DOI: 10.1002/cmdc.200600285

Unique Composite Active Site of the Hepatitis C Virus NS2-3 Protease: a New Opportunity for Antiviral Drug Design

Zucai Suo[a] and Mohd Amir F. Abdullah[b]

Hepatitis C virus (HCV), the leading cause of non-A and non-B viral hepatitis, is a plus-stranded RNA about 10 kb in length, and its organization is similar to that of members of the family *Flaviviridae*.[1–3] The genome encodes a single precursor protein that includes structural and nonstructural proteins.[2,4] The precursor protein is proteolytically processed by both host signal peptidases and viral proteases to produce at least 10 viral proteins: Core, E1, E2, p7, NS2, NS3, NS4A, NS4B, NS5A, and NS5B.[4–6] Host signal peptidases cleave at the Core/E1, E1/E2, E2/p7, and p7/NS2 junctions in the precursor protein, while the HCV-encoded NS3-NS4A serine protease complex cleaves the NS3/NS4A junction in cis (on the same polypeptide)[7] and the NS4A/NS4B, NS4B/NS5A, and NS5A/NS5B junctions in trans (on a different polypeptide).[8] The NS2/NS3 junction cleavage is thought to be self-processed by another virally encoded protease, NS2-3, in cis.[9–11] However, this mechanism may need to be changed based on the X-ray crystal structure of the catalytic domain ($NS2^{protease}$, Figure 1a) of the NS2-3 protease[12] published recently by Lorenz, Marcotrigiano, and colleagues.

The NS2-3 protease, a cysteine protease, consists of residues 94–217 of NS2 and residues 1–181 of NS3. The NS2 region that forms the catalytic domain $NS2^{protease}$ does not share any clear sequence homology with known proteolytic enzymes, and its structure shows a novel protein fold.[12] Surprisingly, the authors found that the $NS2^{protease}$ forms a homodimer containing two composite active sites: residues His 143 and Glu 163 from chain A and Cys 184 from chain B

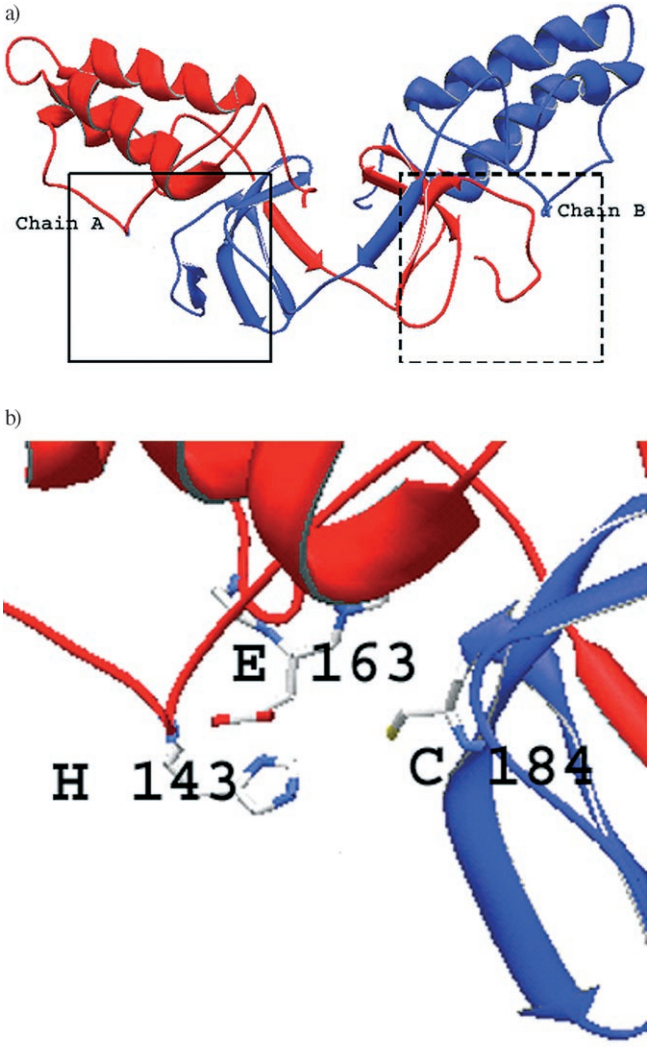

Figure 1. The active site of $NS2^{protease}$.[12] a) The locations of the two composite active sites in the $NS2^{protease}$ dimer (chain A in red, chain B in blue) are shown as boxed regions. The solid-lined box represents the active site displayed in part b. b) The $NS2^{protease}$ active site residues His 143 and Glu 163 (from chain A), and Cys 184 (from chain B) are shown in stick representation. Part of chain B is not shown to clarify the active site region because it overlaps with the active site residues.

[a] *Dr. Z. Suo*
Department of Biochemistry
The Ohio State University
484 W. 12th Ave., Biological Sciences Building
Columbus, OH 43210 (USA)
Fax: (+1) 614-292-6773
E-mail: suo.3@osu.edu

[b] *M. A. F. Abdullah*
Department of Entomology
University of Georgia
125 Cedar St., Athens, GA 30602 (USA)

(Figure 1b). Further biochemical characterization using active site mutants strongly supports these data. In addition, these three NS2^{protease} active site residues share a similar spatial distribution with the active sites from cysteine proteases such as papain[13] and poliovirus 3C protease.[14] Furthermore, the orientation of Cys 184 is similar to those of the catalytic serine residues of Sindbis virus capsid[15] and the cellular protease subtilisin.[16]

Although the crystal structure of NS2^{protease} is instructive, the approaches used in the studies of Lorenz et al.[12] raise some important questions. The truncated NS2 protein used for growing crystals and ultimately used to solve the NS2^{protease} structure is an inactive protease, based on previous studies in which the minimal region required for NS2-3 protease activity was determined.[9–11] The presence of the NS3 subdomain in the NS2-3 protease may change the structure of NS2^{protease} considerably. The biochemical studies carried out by Lorenz et al. using active site mutants (H143A and C184A)[12] seem to support the credibility of the NS2^{protease} structure. However, a previously published study suggests that the active site mutants cannot efficiently form dimers in comparison with the wild-type protease, thus rendering them inactive.[9] Therefore, do the active site mutations only affect the NS2-3 active site by removing critical amino acids involved in the protease catalytic triad (while still allowing proper NS2 dimer formation),[12] or, as previous results suggest, do these mutations affect the overall protein folding?[9] In addition, previous results that show cleavage of the C993A mutant (analogous to C184A) in trans by wild-type NS2-3[10] seem to contradict the hypothesis of Lorenz and co-workers[12] that the Cys→Ala mutation is supposed to prevent same-chain cleavage. Perhaps the in trans cleavage mode of action behaves very differently from the mode of action proposed by Lorenz et al.

Overall, the NS2^{protease} structure contributes greatly to the current knowledge of HCV biology and provides valuable information that might someday help control hepatitis C disease. The current standard treatment for chronic hepatitis C patients includes combination therapies of either PEG–interferon α2-a plus ribavirin or PEG–interferon α2-b plus ribavirin. However, these treatments do not work on all HCV-infected patients, so more drugs are needed. New protease-inhibitor-based drugs (SCH 503034 by Schering–Plough and VX-950 by Vertex Pharmaceuticals), which have shown very promising results in clinical trials, are under development. These inhibitors target the active site of the NS3-NS4A serine protease complex. Processing of the NS2/NS3 junction is thought to be essential for the normal infectious cycle of HCV, suggesting that the NS2-3 protease is an attractive drug target. A cell-based assay has thus been developed to screen for small-molecule inhibitors that target the NS2-3 protease.[17] The NS2^{protease} structure (Figure 1) will provide structural guidance for the development of inhibitors directed against the active site cleft of the NS2-3 protease and other conserved features of the protein.[12] The effectiveness, however, of NS2-3 protease inhibitors in halting HCV replication and infection remains to be shown.

Keywords: crystal structures · hepatitis C virus · NS2-3 protease · protease active site

[1] A. Ishihama, K. Nagata, *Crit. Rev. Biochem.* **1988**, *23*, 27–76.
[2] C. M. Rice, E. M. Lenches, S. R. Eddy, S. J. Shin, R. L. Sheets, J. H. Strauss, *Science* **1985**, *229*, 726–735.
[3] A. Takamizawa, C. Mori, L. Fuke, S. Manabe, S. Murakami, J. Fujita, E. Onishi, T. Andoh, I. Yoshida, H. Okayama, *J. Virol.* **1991**, *65*, 1105–1113.
[4] A. Grakoui, D. W. McCourt, C. Wychowski, S. M. Feinstone, C. M. Rice, *J. Virol.* **1993**, *67*, 2832–2843.
[5] C. Lin, B. D. Lindenbach, B. M. Prágai, D. W. McCourt, C. M. Rice, *J. Virol.* **1994**, *68*, 5063–5073.
[6] H. Mizushima, M. Hijikata, S. Asabe, M. Hirota, K. Kimura, K. Shimotohno, *J. Virol.* **1994**, *68*, 6215–6222.
[7] L. Tomei, C. Failla, E. Santolini, R. D. Francesco, N. L. Monica, *J. Virol.* **1993**, *67*, 4017–4026.
[8] C. Failla, L. Tomei, R. De Francesco, *J. Virol.* **1994**, *68*, 3753–3760.
[9] M. Pallaoro, A. Lahm, G. Biasiol, M. Brunetti, C. Nardella, L. Orsatti, F. Bonelli, S. Orru, F. Narjes, C. Steinkuhler, *J. Virol.* **2001**, *75*, 9939–9946.
[10] A. Grakoui, D. W. McCourt, C. Wychowski, S. M. Feinstone, C. M. Rice, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10583–10587.
[11] M. Hijikata, H. Mizushima, T. Akagi, S. Mori, N. Kakiuchi, N. Kato, T. Tanaka, K. Kimura, K. Shimotohno, *J. Virol.* **1993**, *67*, 4665–4675.
[12] I. C. Lorenz, J. Marcotrigiano, T. G. Dentzer, C. M. Rice, *Nature* **2006**, *442*, 831–835.
[13] R. W. Pickersgill, G. W. Harris, E. Garman, *Acta Crystallogr. Sect. B* **1992**, *48*, 59–62.
[14] S. C. Mosimann, M. M. Cherney, S. Sia, S. Plotch, M. N. James, *J. Mol. Biol.* **1997**, *273*, 1032–1047.
[15] L. Tong, G. Wengler, M. G. Rossmann, *J. Mol. Biol.* **1993**, *230*, 228–247.
[16] C. A. McPhalen, M. N. James, *Biochemistry* **1988**, *27*, 6582–6598.
[17] M. Whitney, J. H. Stack, P. L. Darke, W. Zheng, J. Terzo, J. Inglese, B. Strulovici, L. C. Kuo, B. A. Pollok, *J. Biomol. Screening* **2002**, *7*, 149–154.

Received: December 8, 2006
Published online on February 1, 2007

