Eva Böttcher-Friebertshäuser
Wolfgang Garten · Hans Dieter Klenk
Editors

Activation of Viruses by Host Proteases

Springer

Activation of Viruses by Host Proteases

Eva Böttcher-Friebertshäuser
Wolfgang Garten
Hans Dieter Klenk
Editors

Activation of Viruses by Host Proteases

Springer

Editors
Eva Böttcher-Friebertshäuser
Institut für Virologie
Philipps-Universität
Marburg
Germany

Wolfgang Garten
Institut für Virologie
Philipps-Universität
Marburg
Germany

Hans Dieter Klenk
Institut für Virologie
Philipps-Universität
Marburg
Germany

ISBN 978-3-319-75473-4 ISBN 978-3-319-75474-1 (eBook)
https://doi.org/10.1007/978-3-319-75474-1

Library of Congress Control Number: 2018941240

© Springer International Publishing AG, part of Springer Nature 2018
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Printed on acid-free paper

This Springer imprint is published by the registered company Springer International Publishing AG part of Springer Nature
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

To Purnell W. Choppin, the founder of the field

Preface

More than four decades ago the groups of Purnell Choppin in the United States and of Morio Homma in Japan identified a glycoprotein in paramyxoviruses that induces membrane fusion and initiates infection, and they discovered that its activity depends on posttranslational cleavage by host proteases. At about the same time similar observations were made with the influenza virus hemagglutinin, and proteolytic glycoprotein activation proved to be an important determinant for organ tropism and spread of infection of paramyxoviruses as well as influenza viruses. Thus, proteases were among the first host factors recognized to play a prominent regulatory role in virus infection. Moreover, these studies provided for the first time insight into the molecular mechanisms underlying viral pathogenicity. These early discoveries have been described in detail in previous reviews (Webster RG and Rott R, Cell 50, 665-666, 1987; Nagai Y, Trends Microbiol 1, 81-87, 1993; Klenk HD and Garten G, in: Cellular Receptors for Animal Viruses, Wimmer E, ed, pp241-280, Cold Spring Harbor Press, 1994). Today we know that proteolytic activation of envelope proteins is a characteristic feature of many important viral pathogens including not only influenza virus but also respiratory syncytial virus, Nipah and Hendra viruses, Lassa virus, Marburg and Ebola viruses, SARS and MERS coronaviruses, yellow fever virus, Dengue virus, Zika virus, and tick-borne encephalitis virus, just to name a few. A large number of new activating proteases have also been identified that differ in substrate specificity and expression patterns in tissues and cell compartments, and the concept has been strengthened that these variations have a high impact on the outcome of infection. Of particular interest are activation mutants of viruses that have been used for vaccine design. These and other recent developments will be discussed in this book. Activation by host proteases observed with non-enveloped viruses, such as rotaviruses, will not be addressed.

The first chapters of the book will focus on envelope proteins undergoing proteolytic activation. Most of them induce membrane fusion involving a conformational change that is primed by cleavage of the fusion protein itself or by cleavage of an accessory protein and then triggered by exposure to low pH or interaction with a receptor-binding protein. Based on structural differences, three classes of fusion proteins can be discriminated. Class I fusion proteins will be addressed in contributions by Summer Galloway, Bo Liang, and David Steinhauer on the influenza virus hemagglutinin, by Everett Smith and Rebecca Dutch on the F protein of paramyxoviruses and pneumoviruses, by Antonella Pasquato, Laura Cendron, and Stefan

Kunz on the glycoprotein of arenaviruses, and by Markus Hoffmann, Heike Hofmann-Winkler, and Stefan Pöhlmann on the S protein of SARS and MERS coronaviruses. Activation of the coronavirus S protein is particularly interesting, because it depends on the action of different proteases in concert. As discussed in Chap. 5 the GP protein of filoviruses is also a class I fusion protein undergoing sequential cleavage by two proteases that removes a large carbohydrate-rich segment of GP, thereby exposing the receptor binding region and the fusion loop. The chapter by Franz Heinz and Karin Stiasny illustrates that activation of the class II fusion protein of a flavivirus depends on cleavage of the tightly associated accessory prM protein. Processing by host proteases has also been observed with a few viral proteins that do not induce membrane fusion. Thus, in the chapter on paramyxoviruses the authors will report that, with some strains of Newcastle disease virus, the HN glycoprotein mediating binding to and release from receptors is also activated by proteolytic cleavage.

The section on viral proteins will be followed by reviews of the activating proteases. Obviously, endoproteolytic cleavage is not only necessary for the activation of viruses, but it is also an essential step in the maturation of many cellular proteins with important biological functions, such as peptide hormones, neuropeptides, growth and coagulation factors, cell surface receptors, adhesion molecules, and transcription factors involved in lipid metabolism. Of the large number of vertebrate endoproteases, so far only a fraction has been found to activate viruses. They fall into four major groups, each of which will be reviewed in a separate chapter. The first enzymes identified were soluble trypsin-like serine proteases. As described in the contribution of Hiroshi Kido, these enzymes have several functions in influenza virus pathogenesis of which hemagglutinin activation is only one. In recent years a still increasing number of membrane-anchored serine proteases (MASPs) have been discovered that form the second group (Chap. 8). Proprotein convertases that have the widest spectrum of viral glycoprotein substrates form the third group (Chap. 9). The fourth group of host proteases processing envelope proteins are cathepsins as described in the contribution of Klaudia Brix.

The book will close with a comprehensive overview by Torsten Steinmetzer and Kornelia Hardes on protease inhibitors. Elucidation of the structural details of an increasing number of proteases provides a solid basis for rational inhibitor design. Since host proteases are stable targets, their inhibition should prevent the rapid development of resistance that is observed when viral proteins are addressed. There is experimental evidence for the antiviral potential of such compounds, and there is hope that they may find their way to therapeutic application.

Marburg, Germany Eva Böttcher-Friebertshäuser
Marburg, Germany Wolfgang Garten
Marburg, Germany Hans Dieter Klenk

Contents

Part I Viral Proteins Activated By Host Proteases

Part II Activating Proteases

Part III Protease Inhibitors

Part I

Viral Proteins Activated By Host Proteases

1 Activation of the Hemagglutinin of Influenza Viruses

Summer E. Galloway, Bo Liang, and David A. Steinhauer

Abstract

The hemagglutinin (HA) glycoprotein of influenza viruses is posttranslationally cleaved into the disulfide-linked subunits HA_1 and HA_2, and this proteolytic processing event is critical to the virus life cycle as it is required to activate membrane fusion potential and virus infectivity. High-resolution structures are available for the HA precursor (HA0), the cleaved neutral pH conformation of HA, and the low pH conformation that the HA assumes when triggered by acidification of endosomes to mediate fusion of viral and cellular membrane during virus entry. These structures have provided clues regarding the mechanisms by which proteolytic cleavage activates membrane fusion potential and how subsequent acidification drives the fusion process. It has been known for decades that influenza strains and subtypes can vary with regard to HA cleavage properties and that cleavage site sequences and the proteases that recognize them can represent a major determinant for virus pathogenicity. However, a number of questions remain with respect to the identity and characteristics of the proteases that activate HAs in the various natural hosts and complex ecosystems that constitute the realm of influenza viruses. The continuing study of HA cleavage properties

---

S. E. Galloway
Influenza Division, National Center for Immunization and Respiratory Diseases,
Centers for Disease Control and Prevention, Atlanta, GA, USA
e-mail: nmd3@cdc.gov

B. Liang
Department of Biochemistry, O. Wayne Rollins Research Center,
Emory University School of Medicine, Atlanta, GA, USA
e-mail: bo.liang@emory.edu

D. A. Steinhauer (✉)
Department of Microbiology and Immunology, O. Wayne Rollins Research Center,
Emory University School of Medicine, Atlanta, GA, USA
e-mail: dsteinh@emory.edu

E. Böttcher-Friebertshäuser et al. (eds.), *Activation of Viruses by Host Proteases*,
https://doi.org/10.1007/978-3-319-75474-1_1

and the proteases involved should illuminate our understanding not only of pathogenicity but other aspects of influenza biology including host range, transmission, and interplay with other microorganisms such as bacteria.

There are three genera of influenza viruses in the *Orthomyxoviridae* family: the influenza A, influenza B, and influenza C. All are human pathogens; however, influenza A viruses are generally of greatest concern as they are responsible for most of the annual seasonal epidemics, and their vast natural reservoir in aquatic avian species provides a fertile source for the unpredictable emergence of novel pandemic strains. For the influenza A and B viruses, the initial stage of entry into host cells is mediated by receptor-binding and membrane fusion functions provided by the hemagglutinin (HA) glycoprotein. A separate viral surface protein, the neuraminidase (NA), is responsible for the receptor-destroying sialidase activity that allows for progeny virus dissemination at the end of the replication cycle. For influenza C viruses, both the receptor-binding function and receptor-destroying esterase functions are provided by a single glycoprotein, termed HEF, for hemagglutinin/esterase/fusion. The HA and HEF proteins of all influenza viruses share a common requirement for protease activation in order to potentiate membrane fusion and activate virus infectivity. This chapter will focus on the proteolytic activation of the HA glycoproteins of influenza A viruses, but many of the concepts outlined herein are equally relevant for the HA and HEF proteins of influenza B and C viruses.

The influenza A viruses are distinguished by their complex ecology and dynamic transmission cycles involving more than 100 avian species and a range of mammalian hosts (Munster et al. 2007; Slusher et al. 2014; Webster et al. 1992). In aquatic birds belonging to the order Anseriformes (ducks, geese, and swans) and Charadriiformes (gulls, terns, and shorebirds) that serve as the “natural” hosts of influenza A virus, 16 antigenically distinct HA subtypes and 9 NA subtypes are known to circulate, though they rarely cause symptomatic disease in these species. Under conducive conditions, viruses maintained within the natural reservoir may transmit to Galliformes (chickens, turkeys, and quail) and passerine birds or mammalian species such as swine, horses, dogs, and humans, causing a range of morbidity and mortality. The outbreaks resulting from cross-species transmission are often limited, but when influenza A virus strains become established in a new host, they can perpetuate for varying lengths of time and have devastating consequences to human health and global economies. Cross-species transmission events of influenza A virus are often characterized by the relatively rapid evolution of the virus to facilitate adaptation to the new host. These adaptive changes may be influenced by a variety of selective pressures, including, but not limited to, the immune response of the host; the availability and structure of host cell receptors; variations in host cell replication machinery; differences in the site of replication, which may involve pH and temperature; and environmental persistence to enable continued transmission. While these concepts are generally understood, we are only at the inception of understanding the complex interplay of viral and host factors involved in host range, transmission, and pathogenicity of these viruses. What we do know is that HA cleavage properties constitute one of the most critical determinants of pathogenicity and the environments and biological niches where activating proteases can be found in nature may play an important role in the complex ecology of influenza A viruses.

Proteolytic cleavage is a fundamental biochemical process that activates a multitude of functions for viruses as well as their hosts. A requirement for proteolytic activation of influenza virus infectivity was first discovered in studies examining influenza virus replication in cell culture. For many years after influenza viruses were first isolated in the 1930s (Shope 1931; Smith et al. 1933), embryonated chicken eggs were the substrate of choice for influenza virus propagation, and even as cell culture systems were developed, only a few strains such as A/WSN/33 (H1N1) were able to replicate and form plaques on cell monolayers. Decades passed before it was reported that virus replication and plaquing efficiency could be enhanced appreciably by the addition of proteases such as trypsin or pancreatin to the cell monolayers (Appleyard and Maber 1974; Came et al. 1968; Tobita and Kilbourne 1974) and that the extent of HA cleavage correlated with cytopathology following infection (Lazarowitz et al. 1973a). These studies were extended to show that HA cleavage activation was required for an early stage in the infection process, but was not required for virion assembly, hemagglutination, or virus adsorption functions (Klenk et al. 1975; Lazarowitz and Choppin 1975). Polyacrylamide gel electrophoresis (PAGE) analyses revealed that cleavage activation resulted in the digestion of the HA precursor protein (HA0) into faster migrating polypeptides known as the HA_1 and HA_2 subunits. N-terminal sequencing of the HA_2 polypeptides derived from various infectious viruses revealed a highly conserved sequence, GLFGAIAGFIE (Skehel and Waterfield 1975), which is the N-terminal portion of the domain now universally referred to as the fusion peptide. Subsequent studies clearly demonstrated that the critical function activated by proteolytic cleavage is HA-mediated membrane fusion of the viral and endosomal membranes, facilitating transfer of the viral genome into host cells during virus entry for replication (Huang et al. 1980, 1981; Maeda and Ohnishi 1980; White and Helenius 1980).

1.1 Structural Basis for Activation of HA Fusion Potential

From a structural perspective, the mechanisms by which influenza HA becomes fusogenic upon protease cleavage and the subsequent acid-induced conformational rearrangements that drive the membrane fusion process in endosomes are quite well developed. For the HA of H3N2 subtype viruses that have circulated in humans since 1968, there is high-resolution structural information for the uncleaved HA0 precursor, the cleaved pre-fusion neutral pH HA present on the surface of infectious virions, and the low pH conformation adopted by HA following the acid-induced structural rearrangements required for fusion (Bizebard et al. 1995; Bullough et al. 1994; Chen et al. 1998; Wilson et al. 1981). Therefore, we base most of our structural discussions on the HA of the 1968 human virus A/Aichi/2/68 (H3N2).

For discussion purposes, we generally assume that infection of a new host is initiated by virions that contain cleaved HAs on their surfaces and are therefore fully activated to facilitate membrane fusion; however, infectious virions can also contain mixed ratios of cleaved and uncleaved HAs on their surfaces, and there may be examples for which cleavage can also occur in endosomes during the early stages of entry. In any case, influenza viruses initiate infection by attaching to sialic acid-containing glycan receptors on host cell surfaces, followed by entry via the endocytic pathway. The trigger for fusion of the viral and endosomal membranes is the acidification of the vesicular compartments by cellular proton pumps. When the endosomal pH reaches a critical threshold, usually between pH 6.0 and 5.0 depending on the viral strain, a number of HA structural rearrangements are triggered that coordinately function to drive the fusion process (Bizebard et al. 1995; Bullough et al. 1994; Chen et al. 1999). The pre- and post-fusion structures of HA are shown in Fig. 1.1, and the conformational changes that take place include detrimerization of the membrane-distal head domains, extrusion of the HA_2 N-terminal fusion peptide domains from the trimer interior, extension of the long HA_2 coiled coil by recruitment of the short HA_2 α-helix and connecting polypeptide to the N-terminal end of the long α-helix of each monomer, and 180° reorientation of the C-terminal domain of the long α-helices to "jackknife," this domain against the central coiled coil. The extension of the central coiled coil directs the fusion peptides to the end of the low pH structure (the top, as shown in Fig. 1.1, right panel), and residues C-terminal to the "jackknifed" helices trace along a groove on the outside of the coiled coil, as can be viewed for the polypeptide chain of one monomer to the left of the coiled coil in Fig. 1.1, and this places the HA_2 C-terminal transmembrane domains at the same end of the rod-shaped structure as the fusion peptides (Fig. 1.1, right panel). Based on these structural changes, a mechanism for initiating the membrane fusion process can be envisaged, in which hydrophobic fusion peptides are "harpooned" into the endosomal membrane to link the viral and cellular membranes via the HA_2 subunit for influenza A virus, and the 180° "jackknife" of the long helices bring the viral and endosomal membranes into proximity as a prerequisite for the fusion process.

Collectively, the acid-induced HA conformational changes that initiate membrane fusion are irreversible, and the observation that the rodlike low pH HA structure is considerably more stable than cleaved neutral pH HA (Carr et al. 1997; Chen et al. 1995, 1999; Ruigrok et al. 1988) indicates that the cleaved neutral pH HA is a metastable molecule that transitions to a more energetically stable form during the fusion process. In contrast, the uncleaved HA is relatively unresponsive to acidification with respect to structural rearrangements, suggesting that proteolytic cleavage primes the HA for fusion by allowing it to adopt the metastable conformation that can be triggered by acidification to induce membrane fusion. A comparison of the structures of the HA0 and cleaved neutral pH HA reveals that only 19 residues relocate upon protease cleavage (Chen et al. 1998; Wilson et al. 1981); however, two critical prerequisites for fusion transpire upon cleavage. First, protease cleavage liberates the hydrophobic fusion peptide from its internal position within the HA0 polypeptide chain, becoming the N-terminal domain of the newly generated HA_2

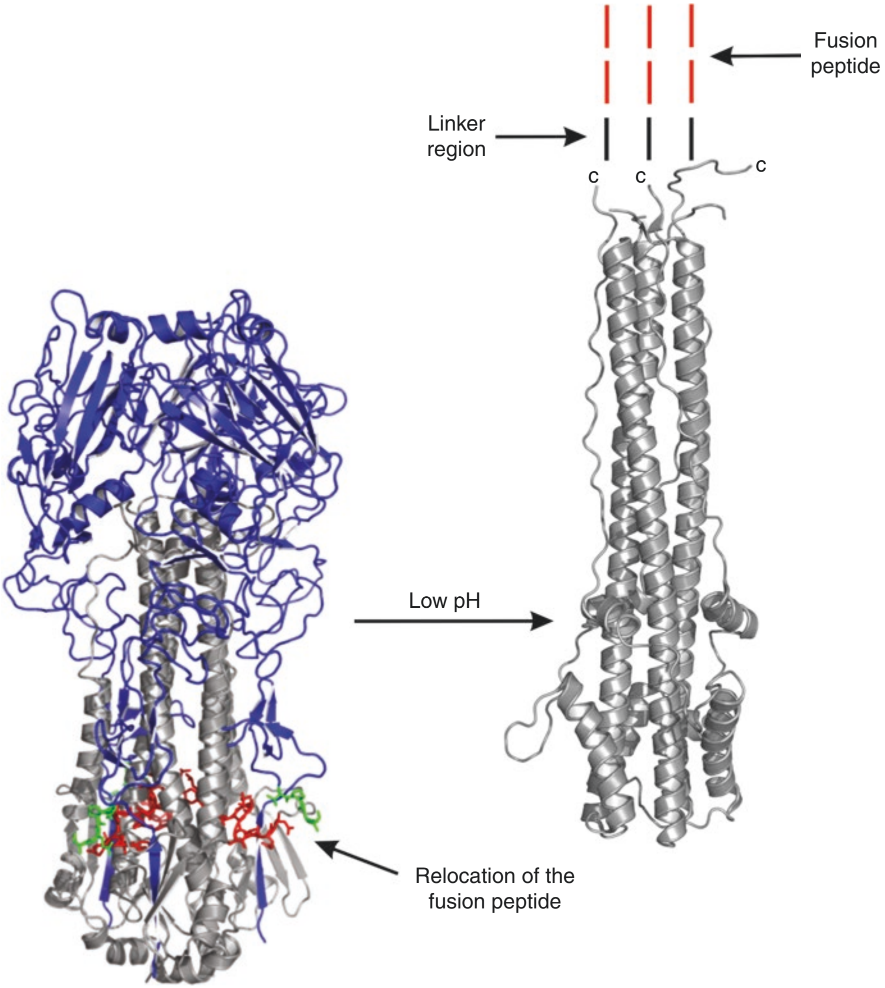

Fig. 1.1 Structures of the neutral pH cleaved HA (left) and the low pH conformation (right) assumed when membrane fusion is triggered by acidification of endosomes. In the neutral pH structure (left), the HA_1 subunit is shown in blue, the HA_2 subunit is in gray, and the C-terminal residues of HA_1 are highlighted in green, and HA_2 N-terminal fusion peptide residues highlighted in red. Keep in mind that the neutral pH structure is derived from HA ectodomains solubilized by proteolytic treatment of virions, which removes the HA transmembrane domains located near the C-terminus of each HA_2 subunit (the viral membrane would be at the bottom of the figure). The low pH rodlike structure (right) is composed entirely of HA_2 residues (gray). Hatched lines indicate regions for which the structure is unknown as they are disordered in the crystal structure (linker region) or were removed by proteolysis to solubilize the protein for crystal preparation (fusion peptide). The rodlike structure illustrates that fusion peptide residues are relocated to the same end of the helical rod as the C-terminal end of the polypeptide chains (labeled C) that are proximal to the membrane anchor domains (removed by protease treatment); therefore, the viral and endosomal membranes are pulled into close proximity with one another by acid-induced conformational changes

subunit. Second, the N-terminal 10 residues of the HA_2 fusion peptide relocate into a “cavity” present in the interior of the trimeric HA0 structure and form new contacts with a number of ionizable residues that line the cavity. These structural changes are depicted in Fig. 1.2, which shows a side view of the HA0 and cleaved HA ectodomains as well as a view down the threefold axes of symmetry. In the uncleaved HA0 structures (panels a and c), the cleavage loop extends out into solution; the conserved arginine residue that serves as the cleavage site of each monomer is indicated by the arrows. The residues colored in green and red are the ones that relocate upon protease cleavage; the relocation of HA_2 N-terminal residues (indicated in red) into the trimer interior is best illustrated when viewed down the threefold axis of cleaved HA (panel d).

A preponderance of evidence suggests that the new contacts formed when the HA_2 N-terminal fusion peptide inserts into the interior of the HA trimer are critical for priming the metastable HA structure for subsequent acid-induced conformational changes. First, let us consider the fusion peptide itself. The N-terminal 11 residues of HA_2 constitute the most highly conserved region of the HA, completely conserved in nearly all strains of the 16 influenza A HA subtype viruses and differing by only one residue in most influenza B strains (Cross et al. 2009; Nobusawa et al. 1991) (Table 1.1). Surprisingly, HAs having mutations in the fusion peptide domain that retain membrane fusion functionality can be selected for in the laboratory and generated easily by reverse genetics and have been documented at nearly every position (Cross et al. 2001, 2009; Daniels et al. 1985; Gething et al. 1986; Korte et al. 2001; Lin et al. 1997; Nobusawa et al. 1995; Orlich and Rott 1994; Qiao et al. 1999; Steinhauer et al. 1995; Yewdell et al. 1993). However, nearly all HA fusion peptide mutants with substitutions in the N-terminal 10 residues mediate membrane fusion at elevated pH relative to wild-type (WT) HA (Cross et al. 2009). Furthermore, when infectious viruses containing such mutants are generated by reverse genetics, minimal passage frequently results in either reversion to the WT HA residue or pseudoreversion at the position that was altered (Cross et al. 2001). These observations suggest an evolutionary pressure on the HA beyond that which operates on functional interactions with target membranes and suggest that fusion peptide contacts in the metastable cleaved neutral pH HA serve as pH “sensors” that are reactive to acidification. As mentioned previously, in the HA0 structure, the cleavage loop is proximal to a cavity lined with ionizable residues, which are subsequently buried by the fusion peptide upon proteolytic cleavage (Chen et al. 1998). Since the relocation of fusion peptide residues constitutes the only structural change that takes place following cleavage, the newly formed contacts are likely to play a role in potentiating fusion activity. An analysis of this region of cleaved HA shows that HA_2 ionizable residues Asp109 and Asp112 are invariantly conserved and form numerous hydrogen bonds with the fusion peptide. Other residues of interest in this region include conserved histidine residues that may serve as potential “trigger” residues for initiating acid-induced conformational changes. The pKa of the histidine side chain is around pH 6.0 in aqueous solution, and therefore, protonation of such residues would occur within the biologically relevant range during endosomal acidification, if accessible to solvent. Of particular interest are HA_1 position 17 and

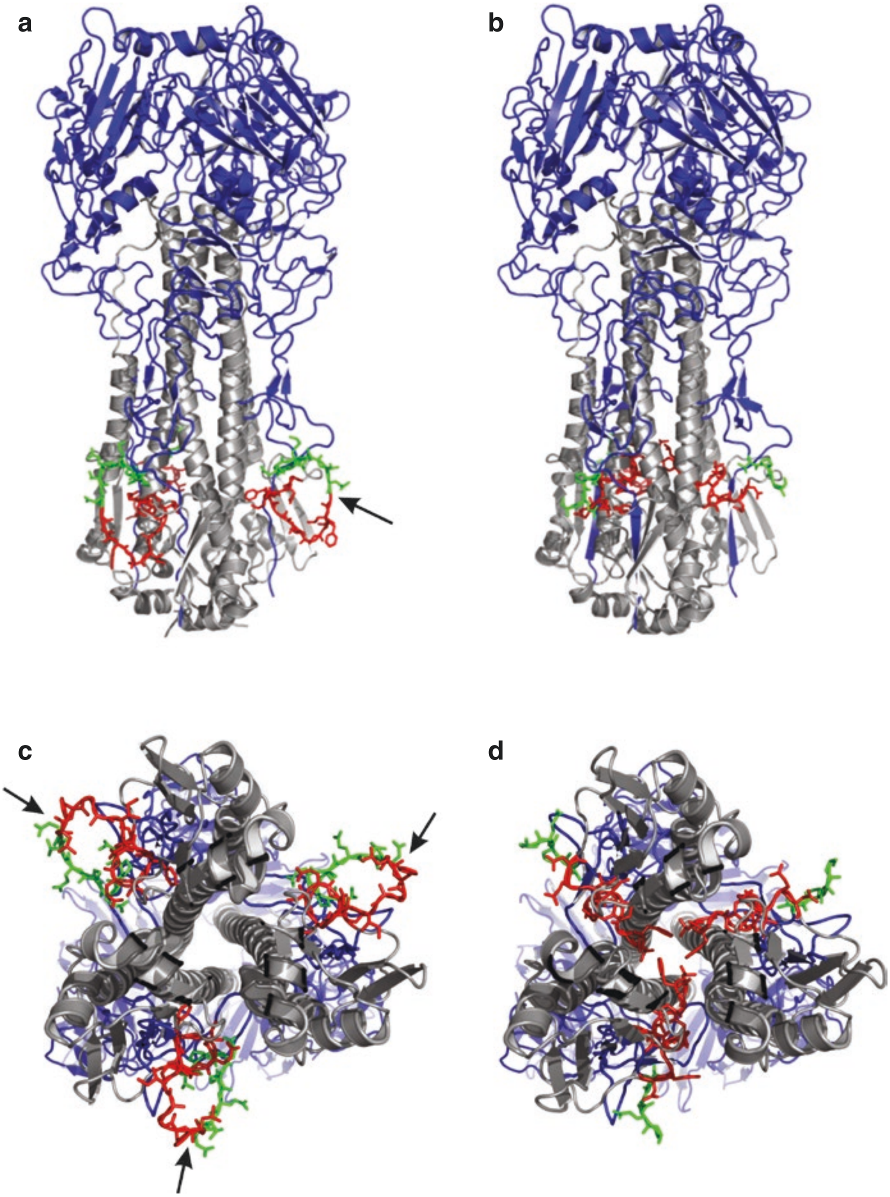

Fig. 1.2 Structures of the HA ectodomains of A/Aichi/2/68 virus (H3N2 subtype). Panels **a** and **c** represent uncleaved HA0 viewed side-on (**a**) or side-down the threefold axis of symmetry (**c**), and panels **b** and **d** represent cleaved neutral pH HA from the same orientations. The residues that constitute the HA_1 subunits are shown in blue and HA_2 in gray, and these residues do not relocate following cleavage (they are superimposable in the two structures). The cleavage loops of HA0 are shown in red and green in panels **a** and **c** with arrows indicating the site of cleavage at residue Arg 329. These are the only residues that relocate following cleavage, with residues that constitute the N-terminus of HA_2 (red) inserting into the trimer interior, as illustrated most clearly in panel **d**

Table 1.1 Cleavage site sequences of HA subtypes known to circulate in wild birds

		HA1									↓	HA2										
		Consensus sequence																				
Subtype	Virus	P						P_4	P_3	P_2	R	G	L	F	G	A	I	A	G	F	I	E
H1	A/duck/Alberta/35/1976	P	S					I	Q	S	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Puerto Rico/8/1934	P	S					I	Q	S	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/WSN/1933	P	S					I	Q	Y	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/California/04/2009	P	S					I	Q	S	R	**G**	-	-	-	-	-	-	-	-	-	-
H2	A/duck/Germany/1215/1973	P	Q					I	E	S	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Japan/305/1957	P	Q					I	E	S	R	**G**	-	-	-	-	-	-	-	-	-	-
H3	A/duck/Ukraine/1/1963	P	E					K	Q	T	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Aichi/2/1968	P	E					K	Q	T	R	**G**	-	-	-	-	-	-	-	-	-	-
H4	A/duck/Czechoslovakia/1/1956	P	E					K	A	S	R	**G**	-	-	-	-	-	-	-	-	-	-
H5	A/tern/South Africa/1961	P	Q	R	E	T	R	R	Q	K	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Vietnam/1203/2004	P	Q	R	E		R	R	R	K	R	**G**	-	-	-	-	-	-	-	-	-	-
H6	A/shearwater/Australia/1/1972	P	Q					I	E	T	R	**G**	-	-	-	-	-	-	-	-	-	-
H7	A/turkey/England/1963	P	K					R	R	R	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/turkey/Oregon/1971	P	E				N	P	K	T	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Shanghai/02/2013 (H7N9)	P	E				I	P	K	G	R	**G**	-	-	-	-	-	-	-	-	-	-
H8	A/turkey/Ontario/6118/1968	P	S					V	E	P	R	**G**	-	-	-	-	-	-	-	-	-	-
H9	A/turkey/Wisconsin/1/1966	P	A					V	S	S	R	**G**	-	-	-	-	-	-	-	-	-	-
H10	A/chicken/Germany/N/1949	P	E				V	V	Q	G	R	**G**	-	-	-	-	-	-	-	-	-	-
H11	A/duck/England/1/1956	P	A					I	A	S	R	**G**	-	-	-	-	-	-	-	-	-	-
H12	A/duck/Alberta/60/1976	P	Q					V	Q	D	R	**G**	-	-	-	-	-	-	-	-	-	-
H13	A/gull/Maryland/704/1977	P	A					I	S	N	R	**G**	-	-	-	-	-	-	-	-	-	-
H14	A/duck/Astrakhan/263/1982	P	G					K	Q	A	K	**G**	-	-	-	-	-	-	-	-	-	-
H15	A/duck/Australia/341/1983	P	E				K	I	H	T	R	**G**	-	-	-	-	-	-	-	-	-	-
H16	A/black-headed gull/Sweden/5/1999	P	S					V	G	E	R	**G**	-	-	-	-	-	-	-	-	-	-

Sequences are anchored by the conserved proline (P) N-terminal to the R329 (H3 numbering) that is the site of proteolytic cleavage (denoted with ↓). The P4, P3, P2 positions are labeled such that P1 is the R329 cleavage site. C-terminal to the R329 is the highly conserved HA2 N-terminal fusion peptide (N-GLFGAIAGFIE-C). The first glycine (G) of the fusion peptide is in bold type. The (-) indicates invariant conservation

HA_2 positions 106 and 111. The 16 HA subtypes that circulate in aquatic birds can be phylogenetically and structurally segregated into two groups (Air 1981; Nobusawa et al. 1991; Russell et al. 2004), Group-1 (H1, H2, H5, H6, H8, H9, H11, H12, H13, H16) and Group-2 (H3, H4, H7, H10, H14, H15), and these three positions (HA_1 17 and HA_2 106 and 111) segregate along evolutionary lines and are well conserved within groups. In Group-1, HA_1 17 is a tyrosine, HA_2 106 is an arginine, and HA_2 111 is a histidine; in Group-2, HA_1 17 is a histidine, HA_2 106 is a histidine, and HA_2 111 is a threonine. Extensive mutagenesis studies on these residues of an H3 subtype (Group-2) HA reveals that HA_1 His17 may play a role in triggering structural changes upon acidification (Thoennes et al. 2008). In Group-1 HAs, HA_1 17 is a tyrosine, but nearby HA_2 residue His111 is invariantly conserved and may serve an equivalent role to Group-2 HA_1 His17, and indeed, mutagenesis studies on a number of Group-1 HAs suggest that changes to His111 result in the inactivation of fusion activity (J. Trost, D.A.S., unpublished).

A number of additional studies support the idea that the contacts made between the fusion peptide and the cavity region may be critical in triggering acid-induced structural rearrangements. HA mutants that mediate membrane fusion at elevated pH relative to WT can be found throughout the HA trimer at domain interfaces that rearrange upon acidification (Byrd-Leotis et al. 2015; Cross et al. 2001; Daniels et al. 1985; Doms et al. 1986; Gething et al. 1986; Lin et al. 1997; Qiao et al. 1999; Steinhauer et al. 1995; Thoennes et al. 2008; Zaraket et al. 2013b), whereas stabilizing mutations, or those resulting in an HA that mediates membrane fusion at a lower pH compared to WT, have been identified at fewer positions to date (Byrd-Leotis et al. 2015; Steinhauer et al. 1991; Thoennes et al. 2008; Xu and Wilson 2011; Zaraket et al. 2013a). The identification of mutants that mediate membrane fusion at higher or lower pH relative to WT provided for the rational design of double mutant HAs, in which mutations known to confer high- and low-pH phenotypes were generated in various combinations to examine the cumulative effects. These studies revealed that when the structural locations of the mutations were proximal, the pH phenotype was additive, whereas when the mutations were distal to one another, the phenotype of one mutation was dominant, with changes in the fusion peptide region being most critical in determining the overall fusion phenotype (Steinhauer et al. 1996). These results supported data on the kinetics of conformational changes, as determined using a panel of conformation-specific monoclonal antibodies, which showed that structural alterations in the stem region preceded changes in the HA head domains (White and Wilson 1987). These interpretations were further supported by experiments on the fusion kinetics of single virions with planar lipid bilayers, which indicated that the “withdrawal” of the fusion peptide from the trimer interior is the rate-limiting step (Ivanovic et al. 2013). In addition, another study involving hydrogen-deuterium exchange and mass spectrometry approaches revealed that release of the fusion peptide occurs prior to the structural changes involving coiled-coil rearrangements (Garcia et al. 2015), and studies on the membrane fusion process by Cryo-EM indicate that extrusion of the fusion peptide and insertion into the target membrane as extended intermediates precedes the foldback into helical rods at fusion “dimples” where membrane merger occurs (Calder and Rosenthal 2016). Overall, the data strongly support a mechanism by which HA0 cleavage and coincident conformational changes prime the cleaved neutral pH structure for fusion by generating functional fusion peptides and critically relocating them to structural regions that can respond to acidification.

1.2 Structure of HA0 Cleavage Sites

Not only is cleavage activation required for infectivity, but the available evidence indicates that the sequence and structure of the HA cleavage loop may reveal clues about the ability of HAs to serve as substrates for specific activating proteases and dictate traits such as the pathogenicity of influenza strains and the ecology and host range of influenza viruses. To date, the HA0 precursor structure of three HA subtypes has been determined: the H3 subtype described above from a human 1968 pandemic H3N2 virus (Chen et al. 1998), the H1 subtype representing 1918 human pandemic H1N1 strains (Stevens et al. 2004), and the H16 subtype from an H16N3 avian virus isolated from a black-headed gull (Lu et al. 2012). Figure 1.3 depicts these three structures as a single monomer of the trimer, viewed from equivalent orientation. A visual comparison of these structures reveals major differences in the structural elements of the cleavage loop, including the location of the Arg329 cleavage site (green). Shown in yellow are the conserved electronegative residues discussed above, HA_2 D109 and D112, HA_2 His 17 for H3, and HA_2 His 111 for H1 and H16. The H3 Arg329 cleavage site is located within a relatively standard loop structure that orients away from the trimer surface, leaving it easily accessible at a distal position in the loop. Unlike the H3 HA0, the loop structure of the H1 HA0 packs against the surface of the trimer, and the Arg329 residue is positioned such that it covers the electronegative cavity, though it orients out into solution. The H16 cleavage loop is somewhat similar to the H1 loop in that it packs against the trimer surface, but unlike the other two, it contains a five-residue α-helix that includes positions 325 through the Arg329 that is cleaved by activating proteases. Remarkably, this helix covers the electronegative cavity and orients the side chain of Arg329 into the cavity where it forms a salt bridge with highly conserved HA_2 residue D112. The orientation of Arg329 in H16 HA0 is likely a factor in limiting its accessibility to proteases that can activate other subtypes, but not H16 (see below). It is possible that the differences in cleavage loop structures may exist, in part, due to the location of proximal glycosylation sites, which for H1 and H16 is at HA_1 N20 (H3 numbering), whereas for H3 the glycosylation site is slightly farther away at HA_1 N22 (Chen et al. 1998; Lu et al. 2012; Stevens et al. 2004). It should also be noted that the H3 HA0 structure was derived from protein expressed in mammalian cells, and the H1 and H16 structures were based on proteins expressed using baculovirus expression systems. It will be very interesting to extend our knowledge of HA0 precursor structures to additional subtypes and relate these to data that is accumulating on the range of proteases capable of activating individual HA subtypes.

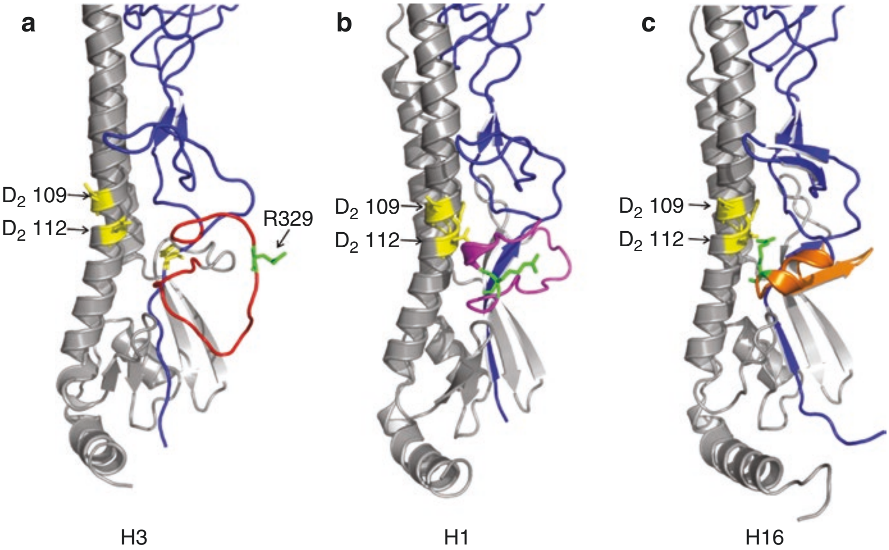

Fig. 1.3 Cleavage site structures showing a single monomer of the HA0 trimer for H3 subtype (**a**), H1 subtype (**b**), and H16 subtype (**c**). (**a**) The cleavage domain of H3 (red) forms a loop structure with R329 (green) oriented out into solution. Shown in yellow are residues that line the electronegative cavity, residues HA_2 D109, HA_2 D112 (subscript denotes the HA_2 subunit in the figure), and HA_1 His 17 (between $D_2$112 and R329). (**b**) The cleavage domain of H1 (magenta) packs more closely against the trimer, but R329 (green) remains oriented out into solution. HA_2 D109, HA_2 D112, and HA_2 H111 are shown in yellow, though H111 is on the opposite face of the long helix and largely hidden in the depicted orientation. (**c**) The cleavage domain of H16 (orange) contains a small helical domain that orients the side chain of R329 (green) into the electronegative cavity, where it forms a salt bridge with HA_2 D112 (again, ionizable residues shown in yellow). In all panels, residues that will constitute the HA_1 and HA_2 subunits following cleavage are shown in blue and gray, respectively

1.3 Activating Proteases

Whereas the residues of the cleavage loop that constitute the N-terminal portion of the fusion peptide are highly conserved, the upstream residues between the highly conserved Pro324 (H3 numbering) and Arg329 cleavage site are quite variable among subtypes and provide important clues regarding cleavage activation and the proteases involved (Tables 1.1 and 1.2). In particular, variations in the number and sequence of arginine and lysine residues at the cleavage site account for major differences in proteolytic activation and proved to be critical determinants of pathogenicity (Bosch et al. 1981; Garten and Klenk 1999; Klenk and Garten 1994; Steinhauer 1999) and host range (Galloway et al. 2013). Other factors that can influence the capacity for particular proteases to activate HAs include the presence or absence of nearby carbohydrates that can alter the accessibility of an activating protease (Deshpande et al. 1987; Kawaoka et al. 1984; Kawaoka and Webster 1989; Ohuchi et al. 1991). Studies on a wide range of influenza strains and subtypes have addressed

Table 1.2 Cleavage site sequences of H7 LPAI and HPAI viruses having insertions derived through recombination

		HA1																									↓	HA2										
		Consensus sequence																																				
		P			P_4	P_3	P_2																				R	**G**	L	F	**G**	A	I	A	**G**	F	I	E
H7 LPAI	A/turkey/ Oregon/1971	P	E	N	P	K	T																				R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Shanghai/ 02/2013	P	E	I	P	K	G																				R	**G**	-	-	-	-	-	-	-	-	-	-
H7 HPAI	A/turkey/ Oregon/1971 (lab)	P	E	N	P	K	T	**S**	**L**	**S**	**P**	**L**	**Y**	**P**	**G**	**R**	**T**	**T**	**D**	**L**	**V**	**Q**	**V**	**P**	**T**	**A**	R	**G**	-	-	-	-	-	-	-	-	-	-
	A/chicken/ Chile/2002	P	E	N	P	K	T	**C**	**S**	**P**	**L**	**S**	**R**	**C**	**R**	**E**	**T**										R	**G**	-	-	-	-	-	-	-	-	-	-
	A/chicken/British Columbia/2004	P	E	N	P	K	**Q**	**A**	**Y**	**Q**	**K**	**R**	**M**	T													R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Mexico/ InDRE7218/2012	P	E	N	P	K	**D**	**R**	**K**	**S**	**R**	**H**	**H**	**R**	T												R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Guangdong/ 17SF003/2016	P	E	V	P	K	R	**K**	**R**	**T**	**A**																R	**G**	-	-	-	-	-	-	-	-	-	-
	A/Taiwan/1/2017	P	E	V	P	K	R	**K**	**R**	**T**	**A**																R	**G**	-	-	-	-	-	-	-	-	-	-

Sequences are anchored by the conserved proline (P) located N-terminal to the R329 (H3 numbering) that is the site of proteolytic cleavage (denoted with ↓). The P4, P3, P2 positions are labeled such that P1 is the R329 cleavage site. HA1 insertions are shown in bold underline type. C-terminal to the R329 is the highly conserved HA2 N-terminal fusion peptide (N-GLFGAIAGFIE-C). The first glycine (G) of the fusion peptide is in bold type. The (-) indicates invariant conservation functional activation of infectivity or fusion potential, and a variety of candidate proteases have been identified that activate low pathogenic avian influenza (LPAI) viruses and mammalian strains at monobasic cleavage sites. They include membrane-anchored serine proteases (MASP) (see Chap. 8) such as human airway tryptase (HAT) (Bottcher et al. 2006; Bertram et al. 2012), transmembrane protease, serine S1 member 2 (TMPRSS2) (Bottcher et al. 2006; Chaipan et al. 2009; Hatesuer et al. 2013; Sakai et al. 2014; Tarnow et al. 2014), TMPRSS4 (Bertram et al. 2010a; Kuhn et al. 2016), and matriptase (Baron et al. 2013; Hamilton et al. 2012), as well as soluble serine proteases (see Chap. 9) such as tryptase Clara (Kido et al. 1992), plasmin (Lazarowitz et al. 1973b), and kallikrein-related peptidases (Hamilton and Whittaker 2013). In considering the numerous studies published to date on an extended panel of proteases, it is apparent that no single universal protease is involved in activating the membrane fusion potential of the HA, as most of the proteases examined activate subsets of viruses in nonoverlapping fashion and with extensive variation in cleavage efficiency. For example, Galloway et al. (2013) characterized cleavage activation from representatives of all 16 avian HA subtypes and several human strains for activation by trypsin, HAT, and TMPRSS2 and observed a variety of phenotypes. Trypsin displayed fairly broad cleavage activation, but showed low or no activity for H12, H13, and H16 subtype HAs, whereas HAT was more limited in its ability to cleave various subtypes, but was able to cleave H9, H11, and H12 HAs. On the other hand, TMPRSS2 was broadly effective but demonstrated no activity against H8 and H12 and was the only one of the three proteases examined in this study that showed specificity for the H13 and H16 HAs. Interestingly, H13 and H16 appear to have a limited host range, generally being isolated only from Charadriiformes, such as gulls and shorebirds, suggesting perhaps that cleavage activation can serve as a host range determinant for particular viruses. Furthermore, a study documenting a repository of Eurasian-lineage reverse genetics vectors and recombinant viruses found that the H13 and H16 recombinant viruses were only able to be recovered when transfected 293T cells were injected into embryonated chicken eggs for further propagation; embryonated chicken eggs are known to express a factor Xa-like serine protease that is capable of cleaving HA0, which may provide the rationale for this observation (Keawcharoen et al. 2010). As mentioned previously, the cleavage loop of the H16 HA appears to be rather structured with the critical Arg329 less exposed than observed in the loop structure of the H3 subtype HA. This may be indicative of host-specific proteases found in Charadriiformes that have greater activity against H13 and H16 HA substrates, but this remains to be determined. Additionally, it was shown that human TMPRSS2 homologues in chicken and swine were capable of activating an influenza virus having a monobasic cleavage site in vitro (Bertram et al. 2012). In another example, studies using matriptase have shown that it can efficiently activate H1 and H9 subtype viruses, but is much less effective against H2 and H3 subtypes. Furthermore, while matriptase was able to cleave H1 subtype HAs, in general, it displayed selective activity against certain H1 subtype strains (Baron et al. 2013; Hamilton et al. 2012). Similarly, studies examining the cleavage activity of the thrombolytic proteases kallikrein (KLK) 5 and 12 with various HA strains belonging to H1, H2, and H3 subtypes found that KLK5 and KLK12 displayed differential activity against HAs within and across subtypes. For example, KLK12, but not KLK5, was able to cleave H2 HA; the opposite was observed with the H3 HAs examined. For the H1 HAs, KLK5 and KLK12 were able to efficiently cleave the HA from A/California/04/2009 (H1N1), but have substantially reduced activity against other H1 strains, such as A/New Caledonia/99, A/South Carolina/18, and the lab-adapted strains A/Puerto Rico/8/34 and A/WSN/33 (Hamilton and Whittaker 2013

Inextricably linked to the recognition of HA as a substrate by activating proteases is the requirement for the protease to generate specific cleavage products containing fusion-competent sequences at the HA_2 N-terminus. As mentioned previously, the HA_2 N-terminal fusion peptide domain that results from cleavage is the most conserved region in HA, with the sequence GLFGAIAGFIE being virtually invariant in natural isolates. Studies have shown that HA digestion by proteases such as thermolysin, which cleaves between the Gly_1 and Leu_2 of authentic HA_2, renders the HA inactive for fusion and results in noninfectious virus (Garten et al. 1981; Lazarowitz and Choppin 1975; Orlich and Rott 1994; Steinhauer et al. 1995). Interestingly, the selection of influenza viruses capable of replicating in the presence of thermolysin revealed HA mutants having a single amino acid insertion just downstream of the Leu_2, effectively restoring the length and spacing of critical residues of the fusion peptide (Orlich and Rott 1994), and fusion peptide length as well as composition has been demonstrated as a requirement for fusion function (Langley et al. 2009; Steinhauer et al. 1995). As mentioned previously, the HAT protease was able to generate HA_1 and HA_2 cleavage products with similar mobility as HAs digested with trypsin, but the HAT-cleaved H12 HA was not capable of mediating membrane fusion (Galloway et al. 2013). It was also shown that a recombinant virus expressing the H12 HA replicated poorly in MDCK cells, but replicated quite well in embryonated chicken eggs (Keawcharoen et al. 2010), which express a factor Xa-like serine protease that is capable of activating HA0 (Gotoh et al. 1990).

An influenza HA cleavage anomaly is observed from studies on the H1N1 strain A/WSN/33. This is a lab-adapted neurovirulent strain selected by passage of the first human influenza A isolate A/WS/33 (Smith et al. 1933) in mouse brain (Francis and Moore 1940) and is efficiently cleaved by the fibrinolytic protease plasmin (Lazarowitz et al. 1973b). The mechanism by which plasmin-mediated HA cleavage occurs was shown to be moderated by the ability of the NA of A/WSN/33 (H1N1) to sequester plasminogen, the zymogen precursor to plasmin, until it is converted to plasmin and able to cleave HA0 into HA_1 and HA_2. The ability of the NA to sequester plasminogen was shown to be dependent on the absence of a carbohydrate at position 146 (N2 numbering) and the presence of a carboxy-terminal lysine (Lys453) (Goto and Kawaoka 1998; Li et al. 1993). This feature has only been observed for the A/WSN/33 NA, but shows specificity for multiple HA subtypes (Goto and Kawaoka 1998). The HA cleavage site sequence of the A/WSN/33 H1 HA contains a serine to tyrosine substitution at the P2 position, resulting in the cleavage site IQY↓R rather than IQS↓R, as is found in most H1 HAs (Table 1.1) (Sun et al. 2010). In addition to the A/WSN/33 H1 HA, a subsequent study reported on the acquisition of a similar cleavage site sequence in a more contemporary seasonal H1N1 strain, A/Beijing/718/2009, that is preferentially cleaved by plasmin

(Tse et al. 2013). These data are consistent with other studies showing that bulky hydrophobic amino acids, such as tyrosine, in the P2 position promote HA cleavage by plasmin. Unlike the A/WSN/33 HA, the ability of the HA to be cleaved by plasmin was found to be independent of the NA. Proteolytic conversion of plasminogen into plasmin is mediated by plasminogen activators. It has recently been shown that interferon stimulates the expression of the plasminogen activator inhibitor PAI-1 which prevents not only activation of plasmin but inhibits also other HA-activating proteases, such as TMPRSS2 and HAT. Inhibition of proteolytic activation of HA may therefore be an important mechanism in innate immunity (Dittmann et al. 2015).

Most studies evaluating the proteolytic activation of HA have been cell-based or in vitro studies; however, following the discovery of the ability of TMPRSS2 to cleave influenza A virus HA (Bottcher et al. 2006), several groups investigated this further in knockout and mutant mice by examining the ability of several different influenza A viruses to replicate and cause disease in *Tmprss2*-deficient mice. These studies showed that H1N1-, H3N2-, and H7N9-infected *Tmprss2*-deficient mice were generally protected from influenza-associated morbidity and mortality (Hatesuer et al. 2013; Sakai et al. 2014; Tarnow et al. 2014). TMPRSS2-mediated influenza virus replication in vivo appears to be limited to influenza A viruses, as a recent study showed that influenza B viruses were able to efficiently replicate and cause disease in *Tmprss2*-deficient mice (Sakai et al. 2016).

Many of the proteases that have been implicated in activating cleavage are secreted and activate HA extracellularly; however, proteases such as TMPRSS2, TMPRSS4, and HAT have membrane-bound forms that may be active during HA transport or at the plasma membrane (Bottcher et al. 2009). In addition, a description of A/WSN/33 HA cleavage in endosomes during virus entry, distinct from the plasmin-directed cleavage described above, has been reported (Boycott et al. 1994). Though this was unique for MDBK cells and the protease involved was not identified, cell-associated proteases have the potential for HA activation during entry as well as during transport following de novo synthesis (Zhirnov et al. 2002). Cumulatively, it appears that among influenza strains, a broad range of proteases have the potential to activate infectivity and that activation can occur in cell-associated fashion at early or late stages of the replication cycle or extracellularly during virus spread in a host or during transmission (see Fig. 1.2 of Bertram et al. (2010b) and Fig. 1.1 of Bottcher-Friebertshauser et al. (2014) for useful illustrations). Proteolytic activation of HA is also an important pathogenic mechanism in viral-bacterial coinfection. It has long been known that various bacteria, including *Staphylococcus aureus*, secrete proteases that cleave HA and promote the development of pneumonia in virus-infected mice (Tashiro et al. 1987; Maeda 1996). Moreover, bacterial proteases may activate host proteases that cleave HA (Scheiblauer et al. 1992).

Perhaps less complicated are the HAs having cleavage site sequences that consist of a stretch of polybasic amino acids, with the general sequence motifs of R-X-(K/R)-R, which can be cleaved in the trans-Golgi by subtilisin-like proteases (see Chap. 10) such as furin and proprotein convertase 5/6 (PC5/6) (Horimoto et al. 1994;