International Journal of Infectious Diseases 41 (2015) 21–28

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

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

Secondary structure conservation of the stem-loop IV sub-domain of internal ribosomal entry sites in human rhinovirus clinical isolates

Hak Kim [a,b], Kisoon Kim [c], Taesoo Kwon [d], Dae-Won Kim [d], Sung Soon Kim [a], You-Jin Kim [a,*]

[a] Division of Respiratory Viruses, Center for Infectious Diseases, Korea National Institute of Health, Cheongju-si, Chungbuk, 363-951, Republic of Korea
[b] Department of Biotechnology, Catholic University of Korea, Bucheon-si, Gyeonggi-do, 420-743, Republic of Korea
[c] Division of Influenza Virus, Center for Infectious Diseases, Korea National Institute of Health, Cheongju-si, Chungbuk, 363-951, Republic of Korea
[d] Division of Biosafety Evaluation and Control, Korea National Institute of Health, Cheongju-si, Chungbuk, 363-951, Republic of Korea

ARTICLE INFO

Article history:
Received 7 May 2015
Received in revised form 17 October 2015
Accepted 21 October 2015

Corresponding Editor: Eskild Petersen, Aarhus, Denmark.

Keywords:
Human rhinovirus
Internal ribosomal entry site
Stem-loop IV
Compensatory substitution

SUMMARY

Objectives: The aim of this study was to investigate the genetic diversity in the stem-loop (SL) IV sub-domain of the human rhinovirus (HRV) internal ribosomal entry site (IRES), which plays key roles in the initiation of viral translation by host protein interaction.
Methods: The primary SL-IV sequences of 194 HRVs, consisting of 97 reference strains and 97 clinical isolates, including the IRES sub-domains SL-IVa, SL-IVb, SL-IVc, and SL-IVd, were analyzed using Lasergene, MEGA 4, and WebLogo. Additionally, secondary structures of SL-IV were predicted and classified by RNAfold and CentroidHomfold-LAST.
Results: The predicted secondary structures of SL-IV showed variations in the position of bulbs, size of the loop, and length of stems. SL-IVc had the most highly conserved nucleotide sequence, with structures classified into two groups by the location of the poly(C) loop. Of the SL-IV sequences analyzed, 74 (79.56%) were classified in the major group and 19 (20.44%) in the minor group. Thirteen compensatory substitution pairs of SL-IVc contributed to maintaining the stem structure.
Conclusions: This study showed that the IRES secondary structures of a large number of reference and clinical HRVs were highly conserved, with several compensatory substitutions. It is expected that these results will facilitate investigations into HRV function based on IRES secondary structures.

1. Introduction

Human rhinoviruses (HRVs), members of the family *Picornaviridae*, were first isolated in 1956. These viruses are 7200 nucleotides in length and have a single-stranded, positive-sense RNA genome. Recently, HRVs have been classified into HRV-A, HRV-B, and HRV-C species, with 150 types of this virus proposed.[1–6] HRVs are not only the most common causative agents of mild upper respiratory tract infections, but are also associated with more serious diseases, including pneumonia and acute wheezing episodes associated with bronchiolitis and acute asthma in children.[7,8]

At the canonical translation initiation site, eukaryotic mRNAs have a 5′ cap structure (methyl[7] Gppp) that binds the eukaryotic initiation factor eIF4E.[9–12] However, some mRNAs, including both cellular and viral, use a cap-independent translation initiation mechanism at the internal ribosome entry site (IRES).[9,11,13–16] The IRES was first identified in 1988 in poliovirus and encephalomyocarditis virus (EMCV), also members of the *Picornaviridae*.[13,16] Since then, other RNA viruses such as HRV, DNA viruses including Kaposi's sarcoma herpesvirus (KSHV), and cellular mRNAs have been discovered to use IRES elements for translation.[9,11,17] The picornavirus IRES is classified into four groups, types I–IV, by sequence similarity and shared secondary RNA structures. Each IRES type has three to seven stem-loops (SLs) in the 5′ non-coding region, and type I IRES, found in HRV and enterovirus, has six SLs and one poly-pyrimidine tract (PPT) located between SL-V and SL-VI.[15,18–20] Translation initiated at the picornavirus IRES is regulated by specific IRES-binding host proteins.[21] Upstream-of-Nras (Unr), a cytoplasmic RNA-binding protein, was identified as a synergistic activator that binds the SL-V and SL-VI region in HRV.[22] The PPT-binding protein (PTB) that binds to the PPT region and the poly(C) binding protein (PCBP) that interacts with SL-IV are also known to be involved in alternative splicing and translation.[9,23–25]

* Corresponding author. Tel.: +82 43 719 8223; fax: +82 43 719 8239.
E-mail address: youjin3693@gmail.com (Y.-J. Kim).

http://dx.doi.org/10.1016/j.ijid.2015.10.015

In this study, the detailed features of the HRV IRES-IV domain were investigated; these have essential roles in HRV translation by interacting with host proteins, thus the dimensional RNA structures are especially important for functions among the other IRES domains. The analysis of the sequences and predicted secondary structures of IRES SL-IV in 97 reference strains and 97 clinical isolates of HRV are reported here.

2. Materials and methods

2.1. Ethics statement

This study was approved by the Institutional Review Board of the Korea Centers for Disease Control and Prevention (KCDC; 2012-09CON-03-4C) and the Yonsei University Health System Institutional Review Board, Seoul, Korea (4-2008-0649). Patients who had parental or legal guardian written consent to participate were enrolled in a respiratory disease surveillance study. Participant data were de-identified, except for age, sex, time of sample collection, reported symptoms, and virus detection results, as described previously.[26]

2.2. Specimen collection and sequencing of the HRV IRES

Nasal aspirate specimens from outpatients with an acute respiratory tract infection (ARI, $n = 3082$) and nasopharyngeal aspirate specimens from inpatients with a severe lower respiratory tract infection (SLRI, $n = 381$) were collected from hospitals as part of two surveillance systems, the Acute Respiratory Infections Network (ARINET) and the SLRI surveillance system, in 2008–2009. The ARINET surveillance system covers about 100 hospitals all over Korea and patients of all ages. The SLRI surveillance system covers four general hospitals in metropolitan areas and children aged less than 5 years.[26,27] Among the HRV-positive samples, including 827 from ARI and 85 from SLRI cases, selected viral RNAs were amplified and sequenced using previously described methods.[26] In total, 44 ARI and 53 SLRI sequences were analyzed.

2.3. Nucleotide and phylogenetic analysis of SL-IV sequences

IRES SL-IV sequences (nt 235–437, based on HRV A41, accession number **DQ473491**, as shown in Figure 1) of the 97 clinical isolates and 97 reference HRVs were aligned, redundant sequences were removed, and nucleotide percentage identity was calculated using the ClustalW method in MegAlign (ver. 8.0.2(13.4)2) of the Lasergene 8 program suite (DNAStar, Madison, WI, USA). The phylogenetic tree of selected unique sequences of the 5′ IRES SL-IVs was estimated using the neighbor-joining method in MEGA 4 (ver. 4.0.2). Human enterovirus D68 (GenBank accession number **AY426531**) was used as an outgroup.[28] The pairwise identity analysis was performed in stand-alone BLAST+. Then, an $n \times n$ matrix of all pairwise identities was produced. The matrix was plotted as a heat map using the default settings for heatmap.2 in the gplots package of R (http://cran.r-project.org/web/packages/gplots/index.html).[29] The sequences were used to create a WebLogo (http://weblogo.threeplusone.com) to examine nucleotide conservation based on the height of a single letter nucleotide.[30]

2.4. Prediction of RNA secondary structures in SL-IV

The secondary structures of 5′ IRES SL-IV from clinical isolates and reference HRVs were predicted by RNAfold in the Vienna RNA package and by CentroidHomfold.[31–33] With RNAfold, the default option was used for dangling energies ('-d2'), and lonely pairs in the structure were disallowed with the '-noLP' option. The resulting secondary structures of a single sequence determined by RNAfold were saved as an mfe (minimum free energy) structure. To predict a single sequence secondary structure with CentroidHomfold, the user-specified RNA sequence option was used and the database was compiled in the pipeline.

2.5. Nucleotide sequence accession numbers

The genome sequences of human rhinoviruses obtained from clinical isolates that are described in this report have been

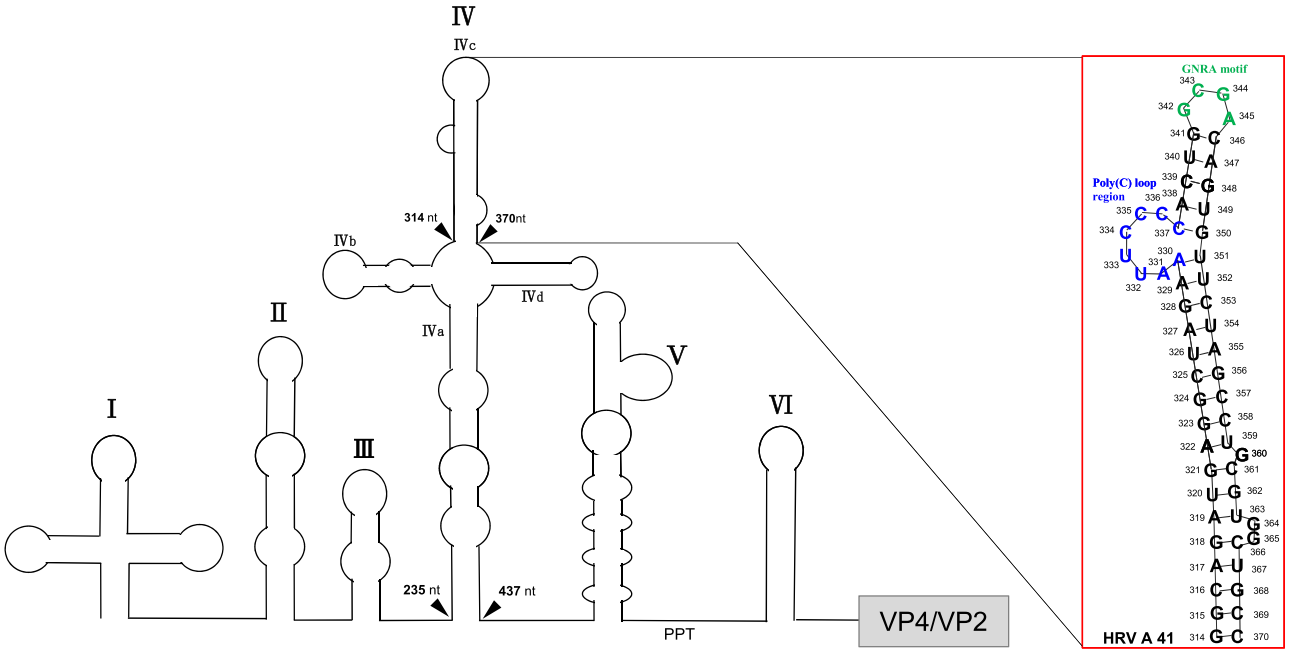

Figure 1. Schematic representation of the RNA secondary structure of human rhinovirus (HRV) A41 IRES SL-IVc. The stem-loop SL-IV of the HRV-A41 IRES was located at nt 235–437, and SL-IVc was located within the SL-IV domain (nt 314–370, based on accession number **DQ473491**). The prediction of RNA secondary structure was performed using RNAfold and CentroidHomfold-LAST.[39–41].