactions with the cellular factors (252, 339). In any case, once in the nucleus, the excess of the solvent-exposed positively charged residues in this mostly basic (pI ~ 9)

protein and its composite organization suggest many opportunities for nonspecific contacts with nucleic acids and communication with potential partner proteins. Thus, armored with the nucleic acid structure-modifying tool and powered by its own energy supply, nsP2 appears to be a potentially very efficient transcription (interfering) factor. Not surprisingly, nucleus-residing nsP2 protein is also credited with antagonism of the induction of the interferon response during viral replication (33, 105, 108, 125, 363). Recent reports also suggest the involvement of nsP2 in interference with the interferon-stimulated Jak-Stat signaling pathway (111, 373).

NsP3 protein (482 aa in SFV) for a long period of time remained rather enigmatic in function. Mostly due to the lack of defined enzymatic functions and in analogy with some proteins from other viruses, the alphaviral nsP3 protein has been viewed simply as a possible platform that links the actively working units of the replicase complex (158, 345). However, this view is rapidly changing, and nsP3 appears to be an indispensable part of the replication complex, rather than an accessory protein for the improvement of viral life (78, 127, 202, 209, 437). As evident from sequence analysis, approximately 320 N-terminal aa residues of nsP3 are conserved among alphaviruses, and this region is predicted to fold into two distinct domains, whereas the C-terminal region of varied length is predicted to contain no regular structure and is actually hypervariable in sequence among different alphaviruses. The first N-terminal domain of nsP3 has a homology with the so-called macro domain found in many different organisms from archaea to mammals (145, 173); besides alphaviruses, macro domains appear in the proteomes of coronaviruses and several members of the alpha-like viruses, such as Rubella virus and HEV (192). Macro domain sequence and therefore overall fold is well conserved (297), accordingly, its structure can be reliably predicted using homologous modelling. Along these lines, the actual structures determined for macro domains, aa 2–158 of nsP3, from VEEV and CHIKV at the resolutions of 2.30 Å and 1.65 Å respectively, yielded few, if any, surprises (247). Biochemically, it was demonstrated that the CHIKV and VEEV macro domains are adenosine di-phosphoribose 1"-phosphate phosphatases, whereas for SFV the catalytic activity of macro domain was at the limit of detection, suggesting that this enzymatic activity is not important for the viral lifecycle (79, 276). The macro domain of alphaviruses was shown to bind ADP-ribose, poly-ADP-ribose (PAR) and RNA. The structure of the macro domain complex with an RNA oligonucleotide revealed that RNA binding occurs slightly differently compared with PAR binding, although the same conserved D10 residue of nsP3 was involved in the specific recognition of the adenine moiety (247, 336). Nevertheless, the surface of the alphaviral macro domain reveals the presence of patches of positively charged residues that allow for the nonspecific binding of negatively charged polymers irrespective of the described adenine crevice. It was therefore proposed that the macro domain could be a non-sequence-specific

RNA recruitment factor or adenine-containing RNA recruitment factor to provide the RNA template to the neighboring ns-proteins (247). It is rather difficult to imagine PAR as a true substrate for alphaviral macro domain because poly-ADP-ribosylation is a form of posttranslational modification synthesized by poly-ADP-ribose polymerases (PARP) in response to certain stress conditions, like chemical, metabolic or radiation-induced DNA damage resulting in single-strand DNA breaks (405). Macro domains, including those of histone macroH2A1.1, are recruited to sites of PARP1 activation, directly sensing poly-ADP-ribosylation and acting as a platform for the other DNA-repairing or chromatin-remodeling enzymes in the nucleus (129, 406).

The crystal structure of the complex of alphaviral macro domain with RNA oligonucleotide suggests that RNA indeed can be truly biologically relevant substrate (247). It is interesting, however, to see possible parallels with the SARS coronavirus macro domain, which is adjacent to SUD (SARS unique domain), whose structure revealed two sub-domains with a macro domain fold; strikingly, this structure was hardly evident from the analysis of the primary sequence (164, 397). SUD, in turn, appears to form dimers, meaning that SARS coronavirus uses the whole set of macro domain-like units. It was proposed that this combination is utilized to recognize specific RNA structures including G-quadruplexes (397). To this end, it is tempting to speculate that the revealed structure of the alphaviral macro domain is only half or less, if accounting for the possible involvement of the adjacent FtsJ-like domain of nsP2, of the full RNA-binding module, whereas another portion may be hidden in the central domain of nsP3. This domain appears to be unique for alphaviruses and has a suspiciously similar size to the macro domain itself, but whether it is folded similarly can only be revealed by future experimentation. In any case, the second domain of nsP3 is supposed to co-evolve with the macro domain and is probably implicated in the regulation of minus-strand synthesis, as concluded from the analysis of several ts mutants (202, 437).

It is of special note that, to fulfill viability requirements, the minimal replicase of an RNA virus in principle should contain a helicase-like activity, a polymerase and a single-strand nucleic acid-binding protein, whereas other functions that aid replication may be hijacked from the cell (discussed in (318)). If it would be a duty of the helicase to catalyze the opening of a dsRNA segment and to drive rearrangements ensuring that the ssRNA product ends up bound to another component of the macromolecular complex, then a set of helping factors are expected to exist (60, 61, 296, 467). A loading factor is expected to facilitate the initiation of the helicase reaction, which is *in vitro* shown to be far more efficient if ssRNA overhangs are present on the dsRNA substrates (82, 220). A trapping factor, a single-strand binding protein, is then obliged to stabilize the single-stranded intermediates as they are formed to facilitate elongation (317). To this end, nsP3 with its accessories represents a very probable candidate to play the role of single-strand-binding protein, thus presumably assisting the RNA helicase in its functions. Such activities would

thus be analogous to those of RNA chaperones, which are dedicated to loosening RNA structures or to disrupting RNA-RNA interactions, which are excessively found in viral untranslated regions as well, in an ATP-binding or hydrolysis-independent manner (55, 319). Curiously, several proteomic studies identified cellular proteins hnRNP A1 (139) and hnRNP K (36) being recruited to the sites of alphaviral replication. These proteins shuttling between the nucleus and the cytoplasm are usually transient components of various ribonucleoprotein complexes that are involved in different aspects of RNA metabolism (44, 181). They are also known for RNA chaperone activities (236, 319) that would then likely be advantageous for viral replication as well (462), although rather obviously, in most vital functions, alphaviruses would prefer to rely on themselves.

The variable region of nsP3 is separated from a folded N-terminal region by a stretch of aa residues rich in serines and threonines, which inside the cells becomes hyperphosphorylated (78, 223, 302, 429), presumably by casein kinase II (430). Several ts-mutations in SINV nsP3, namely ts4 (A268V) and ts113 (A68G), were found to lead to the production of underphosphorylated nsP3 and to specific defects in minus-strand synthesis (78). The complete removal of the hyperphosphorylated region in SFV resulted in a significantly decreased rate of RNA synthesis and greatly reduced the pathogenicity of SFV in mice (429). It is not clear at which stage of alphaviral replication cycle nsP3 becomes phosphorylated, although it should be noted that the introduction of such significant negative charge onto the surface of the replicase complex is hardly incidental and most obviously not without consequences (343, 453). With respect to the acquisition of negative charge by phosphorylation, mostly positively charged surfaces of the replicase could be a source of attractive interactions, whereas negatively charged RNA molecules could provide possibilities for repulsive contacts (161). In either case, the hyperphosphorylated region may be envisioned as an electrostatic switch that dynamically regulates the properties of the replication complex, possibly through the displacement of another negatively charged region, such as the one noted in the C-terminal region of nsP1. An interesting analogy can be found with the proposed system of electrostatic switches in the case of HCV replication, in which a peculiar, negatively charged region of the HCV NS4A protein appeared to act by switching the hyperphosphorylation level of HCV NS5A thus affecting the overall replication success (232). The phosphorylation of HCV NS5A was proposed to cause the dissociation of this protein from the otherwise stable RNA replication complex to promote formation of the NS5A complexes with specific host proteins (80, 155). Consequently, the removal or certain mutations of the phosphorylated region that were introduced to reverse the assumed destabilization of the replication complex appeared to specifically inhibit virion production, suggesting that the increased stability of the replicase complex is inconsistent with the progression to RNA packaging; therefore, the change of phosphorylation status appears to be a prerequisite step for switching between

these closely related events (218, 402). The applicability of such a hypothesis to alphavirus replication events has not yet been pursued but certainly deserves closer inspection in the future.

To further complicate the picture of multifarious functions of nsP3, it should be noted that the apparent lack of regular structure and sequence hypervariability of the non-conserved C-terminal region of nsP3 does not imply that it is unimportant. The aa composition of this region is remarkable, being rich in Ala, Pro, Arg and Glu residues, and, although predicted to be intrinsically disordered (330), the C-terminal part of nsP3 is certainly not resembling normal linker regions. Similar combinations of aa residues can be found in the C-terminal region of nsP3 in different alphaviruses, which hints at their possible employment as linear recognition motifs (67). A growing number of examples in cell biology, although relatively understudied in the world of RNA viruses, provide evidence that much of the molecular recognition of proteins can be attributed to short linear segments rather than to traditional protein-protein interactions involving globular domains (235, 300). These short linear motifs (SLiMs) are continuous regions of several aa residues that are relatively unconstrained in sequence, with few specificity determinants (58). SLiMs can provide micromolar affinity with their recognition modules, ensuring the formation of transient and reversible complexes, and are therefore amenable to the efficient control needed in signaling pathways (65). This alternative mechanism for protein-protein communication should be especially beneficial for viruses, given the spatial constraints imposed by their compact genomes and the almost unlimited possibilities to interfere with well-organized cellular life with just a few properly selected and positioned aa residues (57, 169). Often, the presence of SLiMs is connected to intrinsically unstructured regions of proteins, thus implying the involvement of such regions in the regulatory functions of signal transduction (114). From this standpoint, the low-complexity, non-conserved C-terminal region of nsP3 (and probably the C-terminal region of nsP1) can be viewed as an alphaviral depository of SLiMs. Although bioinformatic methods (67, 361) readily predict plenty of putative motifs that can potentially be recognized by host proteins, their true functionality and relevance of such interactions to alphaviral lifecycle remain to be discovered. To this end, a recent publication confirmed that the predicted SH3-binding motif within alphaviral nsP3 is functional and it can be employed to recruit amphiphysins, which are BAR domain-containing proteins, to the sites of alphavirus replication (277). Importantly, nsP3 was also reported to associate with membranes (303) and was found to be the main determinant of the subcellular localization of the alphaviral replication complexes (127, 345), whereas the region in the C-terminal part of nsP3 may specifically contribute to this activity (420). It can be inferred that the proper understanding of the ability of alphaviruses to use SLiMs to mimic the natural components of cellular complexes and reprogram the regulatory processes in the host cell may represent a gold mine for the future studies of alphavirus biology and antiviral research.

AIMS OF THE STUDIES

A large body of evidence suggests that the RNA template preferences and genomic or subgenomic RNA-synthesizing capabilities of the alphaviral replication complex have very important connections with its composition, which is irreversibly altered by proteolysis during replicase maturation. In the past, numerous research projects using various polyprotein production systems, such as pulse-chase labeling of alphavirus infected cells, vaccinia virus-driven expression and *in vitro* translation, devised the rules for when and how ns-protease acts. Alphaviral ns-polyprotein processing was found to be sequential and with a strict order of specific cleavages. The revealed characteristics of proteolytic processing within SFV P1234 showed that the 3/4 site is swiftly processed first, most probably co-translationally, whereas the 1/2 site is cleaved at a later time *in cis,* leading to almost simultaneous *in trans* cleavage of the 2/3 site. Nevertheless, an understanding of the molecular basis for the processing rules and the triggers of the specific cleavage of a particular site at the right moment of the replication cycle remained rather elusive. Therefore, the aim of this work was to elucidate the determinants of the proteolytic processing of the alphaviral ns-polyprotein using defined experimental systems involving purified recombinant proteins along with the fine positional scanning and to shed light on the possible organization and functionality of the replicase complex.

RESULTS

Molecular determinants of substrate recognition by SFV ns-protease

From prior studies by de Groot and Strauss *et al.* of SINV and by Merits and Vasiljeva *et al.* of SFV ns-polyprotein processing, we inherited a wealth of data that confirmed that nsP2 is the protease that is solely responsible for all three of the cleavages in the ns-polyprotein of alphaviruses (149, 150, 182, 259, 369, 389, 422, 423, 463). Mature nsP2 and active nsP2-containing polyproteins P123, P12 and P23, artificially stabilized in a non-processed form by mutating the penultimate Gly of the P-side of the cleavage site(s) to Val or Glu, were capable of cleaving all of the cleavage sites in presented polyprotein substrates, albeit with different efficiency (423, 463). Hereafter the designations P1-P1' are used to indicate the N-terminal (non-primed or P-side) and C-terminal (primed, P' side) regions of the substrate with respect to the cleavage position. Similar findings were obtained in experiments in which purified full-length SFV nsP2 or its protease domain Pro were challenged to cleave artificial substrates containing a virus-specific 20–20' sequence (20 aa upstream and 20 aa downstream of the scissile bond) fused to a thioredoxin domain. The conclusion was made that the SFV 3/4 site was recognized very efficiently, the 1/2 cleavage site was suboptimal for protease recognition and cleavage of the 2/3 site in short form could not be convincingly detected (422). Several studies also noted that in P1234, the 3/4 and 1/2 sites are cleaved independently from all of the other sites, whereas the cleavage of the 2/3 site requires the preceding cleavage of the 1/2 site to occur (150, 369, 423, 463). Assessing the sensitivity of the processing kinetics to the effects of dilution of the reaction components, it was concluded that the 1/2 site cleavage is insensitive to dilution and therefore occurs preferentially *in cis*, whereas the 2/3 cleavage is a *trans* reaction (150, 423, 463).

In early days of alphaviral ns-protease studies, Strauss *et al.* concluded from the analysis of the aa sequences surrounding the cleavage positions that a protease recognition site can indeed be as small as 3 aa residues, which are the only conserved elements in the protease cleavage motifs (390, 463). From sequence alignment, it can be concluded that the downstream regions of the cleavage sites are well conserved vertically, so that in different alphaviruses, the N-terminal regions of different ns-proteins generated by proteolysis are conserved, suggesting the conservation of the function of those proteins. In contrast, within a single alphaviral polyprotein, these regions have nothing in common (Paper I, Table 1), leading to the conclusion that the protease does not require the recognition of downstream regions. The upstream cleavage regions, in turn, have sequence similarity both vertically and horizontally. The main features of the non-primed region include the absolute conservation of Gly in the P2 position and a preference for small aa residues (Gly, Ala or Cys) in the

P1 and P3 positions. The P4 position in the cleavage sites of many viruses often contained a charged residue (Arg, His, Glu or Asp), and positions P5 and P6 allowed considerable variation (Paper I, Table 1). These observations were in line with a number of studies on many types of viral proteases, indicating that protease recognition sites are typically relatively short (usually 5–7 aa) and normally reside in non-primed regions (243, 279, 392, 418, 419).

Because the SFV 3/4 site can be efficiently cleaved in the form of an artificial substrate containing a short virus-specific sequence, it was used to map the minimal requirements in the substrate for efficient cleavage. To this end, a gradual decrease of the length of the presented virus-derived sequence, point mutagenesis and shuffling of half-sites of the substrates from the other (1/2 and 2/3) sites were employed. It was found that, in the case of the 3/4 site, SFV protease recognized 6 aa residues upstream of the cleavage position (237). The P1' residue had a significant influence on the efficiency of the cleavage: bulky hydrophobic or negatively charged residues were clearly unfavourable, Pro was not tolerated at all, whereas small residues (Gly, Ser), which apparently didn't interfere with the accommodation of the substrate in the recognition pocket, were preferred (Paper I, Figure 3). It also appeared that region P10-P7 has an influence on the efficiency of the cleavage (237), although elucidation of the extent of its effect was not pursued. Within the region composed of the 6 recognized aa residues, we found that positions P1 and P3 allow some variation, and the P5 position contributes less to the recognition and can be extensively varied, but the Gly in the P2 position is invariably constant, and the residue in P4 actually determines the overall efficiency of the cleavage (Paper I, Figure 2 and Figure 5). Experiments with the shuffled substrate half-sites generally confirmed the findings from the mapping and mutagenesis experiments (Paper I, Figure 4). Therefore, it was concluded that the 1/2 site or its analogues were poorly cleaved *in trans* due to the suboptimal His residue in the P4 position, which was partly compensated by the best possible residue (Gly) in the P1' position. In contrast, the 2/3 site was found to be deficient in signature motifs that would meet the minimal requirements for the recognition by the protease. Replacement of the Thr residue in the P4 position for optimal Arg rendered this site cleavable (Paper I, Figure 5); clearly, however, other elements of the recognition sequence were still far from the optimum to allow efficient cleavage. These results suggested that the P4 residue is dominant, yet not the ultimate, factor for cleavage proceeding; in fact, certain combinations of aa residues in the P6-P4 region supported by an optimum residue in the P1' position are the main determinants of proteolytic success.

Experiments with viral genomes containing mutations that were supposed to interfere with the normal regulation of processing by making cleavage sites correspond better or worse to the consensus according to the established requirements for proteolytic activity, in general supported the *in vitro* findings. The mutations that *in vitro* rendered the 3/4 site less efficiently cleaved (the change of the Arg in the P4 position to His or Thr, as in the 1/2 and 2/3 sites,

correspondingly) also slowed polyprotein processing during the infection cycle as evidenced by pulse-chase experiments, and slightly affected RNA infectivity (Paper I, Figure 6). The presumed improvement of the protease recognition sequence in the 2/3 site (through the mutation of the Thr in the P4 position to Arg, as found in the efficiently processed 3/4 site) resulted in a virus that made larger plaques; the mutation also hastened the growth of viral titer, whereas the final titer and processing in particular was rather indistinguishable from those of the wt virus (Paper I, Figure 6). Taking into account previous knowledge that, after the cleavage of the 2/3 site, the replicase changes its activities in favor of subgenomic RNA synthesis (182, 217, 394), it can be concluded that the unusually rapid 2/3 cleavage switched viral replication earlier to structural polyproteins and virions production, thus resulting in a larger plaques. Mutation of the penultimate Gly together with the P1 residue (GRAGA to GRAEV) in the 3/4 site to block its cleavage resulted in non-infectious RNA, emphasizing the ultimate requirement for the liberation of individual RdRp early in infection. Nevertheless, all of the other mutations in the cleavage sites, which were aimed to enhance or reduce their processivity, were well tolerated, and no reversions or compensatory second-site mutations were found in the region studied. This finding also suggests that the requirements in the protease recognition sequences are somewhat relaxed, which may explain their relative non-conservation.

It should, however, be noted that in the *in vitro* experiments, the 3/4 site cleavage was performed *in trans,* whereas in cell culture experiments with infectious virus, both *in cis* and *in trans* modes of processing can be involved. Two substrates containing the 3/4 site can be found in SFV infected cells – P1234 and P34 (Paper I, Figure 6; Paper III, Figure 1 and Figure 3). The *in trans* cleavage is the only way to process P34, and this process occurs efficiently in SFV-infected cells (Paper I, Figure 6); nevertheless, its significance is not obvious (to note, P34 of SINV cannot be processed in late infection and becomes accumulated in infected cells (463). The availability of the 3/4 site for *trans* cleavage does not, however, exclude the possibility that, in the process of P1234 processing, this crucial cleavage is carried out intramolecularly as it was suggested in early studies of SINV P1234. Theoretically, the cleavage rules that are valid for the *in cis* cleavage of the 3/4 site could be different from those described above; however, as the results obtained in the cell-free system and in infected cells were highly coherent, this is unlikely to be the case. Nonetheless, the *cis* mode of cleavage may serve to potentiate the 3/4 processing and contribute to the high infectivity of viruses with mutated 3/4 sites (Paper I, Figure 6). Obtaining direct experimental proof of this mechanism or, conversely, ruling out this possibility proved to be extremely challenging due to the large size of the P1234 precursor which, coupled to the high efficiency of 3/4 processing, did not allow us to perform dilution experiments in the manner previously emloyed for the 1/2 and 2/3 site processing reactions.

The ability of the nsP2 protease of SFV to process the 3/4 site was subsequently used to generate unique SFV marker viruses, in which the sequence encoding the marker protein became inserted between the nsP3 and nsP4 coding regions. The strategy of marker insertion into ns-protein region has been successfully used in alphavirus research, but in the described viruses the reporter or epitope tag has always been expressed as a fusion product with either nsP2, nsP3 or nsP4 (14, 29, 53, 54, 110, 127, 340). In contrast to these vectors, still preserving the four-ns-protein structure of the replicase region, the vectors constructed in our study utilized a five-ns-protein structure of the replicase region: the inserted EGFP marker was flanked by nsP2 recognition sites and is processed away from both nsP3 and nsP4 (Paper II, Figure 1). In these constructs, the cleavage site after the marker and before nsP4 (marker/nsP4) consisted of 30 P-side residues, which should be more than sufficient to achieve optimal cleavage, as described above. Indeed, this site was quite efficiently processed, and only small amounts of EGFP-nsP4 fusion protein were detected in the infected cells (Paper II, Figure 2). The truncation of the P-side of this artificial site to 10 or even 6 SFV-specific residues did not result in any significant decrease of the infectivity of these recombinant viruses (unpublished results), a finding that is highly consistent with the cleavage requirements outlined above. The cleavage site before the marker (nsP3/marker site) consisted of the full P-side of the 3/4 site and had an artificial P' side consisting of either a P1' Gly residue followed by the natural P2'-P7' residues of the 3/4 site, such a site was designated as “H” for “highly efficient cleavage site” (Paper II, Figure 1), or of the Gly-Pro dipeptide with no aa residues from the native P' side of the 3/4 site, such a site was designated as “L” for “lower-efficiency cleavage site” (unpublished data). The recombinant viruses carrying the “H” and “L” versions of the nsP3/marker cleavage sites turned out to be remarkably similar in infectivity as well as in their growth properties, and the processing of the “H” and “L” sites was consistently similar; only a minor reduction of the cleavage efficiency of the “L” site compared with the “H” site was observed (unpublished data). This finding was again fully in line with the revealed cleavage rules and indicates that the principles revealed in cell-free studies could be used in the design of infectious virus genomes.

The publication of the crystal structure of the protease domain of VEEV nsP2 (339) allowed the homologous modeling of the structure of the SFV protease as well. Later, additional structure-based analysis of the substrate recognition pocket was performed (338) by modeling of the accommodation of the consensus substrate peptide at the active site of the protease, using the structure of the Ulp1 cysteine protease co-crystallized with its substrate (267) as a template. Concerning SFV protease specificity, it can be concluded that the substrate recognition subsites are distributed between the two sub-domains of the protease domain Pro region (Figure 4). The model assumed that the alphaviral protease accommodates a recognized peptide in the substrate pocket in an extended beta-strand conformation. In particular, Russo and colleagues

suggested that residues Ser 513, Val 515, Trp 549, Met 707 and Asp 711 cooperatively accommodate the P4 residue of the substrate (aa positions are given in coordinate system of SFV nsP2). Notably, for some alphaviruses, a strong correlation is apparent between the charged S4 subsite residue (Lys 706 in VEEV and Asp 711 in SFV) and the oppositely charged P4 substrate residue (Asp in VEEV and Arg in SFV), which are proposed to form a salt bridge, implying that this interaction is the key determinant for substrate specificity. Therefore, the molecular modeling and structural analysis by Russo *et al*. supported our experimental data concerning the requirements for the specific recognition of a substrate by the alphaviral protease.

Because nsP2 is a single protease that is responsible for all three of the cleavages within the ns-polyprotein (259), it can be concluded that the accommodation of peptides representing other sites than the 3/4 site in a substrate recognition pocket of the protease should proceed in a similar way. Nevertheless, it remained completely unclear what force is driving the efficient processing of the 2/3 site in polyprotein context given that its immediate sequence adjacent to the cleavage position remains incapable of being recognized on its own (422; Paper I). A preliminary search for the elements responsible for the 2/3 site processing indicated that these elements are not localized in the nsP2 protease but instead in the N-terminal half of nsP3. Further experiments clarified that the 2/3 site is cleaved efficiently only if at least 165 aa from the P' region are included in the substrate (Paper III, Figure 1). This fragment constitutes the full macro domain present within the nsP3 protein. The overall fold of this domain is well conserved from archaeal to human proteins that possess it (145, 173). The important feature of this domain is that its C-terminus folds back to meet the N-terminus (79, 247), so that aa residues at the carboxy-terminal end of the macro domain have the possibility of contributing to protease recognition of the 2/3 site. In particular, the charged aa residues located in the C-terminal region of the macro domain (Arg 159 and Glu 163) were found to be necessary for the cleavage efficiency (Paper III, Figure 2 and Figure 3). Additionally, it was found that the N-terminal-most residues of the nsP2 protein are involved in the substrate recognition because even the most subtle deletions or extensions in this region severely affected 2/3 site cleavage success, whereas the enzyme represented only by the protease domain was completely impotent to perform this cleavage (Paper III, Figure 5). Taking into account the previously identified intramolecular mode of 1/2 site cleavage that obligatorily precedes 2/3 site processing (423, 463), it is immediately and appealingly obvious that the N-terminus of nsP2, which simultaneously represents the P' portion of the 1/2 site, has an opportunity to become placed in proximity with the macro domain, which comprises the P' side of the 2/3 site, and this increases the likelihood of their interaction. Given that a short sequence from the P-side of the 2/3 site apparently does not contain elements that permit it to be efficiently recognized by the protease substrate pocket, we concluded that interactions between the different domains of nsP2 and nsP3 result in an

assembly that directs the 2/3 cleavage sequence to the protease active center, thus overriding the lack of sufficiency of affinity-determining residues in the primary sequence preceding the cleavage position (Paper III).

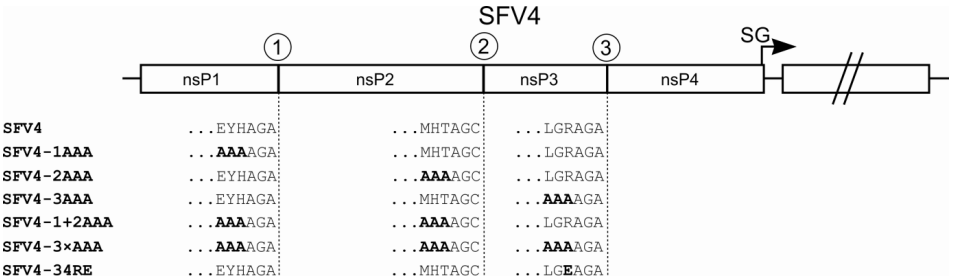
After finding that polyprotein processing is a complex function that uses parameters of two different types, such as the recognized sequence itself and its accessibility through the presentation implied by the configuration of the macromolecular complex, we next decided to assess the specific contribution of each constituent that influences cleavage efficiency for every site in the polyprotein. To this end, alanine scanning was employed to zero out the information needed for protease recognition that is contained in the side chains of the aa residues of the substrates, thus revealing the weight of the macromolecular complex-dependent component. Rather surprisingly, it was found that the placement of alanines in positions P6-P4 in the 1/2 and 2/3 cleavage sites, making the cleavage sites look like Ala-Ala-Ala-Ala-Gly-Ala and Ala-Ala-Ala-Ala-Gly-Cys, respectively, in the context of the P123 polyprotein had very little if any impact on the cleavage efficiency of the 2/3 site and only mildly affected the cleavage of the 1/2 site in the cell-free reaction (manuscript in preparation), suggesting that these cleavage sites are processed in a rather sequence-independent manner and that the factor of complex assembly in their recognition is of high value. In the same conditions the efficiency of the 3/4 site cleavage in the context of P1234 was affected more profoundly, yet even 3/4 site with the sequence changed to Ala-Ala-Ala-Ala-Gly-Ala, identical to the mutated 1/2 site, was also cleaved. Similarly, even the replacement of the P4 Arg with Glu, thus reversing the charge in the position previously shown to have a significant role in 3/4 site recognition, still did not completely block the cleavage, suggesting that an assembly-dependent driving force can, at least to some extent, direct even an inappropriate 3/4 site for proteolytic cleavage. Importantly, all of the mutant viruses containing cleavage sites affected by alanine mutations were still viable and, with the exception of the 3×AAA mutant, in which all cleavage sites in P1234 contained Ala-Ala-Ala sequences in the P6-P4 region, did not incur substantial losses in initial RNA infectivity, replicated to the same high titers as the wt virus and did not require additional second-site mutations to increase their infectivity (Figure 5A). In turn, the 3×AAA mutant revealed a severe drop in the initial infectivity but was able to restore its fitness to a great extent by acquiring just a single change in the protease recognition pocket itself, employing the mutation of Gln 706 to Arg (Figure 5A). The Gln 706 residue is immediately adjacent to the Met 707, which is the residue that was earlier proposed to be one of the subsites that are specifically involved in establishing contacts with the P4 residue of the substrate at the protease recognition pocket (338, 339). To this end, we hypothesized that the defect caused by 3×AAA mutation mostly originated from the disturbed cleavage of the 3/4 site, which processing is known to be greatly dependent on the cleavage sequence composition. To test this hypothesis, the P4 Arg, the main determinant of processing of the 3/4 site (Paper I), was mutated

to Glu in the context of the infectious clone of SFV in order to introduce presumably the most unfavorable aa residue at the P4 position, the opposite charge of which would cause electrostatic repulsion during cleavage sequence accomodation at the protease active site. Rather expectedly, the genome of the mutant harboring the P4 Arg to Glu mutation in the 3/4 site had an initial infectivity only twofold lower than the genome of wt SFV, which is in line with our previous findings (Paper I, Figure 6). The analysis of the sequences of rescued viruses revealed that, in contrast to the mutations described in Paper I, the P4-Glu was poorly tolerated: in approximately 50% of the rescued viruses the pseudo-reversion to Lys occured. Most importantly, it was found that the pseudoreversion was not the only way to overcome the defect – in several clones it was compensated in the similar way as in the case of the 3×AAA mutant: the same residue in the protease recognition pocket was changed, but in this case the Gln 706 was replaced with Lys (Figure 5A). This result suggests that, although the structure of the replication complex assists in the recognition of the 3/4 site in the context of the polyprotein and helps to tolerate unwanted changes in it, the cleavage of the 3/4 site is certainly more sequence-dependent than assembly-dependent.

Interestingly, although apparently sophisticated mechanisms exist to regulate the sequence of proteolytic events during the maturation of the ns-polyprotein of alphaviruses, the 1/2 and 2/3 cleavages themselves are still rather fitness-improving than vitally important and indispensable, because viruses with blocked 1/2 and/or 2/3 sites were found to be viable and, with the help of an adaptive Glu 452 to Ala mutation in nsP4 (a conserved residue among alphaviruses), replicated with essentially the same efficiency as the wt virus (Figure 5B; 125, 246); in contrast, genomes with a blocked 3/4 site are completely non-infectious (Paper I, Figure 6). This observation is consistent with the view that P123 (or its multimeric complex) constitutes the core of the viral replicative ribonucleoprotein; accordingly, the 1/2 and 2/3 site cleavages can be viewed as in-core cleavages, being rather immediate sequence-independent and replicase assembly-dependent. Therefore, their cleavage results from and/or reflects changes in the structure of the replication complex itself. At the same time, clearly, the nsP4 component of the replication complex is stoichiometrically underrepresented and can be viewed as an external, out-of-core component (175, 356, 464). This view is consistent with the experimentally observed properties of the reporter viruses described above: duplication of the out-of-core cleavage sites did not have detrimental effects on the ns-polyprotein processing and consequently on the virus viability (Paper II), whereas the addition of extra highly efficient cleavage sites into the in-core region is deleterious for the virus (Figure 5C; our unpublished results). Therefore, the liberation of nsP4 is more cleavage sequence-dependent and less assembly-dependent in case of SFV polyprotein, although the assembly-dependent component is much more profound in case of 3/4 site of SINV polyprotein (463, our unpublished data). Undoubtedly, however, that because

3/4 site processing signifies the success in the initial replicase assemblage, resulting in nsP4 stabilization by proper installation, and starts the engine of replication, both parameters of processing are of greatest importance.

A

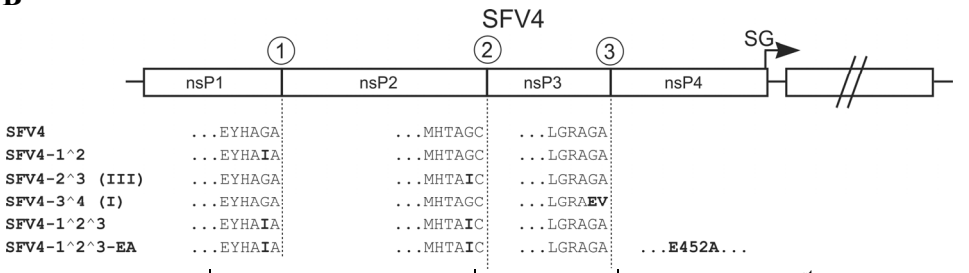


Construct	RNA infectivity, PFU/μg	Plaque size	Second-site mutations
SFV4 (wt)	$5.0 \cdot 10^5$	normal	not applicable
SFV4-1AAA	$2.4 \cdot 10^5$	small	none
SFV4-2AAA	$1.6 \cdot 10^5$	small	none
SFV4-3AAA	$5.0 \cdot 10^5$	normal	none
SFV4-1+2AAA	$8.0 \cdot 10^4$	small	multiple (found in nsP2, nsP3, nsP4)
SFV4-3×AAA	$3.0 \cdot 10^2$	minute	Q706R in nsP2, also in nsP4
SFV4-3×AAA-Q706R	$8.0 \cdot 10^4$	normal	not analyzed
SFV4-34RE	$2.5 \cdot 10^5$	normal	Q706L in nsP2*

** – or pseudoreversion of P4-Glu to Lys*
From L. Karo-Astover, V. Lulla, K. Rausalu, A. Merits and A. Lulla. Probing the plasticity of the alphaviral proteolytic activity by mutagenic analysis. Manuscript in preparation.

Figure 5. Analysis of the properties of recombinant viruses. **(A)** Analysis of viruses with triple alanine mutations in the cleavage sites and their derivatives.

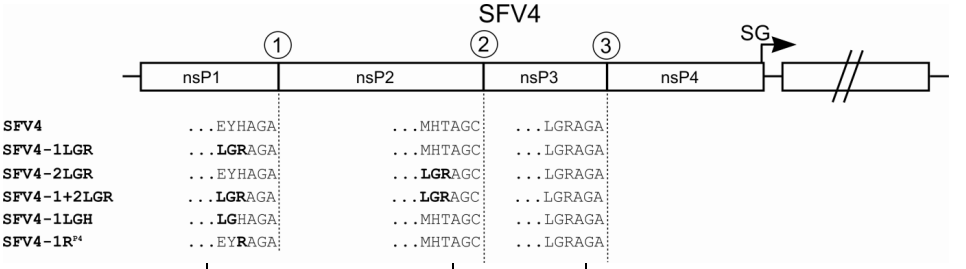
B



Construct	RNA infectivity, PFU/μg	Plaque size	Titer of the 1st passage, PFU/ml
SFV4 (wt)	$5.0 \cdot 10^5$	normal	$3.0 \cdot 10^9$
SFV4-1^2	< 1	n/a	$3.2 \cdot 10^3$
SFV4-2^3	16	small	$5.0 \cdot 10^7$
SFV4-3^4	0	n/a	0
SFV4-1^2^3	< 1	n/a	$1.1 \cdot 10^3$
SFV4-1^2^3-EA	96	small	$4.5 \cdot 10^7$

^ – indicates blocked cleavage site; n/a – not applicable; E452A mutation was copied from analogous SINV mutant (125).
*From **I**, **III** and unpublished data by V. Lulla and A. Lulla.*

C



Construct	RNA infectivity, PFU/μg	Plaque size	Second-site mutations / reversion / pseudoreversion
SFV4 (wt)	$5.0 \cdot 10^5$	normal	not applicable
SFV4-1LGR	$5.0 \cdot 10^2$	small	(pseudo)reversion of $1R^{P4}$ to H/C/S/L
SFV4-2LGR	$5.0 \cdot 10^5$	normal	none
SFV4-1+2LGR	$1.0 \cdot 10^2$	small	pseudoreversion of $1R^{P4}$ to C or 1LGRAGA → 1LGRAGV + V515M in nsP2
SFV4-1LGH	$4.0 \cdot 10^5$	normal	1LGH is preserved
SFV4-$1R^{P4}$	$5.0 \cdot 10^4$	medium	$1R^{P4}$ is preserved

From V. Lulla, L. Karo-Astover, A. Merits and A. Lulla. Intramolecular processing timer controls the replication of alphaviruses. Manuscript in preparation.

Figure 5. Analysis of the properties of recombinant viruses. **(B)** Analysis of viruses with blocked cleavage sites. **(C)** Analysis of viruses with processing-affecting mutations.

The relative insensitivity of the 1/2 site cleavage to substitutions in its sequence, however, does not imply that this sequence itself can be random. Our unpublished data suggest that reprogramming of the mode of cleavage of this site to be more sequence-dependent by introducing P6-P4 residues from the 3/4 site that are favorable for protease recognition dramatically affects the viral RNA infectivity (Figure 5C), possibly through deregulation of the viral lifecycle. For such mutated virus the only means of survival turned out to be in the acquisition of second-site mutations in the protease recognition pocket, like Val 515 to Met mutation of the residue proposed to be again involved in the recognition of the residue in the P4 position of the substrate, and/or different mutations in the 1/2 cleavage site itself, most often in the P4 position, to slow down the artificially accelerated processing (Figure 5C). Because premature cleavage was found to be highly detrimental, this observation suggests that 1/2 site cleavage should be viewed as a sort of molecular timer that controls the succession of specific steps in viral replication, most probably minus-strand synthesis. Completion of this particular step should potentially lead to corresponding conformational changes in the replicase core that, in turn, result in the proper orientation of the 1/2 cleavage sequence in the protease active site pocket and in cleavage occurrence. The 1/2 junction is thus a sensor in the transduction of a signal about the replication state, and its cleavage in its turn further transmits the signal that transformation of replicase core is being induced. This transmission is probably reflected in the new contacts of the nascent N-terminal end of nsP2 within the replicase core that result in concomitant 2/3 site cleavage that should then serve to stabilize new replicase construction, making its reconfiguration essentially irreversible and influencing the promoter-recognition properties of the viral polymerase subunit (182, 217, 359).

DISCUSSION

"For the marriage of proteolysis, a protease and a substrate are needed"

Proteolytic reactions are essential to all living systems, and proteases may constitute up to 4% of the gene products of an organism. Different sources of information report about 550 defined members and a further 100 or so predicted human proteases (243, 311, 312, 322). The relative abundance of proteases is especially striking in many small RNA viruses that encode approximately 10 gene products, of which 2–3 have protease functions (21, 71, 367, 384, 411, 427, 460). Essentially irreversible nature of peptide bond cleavage dictates that, although "for the marriage of proteolysis, a protease and a substrate are needed", as metaphorically suggested by Aaron Ciechanover (51), the proteolytic reactions must be tightly controlled. The relationships between a protease and its substrates therefore require a high level of trust, whereas the dysregulation of protease activity, causing disruptions in a balanced life- or infection cycle, may thus have severe biological outcomes incompatible with the very existence of the biological object. The question of greatest biological significance is, therefore, what exactly is required for a protease to exert its proteolytic activity towards a substrate?

Conventionally, a protease substrate contains a scissile bond, which resides between aa residues designated as P1–P1' (360). Proteases, in turn, contain a number of subsites, (respectively, S1, S2, etc. and S1', S2', etc.) that form a recognition pocket near the active site and provide varying degree of selectivity against particular aa residues in positions adjacent to the cleavage site. The intrinsic specificity of a protease is thus defined by the optimal peptide sequence that binds to the recognition pocket. Although certain positions in the vicinity of the cleaved bond may have mostly spacer value and need just to fit into the cleft, other substrate residues may have more prominent, affinity-determining roles. This relative importance (conservation) of certain aa residues in particular locations in the substrate provides a basis for the definition of the recognition motifs that are preferred in substrates (43, 230, 241, 342). Digestive enzymes, such as trypsin, apparently have relaxed requirements for substrate recognition to fulfill their task. The apparent non-specificity of processive proteases, such as Clp, Lon, etc., whose task is to complete the degradation of proteins, is, in turn, compensated by their incorporation into multi-subunit complexes so that specific targeting is imposed by separate recognition factors, the activity and presence of which can be controlled by intra- and/or extracellular signals (70, 134, 159, 353). In contrast, non-processive and, in extreme cases, single-hit proteases are believed to possess high specificity thresholds that are implemented into their structure. An important aspect of specificity that also should not be underestimated is the spatiotemporal co-localization of a protease and its potential substrate. To this end, it should be

noted that approximately 20% of human proteases are membrane-bound (243). However, the recognition of substrates by proteases is believed to be a more complex process than the stochastic, diffusion-determined search for potential short cleavage sequences fitting into the recognition cleft with an affinity to match (72, 83).

Theoretical explanations of enzyme selectivity have advanced greatly in transformations trying to cover different aspects of recognition specificity. The early "lock-and-key" concept of the regulation of enzymatic activity suggested the presence of rather uniquely compatible configurations (complementary shapes) in both the enzyme active center and the substrate for productive complex formation (89, 213). For practical reasons, simplified models of protease substrate recognition still utilize this "lock-and-key" concept, which implies that the side chains of aa residues constituting cleaved peptide act like the teeth of a key being accommodated in the keyway of the protease recognition pocket, which is formed by a number of subsite aa residues. This assumption of the association of predetermined rigid shapes later transformed into the induced fit concept (194), suggesting that the binding of the substrate promotes the conformational reshaping of the enzyme active center. The substantiation of this idea led to its evolution into the now generally appreciated "conformational selection" paradigm of molecular recognition (31), which postulates that all protein conformations pre-exist and the ligand initially selects the most favored conformation and subsequent optimization of the backbone and side-chains interactions proceeds via an induced fit mechanism. The fundamental requirement of conformational flexibility for this theoretical model has found support from numerous experimental findings of the structural plasticity of enzymes and their substrates (141, 294).

The description of protease specificity through the definition of a consensus recognition motif accounting for optimal, acceptable and disallowed residues in an aa sequence pattern is not a trivial task because highly specific proteases normally have very few natural substrates and this set of data is often insufficient to produce valuable guidelines for the prediction of proteolytic activity towards yet-unidentified substrates (20, 76, 379, 424, 428). Fortunately, for several proteases of the highest biological importance, the accumulated set of data allows for far-reaching conclusions that are believed to have broad applicability (188, 189). Viral proteases are excellent examples of enzymes with the highest specificity, and the HIV-1 protease, in particular, is an enzyme that has been studied in great detail. In addition to the 10 native cleavage sites present in the HIV gag and gag-pol polyproteins, libraries of hundreds to thousands of artificial sequences were screened to probe for specificity determinants and to define the extent of the promiscuity of HIV-1 protease (189, 456). The analysis of a number of structures of HIV-1 protease co-crystallized with inhibitors or peptides representing native substrates led to the proposal of the "substrate envelope" hypothesis to explain HIV-1 protease selectivity (292, 309). It was concluded that, despite significant diversity in the

primary sequences, the native protease substrates tended to adopt a particular shape and conform to a common volume within the substrate recognition crevice. Accordingly, it was proposed that HIV-1 protease recognizes the shape of the substrate instead of particular aa residues and that those protease inhibitors that have atoms protruding from the consensus volume lead to the appearance of drug-resistance mutations in the corresponding specificity subsites. This appealing concept was successfully implemented in the further development of tightly binding inhibitors that also showed a reduced propensity to induce new resistance mutations and a high tolerance to the existing mutations (11, 46, 170, 186, 273, 274, 308). Thus, functional and structural studies of the basic properties of viral proteases were successfully used for the development of improved antiviral compounds and likely can be used in the case of other viruses as well (170, 328, 329, 452, 468). However, the situation with real substrates, which, unlike the protease inhibitors, must not only be bound but also cleaved, is even more complex (211). More extensive modeling of substrate discrimination principles using computational peptide docking of a large set of previously identified cleavable and non-cleavable substrates into the crystal structure of a protease allowed for the further evaluation of the substrate envelope hypothesis and in general confirmed its usefulness, although pointing to several limitations as well (45, 273). It was found that, indeed, the structural mechanism of substrate discrimination was primarily through steric interactions within the protease-substrate complex; however, although staying within the substrate envelope is necessary, it is not sufficient to determine the cleavability of a particular peptide. To this end, it was found that, aside from binding energetics alone, the ability of a peptide to adopt the appropriate catalytic geometry (correct angles, especially at the cleavage position between P1 and P1') is crucial for cleavage success (273).

The structural presentation of the cleavage sequence to the protease active center is another crucial aspect of recognition because the cleavage position should become available for the proteolytic reaction to occur and, not surprisingly, proteases prefer cutting in inter-domain regions represented by flexible, solvent-exposed loops. Such properties of proteases were extensively used historically for mapping domain boundaries by performing limited proteolysis analysis (68, 92, 156). Unfortunately, flexibility and intrinsic disorder also represent major inherent obstacles for protein crystallization and/or assigning electron density information, thus complicating the comparative analysis of the conformations of substrate residues in their native macromolecular context before and after interaction with a protease active site. However, a thorough analysis of more than 1500 crystal and solution-state nuclear magnetic resonance (NMR) structures of substrates, products and inhibitors bound to the active sites of all classes of proteases led to the conclusion that, universally, a protease substrate pocket accommodates ligands in an extended beta-strand conformation (243, 418). It is intuitively obvious that the structural presentation of substrates in an extended conformation should be

important for proteolytic enzymes with deep, narrow active site clefts which, along these lines, would sterically restrict access for bulkier substrates, whereas this requirement may be alleviated for proteases with broad and exposed substrate pockets. Nevertheless, several existing exceptions contradict this central dogma of proteolysis, and the most notable of these exceptions document the observed cleavages of apparently well-structured alpha-helical substrates identified in whole-proteome limited proteolysis (407). Currently, it is not quite clear whether a local unfolding of the helical substrate or a reshaping of the active site itself grants access to the catalytic center, although it is likely that both processes contribute in this dynamic conformer selection process.

The very conformation of an extended beta strand suggests that the residues on one side of the substrate will have more intimate contacts with the pocket subsites. Not surprisingly, therefore, that apparent periodicity can be observed in the conservation of the preference of substrate residues at even or non-even positions (279). For example, in the case of papain-like proteases, positions P2 and P4 are more conserved and contribute more to the binding efficiency than the other residues in the substrate (49, 136, 360, 416, 417). However, despite few salient features in the sequences of cleavable substrates, there are strong indications of a high degree of interdependence of various substrate residues (279, 293). This interdependence makes the combinatorial positional screening of potent residues context-dependent and thus decreases the accuracy of predictions of potent substrates based on its results (188, 190). Although the aa distribution may obviously affect the cleavage site conformation, the inter-dependence of the substrate residues is underlined by the fact that the specificity subsites in the protease act cooperatively, and usually several residues with high conformational flexibility in the protease active site pocket constitute a subsite for the recognition of a particular substrate position (25, 32, 137, 283). Therefore, the replacement of one residue in the substrate may lead to compen-sation for the loss of the previously preferred links by a redistribution of van der Waals contacts so that another residue forming a subsite pocket of the protease takes more responsibility for binding the mutated substrate residue (292, 468). Additional evidence suggests that, for a particular protease that is entitled to cleave similar yet distinct cleavage site sequences, substrate cleavage may induce changes in the protease active site conformation that may impact the subsequent recognition of another substrate (174, 294). Thus, with the exception of extreme cases, it can be expected that it would be rather difficult to predict the extent of cleavability of a modeled substrate from protease crystal structure alone (198, 293).

The analysis of substrate specificity becomes even more complicated when the protease being embedded in the polyprotein is cleaving sites within it, as is common for viral proteases. It is a common practice for plus-strand RNA viruses (21, 455, 460), retroviruses (306), pararetroviruses (413) and even some DNA viruses, *e.g.,* African swine fever virus (374), to employ proteases whose

working responsibilities include the release of stoichiometric quantities of virus-encoded proteins and control of the maturation of virus-produced complexes involved in genome replication and packaging (384). Not all of the proteases involved in the processing of viral polyproteins are virus-encoded; however, host-encoded proteases, so-called proprotein convertases (13, 52, 272) are largely involved in the processing of virus-encoded structural polyproteins or the structural parts of the single polyprotein (266, 351), whereas the crucial cleavages within the ns-polyproteins are usually carried out by the virus's own enzymes (71, 411).

In the case of plus-strand RNA viruses, ns-polyprotein precursors produced directly from the genomic RNA are processed in an ordered cascade of cleavages; certain cleavage sites are co-translationally processed immediately, others are post-translationally processed in a delayed time frame by one or several proteases, depending on the virus. Viral ns-polyprotein processing may release up to sixteen mature proteins and a corresponding number of processing intermediates, some of which may have functions that are different from those of the mature proteins and are, therefore, biologically significant (388). Historically, the view of the importance of the polyprotein-based expression strategy has evolved considerably since its discovery in 1968, when it was presumed to represent a method by which viruses avoid the premature activation of their proteins, by analogy to the activation of zymogens, or a way in which mammalian viruses compensate for the lack of internal initiation of translation or overcome the restrictions resulting from the monocistronic nature of mammalian mRNAs (160, 393). Since then, studies carried out over four decades have shown that, although both of these assumptions were correct, many other important aspects exist.

Alphavirus nonstructural protease – viral "signaling scissors"

The results that form the basis of this thesis, corroborated by numerous studies by others, lead to the suggestion that the degree of mutual interdependence between a viral protease and its cleavage sites embedded in the viral polyprotein extends beyond the standard enzyme-substrate relationships, and thus, the complexity of its origin and purpose should be taken into account to explicate protease functionality, which should thus not be explained in the simple terms appropriate for a binary system (18, 446, 449). These complications arise from the mere fact of the intimate co-evolution of the protease and its substrates in a single body of a polyprotein (186, 376, 468), which apparently led to increased numbers of inter-protein contacts and enlarged areas of interacting and interpenetrating surfaces. To this end, the mode of proteolytic cleavage in such a complicated system as the viral replicase complex can hardly be strictly defined in simple *cis*/*trans* terms because the *trans* cleavage of one ns-polyprotein

by another can be regarded as *cis*, if viewed in the context of the whole multi-protein complex, in which every single polyprotein or its intermediate processing products form only a fraction of the apparently multimeric assembly residing on the viral RNA template (22, 48, 383, 400, 401). Importantly, this incorporation into a complex structure, which is usually membrane-tethered or even membrane-enwrapped (275, 282, 344, 465), also implies that protease functionality is apparently restricted in its independent decision making. Instead, the possibility of establishing multi-point intra-complex contacts, many of which can be viewed as exosite interactions, providing secondary binding platforms in addition to the confined area of contacts between the protease active site and the cleaved sequence (30, 43, 83, 113, 195, 290), provides vast possibilities for the allosteric regulation of enzyme conformation and thus its activity, as well as for the steric prevention or, on the contrary, increase in affinity of substrate binding. Apparently, most of the regulation triggers can be non-obvious and may arise from subtle changes in a distant part of the complex, inducing sufficient changes in the protease configuration to transduce slight conformational fluctuations into a more powerful signal through the structural transformation imposed by the proteolytic reaction. Therefore, the pattern of proteolytic events during viral ns-polyprotein maturation by processing can hardly be explained simply by the differences in the cleavage efficiencies of the short peptides representing the cleavage sites (211, 305). Somewhat counter-intuitively, it becomes evident from this work that natural virus-derived immediate cleavage sequences, instead of being expected to provide researchers with hints about protease recognition preferences, are leading towards the opposite direction, as they appear to be fairly inefficient when taken out of context. In this regard, the cleavage sites within SINV ns-polyprotein represent apparently extreme case, since all of them appear to be invisible to their own protease in a short-sequence context (458). However, this situation should not be considered surprising, given that, in many cases, the substitution of a sequence-dependent mechanism of substrate recognition for an assembly-dependent strategy could potentially lead to the extinction of recognition signatures in cleavage sites as becoming redundant. Additionally, the inability to perform sequence-dependent cleavage emerges as an additional precautionary measure to secure the improbability of premature accidental cleavages and appears to be not only evolutionarily justified but rather actively naturally selected.

Proteolytic regulation was once named “signaling scissors” (269), reflecting the ultimate role that proteases play in cellular signal transduction. Viruses, although much more compact, are no less complicated, and it would be naive to consider the viral protease as a protein slicing machine instead of regarding it as an integral part of a sophisticated regulatory mechanism in signaling network. In line with these considerations, an appealing analogy was proposed to consider ns-protease in the context of viral replicase as a watchdog awaiting a command to bite (Paper III). This virus wellness guard is kept on a rather short

"chain," being embedded into the multi-protein complex, and is fed on schedule when further transitions in the configuration of the replication complex become necessary, and thus, the proper signal to acquire the active state for the protease is given (e.g., by allosteric influence or by relief from inhibitory interactions or steric hindrance) or when the desired protein segment is forcedly targeted for proteolysis through exosite-driven interactions that virtually put the protein piece into the "mouth" of the protease to take a bite. To recapitulate, the replication complex is a dynamic system with many moving parts thus, it can be envisioned that the regulatory role of the protease is to monitor the succession and completion of the events of viral infection and to respond with a cleavage once the respective conformational changes in the complex allow the presentation of the scissile bond and other essential determinants. In this sense, the correct completion of the translation of the viral polyprotein, the capture of viral messenger RNA and its reassignment for genome replication, the proper assembly of the components of the ribonucleoprotein complex as it acquires the desired quaternary structure, the association with cellular membranes and subsequent enwrapping of the replication complexes into spherular structures, the completion of the replicative intermediate RNA synthesis and other similar events, which are most certainly accompanied by conformational transitions, may thus serve as valid signals for the protease to perform cleavage and to shift the whole system to a new level.

Alphavirus ns-protease is designed to function in the context of an ever-changing genetic environment

The readiness for reshaping of its configuration and the amenability to external regulation of a viral protease should be inevitably reflected in its structural properties. Although direct evidence for possibilities of significant conformational changes of the alphaviral protease is still lacking, some analogies can be drawn from the studies of other protease-substrate relations which suggest that proteases may even "perform yoga" (291) for substrate quality control, whereas "inchworm" (332) or "ratcheting" (28) model mechanisms were sometimes required to explain protease preparedness to perform distinct and/or sequential cleavages within the polyprotein (142). Viral genomes have evolved to withstand the constant mutagenic pressure resulting from high error rate replication (69, 210, 287). They have also adapted not only to tolerate these random changes but to make use of them (152, 251): too little variability is as unfavorable as too much variability (error catastrophe, (56)). Recent observations suggest that viral proteins, especially those of RNA viruses, tend to possess unique biophysical features: they have many disordered regions and loosely packed cores, which increase their structural flexibility (451). It has been proposed that these properties are important for rapidly evolving viral sequences, providing "gradient robustness" to withstand mutational pressure

(408). This same flexibility may also be necessary to allow a protein to be involved in numerous interactions with other proteins (408, 451). Important properties are thus implemented in the very structures of viral proteins, and manifestation of this strategy in alphaviral protease seems apparent in the shallowness of the substrate-binding pocket that is shared between two subdomains constituting the protease and therefore has a high potential for conformational flexibility.

Although macromolecular complex-imposed interactions apparently contribute to the possibility of primary sequence-independence of substrate recognition by alphaviral protease, it is still vital to know about the intrinsic preferences of a protease specificity pocket to bind specific side chains, because potent protease inhibitors are most frequently mimicking the structure of a peptide substrate (12, 27, 230, 263, 388). It can be expected that, in principle, high-affinity competitive protease inhibitors can to some extent control the stages of viral replication that are dependent on the proteolytic cleavages occurring outside of the replicase core; however, the only such cleavage site in the alphaviral ns-polyprotein does not appear to be attractive enough in this respect because even the minute amounts of the matured polymerase may be sufficient to produce inhibitor-resistant mutant genomes. Importantly, however, viral protease functionality is commonly not restricted in performing cleavages only within viral polyprotein, but proteases are extensively exploited for efficient subversive actions against innate immune defense mechanisms and for the restriction of host macromolecular synthesis to gain an advantage for the viral activities (42, 91, 163, 199, 226, 260, 295, 299, 314, 439, 454). In performing these actions, free-range protease liberated from the mature ns-polyprotein should be considerably more dependent on its recognition pocket abilities to select potential substrates, as macromolecular assembly-driven mechanisms are not applicable, and even the exosite-dependent component of recognition becomes fairly unreliable, although several exceptions arguing against this assumption are known, for example, Lb^{pro} of picornaviruses (90, 91). To this end, even if the peptide mimetics that are being developed based on the intrinsic properties of a substrate pocket, and therefore on the substrate recognition preferences of the protease, will predictably be rather inefficient for the control of intra-core cleavages within the viral replicase, they may be indirectly efficient antivirals, acting through the relief of the host from viral protease-dependent shutdown and providing the host defense a window in time to fight back (105). In this context, it should be noted that even with inhibitors that are created using comprehensive approaches of molecular design with best-developed concepts like “substrate envelope” in mind to avoid protrusion from the consensus volume, the practice of inhibitor use suggests that, sooner rather than later, viral resistance eventually emerges. Resistance mutations tend to arise first in the protease active site pocket, resulting in changes in substrate recognition specificity and thus serving to exclude or diminish inhibitor binding (10, 144). Importantly, it appears that the perturbations do not remain local, as it

appears from analysis of HIV-1 protease-inhibitor complex structures that the active site is in fact a highly plastic and interdependent binding site, so that most protease residues seem to adapt to the shape of inhibitor by combining side-chains and backbone rearrangements throughout the enzyme, thus influencing both specificity and affinity, sometimes in an unexpected manner (86, 97, 270, 273, 286, 366). Subsequently, adjusted protease sequence specificity often requires the optimization of viral cleavage sequences through cleavage site mutations (19, 185, 186, 328). The data presented in this thesis (including the unpublished data) indicate that this will more than likely be the case also for alphavirus protease.

Recent research has highlighted an intrinsic capacity of even exquisitely specific enzymes for promiscuous functions, which form a basis for evolutionary adaptability (17, 157, 177, 381). The promiscuous activity of protease shares the catalytic mechanism and the active site features with native activity but differs in substrate accommodation, thus allowing substrate ambiguity. Directed evolution experiments, aiming at further evolving of enzymatic promiscuous function, conclude that enzymes may exhibit high plasticity, whereas a single or few mutations may greatly increase promiscuous activity, leaving the catalytic efficiency largely the same (2, 176, 346). In real life, however, finely tuned protease specificity is evolutionarily shaped by positive selection for turnover rate and negative pressure to suppress otherwise deleterious nonspecificity (86). Our experiments aimed at probing the plasticity of alphaviral substrate recognition by mutagenic analysis suggested that the negative trade-off for evolving altered substrate specificity appears to be low for the alphaviral protease, which often can accept native and mutated cleavage sites without compromising the infectivity of the viral genome (Figure 5A, C; manuscript in preparation). Subsequently, the remarkable plasticity of the protease subsite residues at the perimeter of the active site allows the easy adaptation of proteolytic function against mutated cleavage sites by the acquisition of few mutations of protease subsite residues. Additionally, our data demonstrated that the plasticity not only serves to increase the tolerance to mutations but also provides a means for finding new, effective solutions for the problems created by the changes by introducing few fitness mutations, some changing the conformation of the substrate pocket and others possibly being involved in epistasis. Apparently, the constant balancing of plasticity and robustness is a key to the survival of ever-evolving viral species. The real-life situation actually appears to be even more complicated for protease inhibitors to persist in their actions in the reality of existence of viral quasispecies (327). The theory suggests that, given the error rate of viral polymerase and lack of its proof-reading, every single genome copy of a standard RNA virus of approximately 10,000 nucleotides during replication in the cell may deviate from consensus sequence by at least one mutation (75, 347). To this end, recent massively parallel sequencing studies suggested that potential protease resistance mutations not only can arise upon inhibitor treatment but in fact

already exist in several copies of the replicating genome, thus leading to the conclusion that viruses are already prepared to escape even from inhibitors that are not yet invented (34, 197, 350, 426, 432). However, despite these rather skeptical conclusions, potential alphaviral protease inhibitors may nevertheless become significant, especially in combination therapy approaches.

In conclusion, the results of the studies presented in this thesis point to the existence of yet another, but hardly ultimate, level of complexity in the regulation of alphaviral proteolytic function, which is inevitably linked to the genesis, dynamic transformation and fate of the viral self embodied in the replication complex of the virus. I believe that revealing of ever-increasing complexity in the lifecycle organization of even such a small object as a virus will never be boring, just as long as we continue to ask the right questions

SUMMARY

Alphaviruses from the *Togaviridae* family are plus-stranded, enveloped RNA viruses that are transmitted by insect vectors and infect vertebrate hosts. Among almost thirty catalogued alphaviruses, Semliki Forest virus (SFV) and Sindbis virus are known as the best models for the studies of viral replication strategies and have a well-earned reputation as gene delivery, viral vaccine and expression vectors.

The alphaviral genome is a plus-strand RNA of approximately 12kb, which contains two open reading frames, the first of which is directly employed for translation to produce virus-specific components of the replicase complex. The SFV replication strategy relies on the production of replicase proteins initially in the form of non-structural (ns) polyprotein precursor P1234, which during the course of replication becomes proteolytically processed by the virus-encoded nsP2 protease in a temporally regulated manner to eventually release four ns-proteins. Mature ns-proteins of alphaviruses exhibit multiple activies that are vital for the replication process, whereas polyprotein processing intermediates evidently possess functionalities different from mature proteins and are, therefore, biologically significant as well. Temporally controlled ns-polyprotein processing thus defines regulation of the viral RNA replication presumably by causing rearrangements of the viral replication complex, so that its RNA template preference is changed in favour of minus-strands over plus-strands during the course of viral infection. Numerous studies have devised rules for when and how ns-protease acts at processing of P1234 polyprotein. To this end, it was revealed that during maturation of the replicase complex the 3/4 site cleavage is the first to take place, which later as infection progresses is followed by in *cis* cleavage of the 1/2 site, in turn triggering concomitant in *trans* cleavage of the 2/3 site. The molecular grounds behind the processing rules have been largely unexplored and the true understanding of the essential triggers leading to sequential order of processing remained rather elusive at the starting point of the presented studies.

Initial experiments that were aimed at dissecting the requirements for the efficient in *trans* processing of the SFV 3/4 site utilised assays that employed recombinant protease and substrate variants. These analyses revealed that the main determinants of the cleavage efficiency are located in the region preceding the scissile bond and that the protease recognizes at least the residues P5 to P1′. In this context, the amino acid (aa) residues in positions P5 and P1′ tolerated much of variation without compromising cleavage efficiency, small aa residues were preferred in positions P1 and P3, P2 Gly was concluded to be invariant and the properties of the residue in P4 position modulated the efficiency of processing to the largest extent. The results obtained in the cell-free test system correlated with those obtained using cells infected with the targetedly modified viruses, although the effects of the introduced mutations on the viral replication in the infected cells were significantly less profound. The knowledge about the

SFV protease preference in processing the 3/4 site was subsequently used to generate unique SFV marker viruses, in which the sequence encoding the marker protein flanked with protease cleavage sites became inserted between the nsP3 and nsP4 coding regions. Furthermore, it was found that, besides the optimal short sequence immediately adjacent to the cleavage position, additional factors should exist to assist the spatiotemporal regulation of proteolysis.

Analysis of the requirements for the efficient processing of the 2/3 site, which evidently lacks primary sequence elements in the vicinity of the scissile bond sufficient for efficient cleavage, revealed that the specific factors residing outside of the protease domain determine the substrate recognition and thus decide the outcome of proteolytic reaction. These findings supported a model in which structural elements within the C-terminal region of the macro domain of nsP3 are used for precise positioning of a 2/3 substrate cleavage sequence at the catalytic center of the protease and that this process is coordinated by the exact N-terminal end of nsP2, thus representing a unique regulatory mechanism used by alphaviruses. The conclusion that correct macromolecular assembly as well as transformations within the multi-protein complex, presumably strengthened by exosite interactions rather than the functionality of the individual nsP2 protease, emerges as an indispensable driving force for specific substrate targeting, paved the way to subsequent experiments, in which positional scanning, aimed to zero out cleavage site sequence-dependent effects on efficiency of substrate recognition, was employed to assess the extent of assembly-dependent recognition. The obtained data suggested that the macromolecular assembly-dependent processing undoubtedly contributes to the capability of the virus to survive and adapt even to the situation, in which all affinity-determining residues in all three cleavage sites of P1234 were knocked-out by alanine mutagenesis, essentially preserving the order and timing of the processing or restoring these parameters through adaptive non-cleavage site mutations. Nevertheless, the primary sequences of 1/2 and 3/4 sites appear to be evolutionarily selected to ensure their slow/delayed or swift cleavages, respectively. Therefore, interventions into these cleavage sites' sequences in order to delay 3/4 site cleavage or, especially, to artificially speed up the cleavage of 1/2 site were found to cause seriously detrimental effects on the viral infection. Consequently, such defects were not tolerated in the viral genome and required corrections such as (pseudo)reversions, mutations in the cleavage sites or in the subsites of the substrate recognition cleft in the protease domain.

Combined, the results of the studies described in this thesis pointed to the conclusion that the processing of the SFV ns-polyprotein is a complex process regulated by the multi-layered control system that prevents from premature proteolytic events and guarantees timely cleavages according to the requirements of the particular stage of viral replication. The alphaviral protease therefore emerges as an integral part of the sophisticated signaling mechanism, in which the regulatory task of the protease consists of monitoring the

succession and completion of the events of viral infection. Once the respective replication status-induced conformational changes within replicase allow the presentation of the scissile bond and/or other essential determinants of substrate recognition like exosites, the local protease signaling is initiated, which apparently leads to reconfiguration of the viral replication complex. Importantly, the studies unveiled the decisive role played by the macromolecular assembly-dependent component of substrate recognition in addition to the sequence-dependent component, the combination of which may be expected to constitute the basis of regulation in multi-site proteolytic systems in general. Further evaluation of this concept may provide valuable insights into the proteolytic regulation of replication in other viruses, a gene expression strategy of which relies on the production of viral polyprotein precursors. Indeed, this work has confirmed that the intra-replicase contacts are at least of the same importance as the individual properties and activities of matured viral proteins. Described findings and their interpretations are, thus, expected to provide with essential grounds and directions for further studies on the restriction of alphaviral replication through affecting the center of viral proteolytic activity or via intervention with its regulation by targeting intramolecular interactions that determine the stability and functionality of the viral macromolecular assemblies.

SUMMARY IN ESTONIAN

Alfaviiruse mittestruktuurne proteaas ja tema liitvalgust substraat: täiuslikult korraldatud kooselu reeglid

Alfaviirused (sugukond *Togaviridae*) on positiivse RNA genoomi ja lipiidse ümbrisega sfääriliste virionidega viirused, mis peale oma selgroogsete peremeeste nakatavad ka putukatest ülekandevektoreid. Ligikaudu kolmekümnest teadaolevast alfaviirusest on enim uuritud Semliki Forest viirust (SFV) ja Sindbis viirust, mis on kujunenud olulisteks mudelsüsteemideks ning tööriistadeks geeniülekandes ja -vaktsineerimisel.

Ligikaudu 12 kb pikkune genoomne RNA sisaldab kahte avatud lugemisraami, millest esimest transleeritakse otse genoomselt RNA-lt ja mis kodeerib replikaasi viirus-spetsiifilisi komponente. SFV puhul ekspresseritakse need komponendid ühe liitvalgu, P1234 polüproteiini, kujul ja see lõigatakse polüproteiinis paikneva nsP2 proteaasi poolt ajaliselt täpselt reguleeritud protsessingu käigus valmis mittestruktuurseteks (ns) valkudeks (nsP1-nsP4). Kõik ns-valgud on multifunktsionaalsed ning olulised viiruse infektsiooniks. Peale selle on varasemad uurimised näidanud, et ka P1234 polüproteiin ja selle lõikamise vaheproduktid omavad bioloogilisi aktiivsusi, mis valmis ns-valkudel puuduvad, ning on samuti infektsiooniks olulised. Seetõttu on erinevaid katsesüsteeme kasutades ns-valkude ekspressiooni ja valmimist põhjalikult analüüsitud. Need tööd on näidanud, et ns-polüproteiinide ajaliselt kontrollitud protsessing koordineerib viiruse infektsioonitsükli kulgemist, aktiveerides alguses viiruse RNA replikaasi ja muutes seejärel tema spetsiifilisust nii, et replikaas hakkab sünteesima viiruse positiivseid RNA-sid. Nende tööde tulemusena on kindlaks tehtud ka P1234 lõikamise reeglid. Nii on näidatud, et replikaasi moodustumisele viiv protsessing algab viimase (3/4) saidi lõikamisest, millele järgneb 1/2 saidi *in cis* lõikamine ning et see sündmus käivitab viimasena toimuva 2/3 saidi *in trans* protsessimise. Samas olid sellise lõikamiste järjekorra taga seisvad molekulaarsed mehhanismid käesoleva töö alustamise ajal suurel määral tundmatud.

Käesolevas väitekirjas kirjeldatud tööde käigus iseloomustati kõigepealt SFV 3/4 saidi protsessimise mehhanismi. Rekombinantseid substraate ja ensüüme kasutades näidati, et selle saidi lõikamise peamised determinandid paiknevad lõigatavast peptiidsidemest ülalpool. 3/4 saidi puhul tunneb proteaas ära vähemalt kuute aminohappe (ah) jääki, mis paiknevad positsioonides P5-P1'. Nende ah-jääkide roll ei ole aga ühesugune: kui P1' positsioonis (aga ka P5 positsioonis) paiknev ah-jääk mõjutas lõikamise efektiivsust suhteliselt vähe, siis P1 ja P3 positsioonides olid tugevasti eelistatud väikesed ah-jäägid. P2 positsioonis oli absoluutselt vajalik glütsiinijääk ja protsessingu efektiivsuse peamiseks mõjutajaks osutus P4 ah-jääk. Saadud tulemuste kontrollimisel

nakatatud rakkudes selgus, et, ehkki viiruse genoomi viidud mutatsioonide efektid P1234 protsessingule olid sarnased rakuvabas süsteemis nähtutele, mõjutasid need mutatsioonid viiruse infektsioonilisust üllatavalt vähe. Samuti osutus võimalikuks lisada SFV nsP3 ja nsP4 regioonide vahele täiendav valk koos ülal toodud reeglite alusel konstrueeritud kunstlike lõikamissaitidega. Need andmed viitavad sellele, et P1234 protsessingus on oluline roll ka mehhanismidel, mis ei ole otseselt seotud lühikese ah-järjestuse äratundmisega proteaasi poolt.

Selliseid mehhanisme kirjeldati esmalt analüüsides detailselt 2/3 saidi, milles puuduvad selle äratundmist võimaldava ah-järjestused, protsessimist. See uurimus näitas, et nimetatud saidi lõikamine ei põhine lühikese ah-järjestuse äratundmisel nsP2 proteaasi domääni poolt vaid et selles osaleb ka nsP2 N-terminaalne regioon, ning, mis eriti oluline, lõikamiskohast ligi 160 ah jäägi kaugusel nsP3 macro-domääni C-terminaalses alas paiknev regioon – eksosait. Ilmnes, et see regioon osaleb 2/3 saiti sisaldava substraadi täpses paigutamises nsP2 aktiivsaiti ja et seda protsessi koordineerib nsP2 N-terminus: ainult juhul, kui kõik nimetatud komponendid paiknesid üksteise suhtes õiges konfiguratsioonis, toimus 2/3 saidi efektiivne lõikamine; juhul, kui mõni element puudus, oli defektne või oli rikutud nende vastastikune paiknemine, lõikamist ei toimunud. Seega osutus 2/3 saidi lõikamise mehhanism uudseks ja selle saidi protsessing sõltuvaks viiruse valkude poolt moodustatud makromolekulaarsest struktuurist. Järeldust, et just korrektsete makromolekulaarsete struktuuride moodustamine (aga ka nendes struktuurides toimuvad muutused) ja interaktsioonid eksosaidiga omavad proteaasi toimimise reguleerimisel peamist rolli kinnitasid ka järgnevad uurimused. Nii näitas läbiviidud positsiooniline skanneerimine, et isegi juhul kui kogu lõikamissaitide järjestuses paiknev informatsioon ära nullida (muuta kõik saidid ühesugusteks järjestusega Ala-Ala-Ala-Ala-Gly positsioonides P6-P2) suudab viirus adapteeruda ja korrektse lõikamisjärjekorra säilitada või taastada kas pseudoreversioonide või kompensatoorsete mutatsioonide tekkimise abil. Peale selle selgus, et 3/4 ja 1/2 saitide järjestused on evolutsiooni käigus välja kujunud selliselt, et tagavad vastavalt nende kiire (3/4) ning, vastupidi, piisavalt aeglase (1/2) lõikamise. Viimane on ilmselt vajalik võimaldamaks viirusele oluliste bioloogiliste protsesside (tõenäoliselt negatiivse RNA sünteesi) läbiviimist, mistõttu on selle kiirendamine, sarnaselt 3/4 saidi lõikamise blokeerimisele, viirusele fataalne. Siiski olid ka sellised mutandid võimelised infektsioonilisust taastama, tehes seda taas mutatsioonide tekkimise abil kas proteaasi lõikesaiti ja/või vastavas subsaidis proteaasi substraati äratundvas ja seondavas regioonis.

Kokkuvõttes näitasid käesoleva väitekirja aluseks olevad uurimised, et SFV P1234 protsessimine on kompleksne protsess, mida reguleerib mitmetasandiline kontrollsüsteem, mis tagab selle, et ennetähtaegsed (ja muud viiruse eksisteerimisega mittekokkusobivad) lõikamised on välistatud, samal ajal kui on tagatud viirusele oluliste lõikamiste efektiivne ja õigeaegne toimumine replikatsioonitsükli kindlates faasides. Saadud andmed lubavad väita, et alfaviiruse proteaas

kujutab endast osa elegantsest signaali ülekandmise mehhanismist ja et tema peamine roll seisneb viiruse infektsioonitsükli käigu "jälgimises". Selle protsessi käigus tunneb proteaas ära viirusele oluliste sündmuste toimumisega kaasnevad muudatused replikaasi struktuuris, millega kaasnevad kindlate lõikesaitide ja/või muude oluliste elementide (nagu eksosaidid) kättesaadavaks muutumine. Sellistele sündmustele järgnev lõikamine muudab toimunud ümberkorraldused replikaasis praktiliselt pöördumatuks. Saadud andmed näitasid selgelt, et lisaks lõikesaitide järjestustele on selles protsessis oluline roll replikaasi moodustavatel makromolekulaarsetel struktuuridel; nende faktorite omavaheline koostoime kujutab endast „jõudu", mis tagab paljudest lõikamistest koosneva protsessingukaskaadi toimumise ja regulatsiooni. Kuna ka teiste polüproteiine ekspresseerivate viiruste proteaasid võivad nende viiruste infektsiooniprotsessides omada sarnast rolli, omavad saadud tulemused tähtsust nii erinevate viiruse infektsiooniprotsessiga kaasnevate sündmuste lahtimõtestamisel, kui ka viiruse proteaaside vastu suunatud inhibiitorite kavandamisel. Peale selle näitavad saadud tulemused, et lisaks traditsioonilistele proteaasi inhibiitoritele eksisteerib võimalus kavandada ka uut tüüpi inhibiitoried, mille toime seisneks proteaasi töö regulatsiooni rikkumises ja selle kaudu aktiivse replikaasi moodustamise takistamises.

REFERENCES

1. **Agapov, E. V., I. Frolov, B. D. Lindenbach, B. M. Prágai, S. Schlesinger, and C. M. Rice**. 1998. Noncytopathic Sindbis virus RNA vectors for heterologous gene expression. *Proc Natl Acad Sci USA,* **95**:12989–94.
2. **Aharoni, A., L. Gaidukov, O. Khersonsky, S. McQ Gould, C. Roodveldt, and D. S. Tawfik**. 2005. The "evolvability" of promiscuous protein functions. *Nature Genetics*, **37**:73–6.
3. **Ahola, T., and P. Ahlquist**. 1999. Putative RNA capping activities encoded by brome mosaic virus: methylation and covalent binding of guanylate by replicase protein 1a. *J Virol*, **73**:10061–9.
4. **Ahola, T., P. Kujala, M. Tuittila, T. Blom, P. Laakkonen, A. Hinkkanen, and P. Auvinen**. 2000. Effects of palmitoylation of replicase protein nsP1 on alphavirus infection. *J Virol*, **74**:6725–33.
5. **Ahola, T., and L. Kääriäinen**. 1995. Reaction in alphavirus mRNA capping: formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc Natl Acad Sci USA,* **92**:507–11.
6. **Ahola, T., P. Laakkonen, H. Vihinen, and L. Kääriäinen**. 1997. Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J Virol*, **71**:392–7.
7. **Ahola, T., A. Lampio, P. Auvinen, and L. Kääriäinen**. 1999. Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. *EMBO J*, **18**:3164–72.
8. **Ahola, T., J. A. den Boon, and P. Ahlquist**. 2000. Helicase and capping enzyme active site mutations in brome mosaic virus protein 1a cause defects in template recruitment, negative-strand RNA synthesis, and viral RNA capping. *J Virol*, **74**:8803–11.
9. **Alcover, A., and M.-I. Thoulouze**. 2010. Vesicle traffic to the immunological synapse: a multifunctional process targeted by lymphotropic viruses. *Curr Top Microbiol Immunol,* **340**:191–207.
10. **Ali, A., R. M. Bandaranayake, Y. Cai, N. M. King, M. Kolli, S. Mittal, J. F. Murzycki, M. N. L. Nalam, E. A. Nalivaika, A. Ozen, M. M. Prabu-Jeyabalan, K. Thayer, and C. A. Schiffer**. 2010. Molecular Basis for Drug Resistance in HIV-1 Protease. *Viruses*, **2**:2509–35.
11. **Altman, M. D., A. Ali, G. S. K. K. Reddy, M. N. L. Nalam, S. G. Anjum, H. Cao, S. Chellappan, V. Kairys, M. X. Fernandes, M. K. Gilson, C. A. Schiffer, T. M. Rana, and B. Tidor**. 2008. HIV-1 protease inhibitors from inverse design in the substrate envelope exhibit subnanomolar binding to drug-resistant variants. *J Am Chem Soc*, **130**:6099–113.
12. **Anderson, J., C. Schiffer, S.-K. Lee, and R. Swanstrom**. 2009. Viral protease inhibitors. *Handbook of experimental pharmacology*, 85–110.
13. **Artenstein, A. W., and S. M. Opal**. 2011. Proprotein convertases in health and disease. *New Engl J Med*, **365**:2507–18.
14. **Atasheva, S., R. Gorchakov, R. English, I. Frolov, and E. Frolova**. 2007. Development of Sindbis viruses encoding nsP2/GFP chimeric proteins and their application for studying nsP2 functioning. *J Virol*, **81**:5046–57.
15. **Atkins, G. J., M. N. Fleeton, and B. J. Sheahan**. 2008. Therapeutic and prophylactic applications of alphavirus vectors. *Expert Rev Mol Med*, **10**:e33.

16. **Atkinson, H. J., P. C. Babbitt, and M. Sajid**. 2009. The global cysteine peptidase landscape in parasites. *Trends Parasitol*, **25**:573–81.
17. **Babtie, A., N. Tokuriki, and F. Hollfelder**. 2010. What makes an enzyme promiscuous? *Curr Opin Chem Biol*, **14**:200–7.
18. **Bamford, D. H., R. J. Gilbert, J. M. Grimes, and D. I. Stuart**. 2001. Macromolecular assemblies: greater than their parts. *Curr Opin Struc Biol*, **11**:107–13.
19. **Banke, S., M. R. Lillemark, J. Gerstoft, N. Obel, and L. B. Jørgensen**. 2009. Positive selection pressure introduces secondary mutations at Gag cleavage sites in human immunodeficiency virus type 1 harboring major protease resistance mutations. *J Virol*, **83**:8916–24.
20. **Barkan, D. T., D. R. Hostetter, S. Mahrus, U. Pieper, J. A. Wells, C. S. Craik, and A. Sali**. 2010. Prediction of protease substrates using sequence and structure features. *Bioinformatics (Oxford, England)*, **26**:1714–22.
21. **Bedard, K. M., and B. L. Semler**. 2004. Regulation of picornavirus gene expression. *Microbes Infect*, **6**:702–13.
22. **Bellón-Echeverría, I., A. J. López-Jiménez, P. Clemente-Casares, and A. Mas**. 2010. Monitoring hepatitis C virus (HCV) RNA-dependent RNA polymerase oligomerization by a FRET-based in vitro system. *Antiviral Res*, **87**:57–66.
23. **Belov, G. A., V. Nair, B. T. Hansen, F. H. Hoyt, E. R. Fischer, and E. Ehrenfeld**. 2011. Complex Dynamic Development of Poliovirus Membranous Replication Complexes. *J Virol*, **86**:302–12.
24. **Benarroch, D., B. Selisko, G. A. Locatelli, G. Maga, J.-L. Romette, and B. Canard**. 2004. The RNA helicase, nucleotide 5'-triphosphatase, and RNA 5'-triphosphatase activities of Dengue virus protein NS3 are Mg2+-dependent and require a functional Walker B motif in the helicase catalytic core. *Virology*, **328**:208–18.
25. **Berti, P. J., C. H. Faerman, and A. C. Storer**. 1991. Cooperativity of papain-substrate interaction energies in the S2 to S2' subsites. *Biochemistry*, **30**:1394–402.
26. **Bhatia, V. K., N. S. Hatzakis, and D. Stamou**. 2010. A unifying mechanism accounts for sensing of membrane curvature by BAR domains, amphipathic helices and membrane-anchored proteins. *Semin Cell Dev Biol*, **21**:381–90.
27. **Bianchi, E., and A. Pessi**. 2002. Inhibiting viral proteases: challenges and opportunities. *Biopolymers*, **66**:101–14.
28. **Bianchini, E. P., S. J. Orcutt, P. Panizzi, P. E. Bock, and S. Krishnaswamy**. 2005. Ratcheting of the substrate from the zymogen to proteinase conformations directs the sequential cleavage of prothrombin by prothrombinase. *Proc Natl Acad Sci USA,* **102**:10099–104.
29. **Bick, M. J., J.-W. N. Carroll, G. Gao, S. P. Goff, C. M. Rice, and M. R. MacDonald**. 2003. Expression of the zinc-finger antiviral protein inhibits alphavirus replication. *J Virol*, **77**:11555–62.
30. **Bock, P. E., P. Panizzi, and I. M. A. Verhamme**. 2007. Exosites in the substrate specificity of blood coagulation reactions. *J Thromb Haemost*, 5 Suppl **1**:81–94.
31. **Boehr, D. D., R. Nussinov, and P. E. Wright**. 2009. The role of dynamic conformational ensembles in biomolecular recognition. *Nature Chem Biol*, **5**:789–96.
32. **Boyd, S. E., F. K. Kerr, D. W. Albrecht, M. G. de la Banda, N. Ng, and R. N. Pike**. 2009. Cooperative effects in the substrate specificity of the complement protease C1s. *Biol Chem*, **390**:503–7.

33. **Breakwell, L., P. Dosenovic, G. B. Karlsson Hedestam, M. D'Amato, P. Liljeström, J. Fazakerley, and G. M. McInerney**. 2007. Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J Virol*, **81**:8677–84.
34. **Briones, C., and E. Domingo**. 2008. Minority report: hidden memory genomes in HIV-1 quasispecies and possible clinical implications. *AIDS rev*, **10**:93–109.
35. **Bronze, M. S., M. M. Huycke, L. J. Machado, G. W. Voskuhl, and R. A. Greenfield**. 2002. Viral agents as biological weapons and agents of bioterrorism. *Am J Med Sci*, **323**:316–25.
36. **Burnham, A. J., L. Gong, and R. W. Hardy**. 2007. Heterogeneous nuclear ribonuclear protein K interacts with Sindbis virus nonstructural proteins and viral subgenomic mRNA. *Virology*, **367**:212–21.
37. **Byrnes, A. P., and D. E. Griffin**. 1998. Binding of Sindbis virus to cell surface heparan sulfate. *J Virol*, **72**:7349–56.
38. **Bügl, H., E. B. Fauman, B. L. Staker, F. Zheng, S. R. Kushner, M. A. Saper, J. C. Bardwell, and U. Jakob**. 2000. RNA methylation under heat shock control. *Mol Cell*, **6**:349–60.
39. **Cameron, C. E., H. Suk Oh, and I. M. Moustafa**. 2010. Expanding knowledge of P3 proteins in the poliovirus lifecycle. *Future Microbiol*, **5**:867–81.
40. **Caruthers, J. M., and D. B. McKay**. 2002. Helicase structure and mechanism. *Curr Opin Struct Biol*, **12**:123–33.
41. **Casales, E., J. R. Rodriguez-Madoz, M. Ruiz-Guillen, N. Razquin, Y. Cuevas, J. Prieto, and C. Smerdou**. 2008. Development of a new noncytopathic Semliki Forest virus vector providing high expression levels and stability. *Virology*, **376**:242–51.
42. **Castelló, A., E. Alvarez, and L. Carrasco**. 2011. The multifaceted poliovirus 2A protease: regulation of gene expression by picornavirus proteases. *J Biomed Biotechnol*, Epub 2011 Apr 14.
43. **Castro, H. C., P. A. Abreu, R. B. Geraldo, R. C. A. Martins, R. dos Santos, N. I. V. Loureiro, L. M. Cabral, and C. R. Rodrigues**. 2011. Looking at the proteases from a simple perspective. *J Mol Recognit*, **24**:165–81.
44. **Chaudhury, A., P. Chander, and P. H. Howe**. 2010. Heterogeneous nuclear ribonucleoproteins (hnRNPs) in cellular processes: Focus on hnRNP E1's multifunctional regulatory roles. *RNA*, **16**:1449–62.
45. **Chaudhury, S., and J. J. Gray**. 2009. Identification of structural mechanisms of HIV-1 protease specificity using computational peptide docking: implications for drug resistance. *Structure*, **17**:1636–48.
46. **Chellappan, S., V. Kairys, M. X. Fernandes, C. Schiffer, and M. K. Gilson**. 2007. Evaluation of the substrate envelope hypothesis for inhibitors of HIV-1 protease. *Proteins*, **68**:561–7.
47. **Cheng, R. H., R. J. Kuhn, N. H. Olson, M. G. Rossmann, H. K. Choi, T. J. Smith, and T. S. Baker**. 1995. Nucleocapsid and glycoprotein organization in an enveloped virus. *Cell*, **80**:621–30.
48. **Chinnaswamy, S., A. Murali, P. Li, K. Fujisaki, and C. C. Kao**. 2010. Regulation of de novo-initiated RNA synthesis in hepatitis C virus RNA-dependent RNA polymerase by intermolecular interactions. *J Virol*, **84**:5923–35.
49. **Choe, Y., F. Leonetti, D. C. Greenbaum, F. Lecaille, M. Bogyo, D. Brömme, J. A. Ellman, and C. S. Craik**. 2006. Substrate profiling of cysteine proteases using

a combinatorial peptide library identifies functionally unique specificities. *J Biol Chem*, **281**:12824–32.
50. **Choi, K. H., and M. G. Rossmann**. 2009. RNA-dependent RNA polymerases from Flaviviridae. *Curr Opin Struct Biol*, **19**:746–51.
51. **Ciechanover, A.** 2006. The ubiquitin proteolytic system: from an idea to the patient bed. *Proc Am Thorac Soc*, **3**:21–31.
52. **Couture, F., F. D'Anjou, and R. Day**. 2011. On the cutting edge of proprotein convertase pharmacology: from molecular concepts to clinical applications. *Biomol Concepts*, **2**:421–438.
53. **Cristea, I. M., J.-W. N. Carroll, M. P. Rout, C. M. Rice, B. T. Chait, and M. R. MacDonald**. 2006. Tracking and elucidating alphavirus-host protein interactions. *J Biol Chem*, **281**:30269–78.
54. **Cristea, I. M., H. Rozjabek, K. R. Molloy, S. Karki, L. L. White, C. M. Rice, M. P. Rout, B. T. Chait, and M. R. MacDonald**. 2010. Host factors associated with the Sindbis virus RNA-dependent RNA polymerase: role for G3BP1 and G3BP2 in virus replication. *J Virol*, **84**:6720–32.
55. **Cristofari, G., and J.-L. Darlix**. 2002. The ubiquitous nature of RNA chaperone proteins. *Prog Nucleic Acid Res Mol Biol*, **72**:223–68.
56. **Crotty, S., C. E. Cameron, and R. Andino**. 2001. RNA virus error catastrophe: direct molecular test by using ribavirin. *Proc Natl Acad Sci USA,* **98**:6895–900.
57. **Davey, N. E., G. Travé, and T. J. Gibson**. 2011. How viruses hijack cell regulation. *Trends Biochem Sci*, **36**:159–69.
58. **Davey, N. E., K. Van Roey, R. J. Weatheritt, G. Toedt, B. Uyar, B. Altenberg, A. Budd, F. Diella, H. Dinkel, and T. J. Gibson**. 2011. Attributes of short linear motifs. *Mol Biosyst*, **8**:268–81.
59. **Decroly, E., F. Ferron, J. Lescar, and B. Canard**. 2011. Conventional and unconventional mechanisms for capping viral mRNA. *Nat Rev Microbiol*, **10**:51–65.
60. **Delagoutte, E., and P. H. von Hippel**. 2002. Helicase mechanisms and the coupling of helicases within macromolecular machines. Part I: Structures and properties of isolated helicases. *Q Rev Biophys*, **35**:431–78.
61. **Delagoutte, E., and P. H. von Hippel**. 2003. Helicase mechanisms and the coupling of helicases within macromolecular machines. Part II: Integration of helicases into cellular processes. *Q Rev Biophys,* **36**:1–69.
62. **Denison, M. R.** 2008. Seeking membranes: positive-strand RNA virus replication complexes. *PLoS Biol*, **6**:e270.
63. **Diaz, A., A. Gallei, and P. Ahlquist**. 2012. Bromovirus RNA replication compartment formation requires concerted action of 1a's self-interacting RNA capping and helicase domains. *J Virol*, **86**:821–34.
64. **Diaz, A., X. Wang, and P. Ahlquist**. 2010. Membrane-shaping host reticulon proteins play crucial roles in viral RNA replication compartment formation and function. *Proc Natl Acad Sci USA,* **107**:16291–6.
65. **Diella, F., N. Haslam, C. Chica, A. Budd, S. Michael, N. P. Brown, G. Trave, and T. J. Gibson**. 2008. Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Front Biosci*, **13**:6580–603.
66. **Ding, M. X., and M. J. Schlesinger**. 1989. Evidence that Sindbis virus NSP2 is an autoprotease which processes the virus nonstructural polyprotein. *Virology*, **171**:280–4.

67. **Dinkel, H., S. Michael, R. J. Weatheritt, N. E. Davey, K. Van Roey, B. Altenberg, G. Toedt, B. Uyar, M. Seiler, A. Budd, L. Jödicke, M. A. Dammert, C. Schroeter, M. Hammer, T. Schmidt, P. Jehl, C. McGuigan, M. Dymecka, C. Chica, K. Luck, A. Via, A. Chatr-Aryamontri, N. Haslam, G. Grebnev, R. J. Edwards, M. O. Steinmetz, H. Meiselbach, F. Diella, and T. J. Gibson**. 2011. ELM – the database of eukaryotic linear motifs. *Nucleic Acids Res*, **40**:D242–51.
68. **Dokudovskaya, S., R. Williams, D. Devos, A. Sali, B. T. Chait, and M. P. Rout**. 2006. Protease accessibility laddering: a proteomic tool for probing protein structure. *Structure*, **14**:653–60.
69. **Domingo, E., V. Martin, C. Perales, A. Grande-Pérez, J. García-Arriaza, and A. Arias**. 2006. Viruses as quasispecies: biological implications. *Curr Top Microbiol Immunol*, **299**:51–82.
70. **Dougan, D. A., D. Micevski, and K. N. Truscott**. 2011. The N-end rule pathway: From recognition by N-recognins, to destruction by AAA+proteases. *Biochim Biophys Acta*, **1823**:83–91.
71. **Dougherty, W. G., and B. L. Semler**. 1993. Expression of virus-encoded proteinases: functional and structural similarities with cellular enzymes. *Microbiol Rev*, **57**:781–822.
72. **Drag, M., and G. S. Salvesen**. 2010. Emerging principles in protease-based drug discovery. *Nat Rev Drug Discov*, **9**:690–701.
73. **Drin, G., and B. Antonny**. 2010. Amphipathic helices and membrane curvature. *FEBS Lett*, **584**:1840–7.
74. **Dryga, S. A., O. A. Dryga, and S. Schlesinger**. 1997. Identification of mutations in a Sindbis virus variant able to establish persistent infection in BHK cells: the importance of a mutation in the nsP2 gene. *Virology*, **228**:74–83.
75. **Duffy, S., L. A. Shackelton, and E. C. Holmes**. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat Rev Genet*, **9**:267–76.
76. **Duverle, D. A., and H. Mamitsuka**. 2011. A review of statistical methods for prediction of proteolytic cleavage. *Brief Bioinform*, Epub 2011 Dec 2.
77. **Dé, I., S. G. Sawicki, and D. L. Sawicki**. 1996. Sindbis virus RNA-negative mutants that fail to convert from minus-strand to plus-strand synthesis: role of the nsP2 protein. *J Virol*, **70**:2706–19.
78. **Dé, I., C. Fata-Hartley, S. G. Sawicki, and D. L. Sawicki**. 2003. Functional analysis of nsP3 phosphoprotein mutants of Sindbis virus. *J Virol*, **77**:13106–16.
79. **Egloff, M.-P., H. Malet, A. Putics, M. Heinonen, H. Dutartre, A. Frangeul, A. Gruez, V. Campanacci, C. Cambillau, J. Ziebuhr, T. Ahola, and B. Canard**. 2006. Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. *J Virol*, **80**:8493–502.
80. **Evans, M. J., C. M. Rice, and S. P. Goff**. 2004. Phosphorylation of hepatitis C virus nonstructural protein 5A modulates its protein interactions and viral RNA replication. *Proc Natl Acad Sci USA,* **101**:13038–43.
81. **Fairman-Williams, M. E., U.-P. Guenther, and E. Jankowsky**. 2010. SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol*, **20**:313–24.
82. **Fairman-Williams, M. E., and E. Jankowsky**. 2011. Unwinding Initiation by the Viral RNA Helicase NPH-II. *J Mol Biol*, **415**:819–32.
83. **Farady, C. J., and C. S. Craik**. 2010. Mechanisms of macromolecular protease inhibitors. *Chembiochem*, **11**:2341–6.

84. **Fata, C. L., S. G. Sawicki, and D. L. Sawicki**. 2002. Modification of Asn374 of nsP1 suppresses a Sindbis virus nsP4 minus-strand polymerase mutant. *J Virol*, **76**:8641–9.
85. **Fayzulin, R., and I. Frolov**. 2004. Changes of the secondary structure of the 5' end of the Sindbis virus genome inhibit virus growth in mosquito cells and lead to accumulation of adaptive mutations. *J Virol*, **78**:4953–64.
86. **Fernández, A., D. S. Tawfik, B. Berkhout, R. Sanders, A. Kloczkowski, T. Sen, and B. Jernigan**. 2005. Protein promiscuity: drug resistance and native functions – HIV-1 case. *J Biomol Struct Dyn*, **22**:615–24.
87. **Ferrer-Orta, C., A. Arias, C. Escarmís, and N. Verdaguer**. 2006. A comparison of viral RNA-dependent RNA polymerases. *Curr Opin Struct Biol*, **16**:27–34.
88. **Firth, A. E., N. M. Wills, R. F. Gesteland, and J. F. Atkins**. 2011. Stimulation of stop codon readthrough: frequent presence of an extended 3' RNA structural element. *Nucleic Acids Res*, **39**:6679–91.
89. **Fischer, E.** 1894. Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der deutschen chemischen Gesellschaft*, **27**:2985–2993.
90. **Foeger, N., E. Kuehnel, R. Cencic, and T. Skern**. 2005. The binding of foot-and-mouth disease virus leader proteinase to eIF4GI involves conserved ionic interactions. *FEBS J*, **272**:2602–11.
91. **Foeger, N., E. M. Schmid, and T. Skern**. 2003. Human rhinovirus 2 2Apro recognition of eukaryotic initiation factor 4GI. Involvement of an exosite. *J Biol Chem*, **278**:33200–7.
92. **Fontana, A., P. Polverino de Laureto, V. De Filippis, E. Scaramella, and M. Zambonin**. 1997. Probing the partly folded states of proteins by limited proteolysis. *Fold Des*, **2**:R17–26.
93. **Fontana, J., C. López-Iglesias, W.-P. Tzeng, T. K. Frey, J. J. Fernández, and C. Risco**. 2010. Three-dimensional structure of Rubella virus factories. *Virology*, **405**:579–91.
94. **Fontana, J., W.-P. Tzeng, G. Calderita, A. Fraile-Ramos, T. K. Frey, and C. Risco**. 2007. Novel replication complex architecture in rubella replicon-transfected cells. *Cell Microbiol*, **9**:875–90.
95. **Foo, S. S., W. Chen, L. Herrero, J. Bettadapura, J. Narayan, L. Dar, S. Broor, and S. Mahalingam**. 2011. The genetics of alphaviruses. Future Virology. *Future Medicine*, **6**:16.
96. **Forrester, N. L., G. Palacios, R. B. Tesh, N. Savji, H. Guzman, M. Sherman, S. C. Weaver, and W. I. Lipkin**. 2012. Genome-scale phylogeny of the alphavirus genus suggests a marine origin. *J Virol*, **86**:2729–38.
97. **Foulkes-Murzycki, J. E., W. R. P. Scott, and C. A. Schiffer**. 2007. Hydrophobic sliding: a possible mechanism for drug resistance in human immunodeficiency virus type 1 protease. *Structure*, **15**:225–33.
98. **Frey, T. K., D. L. Gard, and J. H. Strauss**. 1979. Biophysical studies on circle formation by Sindbis virus 49 S RNA. *J Mol Biol*, **132**:1–18.
99. **Frick, D. N., and A. M. I. Lam**. 2006. Understanding helicases as a means of virus control. *Curr Pharm Des*, **12**:1315–38.
100. **Frick, D. N.** 2007. The hepatitis C virus NS3 protein: a model RNA helicase and potential drug target. *Curr Issues Mol Biol*, **9**:1–20.
101. **Frolov, I.** 2004. Persistent infection and suppression of host response by alphaviruses. *Arch Virol Suppl*, 18:139–47.

102. **Frolov, I., E. Agapov, T. A. Hoffman, B. M. Prágai, M. Lippa, S. Schlesinger, and C. M. Rice**. 1999. Selection of RNA replicons capable of persistent noncytopathic replication in mammalian cells. *J Virol*, **73**:3854–65.
103. **Frolov, I., R. Hardy, and C. M. Rice**. 2001. Cis-acting RNA elements at the 5' end of Sindbis virus genome RNA regulate minus- and plus-strand RNA synthesis. *RNA*, **7**:1638–51.
104. **Frolov, I., and S. Schlesinger**. 1994. Translation of Sindbis virus mRNA: effects of sequences downstream of the initiating codon. *J Virol*, **68**:8111–7.
105. **Frolov, I., M. Akhrymuk, I. Akhrymuk, S. Atasheva, and E. I. Frolova**. 2012. Early events in alphavirus replication determine the outcome of the infection. *J Virol*, Epub 2012 Feb 15.
106. **Frolov, I., N. Garmashova, S. Atasheva, and E. I. Frolova**. 2009. Random insertion mutagenesis of sindbis virus nonstructural protein 2 and selection of variants incapable of downregulating cellular transcription. *J Virol*, **83**:9031–44.
107. **Frolova, E., I. Frolov, and S. Schlesinger**. 1997. Packaging signals in alphaviruses. *J Virol*, **71**:248–58.
108. **Frolova, E. I., R. Z. Fayzulin, S. H. Cook, D. E. Griffin, C. M. Rice, and I. Frolov**. 2002. Roles of nonstructural protein nsP2 and Alpha/Beta interferons in determining the outcome of Sindbis virus infection. *J Virol*, **76**:11254–64.
109. **Frolova, E. I., R. Gorchakov, L. Pereboeva, S. Atasheva, and I. Frolov**. 2010. Functional Sindbis virus replicative complexes are formed at the plasma membrane. *J Virol*, **84**:11679–95.
110. **Frolova, E., R. Gorchakov, N. Garmashova, S. Atasheva, L. A. Vergara, and I. Frolov**. 2006. Formation of nsP3-specific protein complexes during Sindbis virus replication. *J Virol*, **80**:4122–34.
111. **Fros, J. J., W. J. Liu, N. A. Prow, C. Geertsema, M. Ligtenberg, D. L. Vanlandingham, E. Schnettler, J. M. Vlak, A. Suhrbier, A. A. Khromykh, and G. P. Pijlman**. 2010. Chikungunya virus nonstructural protein 2 inhibits type I/II interferon-stimulated JAK-STAT signaling. *J Virol*, **84**:10877–87.
112. **Froshauer, S., J. Kartenbeck, and A. Helenius**. 1988. Alphavirus RNA replicase is located on the cytoplasmic surface of endosomes and lysosomes. *J Cell Biol*, **107**:2075–86.
113. **Fuentes-Prior, P., and G. S. Salvesen**. 2004. The protein structures that shape caspase activity, specificity, activation and inhibition. *Biochem J*, **384**:201–32.
114. **Fuxreiter, M., P. Tompa, and I. Simon**. 2007. Local structural disorder imparts plasticity on linear motifs. *Bioinformatics*, **23**:950–6.
115. **Gallie, D. R.** 1998. A tale of two termini: a functional interaction between the termini of an mRNA is a prerequisite for efficient translation initiation. *Gene*, **216**:1–11.
116. **Garmashova, N., R. Gorchakov, E. Frolova, and I. Frolov**. 2006. Sindbis virus nonstructural protein nsP2 is cytotoxic and inhibits cellular transcription. *J Virol*, **80**:5686–96.
117. **Garmashova, N., R. Gorchakov, E. Volkova, S. Paessler, E. Frolova, and I. Frolov**. 2007. The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. *J Virol*, **81**:2472–84.
118. **Garoff, H., J. Wilschut, P. Liljeström, J. M. Wahlberg, R. Bron, M. Suomalainen, J. Smyth, A. Salminen, B. U. Barth, and H. Zhao**. 1994. Assembly and entry mechanisms of Semliki Forest virus. *Arch Virol Suppl*, **9**:329–38.

119. **Garoff, H., M. Sjöberg, and R. H. Cheng**. 2004. Budding of alphaviruses. *Virus Res*, **106**:103–16.
120. **George, J., and R. Raju**. 2000. Alphavirus RNA genome repair and evolution: molecular characterization of infectious sindbis virus isolates lacking a known conserved motif at the 3' end of the genome. *J Virol*, **74**:9776–85.
121. **Ghosh, A., and C. D. Lima**. 2010. Enzymology of RNA cap synthesis. *Wiley Interdiscip Rev RNA*, **1**:152–72.
122. **Gomez de Cedrón, M., N. Ehsani, M. L. Mikkola, J. A. García, and L. Kääriäinen**. 1999. RNA helicase activity of Semliki Forest virus replicase protein NSP2. *FEBS Lett*, **448**:19–22.
123. **Gong, P., and O. B. Peersen**. 2010. Structural basis for active site closure by the poliovirus RNA-dependent RNA polymerase. *Proc Natl Acad Sci USA,* **107**:22505–10.
124. **Gorchakov, R., E. Frolova, and I. Frolov**. 2005. Inhibition of transcription and translation in Sindbis virus-infected cells. *J Virol*, **79**:9397–409.
125. **Gorchakov, R., E. Frolova, S. Sawicki, S. Atasheva, D. Sawicki, and I. Frolov**. 2008. A new role for ns polyprotein cleavage in Sindbis virus replication. *J Virol*, **82**:6218–31.
126. **Gorchakov, R., E. Frolova, B. R. G. Williams, C. M. Rice, and I. Frolov**. 2004. PKR-dependent and -independent mechanisms are involved in translational shutoff during Sindbis virus infection. *J Virol*, **78**:8455–67.
127. **Gorchakov, R., N. Garmashova, E. Frolova, and I. Frolov**. 2008. Different types of nsP3-containing protein complexes in Sindbis virus-infected cells. *J Virol*, **82**:10088–101.
128. **Gorchakov, R., R. Hardy, C. M. Rice, and I. Frolov**. 2004. Selection of functional 5' cis-acting elements promoting efficient sindbis virus genome replication. *J Virol*, **78**:61–75.
129. **Gottschalk, A. J., G. Timinszky, S. E. Kong, J. Jin, Y. Cai, S. K. Swanson, M. P. Washburn, L. Florens, A. G. Ladurner, J. W. Conaway, and R. C. Conaway**. 2009. Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler. *Proc Natl Acad Sci USA,* **106**:13770–4.
130. **Gould, E. A., B. Coutard, H. Malet, B. Morin, S. Jamal, S. Weaver, A. Gorbalenya, G. Moureau, C. Baronti, I. Delogu, N. Forrester, M. Khasnatinov, T. Gritsun, X. de Lamballerie, and B. Canard**. 2010. Understanding the alphaviruses: recent research on important emerging pathogens and progress towards their control. *Antiviral Res*, **87**:111–24.
131. **Griffiths, G., R. Brands, B. Burke, D. Louvard, and G. Warren**. 1982. Viral membrane proteins acquire galactose in trans Golgi cisternae during intracellular transport. *J Cell Biol*, **95**:781–92.
132. **Grimley, P. M., I. K. Berezesky, and R. M. Friedman**. 1968. Cytoplasmic structures associated with an arbovirus infection: loci of viral ribonucleic acid synthesis. *J Virol*, **2**:1326–38.
133. **Grimley, P. M., J. G. Levin, I. K. Berezesky, and R. M. Friedman**. 1972. Specific membranous structures associated with the replication of group A arboviruses. *J Virol*, **10**:492–503.
134. **Groll, M., and T. Clausen**. 2003. Molecular shredders: how proteasomes fulfill their role. *Curr Opin Struct Biol*, **13**:665–73.

135. **Gros, C., and G. Wengler**. 1996. Identification of an RNA-stimulated NTPase in the predicted helicase sequence of the Rubella virus nonstructural polyprotein. *Virology*, **217**:367–72.
136. **Groves, M. R., R. Coulombe, J. Jenkins, and M. Cygler**. 1998. Structural basis for specificity of papain-like cysteine protease proregions toward their cognate enzymes. *Proteins*, **32**:504–14.
137. **Grøn, H., and K. Breddam**. 1992. Interdependency of the binding subsites in subtilisin. *Biochemistry*, **31**:8967–71.
138. **Gu, B., C. Liu, J. Lin-Goerke, D. R. Maley, L. L. Gutshall, C. A. Feltenberger, and A. M. Del Vecchio**. 2000. The RNA Helicase and Nucleotide Triphosphatase Activities of the Bovine Viral Diarrhea Virus NS3 Protein Are Essential for Viral Replication. *J Virol*, **74**:1794–1800.
139. **Gui, H., C.-W. Lu, S. Adams, V. Stollar, and M.-L. Li**. 2010. hnRNP A1 interacts with the genomic and subgenomic RNA promoters of Sindbis virus and is required for the synthesis of G and SG RNA. *J Biomed Sci*, **17**:59.
140. **Gurke, S., J. F. V. Barroso, and H.-H. Gerdes**. 2008. The art of cellular communication: tunneling nanotubes bridge the divide. *Histochem Cell Biol*, **129**:539–50.
141. **Gáspári, Z., P. Várnai, B. Szappanos, and A. Perczel**. 2010. Reconciling the lock-and-key and dynamic views of canonical serine protease inhibitor action. *FEBS Lett*, **584**:203–6.
142. **Hacisalihoglu, A., P. Panizzi, P. E. Bock, R. M. Camire, and S. Krishnaswamy**. 2007. Restricted active site docking by enzyme-bound substrate enforces the ordered cleavage of prothrombin by prothrombinase. *J Biol Chem*, **282**:32974–82.
143. **Hahn, Y. S., E. G. Strauss, and J. H. Strauss**. 1989. Mapping of RNA-temperature-sensitive mutants of Sindbis virus: assignment of complementation groups A, B, and G to nonstructural proteins. *J Virol*, **63**:3142–50.
144. **Halfon, P., and S. Locarnini**. 2011. Hepatitis C virus resistance to protease inhibitors. *J Hepatol*, **55**:192–206.
145. **Han, W., X. Li, and X. Fu**. 2011. The macro domain protein family: structure, functions, and their potential therapeutic implications. *Mutat Res*, **727**:86–103.
146. **Hansen, J. L., A. M. Long, and S. C. Schultz**. 1997. Structure of the RNA-dependent RNA polymerase of poliovirus. *Structure*, **5**:1109–22.
147. **Hardy, R. W.** 2006. The role of the 3' terminus of the Sindbis virus genome in minus-strand initiation site selection. *Virology*, **345**:520–31.
148. **Hardy, R. W., and C. M. Rice**. 2005. Requirements at the 3' end of the sindbis virus genome for efficient synthesis of minus-strand RNA. *J Virol*, **79**:4630–9.
149. **Hardy, W. R., Y. S. Hahn, R. J. de Groot, E. G. Strauss, and J. H. Strauss**. 1990. Synthesis and processing of the nonstructural polyproteins of several temperature-sensitive mutants of Sindbis virus. *Virology*, **177**:199–208.
150. **Hardy, W. R., and J. H. Strauss**. 1989. Processing the nonstructural polyproteins of sindbis virus: nonstructural proteinase is in the C-terminal half of nsP2 and functions both in cis and in trans. *J Virol*, **63**:4653–64.
151. **Harris, K. S., S. R. Reddigari, M. J. Nicklin, T. Hämmerle, and E. Wimmer**. 1992. Purification and characterization of poliovirus polypeptide 3CD, a proteinase and a precursor for RNA polymerase. *J Virol*, **66**:7481–9.

152. **Herrera, M., J. García-Arriaza, N. Pariente, C. Escarmís, and E. Domingo**. 2007. Molecular basis for a lack of correlation between viral fitness and cell killing capacity. *PLoS Pathog*, **3**:e53.
153. **Hsu, N.-Y., O. Ilnytska, G. Belov, M. Santiana, Y.-H. Chen, P. M. Takvorian, C. Pau, H. van der Schaar, N. Kaushik-Basu, T. Balla, C. E. Cameron, E. Ehrenfeld, F. J. M. van Kuppeveld, and N. Altan-Bonnet**. 2010. Viral reorganization of the secretory pathway generates distinct organelles for RNA replication. *Cell*, **141**:799–811.
154. **Huang, Y.-L., Y.-T. Han, Y.-T. Chang, Y.-H. Hsu, and M. Meng**. 2004. Critical Residues for GTP Methylation and Formation of the Covalent m7GMP-Enzyme Intermediate in the Capping Enzyme Domain of Bamboo Mosaic Virus. *J Virol*, **78**:1271–1280.
155. **Huang, Y., K. Staschke, R. De Francesco, and S.-L. Tan**. 2007. Phosphorylation of hepatitis C virus NS5A nonstructural protein: a new paradigm for phosphorylation-dependent viral RNA replication? *Virology*, **364**:1–9.
156. **Hubbard, S.** 1998. The structural aspects of limited proteolysis of native proteins. *Biochim Biophys Acta*, **1382**:191–206.
157. **Hult, K., and P. Berglund**. 2007. Enzyme promiscuity: mechanism and applications. *Trends Biotechnol*, **25**:231–8.
158. **Imbert, I., E. J. Snijder, M. Dimitrova, J.-C. Guillemot, P. Lécine, and B. Canard**. 2008. The SARS-Coronavirus PLnc domain of nsp3 as a replication/ transcription scaffolding protein. *Virus Res*, **133**:136–48.
159. **Ito, K., and Y. Akiyama**. 2005. Cellular functions, mechanism of action, and regulation of FtsH protease. *Annu Rev Microbiol*, **59**:211–31.
160. **Jacobson, M. F., and D. Baltimore**. 1968. Polypeptide cleavages in the formation of poliovirus proteins. *Proc Natl Acad Sci USA,* **61**:77–84.
161. **Jakubiec, A., and I. Jupin**. 2007. Regulation of positive-strand RNA virus replication: the emerging role of phosphorylation. *Virus Res*, **129**:73–9.
162. **Jankowsky, E.** 2011. RNA helicases at work: binding and rearranging. *Trends Biochem Sci*, **36**:19–29.
163. **Joachims, M., P. C. Van Breugel, and R. E. Lloyd**. 1999. Cleavage of poly(A)-binding protein by enterovirus proteases concurrent with inhibition of translation in vitro. *J Virol*, **73**:718–27.
164. **Johnson, M. A., A. Chatterjee, B. W. Neuman, and K. Wüthrich**. 2010. SARS coronavirus unique domain: three-domain molecular architecture in solution and RNA binding. *J Mol Biol*, **400**:724–42.
165. **Jolly, C., and Q. J. Sattentau**. 2004. Retroviral spread by induction of virological synapses. *Traffic*, **5**:643–50.
166. **Jose, J., L. Przybyla, T. J. Edwards, R. Perera, J. W. Burgner, and R. J. Kuhn**. 2011. Interactions of the cytoplasmic domain of Sindbis virus E2 with nucleocapside cores promote alphavirus budding. *J Virol*, **86**:2585–99.
167. **Jose, J., J. E. Snyder, and R. J. Kuhn**. 2009. A structural and functional perspective of alphavirus replication and assembly. *Future Microbiol*, **4**:837–56.
168. **Kadaré, G., C. David, and A. L. Haenni**. 1996. ATPase, GTPase, and RNA binding activities associated with the 206-kilodalton protein of turnip yellow mosaic virus. *J Virol*, **70**:8169–74.
169. **Kadaveru, K., J. Vyas, and M. R. Schiller**. 2008. Viral infection and human disease – insights from minimotifs. *Front Biosci*, **13**:6455–71.

170. **Kairys, V., M. K. Gilson, V. Lather, C. A. Schiffer, and M. X. Fernandes**. 2009. Toward the design of mutation-resistant enzyme inhibitors: further evaluation of the substrate envelope hypothesis. *Chem Biol Drug Des*, **74**:234–45.
171. **Kamer, G., and P. Argos**. 1984. Primary structural comparison of RNA-dependent polymerases from plant, animal and bacterial viruses. *Nucleic Acids Res*, **12**:7269–82.
172. **Karpe, Y. A., P. P. Aher, and K. S. Lole**. 2011. NTPase and 5'-RNA triphosphatase activities of Chikungunya virus nsP2 protein. *PloS One*, **6**:e22336.
173. **Karras, G. I., G. Kustatscher, H. R. Buhecha, M. D. Allen, C. Pugieux, F. Sait, M. Bycroft, and A. G. Ladurner**. 2005. The macro domain is an ADP-ribose binding module. *EMBO J*, **24**:1911–20.
174. **Karthik, S., and S. Senapati**. 2011. Dynamic flaps in HIV-1 protease adopt unique ordering at different stages in the catalytic cycle. *Proteins*, **79**:1830–40.
175. **Keränen, S., and L. Ruohonen**. 1983. Nonstructural proteins of Semliki Forest virus: synthesis, processing, and stability in infected cells. *J Virol*, **47**:505–15.
176. **Khersonsky, O., C. Roodveldt, and D. S. Tawfik**. 2006. Enzyme promiscuity: evolutionary and mechanistic aspects. *Curr Opin Chem Biol*, **10**:498–508.
177. **Khersonsky, O., and D. S. Tawfik**. 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. *Annu Rev Biochem*, **79**:471–505.
178. **Kielian, M., C. Chanel-Vos, and M. Liao**. 2010. Alphavirus Entry and Membrane Fusion. *Viruses*, **2**:796–825.
179. **Kielian, M., and F. A. Rey**. 2006. Virus membrane-fusion proteins: more than one way to make a hairpin. *Nat Rev Microbiol*, **4**:67–76.
180. **Kim, D. Y., A. E. Firth, S. Atasheva, E. I. Frolova, and I. Frolov**. 2011. Conservation of a packaging signal and the viral genome RNA packaging mechanism in alphavirus evolution. *J Virol*, **85**:8022–36.
181. **Kim, J. H., B. Hahm, Y. K. Kim, M. Choi, and S. K. Jang**. 2000. Protein-protein interaction among hnRNPs shuttling between nucleus and cytoplasm. *J Mol Biol*, **298**:395–405.
182. **Kim, K. H., T. Rümenapf, E. G. Strauss, and J. H. Strauss**. 2004. Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology*, **323**:153–63.
183. **Knight, R. L., K. L. W. Schultz, R. J. Kent, M. Venkatesan, and D. E. Griffin**. 2009. Role of N-linked glycosylation for sindbis virus infection and replication in vertebrate and invertebrate systems. *J Virol*, **83**:5640–7.
184. **Knoops, K., M. Kikkert, S. H. E. van den Worm, J. C. Zevenhoven-Dobbe, Y. van der Meer, A. J. Koster, A. M. Mommaas, and E. J. Snijder**. 2008. SARS-coronavirus replication is supported by a reticulovesicular network of modified endoplasmic reticulum. *PLoS Biol*, **6**:e226.
185. **Kolli, M., S. Lastere, and C. A. Schiffer**. 2006. Co-evolution of nelfinavir-resistant HIV-1 protease and the p1-p6 substrate. *Virology*, **347**:405–9.
186. **Kolli, M., E. Stawiski, C. Chappey, and C. A. Schiffer**. 2009. Human immunodeficiency virus type 1 protease-correlated cleavage site mutations enhance inhibitor resistance. *J Virol*, **83**:11027–42.
187. **Kononchik, J. P., R. Hernandez, and D. T. Brown**. 2011. An alternative pathway for alphavirus entry. *Virol J*, **8**:304.
188. **Kontijevskis, A., R. Petrovska, S. Yahorava, J. Komorowski, and J. E. S. Wikberg**. 2009. Proteochemometrics mapping of the interaction space for retroviral proteases and their substrates. *Bioorg Med Chem*, **17**:5229–37.

189. **Kontijevskis, A., P. Prusis, R. Petrovska, S. Yahorava, F. Mutulis, I. Mutule, J. Komorowski, and J. E. S. Wikberg**. 2007. A look inside HIV resistance through retroviral protease interaction maps. *PLoS Comput Biol*, **3**:e48.
190. **Kontijevskis, A., J. E. S. Wikberg, and J. Komorowski**. 2007. Computational proteomics analysis of HIV-1 protease interactome. *Proteins*, **68**:305–12.
191. **Koonin, E. V.** 1991. The phylogeny of RNA-dependent RNA polymerases of positive-strand RNA viruses. *J Gen Virol*, **72**:2197–206.
192. **Koonin, E. V., and V. V. Dolja**. 1993. Evolution and taxonomy of positive-strand RNA viruses: implications of comparative analysis of amino acid sequences. *Crit Rev Biochem Mol Biol*, **28**:375–430.
193. **Kopek, B. G., G. Perkins, D. J. Miller, M. H. Ellisman, and P. Ahlquist**. 2007. Three-dimensional analysis of a viral RNA replication complex reveals a virus-induced mini-organelle. *PLoS Biol*, **5**:e220.
194. **Koshland, D. E.** 1958. Application of a Theory of Enzyme Specificity to Protein Synthesis. *Proc Natl Acad Sci USA,* **44**:98–104.
195. **Krishnaswamy, S.** 2005. Exosite-driven substrate specificity and function in coagulation. *J Thromb Haemost*, **3**:54–67.
196. **Kujala, P., A. Ikäheimonen, N. Ehsani, H. Vihinen, P. Auvinen, and L. Kääriäinen**. 2001. Biogenesis of the Semliki Forest virus RNA replication complex. *J Virol*, **75**:3873–84.
197. **Kuntzen, T., J. Timm, A. Berical, N. Lennon, A. M. Berlin, S. K. Young, B. Lee, D. Heckerman, J. Carlson, L. L. Reyor, M. Kleyman, C. M. McMahon, C. Birch, J. Schulze Zur Wiesch, T. Ledlie, M. Koehrsen, C. Kodira, A. D. Roberts, G. M. Lauer, H. R. Rosen, F. Bihl, A. Cerny, U. Spengler, Z. Liu, A. Y. Kim, Y. Xing, A. Schneidewind, M. A. Madey, J. F. Fleckenstein, V. M. Park, J. E. Galagan, C. Nusbaum, B. D. Walker, G. V. Lake-Bakaar, E. S. Daar, I. M. Jacobson, E. D. Gomperts, B. R. Edlin, S. M. Donfield, R. T. Chung, A. H. Talal, T. Marion, B. W. Birren, M. R. Henn, and T. M. Allen**. 2008. Naturally occurring dominant resistance mutations to hepatitis C virus protease and polymerase inhibitors in treatment-naïve patients. *Hepatology*, **48**:1769–78.
198. **Kurt, N., T. Haliloglu, and C. A. Schiffer**. 2003. Structure-based prediction of potential binding and nonbinding peptides to HIV-1 protease. *Biophys J*, **85**:853–63.
199. **Kuyumcu-Martinez, N. M., M. Joachims, and R. E. Lloyd**. 2002. Efficient cleavage of ribosome-associated poly(A)-binding protein by enterovirus 3C protease. *J Virol*, **76**:2062–74.
200. **Kääriäinen, L., and T. Ahola**. 2002. Functions of alphavirus nonstructural proteins in RNA replication. *Prog Nucleic Acid Res Mol Biol*, **71**:187–222.
201. **La Linn, M., J. Gardner, D. Warrilow, G. A. Darnell, C. R. McMahon, I. Field, A. D. Hyatt, R. W. Slade, and A. Suhrbier**. 2001. Arbovirus of marine mammals: a new alphavirus isolated from the elephant seal louse, Lepidophthirus macrorhini. *J Virol*, **75**:4103–9.
202. **LaStarza, M. W., J. A. Lemm, and C. M. Rice**. 1994. Genetic analysis of the nsP3 region of Sindbis virus: evidence for roles in minus-strand and subgenomic RNA synthesis. *J Virol*, **68**:5781–91.
203. **Laakkonen, P., T. Ahola, and L. Kääriäinen**. 1996. The effects of palmitoylation on membrane association of Semliki forest virus RNA capping enzyme. *J Biol Chem*, **271**:28567–71.

204. **Laakkonen, P., P. Auvinen, P. Kujala, and L. Kääriäinen**. 1998. Alphavirus replicase protein NSP1 induces filopodia and rearrangement of actin filaments. *J Virol*, **72**:10265–9.
205. **Laakkonen, P., M. Hyvönen, J. Peränen, and L. Kääriäinen**. 1994. Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in Escherichia coli. *J Virol*, **68**:7418–25.
206. **Laine, M., R. Luukkainen, and A. Toivanen**. 2004. Sindbis viruses and other alphaviruses as cause of human arthritic disease. *J Intern Med*, **256**:457–71.
207. **Lam, A. M. I., D. Keeney, and D. N. Frick**. 2003. Two novel conserved motifs in the hepatitis C virus NS3 protein critical for helicase action. *J Biol Chem*, **278**:44514–24.
208. **Lampio, A., I. Kilpeläinen, S. Pesonen, K. Karhi, P. Auvinen, P. Somerharju, and L. Kääriäinen**. 2000. Membrane binding mechanism of an RNA virus-capping enzyme. *J Biol Chem*, **275**:37853–9.
209. **LaStarza, M. W., A. Grakoui, and C. M. Rice**. 1994. Deletion and duplication mutations in the C-terminal nonconserved region of Sindbis virus nsP3: effects on phosphorylation and on virus replication in vertebrate and invertebrate cells. *Virology*, **202**:224–32.
210. **Lauring, A. S., and R. Andino**. 2010. Quasispecies theory and the behavior of RNA viruses. *PLoS Pathog*, **6**:e1001005.
211. **Lee, S.-K., M. Potempa, M. Kolli, A. Ozen, C. Schiffer, and R. Swanstrom**. 2012. Context surrounding processing sites is crucial in determining cleavage rate of a subset of the processing sites in the HIV-1 Gag and Gag-Pro-Pol precursors by the viral protease. *J Biol Chem*, Epub 2012 Feb 13.
212. **Lehmann, M., D. S. Nikolic, and V. Piguet**. 2011. How HIV-1 Takes Advantage of the Cytoskeleton during Replication and Cell-to-Cell Transmission. *Viruses*, **3**:1757–76.
213. **Lemieux, R. U., and U. Spohr**. 1994. How Emil Fischer was led to the lock and key concept for enzyme specificity. *Adv Carbohydr Chem Biochem*, **50**:1–20.
214. **Lemm, J. A., A. Bergqvist, C. M. Read, and C. M. Rice**. 1998. Template-dependent initiation of Sindbis virus RNA replication in vitro. *J Virol*, **72**:6546–53.
215. **Lemm, J. A., and C. M. Rice**. 1993. Assembly of functional Sindbis virus RNA replication complexes: requirement for coexpression of P123 and P34. *J Virol*, **67**:1905–15.
216. **Lemm, J. A., and C. M. Rice**. 1993. Roles of nonstructural polyproteins and cleavage products in regulating Sindbis virus RNA replication and transcription. *J Virol*, **67**:1916–26.
217. **Lemm, J. A., T. Rümenapf, E. G. Strauss, J. H. Strauss, and C. M. Rice**. 1994. Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus- and plus-strand RNA synthesis. *EMBO J*, **13**:2925–34.
218. **Lemon, S. M., J. A. McKeating, T. Pietschmann, D. N. Frick, J. S. Glenn, T. L. Tellinghuisen, J. Symons, and P. A. Furman**. 2010. Development of novel therapies for hepatitis C. *Antiviral Res*, **86**:79–92.
219. **Leung, J. Y.-S., M. M.-L. Ng, and J. J. H. Chu**. 2011. Replication of alphaviruses: a review on the entry process of alphaviruses into cells. *Adv Virol*, **2011**:249640.

220. **Levin, M. K., M. Gurjar, and S. S. Patel**. 2005. A Brownian motor mechanism of translocation and strand separation by hepatitis C virus helicase. *Nat Struct Mol Biol*, **12**:429–35.
221. **Levine, B., J. Marie Hardwick, and D. E. Griffin**. 1994. Persistence of alphaviruses in vertebrate hosts. *Trends Microbiol*, **2**:25–28.
222. **Levis, R., S. Schlesinger, and H. V. Huang**. 1990. Promoter for Sindbis virus RNA-dependent subgenomic RNA transcription. *J Virol*, **64**:1726–33.
223. **Li, G. P., M. W. La Starza, W. R. Hardy, J. H. Strauss, and C. M. Rice**. 1990. Phosphorylation of Sindbis virus nsP3 in vivo and in vitro. *Virology*, **179**:416–27.
224. **Li, G. P., and C. M. Rice**. 1989. Mutagenesis of the in-frame opal termination codon preceding nsP4 of Sindbis virus: studies of translational readthrough and its effect on virus replication. *J Virol*, **63**:1326–37.
225. **Li, G., and C. M. Rice**. 1993. The signal for translational readthrough of a UGA codon in Sindbis virus RNA involves a single cytidine residue immediately downstream of the termination codon. *J Virol*, **67**:5062–7.
226. **Li, K., E. Foy, J. C. Ferreon, M. Nakamura, A. C. M. Ferreon, M. Ikeda, S. C. Ray, M. Gale, and S. M. Lemon**. 2005. Immune evasion by hepatitis C virus NS3/4A protease-mediated cleavage of the Toll-like receptor 3 adaptor protein TRIF. *Proc Natl Acad Sci USA,* **102**:2992–7.
227. **Li, L., J. Jose, Y. Xiang, R. J. Kuhn, and M. G. Rossmann**. 2010. Structural changes of envelope proteins during alphavirus fusion. *Nature*, **468**:705–8.
228. **Li, M.-L., and V. Stollar**. 2007. Distinct sites on the Sindbis virus RNA-dependent RNA polymerase for binding to the promoters for the synthesis of genomic and subgenomic RNA. *J Virol*, **81**:4371–3.
229. **Li, M.-L., and V. Stollar**. 2004. Identification of the amino acid sequence in Sindbis virus nsP4 that binds to the promoter for the synthesis of the subgenomic RNA. *Proc Natl Acad Sci USA,* **101**:9429–34.
230. **Lim, M. D., and C. S. Craik**. 2009. Using specificity to strategically target proteases. *Bioorg Med Chem*, **17**:1094–100.
231. **Limpens, R. W. A. L., H. M. van der Schaar, D. Kumar, A. J. Koster, E. J. Snijder, F. J. M. van Kuppeveld, and M. Bárcena**. 2011. The transformation of enterovirus replication structures: a three-dimensional study of single- and double-membrane compartments. *MBio*, 2(5).
232. **Lindenbach, B. D., B. M. Prágai, R. Montserret, R. K. F. Beran, A. M. Pyle, F. Penin, and C. M. Rice**. 2007. The C terminus of hepatitis C virus NS4A encodes an electrostatic switch that regulates NS5A hyperphosphorylation and viral replication. *J Virol*, **81**:8905–18.
233. **Linger, B. R., L. Kunovska, R. J. Kuhn, and B. L. Golden**. 2004. Sindbis virus nucleocapsid assembly: RNA folding promotes capsid protein dimerization. *RNA*, **10**:128–38.
234. **Liu, L., H. Dong, H. Chen, J. Zhang, H. Ling, Z. Li, P.-Y. Shi, and H. Li**. 2010. Flavivirus RNA cap methyltransferase: structure, function, and inhibition. *Front Biol*, **5**:286–303.
235. **London, N., D. Movshovitz-Attias, and O. Schueler-Furman**. 2010. The structural basis of peptide-protein binding strategies. *Structure*, **18**:188–99.
236. **Lorsch, J. R.** 2002. RNA Chaperones Exist and DEAD Box Proteins Get a Life. *Cell*, **109**:797–800.

237. **Lulla, A., K. Tints, and S. Tobi**. 2007. Optimized recognition site of the alphavirus non-structural protease for tag removal and specific processing of recombinant proteins. U.S. Patent №7189540.
238. **Lulla, V., A. Merits, P. Sarin, L. Kääriäinen, S. Keränen, and T. Ahola**. 2006. Identification of mutations causing temperature-sensitive defects in Semliki Forest virus RNA synthesis. *J Virol*, **80**:3108–11.
239. **Lulla, V., D. L. Sawicki, S. G. Sawicki, A. Lulla, A. Merits, and T. Ahola**. 2008. Molecular defects caused by temperature-sensitive mutations in Semliki Forest virus nsP1. *J Virol*, **82**:9236–44.
240. **Lundstrom, K.** 2005. Biology and application of alphaviruses in gene therapy. *Gene Ther*, 12 Suppl 1:S92–7.
241. **López-Otín, C., and J. S. Bond**. 2008. Proteases: multifunctional enzymes in life and disease. *J Biol Chem*, **283**:30433–7.
242. **Mackenzie, J.** 2005. Wrapping things up about virus RNA replication. *Traffic*, **6**:967–77.
243. **Madala, P. K., J. D. A. Tyndall, T. Nall, and D. P. Fairlie**. 2010. Update 1 of: Proteases universally recognize beta strands in their active sites. *Chem Rev*, **110**:PR1–31.
244. **Magden, J., N. Takeda, T. Li, P. Auvinen, T. Ahola, T. Miyamura, A. Merits, and L. Kääriäinen**. 2001. Virus-specific mRNA capping enzyme encoded by hepatitis E virus. *J Virol*, **75**:6249–55.
245. **Magliano, D., J. A. Marshall, D. S. Bowden, N. Vardaxis, J. Meanger, and J. Y. Lee**. 1998. Rubella virus replication complexes are virus-modified lysosomes. *Virology*, **240**:57–63.
246. **Mai, J., S. G. Sawicki, and D. L. Sawicki**. 2009. Fate of minus-strand templates and replication complexes produced by a p23-cleavage-defective mutant of Sindbis virus. *J Virol*, **83**:8553–64.
247. **Malet, H., B. Coutard, S. Jamal, H. Dutartre, N. Papageorgiou, M. Neuvonen, T. Ahola, N. Forrester, E. A. Gould, D. Lafitte, F. Ferron, J. Lescar, A. E. Gorbalenya, X. de Lamballerie, and B. Canard**. 2009. The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *J Virol*, **83**:6534–45.
248. **Mancini, E. J., M. Clarke, B. E. Gowen, T. Rutten, and S. D. Fuller**. 2000. Cryo-electron microscopy reveals the functional organization of an enveloped virus, Semliki Forest virus. *Mol Cell*, **5**:255–66.
249. **Marcotte, L. L., A. B. Wass, D. W. Gohara, H. B. Pathak, J. J. Arnold, D. J. Filman, C. E. Cameron, and J. M. Hogle**. 2007. Crystal structure of poliovirus 3CD protein: virally encoded protease and precursor to the RNA-dependent RNA polymerase. *J Virol*, **81**:3583–96.
250. **Marsh, M., and A. Helenius**. 2006. Virus entry: open sesame. *Cell*, **124**:729–40.
251. **Martín, V., and E. Domingo**. 2008. Influence of the mutant spectrum in viral evolution: focused selection of antigenic variants in a reconstructed viral quasispecies. *Mol Biol Evol*, **25**:1544–54.
252. **Mayuri, T. W. Geders, J. L. Smith, and R. J. Kuhn**. 2008. Role for conserved residues of sindbis virus nonstructural protein 2 methyltransferase-like domain in regulation of minus-strand synthesis and development of cytopathic infection. *J Virol*, **82**:7284–97.

253. **McInerney, G. M., N. L. Kedersha, R. J. Kaufman, P. Anderson, and P. Liljeström**. 2005. Importance of eIF2alpha phosphorylation and stress granule assembly in alphavirus translation regulation. *Mol Biol Cell*, **16**:3753–63.
254. **McLoughlin, M. F., and D. A. Graham**. 2007. Alphavirus infections in salmonids – a review. *J Fish Dis*, **30**:511–31.
255. **McMahon, H. T., and J. L. Gallop**. 2005. Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature*, **438**:590–6.
256. **Melton, J. V., G. D. Ewart, R. C. Weir, P. G. Board, E. Lee, and P. W. Gage**. 2002. Alphavirus 6K proteins form ion channels. *J Biol Chem*, **277**:46923–31.
257. **Mercer, J., M. Schelhaas, and A. Helenius**. 2010. Virus entry by endocytosis. *Annu Rev Biochem*, **79**:803–33.
258. **Merits, A., R. Kettunen, K. Mäkinen, A. Lampio, P. Auvinen, L. Kääriäinen, and T. Ahola**. 1999. Virus-specific capping of tobacco mosaic virus RNA: methylation of GTP prior to formation of covalent complex p126-m7GMP. *FEBS Lett*, **455**:45–8.
259. **Merits, A., L. Vasiljeva, T. Ahola, L. Kääriäinen, and P. Auvinen**. 2001. Proteolytic processing of Semliki Forest virus-specific non-structural polyprotein by nsP2 protease. *J Gen Virol*, **82**:765–73.
260. **Meylan, E., J. Curran, K. Hofmann, D. Moradpour, M. Binder, R. Bartenschlager, and J. Tschopp**. 2005. Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. *Nature*, **437**:1167–72.
261. **Mi, S., and V. Stollar**. 1991. Expression of Sindbis virus nsP1 and methyltransferase activity in Escherichia coli. *Virology*, **184**:423–7.
262. **Miller, S., and J. Krijnse-Locker**. 2008. Modification of intracellular membrane structures for virus replication. *Nat Rev Microbiol*, **6**:363–74.
263. **Mittl, P. R., and M. G. Grütter**. 2006. Opportunities for structure-based design of protease-directed drugs. *Curr Opin Struct Biol*, **16**:769–75.
264. **Montgomery, S. A., P. Berglund, C. W. Beard, and R. E. Johnston**. 2006. Ribosomal protein S6 associates with alphavirus nonstructural protein 2 and mediates expression from alphavirus messages. *J Virol*, **80**:7729–39.
265. **Montgomery, S. A., and R. E. Johnston**. 2007. Nuclear import and export of Venezuelan equine encephalitis virus nonstructural protein 2. *J Virol*, **81**:10268–79.
266. **Moradpour, D., F. Penin, and C. M. Rice**. 2007. Replication of hepatitis C virus. Nature reviews. *Microbiology*, **5**:453–63.
267. **Mossessova, E., and C. D. Lima**. 2000. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Mol Cell*, **5**:865–76.
268. **Mukhopadhyay, S., W. Zhang, S. Gabler, P. R. Chipman, E. G. Strauss, J. H. Strauss, T. S. Baker, R. J. Kuhn, and M. G. Rossmann**. 2006. Mapping the structure and function of the E1 and E2 glycoproteins in alphaviruses. *Structure*, **14**:63–73.
269. **Murphy, G.** 2008. The ADAMs: signalling scissors in the tumour microenvironment. Nature reviews. *Cancer*, **8**:929–41.
270. **Muzammil, S., P. Ross, and E. Freire**. 2003. A major role for a set of non-active site mutations in the development of HIV-1 protease drug resistance. *Biochemistry*, **42**:631–8.
271. **Myles, K. M., C. L. H. Kelly, J. P. Ledermann, and A. M. Powers**. 2006. Effects of an opal termination codon preceding the nsP4 gene sequence in the

O'Nyong-Nyong virus genome on Anopheles gambiae infectivity. *J Virol*, **80**:4992–7.

272. **Nakayama, K.** 1997. Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins. *Biochem J*, **327**:625–35.
273. **Nalam, M. N. L., A. Ali, M. D. Altman, G. S. K. K. Reddy, S. Chellappan, V. Kairys, A. Ozen, H. Cao, M. K. Gilson, B. Tidor, T. M. Rana, and C. A. Schiffer**. 2010. Evaluating the substrate-envelope hypothesis: structural analysis of novel HIV-1 protease inhibitors designed to be robust against drug resistance. *J Virol*, **84**:5368–78.
274. **Nalam, M. N. L., and C. A. Schiffer**. 2008. New approaches to HIV protease inhibitor drug design II: testing the substrate envelope hypothesis to avoid drug resistance and discover robust inhibitors. *Curr Opin HIV AIDS*, **3**:642–6.
275. **Netherton, C., K. Moffat, E. Brooks, and T. Wileman**. 2007. A guide to viral inclusions, membrane rearrangements, factories, and viroplasm produced during virus replication. *Adv Virus Res*, **70**:101–82.
276. **Neuvonen, M., and T. Ahola**. 2009. Differential activities of cellular and viral macro domain proteins in binding of ADP-ribose metabolites. *J Mol Biol*, **385**:212–25.
277. **Neuvonen, M., A. Kazlauskas, M. Martikainen, A. Hinkkanen, T. Ahola, and K. Saksela**. 2011. SH3 domain-mediated recruitment of host cell amphiphysins by alphavirus nsP3 promotes viral RNA replication. *PLoS Pathog*, **7**:e1002383.
278. **Ng, K. K. S., J. J. Arnold, and C. E. Cameron**. 2008. Structure-function relationships among RNA-dependent RNA polymerases. *Curr Top Microbiol Immunol*, **320**:137–56.
279. **Ng, N. M., R. N. Pike, and S. E. Boyd**. 2009. Subsite cooperativity in protease specificity. *Biol Bhem*, **390**:401–7.
280. **Nickens, D. G., and R. W. Hardy**. 2008. Structural and functional analyses of stem-loop 1 of the Sindbis virus genome. *Virology*, **370**:158–72.
281. **Niesters, H. G., and J. H. Strauss**. 1990. Mutagenesis of the conserved 51-nucleotide region of Sindbis virus. *J Virol*, **64**:1639–47.
282. **Novoa, R. R., G. Calderita, R. Arranz, J. Fontana, H. Granzow, and C. Risco**. 2005. Virus factories: associations of cell organelles for viral replication and morphogenesis. *Biol Cell*, **97**:147–72.
283. **Nägler, D. K., W. Tam, A. C. Storer, J. C. Krupa, J. S. Mort, and R. Ménard**. 1999. Interdependency of sequence and positional specificities for cysteine proteases of the papain family. *Biochemistry*, **38**:4868–74.
284. **Obeid, S., N. Blatter, R. Kranaster, A. Schnur, K. Diederichs, W. Welte, and A. Marx**. 2010. Replication through an abasic DNA lesion: structural basis for adenine selectivity. *EMBO J*, **29**:1738–47.
285. **Obeid, S., W. Welte, K. Diederichs, and A. Marx**. 2012. Amino acid templating mechanisms in selection of nucleotides opposite abasic sites by a family A DNA polymerase. *J Biol Chem*, Epub 2012 Feb 7.
286. **Ohtaka, H., A. Schön, and E. Freire**. 2003. Multidrug resistance to HIV-1 protease inhibition requires cooperative coupling between distal mutations. *Biochemistry*, **42**:13659–66.
287. **Ojosnegros, S., C. Perales, A. Mas, and E. Domingo**. 2011. Quasispecies as a matter of fact: Viruses and beyond. *Virus Res*, **162**:203–15.

288. **Ou, J. H., E. G. Strauss, and J. H. Strauss**. 1983. The 5'-terminal sequences of the genomic RNAs of several alphaviruses. *J Mol Biol*, **168**:1–15.
289. **Ou, J. H., D. W. Trent, and J. H. Strauss**. 1982. The 3'-non-coding regions of alphavirus RNAs contain repeating sequences. *J Mol Biol*, **156**:719–30.
290. **Overall, C. M.** 2002. Molecular determinants of metalloproteinase substrate specificity: matrix metalloproteinase substrate binding domains, modules, and exosites. *Mol Biotechnol*, **22**:51–86.
291. **Overall, C. M., and G. S. Butler**. 2007. Protease yoga: extreme flexibility of a matrix metalloproteinase. *Structure*, **15**:1159–61.
292. **Ozen, A., T. Haliloğlu, and C. A. Schiffer**. 2011. Dynamics of preferential substrate recognition in HIV-1 protease: redefining the substrate envelope. *J Mol Biol*, **410**:726–44.
293. **Ozer, N., T. Haliloglu, and C. A. Schiffer**. 2006. Substrate specificity in HIV-1 protease by a biased sequence search method. *Proteins*, **64**:444–56.
294. **Pan, R., I. Image, X.-J. Zhang, I. I. Image, Z.-J. Zhang, Y. Zhou, W.-X. Tian, and R.-Q. He**. 2010. Substrate-induced changes in protease active site conformation impact on subsequent reactions with substrates. *J Biol Chem*, **285**:22950–6.
295. **Park, N., T. Skern, and K. E. Gustin**. 2010. Specific cleavage of the nuclear pore complex protein Nup62 by a viral protease. *J Biol Chem,* **285**:28796–805.
296. **Patel, S. S., and I. Donmez**. 2006. Mechanisms of helicases. *J Biol Chem*, **281**:18265–8.
297. **Pehrson, J. R., and R. N. Fuji**. 1998. Evolutionary conservation of histone macroH2A subtypes and domains. *Nucleic Acids Res*, **26**:2837–42.
298. **Pelkmans, L., and A. Helenius**. 2003. Insider information: what viruses tell us about endocytosis. *Curr Opin Cell Biol*, **15**:414–22.
299. **Perera, R., S. Daijogo, B. L. Walter, J. H. C. Nguyen, and B. L. Semler**. 2007. Cellular protein modification by poliovirus: the two faces of poly(rC)-binding protein. *J Virol*, **81**:8919–32.
300. **Perkins, J. R., I. Diboun, B. H. Dessailly, J. G. Lees, and C. Orengo**. 2010. Transient protein-protein interactions: structural, functional, and network properties. *Structure*, **18**:1233–43.
301. **Perri, S., D. A. Driver, J. P. Gardner, S. Sherrill, B. A. Belli, T. W. Dubensky, and J. M. Polo**. 2000. Replicon vectors derived from Sindbis virus and Semliki forest virus that establish persistent replication in host cells. *J Virol*, **74**:9802–7.
302. **Peränen, J.** 1991. Localization and phosphorylation of Semliki Forest virus non-structural protein nsP3 expressed in COS cells from a cloned cDNA. *J Gen Virol*, **72**:195–9.
303. **Peränen, J., and L. Kääriäinen**. 1991. Biogenesis of type I cytopathic vacuoles in Semliki Forest virus-infected BHK cells. *J Virol*, **65**:1623–7.
304. **Peränen, J., P. Laakkonen, M. Hyvönen, and L. Kääriäinen**. 1995. The alphavirus replicase protein nsP1 is membrane-associated and has affinity to endocytic organelles. *Virology*, **208**:610–20.
305. **Pettit, S. C., J. N. Lindquist, A. H. Kaplan, and R. Swanstrom**. 2005. Processing sites in the human immunodeficiency virus type 1 (HIV-1) Gag-Pro-Pol precursor are cleaved by the viral protease at different rates. *Retrovirology*, **2**:66.
306. **Pettit, S. C., J. C. Clemente, J. A. Jeung, B. M. Dunn, and A. H. Kaplan**. 2005. Ordered processing of the human immunodeficiency virus type 1 GagPol

precursor is influenced by the context of the embedded viral protease. *J Virol*, **79**:10601–7.
307. **Powers, A. M., A. C. Brault, Y. Shirako, E. G. Strauss, W. Kang, J. H. Strauss, and S. C. Weaver**. 2001. Evolutionary relationships and systematics of the alphaviruses. *J Virol*, **75**:10118–31.
308. **Prabu-Jeyabalan, M., E. A. Nalivaika, K. Romano, and C. A. Schiffer**. 2006. Mechanism of substrate recognition by drug-resistant human immunodeficiency virus type 1 protease variants revealed by a novel structural intermediate. *J Virol*, **80**:3607–16.
309. **Prabu-Jeyabalan, M., E. Nalivaika, and C. A. Schiffer**. 2002. Substrate shape determines specificity of recognition for HIV-1 protease: analysis of crystal structures of six substrate complexes. *Structure*, **10**:369–81.
310. **Preugschat, F., D. R. Averett, B. E. Clarke, and D. J. Porter**. 1996. A steady-state and pre-steady-state kinetic analysis of the NTPase activity associated with the hepatitis C virus NS3 helicase domain. *J Biol Chem*, **271**:24449–57.
311. **Puente, X. S., L. M. Sánchez, A. Gutiérrez-Fernández, G. Velasco, and C. López-Otín**. 2005. A genomic view of the complexity of mammalian proteolytic systems. *Biochem Soc Trans*, **33**:331–4.
312. **Puente, X. S., L. M. Sánchez, C. M. Overall, and C. López-Otín**. 2003. Human and mouse proteases: a comparative genomic approach. Nature reviews. *Genetics*, **4**:544–58.
313. **Pyle, A. M.** 2011. RNA helicases and remodeling proteins. *Curr Opin Chem Biol*, **15**:636–42.
314. **Qu, L., Z. Feng, D. Yamane, Y. Liang, R. E. Lanford, K. Li, and S. M. Lemon**. 2011. Disruption of TLR3 Signaling Due to Cleavage of TRIF by the Hepatitis A Virus Protease-Polymerase Processing Intermediate, 3CD. *PLoS Pathog*, **7**:e1002169.
315. **Quetglas, J. I., M. Ruiz-Guillen, A. Aranda, E. Casales, J. Bezunartea, and C. Smerdou**. 2010. Alphavirus vectors for cancer therapy. *Virus Res*, **153**:179–96.
316. **Rahmeh, A. A., A. D. Schenk, E. I. Danek, P. J. Kranzusch, B. Liang, T. Walz, and S. P. J. Whelan**. 2010. Molecular architecture of the vesicular stomatitis virus RNA polymerase. *Proc Natl Acad Sci USA,***107**:20075–80.
317. **Rajagopal, V., and S. S. Patel**. 2008. Single strand binding proteins increase the processivity of DNA unwinding by the hepatitis C virus helicase. *J Mol Biol*, **376**:69–79.
318. **Rajagopal, V., S. S. Patel, K. D. Raney, M. Gotte, and C. E. Cameron**. 2009. Viral Genome Replication. Springer US, Boston, MA.
319. **Rajkowitsch, L., D. Chen, S. Stampfl, K. Semrad, C. Waldsich, O. Mayer, M. F. Jantsch, R. Konrat, U. Bläsi, and R. Schroeder**. 2007. RNA chaperones, RNA annealers and RNA helicases. *RNA Biol*, **4**:118–30.
320. **Raju, R., M. Hajjou, K. R. Hill, V. Botta, and S. Botta**. 1999. In vivo addition of poly(A) tail and AU-rich sequences to the 3' terminus of the Sindbis virus RNA genome: a novel 3'-end repair pathway. *J Virol*, **73**:2410–9.
321. **Raju, R., and H. V. Huang**. 1991. Analysis of Sindbis virus promoter recognition in vivo, using novel vectors with two subgenomic mRNA promoters. *J Virol*, **65**:2501–10.
322. **Rawlings, N. D., A. J. Barrett, and A. Bateman**. 2011. MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. *Nucleic Acids Res*, **40**:D343–50.

323. **Reiss, S., I. Rebhan, P. Backes, I. Romero-Brey, H. Erfle, P. Matula, L. Kaderali, M. Poenisch, H. Blankenburg, M.-S. Hiet, T. Longerich, S. Diehl, F. Ramirez, T. Balla, K. Rohr, A. Kaul, S. Bühler, R. Pepperkok, T. Lengauer, M. Albrecht, R. Eils, P. Schirmacher, V. Lohmann, and R. Bartenschlager**. 2011. Recruitment and activation of a lipid kinase by hepatitis C virus NS5A is essential for integrity of the membranous replication compartment. *Cell Host Microbe*, **9**:32–45.
324. **Rikkonen, M., J. Peränen, and L. Kääriäinen**. 1994. ATPase and GTPase activities associated with Semliki Forest virus nonstructural protein nsP2. *J Virol*, **68**:5804–10.
325. **Rikkonen, M., J. Peränen, and L. Kääriäinen**. 1992. Nuclear and nucleolar targeting signals of Semliki Forest virus nonstructural protein nsP2. *Virology*, **189**:462–73.
326. **Rikkonen, M., J. Peränen, and L. Kääriäinen**. 1994. Nuclear targeting of Semliki Forest virus nsP2. *Arch Virol*, Suppl **9**:369–77.
327. **Robinson, M., Y. Tian, W. E. Delaney, and A. E. Greenstein**. 2011. Preexisting drug-resistance mutations reveal unique barriers to resistance for distinct antivirals. *Proc Natl Acad Sci USA,***108**:10290–5.
328. **Romano, K. P., A. Ali, W. E. Royer, and C. A. Schiffer**. 2010. Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding. *Proc Natl Acad Sci USA,***107**:20986–91.
329. **Romano, K. P., J. M. Laine, L. M. Deveau, H. Cao, F. Massi, and C. A. Schiffer**. 2011. Molecular mechanisms of viral and host cell substrate recognition by hepatitis C virus NS3/4A protease. *J Virol*, **85**:6106–16.
330. **Romero, P., Z. Obradovic, and A. K. Dunker**. 2004. Natively disordered proteins: functions and predictions. *Appl Bioinform*, **3**:105–13.
331. **Rose, P. P., S. L. Hanna, A. Spiridigliozzi, N. Wannissorn, D. P. Beiting, S. R. Ross, R. W. Hardy, S. A. Bambina, M. T. Heise, and S. Cherry**. 2011. Natural resistance-associated macrophage protein is a cellular receptor for sindbis virus in both insect and Mammalian hosts. *Cell Host Microbe*, **10**:97–104.
332. **Rosenblum, G., P. E. Van den Steen, S. R. Cohen, J. G. Grossmann, J. Frenkel, R. Sertchook, N. Slack, R. W. Strange, G. Opdenakker, and I. Sagi**. 2007. Insights into the structure and domain flexibility of full-length pro-matrix metalloproteinase-9/gelatinase B. *Structure*, **15**:1227–36.
333. **Rost, B., G. Yachdav, and J. Liu**. 2004. The PredictProtein server. *Nucleic Acids Res*, **32**:W321–6.
334. **Rozanov, M. N., E. V. Koonin, and A. E. Gorbalenya**. 1992. Conservation of the putative methyltransferase domain: a hallmark of the "Sindbis-like" supergroup of positive-strand RNA viruses. *J Gen Virol*, **73**:2129–34.
335. **Rubach, J. K., B. R. Wasik, J. C. Rupp, R. J. Kuhn, R. W. Hardy, and J. L. Smith**. 2009. Characterization of purified Sindbis virus nsP4 RNA-dependent RNA polymerase activity in vitro. *Virology*, **384**:201–8.
336. **Rungrotmongkol, T., N. Nunthaboot, M. Malaisree, N. Kaiyawet, P. Yotmanee, A. Meeprasert, and S. Hannongbua**. 2010. Molecular insight into the specific binding of ADP-ribose to the nsP3 macro domains of chikungunya and Venezuelan equine encephalitis viruses: molecular dynamics simulations and free energy calculations. *J Mol Graph Model*, **29**:347–53.
337. **Rupp, J. C., N. Jundt, and R. W. Hardy**. 2011. Requirement for the amino-terminal domain of sindbis virus nsP4 during virus infection. *J Virol*, **85**:3449–60.

338. **Russo, A. T., R. D. Malmstrom, M. A. White, and S. J. Watowich**. 2010. Structural basis for substrate specificity of alphavirus nsP2 proteases. *J Mol Graph Model*, **29**:46–53.
339. **Russo, A. T., M. A. White, and S. J. Watowich**. 2006. The crystal structure of the Venezuelan equine encephalitis alphavirus nsP2 protease. *Structure*, **14**:1449–58.
340. **Ryman, K. D., K. C. Meier, E. M. Nangle, S. L. Ragsdale, N. L. Korneeva, R. E. Rhoads, M. R. MacDonald, and W. B. Klimstra**. 2005. Sindbis virus translation is inhibited by a PKR/RNase L-independent effector induced by alpha/beta interferon priming of dendritic cells. *J Virol*, **79**:1487–99.
341. **Ryman, K. D., C. L. Gardner, C. W. Burke, K. C. Meier, J. M. Thompson, and W. B. Klimstra**. 2007. Heparan sulfate binding can contribute to the neurovirulence of neuroadapted and nonneuroadapted Sindbis viruses. *J Virol*, **81**:3563–73.
342. **Rögnvaldsson, T., T. A. Etchells, L. You, D. Garwicz, I. Jarman, and P. J. G. Lisboa**. 2009. How to find simple and accurate rules for viral protease cleavage specificities. *BMC Bioinform*, **10**:149.
343. **Salazar, C., and T. Höfer**. 2009. Multisite protein phosphorylation – from molecular mechanisms to kinetic models. *FEBS J*, **276**:3177–98.
344. **Salonen, A., T. Ahola, and L. Kääriäinen**. 2005. Viral RNA replication in association with cellular membranes. *Curr Top Microbiol Immunol*, **285**:139–73.
345. **Salonen, A., L. Vasiljeva, A. Merits, J. Magden, E. Jokitalo, and L. Kääriäinen**. 2003. Properly folded nonstructural polyprotein directs the semliki forest virus replication complex to the endosomal compartment. *J Virol*, **77**:1691–702.
346. **Salverda, M. L. M., E. Dellus, F. A. Gorter, A. J. M. Debets, J. van der Oost, R. F. Hoekstra, D. S. Tawfik, and J. A. G. M. de Visser**. 2011. Initial mutations direct alternative pathways of protein evolution. *PLoS Genet*, **7**:e1001321.
347. **Sanjuán, R., M. R. Nebot, N. Chirico, L. M. Mansky, and R. Belshaw**. 2010. Viral mutation rates. *J Virol*, **84**:9733–48.
348. **Santagati, M. G., P. V. Itäranta, P. R. Koskimies, J. A. Määttä, A. A. Salmi, and A. E. Hinkkanen**. 1994. Multiple repeating motifs are found in the 3'-terminal non-translated region of Semliki Forest virus A7 variant genome. *J Gen Virol*, **75**:1499–504.
349. **Sanz, M. A., A. Castelló, I. Ventoso, J. J. Berlanga, and L. Carrasco**. 2009. Dual mechanism for the translation of subgenomic mRNA from Sindbis virus in infected and uninfected cells. *PloS One*, **4**:e4772.
350. **Sanz-Ramos, M., F. Díaz-San Segundo, C. Escarmís, E. Domingo, and N. Sevilla**. 2008. Hidden virulence determinants in a viral quasispecies in vivo. *J Virol*, **82**:10465–76.
351. **Sariola, M., J. Saraste, and E. Kuismanen**. 1995. Communication of post-Golgi elements with early endocytic pathway: regulation of endoproteolytic cleavage of Semliki Forest virus p62 precursor. *J Cell Sci*, **108**:2465–75.
352. **Sasvari, Z., and P. D. Nagy**. 2010. Making of viral replication organelles by remodeling interior membranes. *Viruses*, **2**:2436–42.
353. **Sauer, R. T., and T. A. Baker**. 2011. AAA+ Proteases: ATP-Fueled Machines of Protein Destruction. *Annu Rev Biochem*, **80**:587–612.

354. **Sawicki, D. L., and S. G. Sawicki**. 1993. A second nonstructural protein functions in the regulation of alphavirus negative-strand RNA synthesis. *J Virol*, **67**:3605–10.
355. **Sawicki, D. L., and S. G. Sawicki**. 1994. Alphavirus positive and negative strand RNA synthesis and the role of polyproteins in formation of viral replication complexes. *Arch Virol*, Suppl **9**:393–405.
356. **Sawicki, D. L., and S. G. Sawicki**. 1985. Functional analysis of the A complementation group mutants of Sindbis HR virus. *Virology*, **144**:20–34.
357. **Sawicki, D. L., and S. G. Sawicki**. 1980. Short-lived minus-strand polymerase for Semliki Forest virus. *J Virol*, **34**:108–18.
358. **Sawicki, D. L., S. G. Sawicki, S. Keränen, and L. Kääriäinen**. 1981. Specific Sindbis virus-coded function for minus-strand RNA synthesis. *J Virol*, **39**:348–58.
359. **Sawicki, D. L., S. Perri, J. M. Polo, and S. G. Sawicki**. 2006. Role for nsP2 proteins in the cessation of alphavirus minus-strand synthesis by host cells. *J Virol*, **80**:360–71.
360. **Schechter, I., and A. Berger**. 1967. On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun*, **27**:157–62.
361. **Schiller, M. R., T. Mi, J. C. Merlin, S. Deverasetty, M. R. Gryk, T. J. Bill, A. W. Brooks, L. Y. Lee, V. Rathnayake, C. A. Ross, D. P. Sargeant, C. L. Strong, P. Watts, and S. Rajasekaran**. 2011. Minimotif Miner 3.0: database expansion and significantly improved reduction of false-positive predictions from consensus sequences. *Nucleic Acids Res*, **40**:D252–60.
362. **Schlesinger, S.** 2001. Alphavirus vectors: development and potential therapeutic applications. *Exp Opin Biol Ther*, **1**:177–91.
363. **Schulz, O., A. Pichlmair, J. Rehwinkel, N. C. Rogers, D. Scheuner, H. Kato, O. Takeuchi, S. Akira, R. J. Kaufman, and C. Reis e Sousa**. 2010. Protein kinase R contributes to immunity against specific viruses by regulating interferon mRNA integrity. *Cell Host Microbe*, **7**:354–61.
364. **Schwartz, M., J. Chen, M. Janda, M. Sullivan, J. den Boon, and P. Ahlquist**. 2002. A positive-strand RNA virus replication complex parallels form and function of retrovirus capsids. *Mol Cell*, **9**:505–14.
365. **Schwartz, O., and M. L. Albert**. 2010. Biology and pathogenesis of chikungunya virus. Nature reviews. *Microbiology*, **8**:491–500.
366. **Scott, W. R., and C. A. Schiffer**. 2000. Curling of flap tips in HIV-1 protease as a mechanism for substrate entry and tolerance of drug resistance. *Structure*, **8**:1259–65.
367. **Seipelt, J., A. Guarné, E. Bergmann, M. James, W. Sommergruber, I. Fita, and T. Skern**. 1999. The structures of picornaviral proteinases. *Virus Res*, **62**:159–68.
368. **Shirako, Y., E. G. Strauss, and J. H. Strauss**. 2000. Suppressor mutations that allow sindbis virus RNA polymerase to function with nonaromatic amino acids at the N-terminus: evidence for interaction between nsP1 and nsP4 in minus-strand RNA synthesis. *Virology*, **276**:148–60.
369. **Shirako, Y., and J. H. Strauss**. 1990. Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology*, **177**:54–64.
370. **Shirako, Y., and J. H. Strauss**. 1994. Regulation of Sindbis virus RNA replication: uncleaved P123 and nsP4 function in minus-strand RNA synthesis,

whereas cleaved products from P123 are required for efficient plus-strand RNA synthesis. *J Virol*, **68**:1874–85.
371. **Shirako, Y., and J. H. Strauss**. 1998. Requirement for an aromatic amino acid or histidine at the N terminus of Sindbis virus RNA polymerase. *J Virol*, **72**:2310–5.
372. **Shirako, Y., E. G. Strauss, and J. H. Strauss**. 2003. Modification of the 5' terminus of Sindbis virus genomic RNA allows nsP4 RNA polymerases with nonaromatic amino acids at the N terminus to function in RNA replication. *J Virol*, **77**:2301–9.
373. **Simmons, J. D., L. J. White, T. E. Morrison, S. A. Montgomery, A. C. Whitmore, R. E. Johnston, and M. T. Heise**. 2009. Venezuelan equine encephalitis virus disrupts STAT1 signaling by distinct mechanisms independent of host shutoff. *J Virol*, **83**:10571–81.
374. **Simón-Mateo, C., G. Andrés, and E. Viñuela**. 1993. Polyprotein processing in African swine fever virus: a novel gene expression strategy for a DNA virus. *EMBO J*, **12**:2977–87.
375. **Singh, I., and A. Helenius**. 1992. Role of ribosomes in Semliki Forest virus nucleocapsid uncoating. *J Virol*, **66**:7049–58.
376. **Singh, M. K., K. Streu, A. J. McCrone, and B. N. Dominy**. 2011. The evolution of catalytic function in the HIV-1 protease. *J Mol Biol*, **408**:792–805.
377. **Singleton, M. R., M. S. Dillingham, and D. B. Wigley**. 2007. Structure and mechanism of helicases and nucleic acid translocases. *Annu Rev Biochem*, **76**:23–50.
378. **Sjöberg, E. M., M. Suomalainen, and H. Garoff**. 1994. A significantly improved Semliki Forest virus expression system based on translation enhancer segments from the viral capsid gene. *Biotechnology (N Y)*, **12**:1127–31.
379. **Song, J., H. Tan, S. E. Boyd, H. Shen, K. Mahmood, G. I. Webb, T. Akutsu, J. C. Whisstock, and R. N. Pike**. 2011. Bioinformatic approaches for predicting substrates of proteases. *J Bioinform Comput Biol*, **9**:149–78.
380. **Soonsawad, P., L. Xing, E. Milla, J. M. Espinoza, M. Kawano, M. Marko, C. Hsieh, H. Furukawa, M. Kawasaki, W. Weerachatyanukul, R. Srivastava, S. W. Barnett, I. K. Srivastava, and R. H. Cheng**. 2010. Structural evidence of glycoprotein assembly in cellular membrane compartments prior to Alphavirus budding. *J Virol*, **84**:11145–51.
381. **Soskine, M., and D. S. Tawfik**. 2010. Mutational effects and the evolution of new protein functions. *Nat Rev Genet*, **11**:572–82.
382. **Sousa, I. P., C. A. M. Carvalho, D. F. Ferreira, G. Weissmüller, G. M. Rocha, J. L. Silva, and A. M. O. Gomes**. 2011. Envelope lipid-packing as a critical factor for the biological activity and stability of alphavirus particles isolated from mammalian and mosquito cells. *J Biol Chem*, **286**:1730–6.
383. **Spagnolo, J. F., E. Rossignol, E. Bullitt, and K. Kirkegaard**. 2010. Enzymatic and nonenzymatic functions of viral RNA-dependent RNA polymerases within oligomeric arrays. *RNA*, **16**:382–93.
384. **Spall V.E., Shanks M., and Lomonossoff G.P.** 1997. Polyprotein Processing as a Strategy for Gene Expression in RNA Viruses. *Semin Virol*, **8**:9.
385. **Spuul, P., G. Balistreri, L. Kääriäinen, and T. Ahola**. 2010. Phosphatidylinositol 3-kinase-, actin-, and microtubule-dependent transport of Semliki Forest Virus replication complexes from the plasma membrane to modified lysosomes. *J Virol*, **84**:7543–57.

386. **Spuul, P., A. Salonen, A. Merits, E. Jokitalo, L. Kääriäinen, and T. Ahola**. 2007. Role of the amphipathic peptide of Semliki forest virus replicase protein nsP1 in membrane association and virus replication. *J Virol*, **81**:872–83.
387. **Stapleford, K. A., and D. J. Miller**. 2010. Role of cellular lipids in positive-sense RNA virus replication complex assembly and function. *Viruses*, **2**:1055–68.
388. **Steuber, H., and R. Hilgenfeld**. 2010. Recent advances in targeting viral proteases for the discovery of novel antivirals. *Curr Top Med Chem*, **10**:323–45.
389. **Strauss, E. G., R. J. De Groot, R. Levinson, and J. H. Strauss**. 1992. Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology*, **191**:932–40.
390. **Strauss, J. H., and E. G. Strauss**. 1994. The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev*, **58**:491–562.
391. **Strauss, J. H., K. S. Wang, A. L. Schmaljohn, R. J. Kuhn, and E. G. Strauss**. 1994. Host-cell receptors for Sindbis virus. *Arch Virol*, Suppl **9**:473–84.
392. **Sukuru, S. C. K., F. Nigsch, J. Quancard, M. Renatus, R. Chopra, N. Brooijmans, D. Mikhailov, Z. Deng, A. Cornett, J. L. Jenkins, U. Hommel, J. W. Davies, and M. Glick**. 2010. A lead discovery strategy driven by a comprehensive analysis of proteases in the peptide substrate space. *Prot Sci*, **19**:2096–109.
393. **Summers, D. F., and J. V. Maizel**. 1968. Evidence for large precursor proteins in poliovirus synthesis. *Proc Natl Acad Sci USA,* **59**:966–71.
394. **Suopanki, J., D. L. Sawicki, S. G. Sawicki, and L. Kääriäinen**. 1998. Regulation of alphavirus 26S mRNA transcription by replicase component nsP2. *J Gen Virol*, **79**:309–19.
395. **Sánchez-San Martín, C., C. Y. Liu, and M. Kielian**. 2009. Dealing with low pH: entry and exit of alphaviruses and flaviviruses. *Trends Microbiol*, **17**:514–21.
396. **Tamm, K., A. Merits, and I. Sarand**. 2008. Mutations in the nuclear localization signal of nsP2 influencing RNA synthesis, protein expression and cytotoxicity of Semliki Forest virus. *J Gen Virol*, **89**:676–86.
397. **Tan, J., C. Vonrhein, O. S. Smart, G. Bricogne, M. Bollati, Y. Kusov, G. Hansen, J. R. Mesters, C. L. Schmidt, and R. Hilgenfeld**. 2009. The SARS-unique domain (SUD) of SARS coronavirus contains two macrodomains that bind G-quadruplexes. *PLoS Pathog*, **5**:e1000428.
398. **Tang, J., J. Jose, P. Chipman, W. Zhang, R. J. Kuhn, and T. S. Baker**. 2011. Molecular links between the E2 envelope glycoprotein and nucleocapsid core in Sindbis virus. *J Mol Biol*, **414**:442–59.
399. **Tanner, J. A., R. M. Watt, Y.-B. Chai, L.-Y. Lu, M. C. Lin, J. S. M. Peiris, L. L. M. Poon, H.-F. Kung, and J.-D. Huang**. 2003. The severe acute respiratory syndrome (SARS) coronavirus NTPase/helicase belongs to a distinct class of 5' to 3' viral helicases. *J Biol Chem*, **278**:39578–82.
400. **Te Velthuis, A. J. W., S. H. E. van den Worm, and E. J. Snijder**. 2011. The SARS-coronavirus nsp7+nsp8 complex is a unique multimeric RNA polymerase capable of both de novo initiation and primer extension. *Nucleic Acids Res*, Epub 2012 Feb 1.
401. **Tellez, A. B., J. Wang, E. J. Tanner, J. F. Spagnolo, K. Kirkegaard, and E. Bullitt**. 2011. Interstitial Contacts in an RNA-Dependent RNA Polymerase Lattice. *J Mol Biol*, **412**:737–50.

402. **Tellinghuisen, T. L., K. L. Foss, and J. Treadaway**. 2008. Regulation of hepatitis C virion production via phosphorylation of the NS5A protein. *PLoS Pathog*, **4**:e1000032.
403. **Thal, M. A., B. R. Wasik, J. Posto, and R. W. Hardy**. 2007. Template requirements for recognition and copying by Sindbis virus RNA-dependent RNA polymerase. *Virology*, **358**:221–32.
404. **Thompson, A. A., and O. B. Peersen**. 2004. Structural basis for proteolysis-dependent activation of the poliovirus RNA-dependent RNA polymerase. *EMBO J*, **23**:3462–71.
405. **Till, S., and A. G. Ladurner**. 2009. Sensing NAD metabolites through macro domains. *Front Biosci*, **14**:3246–58.
406. **Timinszky, G., S. Till, P. O. Hassa, M. Hothorn, G. Kustatscher, B. Nijmeijer, J. Colombelli, M. Altmeyer, E. H. K. Stelzer, K. Scheffzek, M. O. Hottiger, and A. G. Ladurner**. 2009. A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. *Nat Struct Mol Biol*, **16**:923–9.
407. **Timmer, J. C., W. Zhu, C. Pop, T. Regan, S. J. Snipas, A. M. Eroshkin, S. J. Riedl, and G. S. Salvesen**. 2009. Structural and kinetic determinants of protease substrates. *Nat Struct Mol Biol*, **16**:1101–8.
408. **Tokuriki, N., C. J. Oldfield, V. N. Uversky, I. N. Berezovsky, and D. S. Tawfik**. 2009. Do viral proteins possess unique biophysical features? *Trends Biochem Sci*, **34**:53–9.
409. **Tomar, S., R. W. Hardy, J. L. Smith, and R. J. Kuhn**. 2006. Catalytic core of alphavirus nonstructural protein nsP4 possesses terminal adenylyltransferase activity. *J Virol*, **80**:9962–9.
410. **Tomar, S., M. Narwal, E. Harms, J. L. Smith, and R. J. Kuhn**. 2011. Heterologous production, purification and characterization of enzymatically active Sindbis virus nonstructural protein nsP1. *Protein Expr Purif*, **79**:277–84.
411. **Tong, L.** 2002. Viral proteases. *Chem Rev*, **102**:4609–26.
412. **Toribio, R., and I. Ventoso**. 2010. Inhibition of host translation by virus infection in vivo. *Proc Natl Acad Sci USA*, **107**:9837–42.
413. **Torruella, M., K. Gordon, and T. Hohn**. 1989. Cauliflower mosaic virus produces an aspartic proteinase to cleave its polyproteins. *EMBO J*, **8**:2819–25.
414. **Tuittila, M., and A. E. Hinkkanen**. 2003. Amino acid mutations in the replicase protein nsP3 of Semliki Forest virus cumulatively affect neurovirulence. *J Gen Virol*, **84**:1525–33.
415. **Tuomi, K., L. Kääriäinen, and H. Söderlund**. 1975. Quantitation of Semlike Forest virus RNAs in infected cells using 32-P equilibrium labelling. *Nucleic Acids Res*, **2**:555–65.
416. **Turk, B., V. Turk, and D. Turk**. 1997. Structural and functional aspects of papain-like cysteine proteinases and their protein inhibitors. *Biol Chem*, **378**:141–50.
417. **Turk, D., G. Guncar, M. Podobnik, and B. Turk**. 1998. Revised definition of substrate binding sites of papain-like cysteine proteases. *Biol Chem*, **379**:137–47.
418. **Tyndall, J. D. A., T. Nall, and D. P. Fairlie**. 2005. Proteases universally recognize beta strands in their active sites. *Chem Rev*, **105**:973–99.
419. **Van Damme, P., J. Vandekerckhove, and K. Gevaert**. 2008. Disentanglement of protease substrate repertoires. *Biol Chem*, **389**:371–81.
420. **Varjak, M., E. Zusinaite, and A. Merits**. 2010. Novel functions of the alphavirus nonstructural protein nsP3 C-terminal region. *J Virol*, **84**:2352–64.

421. **Vasiljeva, L., A. Merits, P. Auvinen, and L. Kääriäinen**. 2000. Identification of a novel function of the alphavirus capping apparatus. RNA 5'-triphosphatase activity of Nsp2. *J Biol Chem*, **275**:17281–7.
422. **Vasiljeva, L., L. Valmu, L. Kääriäinen, and A. Merits**. 2001. Site-specific protease activity of the carboxyl-terminal domain of Semliki Forest virus replicase protein nsP2. *J Biol Chem*, **276**:30786–93.
423. **Vasiljeva, L., A. Merits, A. Golubtsov, V. Sizemskaja, L. Kääriäinen, and T. Ahola**. 2003. Regulation of the sequential processing of Semliki Forest virus replicase polyprotein. *J Biol Chem*, **278**:41636–45.
424. **Venkatraman, P., S. Balakrishnan, S. Rao, Y. Hooda, and S. Pol**. 2009. A sequence and structure based method to predict putative substrates, functions and regulatory networks of endo proteases. *PloS One*, **4**:e5700.
425. **Ventoso, I., M. A. Sanz, S. Molina, J. J. Berlanga, L. Carrasco, and M. Esteban**. 2006. Translational resistance of late alphavirus mRNA to eIF2alpha phosphorylation: a strategy to overcome the antiviral effect of protein kinase PKR. *Genes Dev*, **20**:87–100.
426. **Verbinnen, T., H. Van Marck, I. Vandenbroucke, L. Vijgen, M. Claes, T.-I. Lin, K. Simmen, J. Neyts, G. Fanning, and O. Lenz**. 2010. Tracking the evolution of multiple in vitro hepatitis C virus replicon variants under protease inhibitor selection pressure by 454 deep sequencing. *J Virol*, **84**:11124–33.
427. **Verchot, J., E. V. Koonin, and J. C. Carrington**. 1991. The 35-kDa protein from the N-terminus of the potyviral polyprotein functions as a third virus-encoded proteinase. *Virology*, **185**:527–35.
428. **Verspurten, J., K. Gevaert, W. Declercq, and P. Vandenabeele**. 2009. SitePredicting the cleavage of proteinase substrates. *Trends Biochem Sci*, **34**:319–23.
429. **Vihinen, H., T. Ahola, M. Tuittila, A. Merits, and L. Kääriäinen**. 2001. Elimination of phosphorylation sites of Semliki Forest virus replicase protein nsP3. *J Biol Chem*, **276**:5745–52.
430. **Vihinen, H., and J. Saarinen**. 2000. Phosphorylation site analysis of Semliki forest virus nonstructural protein 3. *J Biol Chem*, **275**:27775–83.
431. **Voss, J. E., M.-C. Vaney, S. Duquerroy, C. Vonrhein, C. Girard-Blanc, E. Crublet, A. Thompson, G. Bricogne, and F. A. Rey**. 2010. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. *Nature*, **468**:709–12.
432. **Vrancken, B., S. Lequime, K. Theys, and P. Lemey**. 2010. Covering all bases in HIV research: unveiling a hidden world of viral evolution. *AIDS Rev*, **12**:89–102.
433. **Wahlberg, J. M., and H. Garoff**. 1992. Membrane fusion process of Semliki Forest virus. I: Low pH-induced rearrangement in spike protein quaternary structure precedes virus penetration into cells. *J Cell Biol*, **116**:339–48.
434. **Wang, H. L., J. O'Rear, and V. Stollar**. 1996. Mutagenesis of the Sindbis virus nsP1 protein: effects on methyltransferase activity and viral infectivity. *Virology*, **217**:527–31.
435. **Wang, K. S., R. J. Kuhn, E. G. Strauss, S. Ou, and J. H. Strauss**. 1992. High-affinity laminin receptor is a receptor for Sindbis virus in mammalian cells. *J Virol*, **66**:4992–5001.
436. **Wang, X., and M. Ding**. 1993. Nuclear localization of Sindbis virus nonstructural protein nsP2. *Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences*, **3**:27–37.

437. **Wang, Y. F., S. G. Sawicki, and D. L. Sawicki**. 1994. Alphavirus nsP3 functions to form replication complexes transcribing negative-strand RNA. *J Virol*, **68**:6466–75.
438. **Wang, Y. F., S. G. Sawicki, and D. L. Sawicki**. 1991. Sindbis virus nsP1 functions in negative-strand RNA synthesis. *J Virol*, **65**:985–8.
439. **Weidman, M. K., R. Sharma, S. Raychaudhuri, P. Kundu, W. Tsai, and A. Dasgupta**. 2003. The interaction of cytoplasmic RNA viruses with the nucleus. *Virus Res*, **95**:75–85.
440. **Welsch, S., S. Miller, I. Romero-Brey, A. Merz, C. K. E. Bleck, P. Walther, S. D. Fuller, C. Antony, J. Krijnse-Locker, and R. Bartenschlager**. 2009. Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host Microbe*, **5**:365–75.
441. **Wengler, G., and C. Gros**. 1996. Analyses of the role of structural changes in the regulation of uncoating and assembly of alphavirus cores. *Virology*, **222**:123–32.
442. **Wengler, G., and H. S. Gross**. 1979. Replicative form of Semliki Forest virus RNA contains an unpaired guanosine. *Nature*, **282**:754–6.
443. **Wengler, G.** 2009. The regulation of disassembly of alphavirus cores. *Arch Virol*, **154**:381–90.
444. **Wengler, G., A. Koschinski, G. Wengler, and F. Dreyer**. 2003. Entry of alphaviruses at the plasma membrane converts the viral surface proteins into an ion-permeable pore that can be detected by electrophysiological analyses of whole-cell membrane currents. *J Gen Virol*, **84**:173–81.
445. **White, C. L., M. Thomson, and N. J. Dimmock**. 1998. Deletion analysis of a defective interfering Semliki Forest virus RNA genome defines a region in the nsP2 sequence that is required for efficient packaging of the genome into virus particles. *J Virol*, **72**:4320–6.
446. **Whitty, A.** 2008. Cooperativity and biological complexity. *Nat Chem Biol*, **4**:435–9.
447. **Wielgosz, M. M., R. Raju, and H. V. Huang**. 2001. Sequence requirements for Sindbis virus subgenomic mRNA promoter function in cultured cells. *J Virol*, **75**:3509–19.
448. **Wilkinson, T. A., T. L. Tellinghuisen, R. J. Kuhn, and C. B. Post**. 2005. Association of sindbis virus capsid protein with phospholipid membranes and the E2 glycoprotein: implications for alphavirus assembly. *Biochemistry*, **44**:2800–10.
449. **Williamson, J. R.** 2008. Cooperativity in macromolecular assembly. *Nat Chem Biol*, **4**:458–65.
450. **Wu, J., A. K. Bera, R. J. Kuhn, and J. L. Smith**. 2005. Structure of the Flavivirus helicase: implications for catalytic activity, protein interactions, and proteolytic processing. *J Virol*, **79**:10268–77.
451. **Xue, B., R. W. Williams, C. J. Oldfield, G. K.-M. Goh, A. K. Dunker, and V. N. Uversky**. 2010. Viral disorder or disordered viruses: do viral proteins possess unique features? *Protein Pept Lett*, **17**:932–51.
452. **Xue, W., D. Pan, Y. Yang, H. Liu, and X. Yao**. 2011. Molecular modeling study on the resistance mechanism of HCV NS3/4A serine protease mutants R155K, A156V and D168A to TMC435. *Antiviral Res*, **93**:126–37.
453. **Yang, X.-J.** 2005. Multisite protein modification and intramolecular signaling. *Oncogene*, **24**:1653–62.

454. **Yang, Y., Y. Liang, L. Qu, Z. Chen, M. Yi, K. Li, and S. M. Lemon**. 2007. Disruption of innate immunity due to mitochondrial targeting of a picornaviral protease precursor. *Proc Natl Acad Sci USA*, **104**:7253–8.
455. **Yao, N., P. Reichert, S. S. Taremi, W. W. Prosise, and P. C. Weber**. 1999. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure*, **7**:1353–63.
456. **You, L., D. Garwicz, and T. Rögnvaldsson**. 2005. Comprehensive bioinformatic analysis of the specificity of human immunodeficiency virus type 1 protease. *J Virol*, **79**:12477–86.
457. **Zacks, M. A., and S. Paessler**. 2010. Encephalitic alphaviruses. *Vet Microbiol*, **140**:281–6.
458. **Zhang, D., J. Tözsér, and D. S. Waugh**. 2009. Molecular cloning, overproduction, purification and biochemical characterization of the p39 nsp2 protease domains encoded by three alphaviruses. *Protein Expr Purif*, **64**:89–97.
459. **Zhang, R., C. F. Hryc, Y. Cong, X. Liu, J. Jakana, R. Gorchakov, M. L. Baker, S. C. Weaver, and W. Chiu**. 2011. 4.4 Å cryo-EM structure of an enveloped alphavirus Venezuelan equine encephalitis virus. *EMBO J*, **30**:3854–63.
460. **Ziebuhr, J., E. J. Snijder, and A. E. Gorbalenya**. 2000. Virus-encoded proteinases and proteolytic processing in the Nidovirales. *J Gen Virol*, **81**:853–79.
461. **Zimmerberg, J., and M. M. Kozlov**. 2006. How proteins produce cellular membrane curvature. Nature reviews. *Mol Cell Biol*, **7**:9–19.
462. **Zúñiga, S., I. Sola, J. L. G. Cruz, and L. Enjuanes**. 2009. Role of RNA chaperones in virus replication. *Virus Res*, **139**:253–66.
463. **de Groot, R. J., W. R. Hardy, Y. Shirako, and J. H. Strauss**. 1990. Cleavage-site preferences of Sindbis virus polyproteins containing the non-structural proteinase. Evidence for temporal regulation of polyprotein processing in vivo. *EMBO J*, **9**:2631–8.
464. **de Groot, R. J., T. Rümenapf, R. J. Kuhn, E. G. Strauss, and J. H. Strauss**. 1991. Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc Natl Acad Sci USA*, **88**:8967–71.
465. **den Boon, J. A., and P. Ahlquist**. 2010. Organelle-like membrane compartmentalization of positive-strand RNA virus replication factories. *Annu Rev Microbiol*, **64**:241–56.
466. **den Boon, J. A., A. Diaz, and P. Ahlquist**. 2010. Cytoplasmic viral replication complexes. *Cell Host Microbe*, **8**:77–85.
467. **von Hippel, P. H., and E. Delagoutte**. 2001. A general model for nucleic acid helicases and their "coupling" within macromolecular machines. *Cell*, **104**:177–90.
468. **Özen, A., T. Haliloğlu, and C. A. Schiffer**. 2012. HIV-1 Protease and Substrate Coevolution Validates the Substrate Envelope As the Substrate Recognition Pattern. *J Chem Theory Comput*, **8**:703–714.

ACKNOWLEDGMENTS

I would like to express my heartfelt thanks to my family, parents and friends for their love, patience and support. I'm most sincerely grateful to all the people, who passed on valuable experience and knowledge, who kept me motivated and enthusiastic about solving scientific puzzles. I would especially like to thank Professor Andres Merits for his guidance, support and for the possibility to learn his way of doing scientific research. I'm indebted to Lera for the intellectual and moral support and for the help with figures for this thesis. I'm thankful to Professor Juhan Sedman and Dr. Gemma Atkinson for the valuable remarks and suggestions that certainly improved this dissertation. I want to express my special gratitude to all my co-authors and to all the people that were integral for the completion of this thesis. I would also like to thank all the colleagues that I happened to meet in different labs for the supportive atmosphere and lively discussions, you were very encouraging, creative and knowledgeable and I learned a lot of tricks from all of you.

PUBLICATIONS

CURRICULUM VITAE

Name:	Aleksei Lulla
Date and place of birth:	31.03.1978, Narva, Estonia
Citizenship:	Estonian
Marital status:	married, 2 children
Address, phone, e-mail:	University of Tartu, Institute of Technology Nooruse 1, 50411, Tartu, Estonia +372 55 662688 aleksei.lulla@ut.ee
Language skills:	Russian, English, Estonian

Education and professional employment

1993–1996	Narva Pähklimäe Gymnasium, *silver medal*
1996–2001	University of Tartu, *BSc* in bioorganic chemistry, *cum laude*
2001–2003	University of Tartu, *MSc* in molecular biology
2003–	University of Tartu, doctorate studies in molecular biology
2003–2007	Research scientist, Institute of Molecular and Cell Biology, Tartu, Estonia
2007–	Research scientist, Institute of Technology, Tartu, Estonia

Scientific work

Since 2001 I have been working in the group of Prof. Andres Merits. My main research topic has been the replication of alphaviruses.

List of Publications

1. **Lulla, A.**, V. Lulla, K. Tints, T. Ahola, and A. Merits. 2006. Molecular determinants of substrate specificity for Semliki Forest virus protease. *J Virol,* 80(11):5413–22.
2. Tamberg, N., V. Lulla, R. Fragkoudis, **A. Lulla**, J.K. Fazakerley, and A. Merits. 2007. Insertion of EGFP into the replicase gene of Semliki Forest virus results in a novel, genetically stable marker virus. *J Gen Virol,* 88:1225–1230.
3. Lulla, V., D.L. Sawicki, S.G. Sawicki, **A. Lulla**, A. Merits, and T. Ahola. 2008. Molecular defects caused by temperature-sensitive mutations in Semliki Forest virus nsP1. *J Virol,* 82(18):9236–44.
4. Barry, G., R. Fragkoudis, M.C. Ferguson, **A. Lulla**, A. Merits, A. Kohl, and J.K. Fazakerley. 2010. Semliki Forest virus induced endoplasmic reticulum stress accelerates apoptotic death of mammalian cells. *J Virol,* 84(14):7369–77.
5. Pohjala, L., A. Utt, M. Varjak, **A. Lulla**, A. Merits, T. Ahola, and P. Tammela. 2011. Inhibitors of alphavirus entry and replication identified with a stable Chikungunya replicon cell line and virus-based assays. *PLoS One,* 6(12):e28923.

6. **Lulla, A.**, V. Lulla, and A. Merits. 2012. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. *J Virol,* 86(1):553–65.

Patent

Lulla, A., K. Tints, and S. Tobi. 2007. Optimized recognition site of the alphavirus non-structural protease for tag removal and specific processing of recombinant proteins. U.S. Patent №7189540

ELULOOKIRJELDUS

Nimi:	Aleksei Lulla
Sünniaeg ja koht:	31.03.1978, Narva, Eesti
Kodakondsus:	Eesti
Perekonnaseis:	abielus, 2 last
Aadress, telefon, e-post:	Tartu Ülikooli Tehnoloogiainstituut, Nooruse 1, 50411, Tartu, Eesti +372 55 662688 aleksei.lulla@ut.ee
Keelteoskus:	vene, inglise, eesti

Haridus ja erialane teenistuskäik

1993–1996	Narva Pähklimäe Gümnaasium, *lõpetasin hõbemedaliga*
1996–2001	Tartu Ülikool, BSc bioograanilise keemia erialal, *cum laude*
2001–2003	Tartu Ülikool, MSc molekulaar- ja rakubioloogia erialal
2003–	Tartu Ülikool, Molekulaar- ja Rakubioloogia Instutuut, doktoriõpe
2003–2007	Erakorraline teadur, Tartu Ülikooli Molekulaar- ja Rakubioloogia instituut, Tartu
2007–	Teadur, Tartu Ülikooli Tehnoloogiainstituut, Tartu

Teadustegevus

Alates 2001 aastast olen töötanud Prof. Andres Meritsa grupis. Minu põhiliseks uurimisteemaks on olnud alfaviiruste replikatsiooni uurimine.

Publikatsioonide nimekiri

1. **Lulla, A.**, V. Lulla, K. Tints, T. Ahola, and A. Merits. 2006. Molecular determinants of substrate specificity for Semliki Forest virus protease. *J Virol,* 80(11):5413–22.
2. Tamberg, N., V. Lulla, R. Fragkoudis, **A. Lulla**, J.K. Fazakerley, and A. Merits. 2007. Insertion of EGFP into the replicase gene of Semliki Forest virus results in a novel, genetically stable marker virus. *J Gen Virol,* 88:1225–1230.
3. Lulla, V., D.L. Sawicki, S.G. Sawicki, **A. Lulla**, A. Merits, and T. Ahola. 2008. Molecular defects caused by temperature-sensitive mutations in Semliki Forest virus nsP1. *J Virol,* 82(18):9236–44.
4. Barry, G., R. Fragkoudis, M.C. Ferguson, **A. Lulla**, A. Merits, A. Kohl, and J.K. Fazakerley. 2010. Semliki Forest virus induced endoplasmic reticulum stress accelerates apoptotic death of mammalian cells. *J Virol,* 84(14):7369–77.

5. Pohjala, L., A. Utt, M. Varjak, **A. Lulla**, A. Merits, T. Ahola, and P. Tammela. 2011. Inhibitors of alphavirus entry and replication identified with a stable Chikungunya replicon cell line and virus-based assays. *PLoS One,* 6(12):e28923.
6. **Lulla, A.**, V. Lulla, and A. Merits. 2012. Macromolecular assembly-driven processing of the 2/3 cleavage site in the alphavirus replicase polyprotein. *J Virol,* 86(1):553–65.

Patent

Lulla, A., K. Tints, and S. Tobi. 2007. Optimized recognition site of the alphavirus non-structural protease for tag removal and specific processing of recombinant proteins. U.S. Patent №7189540

DISSERTATIONES BIOLOGICAE UNIVERSITATIS TARTUENSIS

1. **Toivo Maimets**. Studies of human oncoprotein p53. Tartu, 1991, 96 p.
2. **Enn K. Seppet**. Thyroid state control over energy metabolism, ion transport and contractile functions in rat heart. Tartu, 1991, 135 p.
3. **Kristjan Zobel**. Epifüütsete makrosamblike väärtus õhu saastuse indikaatoritena Hamar-Dobani boreaalsetes mägimetsades. Tartu, 1992, 131 lk.
4. **Andres Mäe**. Conjugal mobilization of catabolic plasmids by transposable elements in helper plasmids. Tartu, 1992, 91 p.
5. **Maia Kivisaar**. Studies on phenol degradation genes of *Pseudomonas* sp. strain EST 1001. Tartu, 1992, 61 p.
6. **Allan Nurk**. Nucleotide sequences of phenol degradative genes from *Pseudomonas sp.* strain EST 1001 and their transcriptional activation in *Pseudomonas putida.* Tartu, 1992, 72 p.
7. **Ülo Tamm**. The genus *Populus* L. in Estonia: variation of the species biology and introduction. Tartu, 1993, 91 p.
8. **Jaanus Remme**. Studies on the peptidyltransferase centre of the *E.coli* ribosome. Tartu, 1993, 68 p.
9. **Ülo Langel**. Galanin and galanin antagonists. Tartu, 1993, 97 p.
10. **Arvo Käärd**. The development of an automatic online dynamic fluorescense-based pH-dependent fiber optic penicillin flowthrought biosensor for the control of the benzylpenicillin hydrolysis. Tartu, 1993, 117 p.
11. **Lilian Järvekülg**. Antigenic analysis and development of sensitive immunoassay for potato viruses. Tartu, 1993, 147 p.
12. **Jaak Palumets**. Analysis of phytomass partition in Norway spruce. Tartu, 1993, 47 p.
13. **Arne Sellin**. Variation in hydraulic architecture of *Picea abies* (L.) Karst. trees grown under different enviromental conditions. Tartu, 1994, 119 p.
13. **Mati Reeben**. Regulation of light neurofilament gene expression. Tartu, 1994, 108 p.
14. **Urmas Tartes**. Respiration rhytms in insects. Tartu, 1995, 109 p.
15. **Ülo Puurand.** The complete nucleotide sequence and infections *in vitro* transcripts from cloned cDNA of a potato A potyvirus. Tartu, 1995, 96 p.
16. **Peeter Hõrak**. Pathways of selection in avian reproduction: a functional framework and its application in the population study of the great tit (*Parus major*). Tartu, 1995, 118 p.
17. **Erkki Truve**. Studies on specific and broad spectrum virus resistance in transgenic plants. Tartu, 1996, 158 p.
18. **Illar Pata**. Cloning and characterization of human and mouse ribosomal protein S6-encoding genes. Tartu, 1996, 60 p.
19. **Ülo Niinemets**. Importance of structural features of leaves and canopy in determining species shade-tolerance in temperature deciduous woody taxa. Tartu, 1996, 150 p.

20. **Ants Kurg**. Bovine leukemia virus: molecular studies on the packaging region and DNA diagnostics in cattle. Tartu, 1996, 104 p.
21. **Ene Ustav**. E2 as the modulator of the BPV1 DNA replication. Tartu, 1996, 100 p.
22. **Aksel Soosaar**. Role of helix-loop-helix and nuclear hormone receptor transcription factors in neurogenesis. Tartu, 1996, 109 p.
23. **Maido Remm**. Human papillomavirus type 18: replication, transformation and gene expression. Tartu, 1997, 117 p.
24. **Tiiu Kull**. Population dynamics in *Cypripedium calceolus* L. Tartu, 1997, 124 p.
25. **Kalle Olli**. Evolutionary life-strategies of autotrophic planktonic microorganisms in the Baltic Sea. Tartu, 1997, 180 p.
26. **Meelis Pärtel**. Species diversity and community dynamics in calcareous grassland communities in Western Estonia. Tartu, 1997, 124 p.
27. **Malle Leht**. The Genus *Potentilla* L. in Estonia, Latvia and Lithuania: distribution, morphology and taxonomy. Tartu, 1997, 186 p.
28. **Tanel Tenson**. Ribosomes, peptides and antibiotic resistance. Tartu, 1997, 80 p.
29. **Arvo Tuvikene**. Assessment of inland water pollution using biomarker responses in fish *in vivo* and *in vitro*. Tartu, 1997, 160 p.
30. **Urmas Saarma**. Tuning ribosomal elongation cycle by mutagenesis of 23S rRNA. Tartu, 1997, 134 p.
31. **Henn Ojaveer**. Composition and dynamics of fish stocks in the gulf of Riga ecosystem. Tartu, 1997, 138 p.
32. **Lembi Lõugas**. Post-glacial development of vertebrate fauna in Estonian water bodies. Tartu, 1997, 138 p.
33. **Margus Pooga**. Cell penetrating peptide, transportan, and its predecessors, galanin-based chimeric peptides. Tartu, 1998, 110 p.
34. **Andres Saag**. Evolutionary relationships in some cetrarioid genera (Lichenized Ascomycota). Tartu, 1998, 196 p.
35. **Aivar Liiv**. Ribosomal large subunit assembly *in vivo*. Tartu, 1998, 158 p.
36. **Tatjana Oja**. Isoenzyme diversity and phylogenetic affinities among the eurasian annual bromes (*Bromus* L., Poaceae). Tartu, 1998, 92 p.
37. **Mari Moora**. The influence of arbuscular mycorrhizal (AM) symbiosis on the competition and coexistence of calcareous crassland plant species. Tartu, 1998, 78 p.
38. **Olavi Kurina**. Fungus gnats in Estonia (*Diptera: Bolitophilidae, Keroplatidae, Macroceridae, Ditomyiidae, Diadocidiidae, Mycetophilidae*). Tartu, 1998, 200 p.
39. **Andrus Tasa**. Biological leaching of shales: black shale and oil shale. Tartu, 1998, 98 p.
40. **Arnold Kristjuhan.** Studies on transcriptional activator properties of tumor suppressor protein p53. Tartu, 1998, 86 p.

41. **Sulev Ingerpuu.** Characterization of some human myeloid cell surface and nuclear differentiation antigens. Tartu, 1998, 163 p.
42. **Veljo Kisand.** Responses of planktonic bacteria to the abiotic and biotic factors in the shallow lake Võrtsjärv. Tartu, 1998, 118 p.
43. **Kadri Põldmaa.** Studies in the systematics of hypomyces and allied genera (Hypocreales, Ascomycota). Tartu, 1998, 178 p.
44. **Markus Vetemaa.** Reproduction parameters of fish as indicators in environmental monitoring. Tartu, 1998, 117 p.
45. **Heli Talvik.** Prepatent periods and species composition of different *Oesophagostomum* spp. populations in Estonia and Denmark. Tartu, 1998, 104 p.
46. **Katrin Heinsoo.** Cuticular and stomatal antechamber conductance to water vapour diffusion in *Picea abies* (L.) karst. Tartu, 1999, 133 p.
47. **Tarmo Annilo.** Studies on mammalian ribosomal protein S7. Tartu, 1998, 77 p.
48. **Indrek Ots.** Health state indicies of reproducing great tits (*Parus major*): sources of variation and connections with life-history traits. Tartu, 1999, 117 p.
49. **Juan Jose Cantero.** Plant community diversity and habitat relationships in central Argentina grasslands. Tartu, 1999, 161 p.
50. **Rein Kalamees.** Seed bank, seed rain and community regeneration in Estonian calcareous grasslands. Tartu, 1999, 107 p.
51. **Sulev Kõks.** Cholecystokinin (CCK) — induced anxiety in rats: influence of environmental stimuli and involvement of endopioid mechanisms and erotonin. Tartu, 1999, 123 p.
52. **Ebe Sild.** Impact of increasing concentrations of O_3 and CO_2 on wheat, clover and pasture. Tartu, 1999, 123 p.
53. **Ljudmilla Timofejeva.** Electron microscopical analysis of the synaptonemal complex formation in cereals. Tartu, 1999, 99 p.
54. **Andres Valkna.** Interactions of galanin receptor with ligands and G-proteins: studies with synthetic peptides. Tartu, 1999, 103 p.
55. **Taavi Virro.** Life cycles of planktonic rotifers in lake Peipsi. Tartu, 1999, 101 p.
56. **Ana Rebane.** Mammalian ribosomal protein S3a genes and intron-encoded small nucleolar RNAs U73 and U82. Tartu, 1999, 85 p.
57. **Tiina Tamm.** Cocksfoot mottle virus: the genome organisation and translational strategies. Tartu, 2000, 101 p.
58. **Reet Kurg.** Structure-function relationship of the bovine papilloma virus E2 protein. Tartu, 2000, 89 p.
59. **Toomas Kivisild.** The origins of Southern and Western Eurasian populations: an mtDNA study. Tartu, 2000, 121 p.
60. **Niilo Kaldalu.** Studies of the TOL plasmid transcription factor XylS. Tartu 2000. 88 p.

61. **Dina Lepik.** Modulation of viral DNA replication by tumor suppressor protein p53. Tartu 2000. 106 p.
62. **Kai Vellak.** Influence of different factors on the diversity of the bryophyte vegetation in forest and wooded meadow communities. Tartu 2000. 122 p.
63. **Jonne Kotta.** Impact of eutrophication and biological invasionas on the structure and functions of benthic macrofauna. Tartu 2000. 160 p.
64. **Georg Martin.** Phytobenthic communities of the Gulf of Riga and the inner sea the West-Estonian archipelago. Tartu, 2000. 139 p.
65. **Silvia Sepp.** Morphological and genetical variation of *Alchemilla L.* in Estonia. Tartu, 2000. 124 p.
66. **Jaan Liira.** On the determinants of structure and diversity in herbaceous plant communities. Tartu, 2000. 96 p.
67. **Priit Zingel.** The role of planktonic ciliates in lake ecosystems. Tartu 2001. 111 p.
68. **Tiit Teder.** Direct and indirect effects in Host-parasitoid interactions: ecological and evolutionary consequences. Tartu 2001. 122 p.
69. **Hannes Kollist.** Leaf apoplastic ascorbate as ozone scavenger and its transport across the plasma membrane. Tartu 2001. 80 p.
70. **Reet Marits.** Role of two-component regulator system PehR-PehS and extracellular protease PrtW in virulence of *Erwinia Carotovora* subsp. *Carotovora*. Tartu 2001. 112 p.
71. **Vallo Tilgar.** Effect of calcium supplementation on reproductive performance of the pied flycatcher *Ficedula hypoleuca* and the great tit *Parus major,* breeding in Nothern temperate forests. Tartu, 2002. 126 p.
72. **Rita Hõrak.** Regulation of transposition of transposon Tn*4652* in *Pseudomonas putida*. Tartu, 2002. 108 p.
73. **Liina Eek-Piirsoo.** The effect of fertilization, mowing and additional illumination on the structure of a species-rich grassland community. Tartu, 2002. 74 p.
74. **Krõõt Aasamaa.** Shoot hydraulic conductance and stomatal conductance of six temperate deciduous tree species. Tartu, 2002. 110 p.
75. **Nele Ingerpuu.** Bryophyte diversity and vascular plants. Tartu, 2002. 112 p.
76. **Neeme Tõnisson.** Mutation detection by primer extension on oligonucleotide microarrays. Tartu, 2002. 124 p.
77. **Margus Pensa.** Variation in needle retention of Scots pine in relation to leaf morphology, nitrogen conservation and tree age. Tartu, 2003. 110 p.
78. **Asko Lõhmus.** Habitat preferences and quality for birds of prey: from principles to applications. Tartu, 2003. 168 p.
79. **Viljar Jaks.** p53 — a switch in cellular circuit. Tartu, 2003. 160 p.
80. **Jaana Männik.** Characterization and genetic studies of four ATP-binding cassette (ABC) transporters. Tartu, 2003. 140 p.
81. **Marek Sammul.** Competition and coexistence of clonal plants in relation to productivity. Tartu, 2003. 159 p

82. **Ivar Ilves.** Virus-cell interactions in the replication cycle of bovine papillomavirus type 1. Tartu, 2003. 89 p.
83. **Andres Männik.** Design and characterization of a novel vector system based on the stable replicator of bovine papillomavirus type 1. Tartu, 2003. 109 p.
84. **Ivika Ostonen.** Fine root structure, dynamics and proportion in net primary production of Norway spruce forest ecosystem in relation to site conditions. Tartu, 2003. 158 p.
85. **Gudrun Veldre.** Somatic status of 12–15-year-old Tartu schoolchildren. Tartu, 2003. 199 p.
86. **Ülo Väli.** The greater spotted eagle *Aquila clanga* and the lesser spotted eagle *A. pomarina*: taxonomy, phylogeography and ecology. Tartu, 2004. 159 p.
87. **Aare Abroi.** The determinants for the native activities of the bovine papillomavirus type 1 E2 protein are separable. Tartu, 2004. 135 p.
88. **Tiina Kahre.** Cystic fibrosis in Estonia. Tartu, 2004. 116 p.
89. **Helen Orav-Kotta.** Habitat choice and feeding activity of benthic suspension feeders and mesograzers in the northern Baltic Sea. Tartu, 2004. 117 p.
90. **Maarja Öpik.** Diversity of arbuscular mycorrhizal fungi in the roots of perennial plants and their effect on plant performance. Tartu, 2004. 175 p.
91. **Kadri Tali.** Species structure of *Neotinea ustulata*. Tartu, 2004. 109 p.
92. **Kristiina Tambets.** Towards the understanding of post-glacial spread of human mitochondrial DNA haplogroups in Europe and beyond: a phylogeographic approach. Tartu, 2004. 163 p.
93. **Arvi Jõers.** Regulation of p53-dependent transcription. Tartu, 2004. 103 p.
94. **Lilian Kadaja.** Studies on modulation of the activity of tumor suppressor protein p53. Tartu, 2004. 103 p.
95. **Jaak Truu.** Oil shale industry wastewater: impact on river microbial community and possibilities for bioremediation. Tartu, 2004. 128 p.
96. **Maire Peters.** Natural horizontal transfer of the *pheBA* operon. Tartu, 2004. 105 p.
97. **Ülo Maiväli.** Studies on the structure-function relationship of the bacterial ribosome. Tartu, 2004. 130 p.
98. **Merit Otsus.** Plant community regeneration and species diversity in dry calcareous grasslands. Tartu, 2004. 103 p.
99. **Mikk Heidemaa.** Systematic studies on sawflies of the genera *Dolerus, Empria,* and *Caliroa* (Hymenoptera: Tenthredinidae). Tartu, 2004. 167 p.
100. **Ilmar Tõnno.** The impact of nitrogen and phosphorus concentration and N/P ratio on cyanobacterial dominance and N_2 fixation in some Estonian lakes. Tartu, 2004. 111 p.
101. **Lauri Saks.** Immune function, parasites, and carotenoid-based ornaments in greenfinches. Tartu, 2004. 144 p.
102. **Siiri Rootsi.** Human Y-chromosomal variation in European populations. Tartu, 2004. 142 p.

103. **Eve Vedler.** Structure of the 2,4-dichloro-phenoxyacetic acid-degradative plasmid pEST4011. Tartu, 2005. 106 p.
104. **Andres Tover.** Regulation of transcription of the phenol degradation *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 126 p.
105. **Helen Udras.** Hexose kinases and glucose transport in the yeast *Hansenula polymorpha*. Tartu, 2005. 100 p.
106. **Ave Suija.** Lichens and lichenicolous fungi in Estonia: diversity, distribution patterns, taxonomy. Tartu, 2005. 162 p.
107. **Piret Lõhmus.** Forest lichens and their substrata in Estonia. Tartu, 2005. 162 p.
108. **Inga Lips.** Abiotic factors controlling the cyanobacterial bloom occurrence in the Gulf of Finland. Tartu, 2005. 156 p.
109. **Kaasik, Krista.** Circadian clock genes in mammalian clockwork, metabolism and behaviour. Tartu, 2005. 121 p.
110. **Juhan Javoiš.** The effects of experience on host acceptance in ovipositing moths. Tartu, 2005. 112 p.
111. **Tiina Sedman.** Characterization of the yeast *Saccharomyces cerevisiae* mitochondrial DNA helicase Hmi1. Tartu, 2005. 103 p.
112. **Ruth Aguraiuja.** Hawaiian endemic fern lineage *Diellia* (Aspleniaceae): distribution, population structure and ecology. Tartu, 2005. 112 p.
113. **Riho Teras.** Regulation of transcription from the fusion promoters generated by transposition of Tn*4652* into the upstream region of *pheBA* operon in *Pseudomonas putida*. Tartu, 2005. 106 p.
114. **Mait Metspalu.** Through the course of prehistory in india: tracing the mtDNA trail. Tartu, 2005. 138 p.
115. **Elin Lõhmussaar.** The comparative patterns of linkage disequilibrium in European populations and its implication for genetic association studies. Tartu, 2006. 124 p.
116. **Priit Kupper.** Hydraulic and environmental limitations to leaf water relations in trees with respect to canopy position. Tartu, 2006. 126 p.
117. **Heili Ilves.** Stress-induced transposition of Tn*4652* in *Pseudomonas Putida.* Tartu, 2006. 120 p.
118. **Silja Kuusk.** Biochemical properties of Hmi1p, a DNA helicase from *Saccharomyces cerevisiae* mitochondria. Tartu, 2006. 126 p.
119. **Kersti Püssa.** Forest edges on medium resolution landsat thematic mapper satellite images. Tartu, 2006. 90 p.
120. **Lea Tummeleht.** Physiological condition and immune function in great tits (*Parus major* l.): Sources of variation and trade-offs in relation to growth. Tartu, 2006. 94 p.
121. **Toomas Esperk.** Larval instar as a key element of insect growth schedules. Tartu, 2006. 186 p.
122. **Harri Valdmann.** Lynx (*Lynx lynx*) and wolf (*Canis lupus*) in the Baltic region: Diets, helminth parasites and genetic variation. Tartu, 2006. 102 p.

123. **Priit Jõers.** Studies of the mitochondrial helicase Hmi1p in *Candida albicans* and *Saccharomyces cerevisia*. Tartu, 2006. 113 p.
124. **Kersti Lilleväli.** Gata3 and Gata2 in inner ear development. Tartu, 2007. 123 p.
125. **Kai Rünk.** Comparative ecology of three fern species: *Dryopteris carthusiana* (Vill.) H.P. Fuchs, *D. expansa* (C. Presl) Fraser-Jenkins & Jermy and *D. dilatata* (Hoffm.) A. Gray (Dryopteridaceae). Tartu, 2007. 143 p.
126. **Aveliina Helm.** Formation and persistence of dry grassland diversity: role of human history and landscape structure. Tartu, 2007. 89 p.
127. **Leho Tedersoo.** Ectomycorrhizal fungi: diversity and community structure in Estonia, Seychelles and Australia. Tartu, 2007. 233 p.
128. **Marko Mägi.** The habitat-related variation of reproductive performance of great tits in a deciduous-coniferous forest mosaic: looking for causes and consequences. Tartu, 2007. 135 p.
129. **Valeria Lulla.** Replication strategies and applications of Semliki Forest virus. Tartu, 2007. 109 p.
130. **Ülle Reier**. Estonian threatened vascular plant species: causes of rarity and conservation. Tartu, 2007. 79 p.
131. **Inga Jüriado**. Diversity of lichen species in Estonia: influence of regional and local factors. Tartu, 2007. 171 p.
132. **Tatjana Krama.** Mobbing behaviour in birds: costs and reciprocity based cooperation. Tartu, 2007. 112 p.
133. **Signe Saumaa.** The role of DNA mismatch repair and oxidative DNA damage defense systems in avoidance of stationary phase mutations in *Pseudomonas putida.* Tartu, 2007. 172 p.
134. **Reedik Mägi**. The linkage disequilibrium and the selection of genetic markers for association studies in european populations. Tartu, 2007. 96 p.
135. **Priit Kilgas.** Blood parameters as indicators of physiological condition and skeletal development in great tits (*Parus major*): natural variation and application in the reproductive ecology of birds. Tartu, 2007. 129 p.
136. **Anu Albert**. The role of water salinity in structuring eastern Baltic coastal fish communities. Tartu, 2007. 95 p.
137. **Kärt Padari.** Protein transduction mechanisms of transportans. Tartu, 2008. 128 p.
138. **Siiri-Lii Sandre.** Selective forces on larval colouration in a moth. Tartu, 2008. 125 p.
139. **Ülle Jõgar.** Conservation and restoration of semi-natural floodplain meadows and their rare plant species. Tartu, 2008. 99 p.
140. **Lauri Laanisto.** Macroecological approach in vegetation science: generality of ecological relationships at the global scale. Tartu, 2008. 133 p.
141. **Reidar Andreson**. Methods and software for predicting PCR failure rate in large genomes. Tartu, 2008. 105 p.
142. **Birgot Paavel.** Bio-optical properties of turbid lakes. Tartu, 2008. 175 p.

143. **Kaire Torn.** Distribution and ecology of charophytes in the Baltic Sea. Tartu, 2008, 98 p.
144. **Vladimir Vimberg.** Peptide mediated macrolide resistance. Tartu, 2008, 190 p.
145. **Daima Örd.** Studies on the stress-inducible pseudokinase TRB3, a novel inhibitor of transcription factor ATF4. Tartu, 2008, 108 p.
146. **Lauri Saag.** Taxonomic and ecologic problems in the genus *Lepraria* (*Stereocaulaceae*, lichenised *Ascomycota*). Tartu, 2008, 175 p.
147. **Ulvi Karu.** Antioxidant protection, carotenoids and coccidians in greenfinches – assessment of the costs of immune activation and mechanisms of parasite resistance in a passerine with carotenoid-based ornaments. Tartu, 2008, 124 p.
148. **Jaanus Remm.** Tree-cavities in forests: density, characteristics and occupancy by animals. Tartu, 2008, 128 p.
149. **Epp Moks.** Tapeworm parasites *Echinococcus multilocularis* and *E. granulosus* in Estonia: phylogenetic relationships and occurrence in wild carnivores and ungulates. Tartu, 2008, 82 p.
150. **Eve Eensalu.** Acclimation of stomatal structure and function in tree canopy: effect of light and CO_2 concentration. Tartu, 2008, 108 p.
151. **Janne Pullat**. Design, functionlization and application of an *in situ* synthesized oligonucleotide microarray. Tartu, 2008, 108 p.
152. **Marta Putrinš.** Responses of *Pseudomonas putida* to phenol-induced metabolic and stress signals. Tartu, 2008, 142 p.
153. **Marina Semtšenko.** Plant root behaviour: responses to neighbours and physical obstructions. Tartu, 2008, 106 p.
154. **Marge Starast.** Influence of cultivation techniques on productivity and fruit quality of some *Vaccinium* and *Rubus* taxa. Tartu, 2008, 154 p.
155. **Age Tats.** Sequence motifs influencing the efficiency of translation. Tartu, 2009, 104 p.
156. **Radi Tegova.** The role of specialized DNA polymerases in mutagenesis in *Pseudomonas putida.* Tartu, 2009, 124 p.
157. **Tsipe Aavik.** Plant species richness, composition and functional trait pattern in agricultural landscapes – the role of land use intensity and landscape structure. Tartu, 2008, 112 p.
158. **Kaja Kiiver.** Semliki forest virus based vectors and cell lines for studying the replication and interactions of alphaviruses and hepaciviruses. Tartu, 2009, 104 p.
159. **Meelis Kadaja.** Papillomavirus Replication Machinery Induces Genomic Instability in its Host Cell. Tartu, 2009, 126 p.
160. **Pille Hallast.** Human and chimpanzee Luteinizing hormone/Chorionic Gonadotropin beta (*LHB/CGB*) gene clusters: diversity and divergence of young duplicated genes. Tartu, 2009, 168 p.
161. **Ain Vellak.** Spatial and temporal aspects of plant species conservation. Tartu, 2009, 86 p.

162. **Triinu Remmel.** Body size evolution in insects with different colouration strategies: the role of predation risk. Tartu, 2009, 168 p.
163. **Jaana Salujõe.** Zooplankton as the indicator of ecological quality and fish predation in lake ecosystems. Tartu, 2009, 129 p.
164. **Ele Vahtmäe.** Mapping benthic habitat with remote sensing in optically complex coastal environments. Tartu, 2009, 109 p.
165. **Liisa Metsamaa.** Model-based assessment to improve the use of remote sensing in recognition and quantitative mapping of cyanobacteria. Tartu, 2009, 114 p.
166. **Pille Säälik.** The role of endocytosis in the protein transduction by cell-penetrating peptides. Tartu, 2009, 155 p.
167. **Lauri Peil.** Ribosome assembly factors in *Escherichia coli.* Tartu, 2009, 147 p.
168. **Lea Hallik.** Generality and specificity in light harvesting, carbon gain capacity and shade tolerance among plant functional groups. Tartu, 2009, 99 p.
169. **Mariliis Tark.** Mutagenic potential of DNA damage repair and tolerance mechanisms under starvation stress. Tartu, 2009, 191 p.
170. **Riinu Rannap.** Impacts of habitat loss and restoration on amphibian populations. Tartu, 2009, 117 p.
171. **Maarja Adojaan.** Molecular variation of HIV-1 and the use of this knowledge in vaccine development. Tartu, 2009, 95 p.
172. **Signe Altmäe.** Genomics and transcriptomics of human induced ovarian folliculogenesis. Tartu, 2010, 179 p.
173. **Triin Suvi.** Mycorrhizal fungi of native and introduced trees in the Seychelles Islands. Tartu, 2010, 107 p.
174. **Velda Lauringson.** Role of suspension feeding in a brackish-water coastal sea. Tartu, 2010, 123 p.
175. **Eero Talts.** Photosynthetic cyclic electron transport – measurement and variably proton-coupled mechanism. Tartu, 2010, 121 p.
176. **Mari Nelis.** Genetic structure of the Estonian population and genetic distance from other populations of European descent. Tartu, 2010, 97 p.
177. **Kaarel Krjutškov.** Arrayed Primer Extension-2 as a multiplex PCR-based method for nucleic acid variation analysis: method and applications. Tartu, 2010, 129 p.
178. **Egle Köster.** Morphological and genetical variation within species complexes: *Anthyllis vulneraria* s. l. and *Alchemilla vulgaris* (coll.). Tartu, 2010, 101 p.
179. **Erki Õunap.** Systematic studies on the subfamily Sterrhinae (Lepidoptera: Geometridae). Tartu, 2010, 111 p.
180. **Merike Jõesaar.** Diversity of key catabolic genes at degradation of phenol and *p*-cresol in pseudomonads. Tartu, 2010, 125 p.
181. **Kristjan Herkül.** Effects of physical disturbance and habitat-modifying species on sediment properties and benthic communities in the northern Baltic Sea. Tartu, 2010, 123 p.

182. **Arto Pulk.** Studies on bacterial ribosomes by chemical modification approaches. Tartu, 2010, 161 p.
183. **Maria Põllupüü.** Ecological relations of cladocerans in a brackish-water ecosystem. Tartu, 2010, 126 p.
184. **Toomas Silla.** Study of the segregation mechanism of the Bovine Papillomavirus Type 1. Tartu, 2010, 188 p.
185. **Gyaneshwer Chaubey.** The demographic history of India: A perspective based on genetic evidence. Tartu, 2010, 184 p.
186. **Katrin Kepp.** Genes involved in cardiovascular traits: detection of genetic variation in Estonian and Czech populations. Tartu, 2010, 164 p.
187. **Virve Sõber.** The role of biotic interactions in plant reproductive performance. Tartu, 2010, 92 p.
188. **Kersti Kangro.** The response of phytoplankton community to the changes in nutrient loading. Tartu, 2010, 144 p.
189. **Joachim M. Gerhold.** Replication and Recombination of mitochondrial DNA in Yeast. Tartu, 2010, 120 p.
190. **Helen Tammert.** Ecological role of physiological and phylogenetic diversity in aquatic bacterial communities. Tartu, 2010, 140 p.
191. **Elle Rajandu.** Factors determining plant and lichen species diversity and composition in Estonian *Calamagrostis* and *Hepatica* site type forests. Tartu, 2010, 123 p.
192. **Paula Ann Kivistik.** ColR-ColS signalling system and transposition of Tn*4652* in the adaptation of *Pseudomonas putida.* Tartu, 2010, 118 p.
193. **Siim Sõber.** Blood pressure genetics: from candidate genes to genome-wide association studies. Tartu, 2011, 120 p.
194. **Kalle Kipper.** Studies on the role of helix 69 of 23S rRNA in the factor-dependent stages of translation initiation, elongation, and termination. Tartu, 2011, 178 p.
195. **Triinu Siibak.** Effect of antibiotics on ribosome assembly is indirect. Tartu, 2011, 134 p.
196. **Tambet Tõnissoo.** Identification and molecular analysis of the role of guanine nucleotide exchange factor RIC-8 in mouse development and neural function. Tartu, 2011, 110 p.
197. **Helin Räägel.** Multiple faces of cell-penetrating peptides – their intracellular trafficking, stability and endosomal escape during protein transduction. Tartu, 2011, 161 p.
198. **Andres Jaanus.** Phytoplankton in Estonian coastal waters – variability, trends and response to environmental pressures. Tartu, 2011, 157 p.
199. **Tiit Nikopensius.** Genetic predisposition to nonsyndromic orofacial clefts. Tartu, 2011, 152 p.
200. **Signe Värv.** Studies on the mechanisms of RNA polymerase II-dependent transcription elongation. Tartu, 2011, 108 p.
201. **Kristjan Välk.** Gene expression profiling and genome-wide association studies of non-small cell lung cancer. Tartu, 2011, 98 p.

202. **Arno Põllumäe.** Spatio-temporal patterns of native and invasive zooplankton species under changing climate and eutrophication conditions. Tartu, 2011, 153 p.
203. **Egle Tammeleht.** Brown bear (*Ursus arctos*) population structure, demographic processes and variations in diet in northern Eurasia. Tartu, 2011, 143 p.
205. **Teele Jairus.** Species composition and host preference among ectomycorrhizal fungi in Australian and African ecosystems. Tartu, 2011, 106 p.
206. **Kessy Abarenkov.** PlutoF – cloud database and computing services supporting biological research. Tartu, 2011, 125 p.
207. **Marina Grigorova.** Fine-scale genetic variation of follicle-stimulating hormone beta-subunit coding gene (*FSHB*) and its association with reproductive health. Tartu, 2011, 184 p.
208. **Anu Tiitsaar.** The effects of predation risk and habitat history on butterfly communities. Tartu, 2011, 97 p.
209. **Elin Sild.** Oxidative defences in immunoecological context: validation and application of assays for nitric oxide production and oxidative burst in a wild passerine. Tartu, 2011, 105 p.
210. **Irja Saar**. The taxonomy and phylogeny of the genera *Cystoderma* and *Cystodermella* (Agaricales, Fungi). Tartu, 2012, 167 p.
211. **Pauli Saag.** Natural variation in plumage bacterial assemblages in two wild breeding passerines. Tartu, 2012, 113 p.