Virus Research 208 (2015) 56–65

Contents lists available at ScienceDirect

Virus Research

journal homepage: www.elsevier.com/locate/virusres

Prediction and biochemical analysis of putative cleavage sites of the 3C-like protease of Middle East respiratory syndrome coronavirus

Andong Wu[a], Yi Wang[a], Cong Zeng[a], Xingyu Huang[a], Shan Xu[a], Ceyang Su[a], Min Wang[b], Yu Chen[a,**], Deyin Guo[a,b,*]

[a] State Key Laboratory of Virology, College of Life Sciences, Wuhan University, Wuhan, PR China
[b] School of Basic Medical Sciences, Wuhan University, Wuhan, PR China

ARTICLE INFO

Article history:
Received 18 April 2015
Received in revised form 21 May 2015
Accepted 22 May 2015
Available online 31 May 2015

Keywords:
MERS-CoV
3C-like protease
Canonical cleavage sites
Non-canonical cleavage sites
Michaelis constants

ABSTRACT

Coronavirus 3C-like protease (3CLpro) is responsible for the cleavage of coronaviral polyprotein 1a/1ab (pp1a/1ab) to produce the mature non-structural proteins (nsps) of nsp4–16. The nsp5 of the newly emerging Middle East respiratory syndrome coronavirus (MERS-CoV) was identified as 3CLpro and its canonical cleavage sites (between nsps) were predicted based on sequence alignment, but the cleavability of these cleavage sites remains to be experimentally confirmed and putative non-canonical cleavage sites (inside one nsp) within the pp1a/1ab awaits further analysis. Here, we proposed a method for predicting coronaviral 3CLpro cleavage sites which balances the prediction accuracy and false positive outcomes. By applying this method to MERS-CoV, the 11 canonical cleavage sites were readily identified and verified by the biochemical assays. The Michaelis constant of the canonical cleavage sites of MERS-CoV showed that the substrate specificity of MERS-CoV 3CLpro is relatively conserved. Interestingly, nine putative non-canonical cleavage sites were predicted and three of them could be cleaved by MERS-CoV nsp5. These results pave the way for identification and functional characterization of new nsp products of coronaviruses.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Middle East respiratory syndrome coronavirus (MERS-CoV) is an enveloped virus carrying a genome of positive-sense RNA (+ssRNA). It was identified as the pathogen of a new viral respiratory disease outbreak in Saudi Arabia in June 2012, named as Middle East Respiratory Syndrome (MERS). MERS-CoV emerged ten years after severe acute respiratory syndrome coronavirus (SARS-CoV) (Zaki et al., 2012) and quickly spread to several countries in Middle East and Europe (Assiri et al., 2013; Tashani et al., 2014). Soon after the first report, the MERS-CoV genome was sequenced and its genomic organization has been elucidated (van Boheemen et al., 2012). This new coronavirus is classified in the lineage C of *beta coronavirus*, and is close to bat coronavirus HKU4 and HKU5 (de Groot et al., 2013; Lau et al., 2013). Like other coronaviruses (Hussain et al., 2005; Zuniga et al., 2004), MERS-CoV contains a 3′ coterminal, nested set of seven subgenomic RNAs (sgRNAs), enabling translation of at least nine open reading frames (ORFs). The 5′-terminal two thirds of MERS-CoV genome contains a large open reading frame ORF1ab, which encodes polyprotein 1a (pp1a, 4391 amino acids) and polyprotein 1ab (pp1ab, 7078 amino acids), the latter being translated *via* a −1 ribosomal frameshifting at the end of ORF1a. These two polyproteins were predicted to be subsequently processed into 16 non-structural proteins (nsps) by nsp3, a papain-like protease (PLpro), and nsp5, a 3C-like protease (3CLpro) (Kilianski et al., 2013; van Boheemen et al., 2012).

Protease plays a key role during virus life cycle. It is essential for viral replication by mediating the maturation of viral replicases and thus becomes the target of potential antiviral drugs (Thiel et al., 2003; Ziebuhr et al., 2000). Investigating the cleavage sites of coronavirus proteases and the processing of polyproteins pp1a/1ab will benefit to identify the viral proteins and their potential function for viral replication. Some cleavage sites have been identified and confirmed by previous studies, including three cleavage sites of PLpros of human coronavirus 229E (HCoV 229E), mouse hepatitis virus (MHV), SARS-CoV, MERS-CoV and infectious bronchitis virus (IBV), whose cleavages release the first 3 non-structural

* Corresponding author at: College of Life Sciences, Wuhan University, Wuhan 430072, PR China. Tel.: +86 27 68752506.
** Corresponding author at: College of Life Sciences, Wuhan University, Wuhan 430072, PR China. Tel.: +86 27 87884604.
E-mail addresses: chenyu@whu.edu.cn (Y. Chen), dguo@whu.edu.cn (D. Guo).

http://dx.doi.org/10.1016/j.virusres.2015.05.018
0168-1702/© 2015 Elsevier B.V. All rights reserved.

proteins (Bonilla et al., 1995; Kilianski et al., 2013; Lim and Liu, 1998; Ziebuhr et al., 2007). The canonical cleavage sites of 3CLpros, the sites between the recognized nsps, have also been characterized, including all sites of MHV, IBV, SARS-CoV and a fraction of sites of HCoV 229E which release the non-structural proteins from nsp4 to nsp16 (Deming et al., 2007; Grotzinger et al., 1996; Liu et al., 1994, 1997; Lu et al., 1995). For 3CLpro of MERS-CoV, two cleavage sites releasing nsp4 to nsp6 have been identified (Kilianski et al., 2013). However, other cleavage sites remain to be characterized.

Furthermore, efforts have been taken to predict these cleavages sites by sequence comparison. Gorbalenya et al. (1989) made the first systematical prediction on IBV pp1a/1ab according to the substrate specificity of 3C protease of picornaviruses. However, two of their predicted cleavage sites within nsp6 of IBV were proved uncleavable (Liu et al., 1997; Ng and Liu, 2000). Gao et al. (2003) developed a software (ZCURVE_CoV) to predict the nsps as well as gene-encoded ORFs of coronaviruses more accurately based on previous studies of 3CLpros cleavage sites of IBV, MHV and HCoV 229E. Later on, non-orthogonal decision trees were used to mine the coronavirus protease cleavage data and to improve the sensitivity and accuracy of prediction (Yang, 2005). However, while these methods focus on the prediction of the canonical cleavage sites and target more and more on prediction accuracy to avoid false positives, potential non-canonical cleavage sites might be neglected. For example, a cleavage site between nsp7 and 8 of MHV strain A59 is not predicted by above methods, but proved to be physiologically important since it produces a shorter nsp7 that can support the growth of MHV carrying a mutation on nsp7–8 cleavage site (Deming et al., 2007). Therefore, the substrate specificities of coronaviruses 3CLpros are complicated. A 3CLpro substrate library of four coronaviruses (HCoV-NL63, HCoV-OC43, SARS-CoV and IBV) containing 19 amino acids × 8 positions variants was constructed by making single amino acid (aa) substitution at each position from P5 to P3′, and their cleavage efficiencies were measured and analyzed to find out the most preferred residues at each position (Chuck et al., 2011). However, the non-canonical cleavage site with less preferred residues of 3CLpro is adopted by coronaviruses (Deming et al., 2007). Thus we speculate that other potential 3CLpro cleavage sites may still exist in coronaviruses.

In order to set up a more moderate and balanced criteria for protease cleavage site identification, we compared six scanning conditions with different stringency to systematically predict the 3CLpro cleavage sites on pp1a/1ab of five coronaviruses including MERS-CoV. As a representative, the cleavability of the predicted cleavage sites of MERS-CoV 3CLpro was analyzed by the recombinant luciferase cleavage assay and the fluorescence resonance energy transfer (FRET) assay. The results showed that all 11 canonical cleavage sites of MERS-CoV pp1a/1ab were cleavable in our experiments and three of nine predicted non-canonical cleavage sites appeared to be cleavable. Our study points out a new direction regarding the prediction and identification of cleavage sites of proteases and contributes to understanding the mechanism of coronaviral polyprotein processing.

2. Materials and methods

2.1. Information collection of coronavirus 3CLpro cleavage sites

The genome sequences of 28 coronaviruses were downloaded from Genebank database and the sequences of the 3CLpro cleavage sites were collected from P4 to P2′ (Tables S1–S4). The substrate profiles of each coronavirus group and the whole *Coronavirinae* were summarized (Table S5).

2.2. Construction of recombinant 3CLpro expression vectors

The coding sequence of MERS-CoV nsp5 (NC_019843) was synthesized chemically by GenScript and cloned into vectors pET28a and pGEX-6p-1, respectively. The catalytic residue mutation C148A was generated by over lapping PCR with mutagenic primers (Table S6). All the clones and mutations were confirmed by DNA sequencing.

2.3. Expression and purification of recombinant proteins

The expression vectors were transformed into *Escherichia coli* strain BL21 (DE3). The cells were grown at 37 °C in Lysogeny broth (LB) medium with antibiotics and induced with 0.2 mM isopropyl-D-thiogalactopyranoside (IPTG) at 16 °C for 12 h. The cells were harvested and resuspended in lysis buffer (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 0.05% NP40, 0.1 mg/ml lysozyme and 1 mM PMSF) at 4 °C. After incubation for 30 min on ice, 10 mM $MgCl_2$ and 10 μg/ml DNase I (Sigma) were added to digest the genomic DNA. The supernatant of cell lysate was applied to affinity chromatography column after centrifugation. The recombinant protein with His-tag was bound with nickel-nitrilotriacetic acid (Ni-NTA) resin (GenScript) and washed with buffer A (50 mM Tris–HCl, pH 7.5, 150 mM NaCl), buffer B (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 20 mM imidazole) and buffer C (50 mM Tris, pH 7.5, 150 mM NaCl, 50 mM imidazole). Proteins were eluted with buffer D (50 mM Tris, pH 7.5, 150 mM NaCl, 250 mM imidazole). GST-tagged protein was bound with GST resin (GenScript), washed with buffer A and eluted with buffer A supplemented with 10 mM reduced glutathione (GSH). The purified proteins were desalted and concentrated by ultrafiltration using 30 kDa amicon ultra 0.5-ml centrifugal filter (Millipore).

2.4. Luciferase-based biosensor assay

All the cleavage sites (eight residues, ranging from P5 to P3′) were inserted into Glo-Sensor 10F linear vector. Comparing to the wild type firefly luciferase (550 aa), Glo-Sensor luciferase has short truncations at both termini with C- and N-part reversed, resulting in the new 234-aa N- and 233-aa C-terminal region respectively. The inserted sequence and the reversed arrangement of the N- and C-terminal regions reduce the luciferase activity dramatically. After the recognition sequence was cut off by nsp5, the luciferase recover its activity and luminescence in the presence of luciferase substrate. A back to front recombinant firefly luciferase inserted with different cleavage sites was expressed when the recombinant plasmids were co-incubated with a cell-free protein expression system extracted from wheat germ (Promega). After incubation for 2 h at 25 °C, nsp5 was added into the system and the whole system was incubated at 30 °C for 1 h. Then, the reaction system was diluted 20 times and mixed thoroughly with equal volume of luciferase substrate. Luciferase luminescence was measured by a luminometer (Promega) after incubation for 5 min at room temperature.

2.5. Peptide-based FRET assay

All the 11 conserved putative recognition sites were designed from P12 to P8′, synthesized and modified with a typical shorter wavelength FRET pair, N-terminal DABCYL and C-terminal Glu-EDANS by GL Biochem (Shanghai). The peptides were completely dissolved in DMSO and the final concentration of DMSO in the reaction system was 1%. 180 μM substrate peptide and 16.3 μM tagged nsp5 were mixed in the solution of 50 mM Tris, pH 7.5, 1 mM EDTA, 50 μM DTT and incubated at 37 °C for 2 h. To calculate kcat/Km, different amounts (7.2–180 μM) of substrate

Download English Version:

https://daneshyari.com/en/article/3428114

Download Persian Version:

https://daneshyari.com/article/3428114

Daneshyari.com