Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Chapter 504

Sindbis Virus nsP2 Endopeptidase

Databanks

MEROPS name: sindbis virus-type nsP2 peptidase
MEROPS classification: clan CN, family C9, peptidase C09.001
Species distribution: known only from Sindbis virus
Reference sequence from: Sindbis virus (UniProt: P03317)

Name and History

Sindbis virus (family Togaviridae, genus *Alphavirus*) is an animal virus with a single molecule of positive-strand RNA 11 703 nucleotides in length as its genome. The genomic RNA is a message for production of four non-structural proteins which are translated in the cytoplasm of the infected cell as two polyproteins, P123 (1896 amino acids) and P1234 (2513 amino acids, produced by read-through of an opal termination codon at position 1897). The P1234 polyprotein is processed by the ***nsP2 proteinase*** through a series of cleavage intermediates into the four final products nsP1, nsP2, nsP3 and nsP4, named in order 5′ to 3′ along the genome. Since cytoplasmic, nonorganelle-bound proteinases are rare in animal cells, a virally encoded activity was predicted as early as 1981 [1]. The proteinase was localized to the C-terminal half of nsP2 by translation *in vitro* of RNA transcribed from deleted or truncated full-length cDNA clones of Sindbis virus. Deletions within nsP1, nsP3 and nsP4 did not affect proteolysis, but deletions in the C-terminal half of nsP2 abolished proteinase activity, and deletions in the N-terminal half of nsP2 led to aberrant processing [2,3]. At the same time, several mutations that rendered the proteinase temperature sensitive were found to lie in the C-terminal half of nsP2 [4]. Site-specific mutagenesis of Cys and His residues in this domain that are conserved among alphaviruses later showed that Cys481 and His558 are essential for proteolytic activity, and it was proposed that these constitute the catalytic dyad of a cysteine proteinase [5].

Activity and Specificity

The only known substrates for the Sindbis nsP2 proteinase are the Sindbis non-structural polyproteins (P123 and P1234) and their cleavage intermediates (which include P12, P23 and P34). The consensus cleavage site is (Ala/Ile/Val/Ser)-Gly-(Ala/Xaa)↓(Ala/Tyr). Mutational studies have demonstrated that the invariant Gly in the P2 position is essential for efficient cleavage; changing this Gly to Ala led to substantially slower cleavage of the mutated site, and substitution with Val or Glu rendered the site completely uncleavable [6]. Similarly, an Ala or Gly was found to be essential in the P1 position for efficient cleavage in Sindbis virus (Cys was not tested) [7]. The enzyme is relatively insensitive to the identity of the residue in the P1′ position, however [7]. It is thought that additional determinants of specificity must exist, since these are the only sites cleaved by the enzyme. In this regard it has been shown that different polyproteins containing the nsP2 proteinase differ in their preferences for the three cleavage sites, suggesting that cleavage-site recognition involves more extensive interactions than the simple recognition of a short linear sequence by the enzyme [7].

Structural Chemistry

Sindbis nsP2 is a polypeptide of 807 amino acids encoded by nucleotides 1680–4100 of the genome. The N-

terminal half of the protein is believed to be a helicase required for RNA replication. The C-terminal domain between residues 460 and 807 constitutes the proteinase [2]. Both Cys481 and His558, the putative catalytic dyad, are followed by tryptophan residues, and Trp559 has been shown to be essential for proteolysis. These two aromatic residues may interact (stack) to maintain proper conformation of the catalytic pocket [5]. Substitution of Asn614 by Asp enhances cleavage; an Asn residue has been shown previously to be associated with the catalytic pocket in papain [8]. Mutations Phe509Leu, Ala517Thr, Asp522Asn and Gly748Ser have been shown to render the proteinase temperature sensitive *in vivo*; these mutations are believed to alter the structure of the protein [9]. The proteinase has been hypothesized to be related to papain on the basis of alignment studies [10,11], but there is very limited sequence identity between the nsP2 proteinase and papain to support this hypothesis [5]. No structural studies of the enzyme have been produced to date.

Preparation

Sindbis virus has a very wide host range and tissue distribution, and infects many types of cells in culture, including avian, mammalian and mosquito cells. The genome of the infecting virus is a messenger RNA from which the non-structural proteins are translated in the cytoplasm of infected cells. Both nsP2 and polyproteins containing nsP2 have also been expressed by translation in reticulocyte lysates of virion RNA or of *in vitro* transcribed RNA [2,3,5,7,12], and by translation in bacteria or in mammalian cells of appropriate constructs [13,14].

Biological Aspects

Both nsP2 and polyproteins containing nsP2 are active proteinases, but they differ in their site preferences. The activities of the enzyme have been studied both *in vitro* (referenced above) and in infected or transfected cells [6,7,9,12,14,15]. In many of these studies, site-directed mutagenesis was used to destroy cleavage sites or to inactivate the proteinase. These studies have shown that P1234 can cleave itself autoproteolytically in *cis* to give P123 and nsP4, whereas P123 can be cleaved only in *trans*. Cleavage of P123 by another P123 molecule or by P12 produces nsP1 and P23, whereas cleavage by P23 produces predominantly P12 and nsP3. Finally, enzymes that lack nsP3 (that is, nsP2 or P12) cannot cleave the nsP3/nsP4 site. These preferences lead to a temporal regulation of RNA synthesis during the infection cycle. Early in infection P123 and nsP4 are produced, and together with host factors these form a replicase complex that synthesizes negative-strand RNA. Cleavage of P123 in the replicase complex produces a replicase that synthesizes positive-strand RNAs very efficiently, but cannot make negative strands [14,15]. As the concentration of *trans*-acting proteinases builds up, the non-structural polyprotein is cleaved while nascent at the nsP2/nsP3 site, and the products produced are nsP1, nsP2, nsP3 and P34. Negative-strand RNA synthesis is then shut off, because the cleaved products cannot form a negative-strand replicase. The use of the non-structural proteinase to control the RNA developmental pathway represents an interesting evolutionary adaptation of the need of viruses to encode proteinases to process polyproteins in the cell cytoplasm.

Related Peptidases

All alphaviruses encode nsP2 proteinases that share more than 50% amino acid sequence identity. Sequences of nsP2 proteinases that have been deposited in the databases include those for Sindbis virus, eastern equine encephalitis virus, O'nyong-nyong virus, Ross River virus, Semliki Forest virus, Venezuelan equine encephalitis virus, western equine encephalomyelitis virus, Barmah Forest virus, Mayaro virus, and Sagiyama and Aura viruses.

All alphavirus non-structural cleavage sites are of similar form (Table 504.1) (reviewed in Strauss & Strauss [16]).

The extensive sequence similarity between the nsP2 proteinases and the similarity in the cleavage sites suggest that the proteinases must be very similar to one another, but no tests of cross-cleavage have been done to determine how specific each enzyme is for its own substrate.

TABLE 504.1 Substrate sequences of alphavirus nonstructural proteinase cleavage sites.

	P4	P3	P2	P1	P1–	P2–	P3–	P4–
nsP1/nsP2	Xaa	Ala/Ile	Gly	Ala/Glu ↓	Ala/Gly	Xaa	Val	Glu
nsP2/nsP3	Xaa	Ala/Val/Ser	Gly	Ala/Cys/Arg/Ser ↓	Ala	Pro	Ser/Ala	Tyr
nsP3/nsP4	Xaa	Ala/Val	Gly	Ala/Gly ↓	Tyr	Ile	Phe	Ser

Similar papain-like endopeptidases, which may have a common evolutionary origin, have been reported for a number of plant and animal viruses, including rubella virus, a number of coronaviruses, a number of potyviruses, a tymovirus and hypovirulence-associated virus of fungi [5,11].

Further Reading

Ten Dam *et al*. [17] have reviewed the cysteine peptidases of the Togaviridae.

References

[1] Rice, C.M., Strauss, J.H. (1981). Nucleotide sequence of the 26S mRNA of Sindbis virus and deduced sequence of the encoded virus structural proteins. *Proc. Natl. Acad. Sci. USA* 78, 2062–2066.

[2] Hardy, W.R., Strauss, J.H. (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, 4653–4664.

[3] Ding, M., Schlesinger, M.J. (1989). Evidence that Sindbis virus nsP2 is an autoprotease which processes the virus nonstructural polyprotein. *Virology* 171, 280–284.

[4] Hahn, Y.S., Strauss, E.G., Strauss, J.H. (1989). Mapping of RNA-temperature-sensitive mutants of Sindbis virus: Assignment of complementation groups A, B, and G to nonstructural proteins. *J. Virol.* 63, 3142–3150.

[5] Strauss, E.G., de Groot, R.J., Levinson, R., Strauss, J.H. (1992). Identification of the active site residues in the nsP2 proteinase of Sindbis virus. *Virology* 191, 932–940.

[6] Shirako, Y., Strauss, J.H. (1990). Cleavage between nsP1 and nsP2 initiates the processing pathway of Sindbis virus nonstructural polyprotein P123. *Virology* 177, 54–64.

[7] de Groot, R.J., Hardy, W.R., Shirako, Y., Strauss, J.H. (1990). Cleavage-site preferences of Sindbis virus polyproteins containing the nonstructural proteinase: evidence for temporal regulation of polyprotein processing *in vivo*. *EMBO J.* 9, 2631–2638.

[8] Garavito, R.M., Rossmann, M.G., Argos, P., Eventoff, W. (1977). Convergence of active site geometries. *Biochemistry* 16, 5065–5071.

[9] Hardy, W.R., Hahn, Y.S., de Groot, R.J., Strauss, E.G., Strauss, J.H. (1990). Synthesis and processing of the nonstructural polyproteins of several temperature-sensitive mutants of Sindbis virus. *Virology* 177, 199–208.

[10] Gorbalenya, A.E., Koonin, E.V., Lai, M.M.C. (1991). Putative papain-related thiol proteases of positive-strand RNA viruses – identification of rubivirus and aphthovirus proteases and delineation of a novel conserved domain associated with proteases of rubivirus, alpha-, and coronavirus. *FEBS Lett.* 288, 201–205.

[11] Koonin, E.V., Dolja, V.V. (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.

[12] de Groot, R.J., Rümenapf, T., Kuhn, R.J., Strauss, E.G., Strauss, J.H. (1991). Sindbis virus RNA polymerase is degraded by the N-end rule pathway. *Proc. Natl. Acad. Sci. USA* 88, 8967–8971.

[13] Hardy, W.R., Strauss, J.H. (1988). Processing of the nonstructural polyproteins of Sindbis virus: study of the kinetics *in vivo* using monospecific antibodies. *J. Virol.* 62, 998–1007.

[14] Lemm, J.A., Rümenapf, T., Strauss, E.G., Strauss, J.H., Rice, C.M. (1994). Polypeptide requirements for assembly of functional Sindbis virus replication complexes: a model for the temporal regulation of minus-strand and plus-strand RNA-synthesis. *EMBO J.* 13, 2925–2934.

[15] Shirako, Y., Strauss, J.H. (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, 1874–1885.

[16] Strauss, J.H., Strauss, E.G. (1994). The alphaviruses: gene expression, replication, and evolution. *Microbiol. Rev.* 58, 491–562.

[17] Ten Dam, E., Flint, M., Ryan, M.D. (1999). Virus-encoded proteinases of the *Togaviridae*. *J. Gen. Virol.* 80, 1879–1888.

Ellen G. Strauss

Division of Biology, 156-29, California Institute of Technology, Pasadena, CA 91125, USA. Email: strausse@cco.caltech.edu

James H. Strauss

Division of Biology, 156-29, California Institute of Technology, Pasadena, CA 91125, USA. Email: straussj@starbase1.caltech.edu

This article is reproduced from the previous edition, Volume 2, pp. 1269–1271, © 2004, Elsevier Ltd., with revisions made by the Editors.

Handbook of Proteolytic Enzymes, 3rd Edn
ISBN: 978-0-12-382219-2

DOI: http://dx.doi.org/10.1016/B978-0-12-382219-2.00494-4