Virology 454-455 (2014) 206–214

Contents lists available at ScienceDirect

Virology

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

Identification and characterization of multiple conserved nuclear localization signals within adenovirus E1A

Kris S. Marshall [a], Michael J. Cohen [a,1], Greg J. Fonseca [a,1], Biljana Todorovic [a], Cason R. King [a], Ahmed F. Yousef [b], Zhiying Zhang [c], Joe S. Mymryk [a,d,*]

[a] Department of Microbiology & Immunology, Western University, London Regional Cancer Program, London, ON, Canada N6A 4L6
[b] Department of Chemical and Environmental Engineering, Masdar Institute, Abu Dhabi, United Arab Emirates
[c] College of Animal Science and Technologies, Northwest A&F University, Yangling, Shaanxi 712100, China
[d] Department of Oncology, Western University, London Regional Cancer Program, London, ON, Canada N6A 4L6

ARTICLE INFO

Article history:
Received 15 November 2013
Returned to author for revisions
18 December 2013
Accepted 19 February 2014
Available online 13 March 2014

Keywords:
E1A
Adenovirus
Nuclear import
Importin
Qip1
Rch1
NPI1

ABSTRACT

The human adenovirus 5 (HAdV-5) E1A protein has a well defined canonical nuclear localization signal (NLS) located at its C-terminus. We used a genetic assay in the yeast *Saccharomyces cerevisiae* to demonstrate that the canonical NLS is present and functional in the E1A proteins of each of the six HAdV species. This assay also detects a previously described non-canonical NLS within conserved region 3 and a novel active NLS within the N-terminal/conserved region 1 portion of HAdV-5 E1A. These activities were also present in the E1A proteins of each of the other five HAdV species. These results demonstrate that, despite substantial differences in primary sequence, HAdV E1A proteins are remarkably consistent in that they contain one canonical and two non-canonical NLSs. By utilizing independent mechanisms, these multiple NLSs ensure nuclear localization of E1A in the infected cell.

Introduction

The defining characteristic of eukaryotic cells is the separation of the intracellular space into membrane bound compartments, including the nucleus where nucleic acid synthesis and processing occurs. Transport systems exist that allow proteins to be imported into the nucleus from the cytoplasm, as well as the export of proteins and RNA to the cytoplasm. Nuclear import of many cellular and viral proteins is typically mediated by nuclear localization signals (NLS) that physically interact with soluble cytosolic receptor proteins (Macara, 2001). A canonical monopartite NLS contains a short single stretch of at least three basic amino acids (B) with a consensus sequence fitting B_4, P(B_3X), PXX(B_3X) or B_3(H/P), where P is proline, H is histidine, X is any amino acid and letters in parentheses can be in any order (Macara, 2001; Nakai and Horton, 1999). Alternatively, a canonical bipartite NLS contain two short stretches of basic amino acids separated by a non-conserved sequence (Macara, 2001).

Proteins containing canonical NLSs interact in the cytosol with the importin α family of NLS receptors (also known as karyopherin α). Subsequent heterodimerization of importin α with importin β (also known as karyopherin β) and interaction with components of the nuclear pore complex leads to translocation into the nucleus in a GTP dependent fashion (Macara, 2001). The yeast *Saccharomyces cerevisiae* expresses only one importin α, Srp1 (Enenkel et al., 1995). In contrast, there are multiple mammalian importin α proteins, which are divided into three subtypes, represented by Rch1 (human importin α1), Qip1 (human importin α3) and NPI1 (human importin α5) (Miyamoto et al., 1997). The total amount of importin α, as well as the relative content of each importin α subtype, varies among different cell types and during different stages of development (Poon and Jans, 2005). Each importin α isoform has distinct substrate specificities, which may confer a level of regulation to nuclear import (Macara, 2001). Although many thousands of cellular proteins are transported into the nucleus, very few NLSs have been characterized in detail. Existing work has been heavily biased towards canonical monopartite or bipartite signals. Recently, NLSs that are not particularly rich in basic amino acid residues have been identified that also interact with importin α. As these import signals do not

Abbreviations: HAdV, human adenovirus; NLS, nuclear localization signal; E1A, early region 1A; R, residue; DBD, DNA binding domain; AD, activation domain; CR, conserved region; MBP, maltose binding protein; ITT, inducible translocation trap

* Correspondence to: Rm A4-833, 800 Commissioners Road East, London, ON, Canada N6A 4L6. Tel.: +1 519 685 8600x53016; fax: +1 519 685 8616.
E-mail address: jmymryk@uwo.ca (J.S. Mymryk).

[1] These authors contributed equally.

http://dx.doi.org/10.1016/j.virol.2014.02.020
0042-6822

conform to either the monopartite or bipartite types described above, they are referred to as non-canonical. Many of these signals have been identified in viral proteins, such as the influenza virus NP protein, the Borna disease virus p10 protein, cytomegalovirus UL84 protein and the Varicella-Zoster virus ORF29 protein (Wolff et al., 2002; Lischka et al., 2003; Stallings and Silverstein, 2005; Cros et al., 2005).

There are over 50 types of human adenovirus (HAdV), divided into six species. HAdV-5, a species C virus, is the most thoroughly characterized. The first gene expressed during HAdV-5 infection is Early Region 1A (E1A). The largest E1A protein contains 289 amino acid residues (289R) and has four highly conserved regions (CR1-4; Fig. 1A) (Avvakumov et al., 2004). HAdV-5 E1A is localized to the nucleus and contains a well characterized monopartite NLS (KRPRP) located at its C-terminus (Lyons et al., 1987). This NLS mediates nuclear import *in vitro* and *in vivo*, shows a distinct preference for human Qip1 (importin α3) *in vitro* (Kohler et al., 2001), and is regulated by acetylation (Madison et al., 2002). However, not all the E1A proteins of the different types, specifically those encoded by species B viruses, contain this predicted classical NLS (Avvakumov et al., 2004). A second non-canonical NLS, with the consensus sequence $FV(X)_{7\text{-}26}MXSLXYM(X)_4MF$, spans residues 142–182 in CR3 of HAdV-5 E1A, and this sequence is unique to the species C E1A proteins (Slavicek et al., 1989; Standiford and Richter, 1992). This sequence does not resemble other known NLSs and when expressed in *Xenopus* embryos functions only during early neuronal stage, suggesting that it may be developmentally regulated (Standiford and Richter, 1992). At least one other non-canonical NLS may be present within the HAdV-5 E1A protein, as residues 23–120 have been reported to be sufficient to mediate nuclear accumulation in micro-injected *Xenopus laevis* oocytes (Richter et al., 1985).

We have utilized a simple and sensitive genetic method to identify and characterize novel NLSs in E1A. This system is based on the expression of the test protein fused to a modified LexA DNA binding domain (DBD) and the Gal4 transcriptional activation domain (AD) in the yeast *S. cerevisiae* (Marshall et al., 2007). The test protein is too large to passively diffuse into the nucleus and only when fused to a functional NLS will the chimera enter the nucleus and activate transcription of a LexA responsive β-galactosidase reporter gene. Furthermore, and the system is not limited to the identification of yeast NLSs, as the nuclear import apparatus is highly conserved between yeast and higher eukaryotic cells (Macara, 2001; Rhee et al., 2000).

Our previous work has shown that the yeast transcription based assay detects the canonical C-terminal NLS, the non-canonical NLS in CR3 and an N-terminal activity in HAdV-5 E1A (Marshall et al., 2007). In this study, we demonstrate the E1A proteins of all six HAdV species contain a functional canonical NLS at their C-terminus, a non-canonical NLS within CR3 and a third non-canonical NLS within the N-terminal/CR1 region. We also show that the N-terminal/CR1 NLS interacts indirectly with importin α and functions in mammalian nuclear import assays. Thus, despite substantial differences in primary sequence, HAdV E1A proteins are remarkably consistent in that they contain multiple NLSs. This redundancy may ensure nuclear localization of the viral protein during infection regardless of the importin α expression profile.

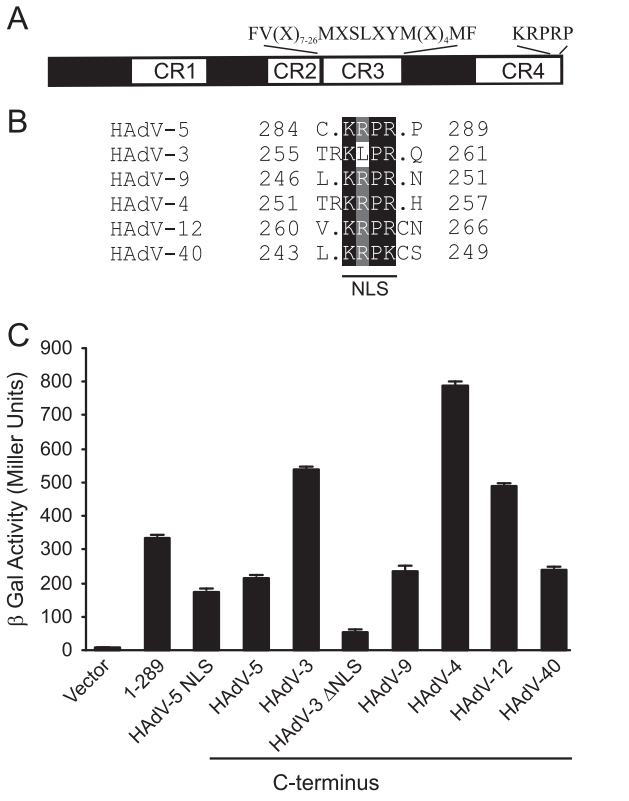

Fig. 1. Map of E1A and analysis of nuclear import mediated by the C-terminus of E1A from representative HAdV types. (A) Map of the 289R E1A protein showing the location of the conserved regions (CR1-4) and known nuclear localization signals. (B) Alignment of the C-terminal E1A NLSs from representative HAdV types used in this study (Avvakumov et al., 2004). (C) Absolute values of ß-galactosidase reporter activity indicating nuclear import in yeast for the C-terminal E1A regions from the indicated HAdV types (means ± SD, $n=3$). Vector—pNIA-CEN-MBP.

Results

The canonical NLS located at the C-terminus of E1A is functional in all HAdV species

We have previously described an improved transcription based assay in yeast for identifying signals that direct nuclear import of a fusion protein that exceeds the diffusion limit of the nuclear pore (Marshall et al., 2007). In this assay, fusion of a functional NLS to the large cytoplasmically localized transcription factor induces the chimera to enter the nucleus and activate transcription of a LexA responsive β-galactosidase reporter gene. This system detects the canonical C-terminal NLS in HAdV-5 E1A (Marshall et al., 2007). Although the E1A proteins from all HAdV species contain a short stretch of basic amino acid residues near their C-terminus, the E1A proteins from many species B viruses, including HAdV-3, do not have a predicted canonical NLS (Fig. 1B). To determine if each of these regions could direct nuclear import, we tested the C-terminal portions of each of the HAdV-3, 4, 5, 9, 12 and 40 E1A proteins using the yeast import assay (Fig. 1C). Expression of the HAdV-5 E1A C-terminus (residues 187–289), which contains the canonical NLS, or just the NLS (residues 282–289) was sufficient to direct nuclear import, although at a reduced level as compared to the full length HAdV-5 E1A. Similarly, the C-terminal fragments from types representative of each other species induced nuclear import as well, or better than HAdV-5 C-terminus (Fig. 1C). Interestingly, the C-terminal fragment of HAdV-3 E1A displayed roughly double the import activity of the HAdV-5 E1A fragment, despite the fact that the RKLPR sequence it contains does not fit that of a canonical NLS. Deletion of the RKLPR sequence (HAdV-3 ΔNLS) greatly reduces import activity conferred by the HAdV-3 C-terminal fragment, suggesting that it does indeed function as an NLS. These results demonstrate that a common feature of all the E1A proteins tested is the presence of a functional canonical, or near canonical, NLS located at their C-terminus.