Life Sciences 149 (2016) 79–95

Contents lists available at ScienceDirect

Life Sciences

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

Review article

Tip110: Physical properties, primary structure, and biological functions

Amanda Whitmill, Khalid Amine Timani, Ying Liu, Johnny J. He *

Department of Cell Biology and Immunology, Graduate School of Biomedical Sciences, University of North Texas Health Science Center, Fort Worth, TX 76107, United States

ARTICLE INFO

Article history:
Received 26 December 2015
Received in revised form 11 February 2016
Accepted 15 February 2016
Available online 16 February 2016

Keywords:
Tip110
Pre-mRNA splicing
Gene transcription
HIV replication
Development
Stem cell biology
Post-translational modifications

ABSTRACT

HIV-1 Tat-interacting protein of 110 kDa (Tip110), also referred to as squamous cell carcinoma antigen recognized by T cells 3 (Sart3), p110 or $p110^{nrb}$, was initially identified as a cDNA clone (KIAA0156) without annotated functions. Over the past twenty years, several functions have been attributed to this protein. The proposed biological functions include roles for Tip110 in pre-mRNA splicing, gene transcription, stem cell biology, and development. Dysregulation of Tip110 is also a contributing factor in the development of cancer and other human diseases. It is clear that our understanding of this protein is rapidly evolving. In this review, we aimed to provide a summary of all the existing literature on this gene/protein and its proposed biological functions.

Contents

* Corresponding author at: Department of Cell Biology and Immunology, Graduate School of Biomedical Sciences, University of North Texas Health Science Center, 3500 Camp Bowie Blvd., Fort Worth, TX 76107, United States.
E-mail address: johnny.he@unthsc.edu (J.J. He).

http://dx.doi.org/10.1016/j.lfs.2016.02.062
0024-3205/

1. Introduction

HIV-1 Tat interacting protein of 110 kDa (Tip110), also referred to as squamous cell carcinoma antigen recognized by T cells 3 (Sart3), p110 or p110^{nrb}, was first discovered as KIAA0156 without any specific functions [1]. The breadth of information that has been published concerning the distinct functions of Tip110 clearly indicate that the ubiquitously expressed nuclear RNA binding protein is necessary for a variety of biological processes including but not limited to pre-mRNA splicing [2–15], regulation of viral and host gene activation and transcription [16–20], regulation of protein degradation [12,20,21], and regulation of cell survival, proliferation and differentiation [8,22,23]. Tip110 has also been extensively studied as a potential antigen for cancer immunotherapy [5,24–31] and has been shown to have roles in skin diseases including disseminated superficial actinic porokeratosis (DSAP) [32–34] and atopic dermatitis (AD) [35,36].

In this review, we will focus on two aspects of Tip110. For the first part, we will discuss the physical properties, the primary structure, and characteristics of Tip110. Specifically, we will explore the possible contribution of post-translational modifications to the apparent differences in the molecular weight between two closely related species: human Tip110 and mouse Tip110. For the second part, we will discuss all the proposed biological functions of Tip110 and their respective molecular mechanisms. By connecting the information in the literature concerning this protein and consolidating what is known about its various functions, we hope to contribute to bettering the understanding of all the functions of Tip110 and to encouraging further investigation into what other biological roles it may have.

2. Structure and characteristics of Tip110

2.1. Physical properties and conserved domains

Tip110 is a 963-amino acid, nuclear RNA-binding protein. This protein has a predicted molecular weight of 110 kDa and a pI of 5.28. The amino terminal portion of Tip110 is more acidic than the carboxyl terminus in human, mouse, and *Caenorhabditis elegans* homologs [37]. Specifically, amino acids 1–321 have a pI of 4.30, amino acids 322–642 have a pI of 6.04, and amino acids 643–963 have a pI of 9.77 in human Tip110 [37]. This particular characteristic distinguishes Tip110 from other ribonucleoprotein (RNP) motif-containing proteins which are not typically acidic and suggests that Tip110 may be a member of a distinct subfamily of RNP proteins [37].

Several conserved domains and putative motifs have been identified within the human Tip110 primary amino acid sequence. The first motif can be found within the amino terminus of the protein and is the nuclear receptor box (NR box), characterized by the sequence LXXLL, where X is any amino acid. The NR box of Tip110 (LIRLL) spans amino acids 118–122 and is an important contributor to the coregulation of nuclear steroid hormone receptors (Figs. 1 & 2A) [16,38–40]. Specifically for Tip110, the NR box confers the ability to negatively regulate the androgen receptor [16]. Next there is a conserved domain, COG5107, also known as the RNA14 motif [41]. This particular domain is important for pre-mRNA 3′-end processing and modification such as cleavage and polyadenylation (Liu et al., unpublished data). There are 7 tetracopeptide repeats (TPR) also known as half-a-TPR (HAT) within and following the RNA14 motif that are necessary for many of the protein-protein interactions involving Tip110 and for *in vitro* splicing activity [2,9,14,19,42]. When this particular domain is deleted using mutagenesis, specific functions of Tip110 are abolished or diminished. A tyrosine phosphorylation kinase site exists at amino acids 309–316. This site is purported to be phosphorylated on the tyrosine at position 316 and as such may play a role in the metabolism of nuclear RNA due in part to the role of Tip110 as an RNA binding protein [4,31]. Unfortunately the phosphorylation state of Tip110 has not been concisely confirmed to be important for any of its cellular functions, with some disputing the functional significance of the tyrosine phosphorylation site altogether [43]. Tip110 is able to function as a nuclear protein through two nuclear localization sequences (NLS), one spanning amino acids 612–619 and the other spanning amino acids 642–646 [4,31,37]. Two RNA recognition motifs (RRM1 and RRM2) can be found near the carboxyl terminus of the protein with one spanning amino acids 705–778 and the other spanning amino acids 799–879 [4,31,37]. The RRM itself consists of two highly conserved stretches of amino acids separated by ~30 amino acids. The first stretch of amino acids are a hydrophobic cement with 6 residues (RNP2) and the second stretch of amino acids are an octapeptide motif (RNP1) [37]. The RRM are important for Tip110 to recognize and bind RNA (*e.g.* U6 snRNA) and for *in vitro* splicing activity. Additional motifs identified within the Tip110 protein sequence include the RGD cell attachment sequence at amino acids 742–744, which contributes to cellular adhesion and possibly caspase-3-mediated apoptosis [31], and the CT10 (C10 or LSm interaction motif) domain spanning the amino acids at the C-terminal end of the protein. The CT10 domain is necessary for Tip110 interaction with LSm proteins [7,13,14]. More concise details concerning the role of these domains will be detailed in ensuing sections.

2.2. Expression and cellular localization

Tip110 is expressed ubiquitously at the mRNA level and can be detected in both normal and malignant cells and tissues. Expression is highest in cells and tissues with cancerous phenotypes. Northern hybridization for mRNA expression patterns in human tissues and cell lines reveal that there is notable expression in a wide variety of cell lines (normal and malignant) and clear expression in the heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, testis, ovary, prostate, small intestine, colon, and in peripheral blood leukocytes [1,31,37]. Regardless of its ubiquitous expression at the mRNA level, the Tip110 protein can only be found in the nucleus of malignant tumor cell lines, cancerous tissues, and the testis [4,31,44]. Tip110 can be found in the cytoplasm of proliferating cells (both normal and malignant), the testis, and in the fetal liver [4]. Within the nucleus the protein is dispersed rather evenly throughout the nucleoplasm with distinct exclusion from nucleoli and prominent expression in Cajal bodies [10,13,14]. The expression in Cajal bodies is associated with its role as a recycling factor in pre-mRNA splicing and has contributed to bettering our understanding of the assembly of spliceosome components in these cellular compartments as detailed in ensuing sections.

There is likely a specific demand for Tip110 in the nucleus of cells with a high metabolic demand (*i.e.* cancerous cells and tissues). This localization may be required for Tip110 to readily bind nuclear RNA, promote transcription of various oncogenes, interact with other nuclear