Published in final edited form as:

Gene. 2012 November 10; 509(2): 215–222. doi:10.1016/j.gene.2012.08.025.

Identification, basic characterization and evolutionary analysis of differentially spliced mRNA isoforms of human YAP1 gene

Christian J Gaffney^a, Tsutomu Oka^a, Virginia Mazack^a, Dror Hilman^b, Uri Gat^b, Tomoki Muramatsu^c, Johji Inazawa^c, Alicia Golden^d, David J Carey^d, Amjad Farooq^e, Gerard Tromp^a, and Marius Sudol^{a,f,*}

^aLaboratory of Signal Transduction and Proteomic Profiling, Weis Center for Research, Geisinger Clinic, 100 North Academy Avenue, Danville, PA 17821-2608, USA

^bDepartment of Cell and Developmental Biology, The Alexander Silberman Institute of Life Sciences, Edmond J Safra Campus, Givat Ram, The Hebrew University, Jerusalem 91904, Israel

^cDepartment of Molecular Cytogenetics, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan

^dLaboratory of Extracellular Matrix, Weis Center for Research, Geisinger Clinic, 100 North Academy Avenue, Danville, PA 17821-2608, USA

^eDept of Biochemistry & Molecular Biology and USylvester Braman Family Breast Cancer Institute Leonard Miller School of Medicine, University of Miami, Miami, FL 33136

^fDepartment of Medicine, Mount Sinai School of Medicine, One Gustave Levy Place, New York, NY 10029-6574, USA

Abstract

The YAP1 gene encodes a potent new oncogene and stem cell factor. However, in some cancers, the YAP1 gene plays a role of tumor suppressor. At present, the gene and its products are intensely studied and its cDNAs are used as transgenes in cellular and animal models. Here, we report 4 new potential mRNA splicing isoforms of the YAP1 gene, bringing the total number of isoforms to 8. We detected all 8 YAP1 isoforms in a panel of human tissues and evaluated the expression of the longest isoform of YAP1 (YAP1-28) using Real Time PCR. All YAP1 isoforms are barely detectable in human leukocytes compared to fair levels of expression found in other human tissues. We analyzed the structure of the genomic region that gave rise to alternatively spliced YAP1 transcripts in different metazoans. We found that YAP1 isoforms, which utilize exon 6 emerged in evolution with the appearance of amniotes. Interestingly, 6 YAP1 isoforms, which contain the exon 5 extension, exon 6 or both would have their leucine zipper region disrupted in the predicted protein product, compared to the intact leucine zipper found in two YAP1 (α) isoforms. This observation has direct functional ramifications for YAP1 signaling. We also propose a normalized nomenclature for the mRNA splice variants of YAP1 gene, which should aid in the characterization of signaling differences among the potential protein products of the YAP1 gene.

Keywords

Alternative splicing; WW domains; Leucine Zipper; Quantitative RT-PCR

*Corresponding author. Laboratory of Signal Transduction and Proteomic Profiling, Weis Center for Research, Geisinger Clinic, M.C. 26-08, 100 North Academy Avenue, Danville, PA 17822-2608, Tel, +1(570) 271-6677, Fax, +1(570) 271-6701, msudol1@geisinger.edu (M. Sudol).

1. Introduction

The YAP1 (Yes associated protein 1) gene encodes a potent oncogenic protein and is one of the two main effectors of the Hippo tumor suppressor pathway (Sudol, 1994; Pan, 2010). However, in some cancers YAP1 plays a role of tumor suppressor (reviewed by Bertini et al., 2009). Originally, the YAP1 cDNA was isolated by screening expression libraries for proteins that associate with Yes and Src protein-tyrosine kinases (Sudol et al., 1995). Soon after, the YAP1 protein was shown to act as a transcriptional co-activator that shuttles between the cytoplasm and the nucleus in a cell density-dependent manner (Komuro et al., 2003; Yagi et al., 1999; Zhao et al., 2008).

The cloning of YAP1 cDNA facilitated the identification of one of the smallest modular protein domains, known as the WW domain (Bork and Sudol, 1994). The WW domain was identified as an imperfect repeat of 38 amino acids present in a differentially spliced isoform of YAP1, named YAP1-2 (Figure 1) (Sudol et al., 1995). The 38 amino acid long WW domain was added by a splicing event to the shorter isoform, named YAP1-1, that already had a single copy of the related sequence. The WW domain mediates protein complexes with partners that contain a PPxY motif (where P is Proline, Y is Tyrosine and x is any amino acid) and these complexes are unusually prevalent among components of the Hippo tumor suppressor pathway (Chen and Sudol, 1995; Sudol and Harvey, 2010). While WW domains are located at the amino terminal half of the YAP1 protein, its carboxy-terminal half harbors transcription activation domain (TAD) (Yagi et al., 1999) within which resides a frank coiled coil domain (Lutenic et al., 2012).

The human YAP1 gene (NCBI GENE ID, 10413) is located on chromosome 11 at position q22 and the gene is frequently amplified in several human cancers including liver, breast, prostate and esophageal cancer (Sudol et al., 1995; Harvey and Tapon, 2007; Pan 2010).

Over the past six years, the number of reports on the YAP1 gene grew exponentially. Several YAP1 cDNA clones, which we and others have isolated, showed that apart from the presence or absence of the second WW domain (Sudol et al., 1995), YAP1 sequences varied in terms of inserted sequences at a position located within the TAD region (transcriptional activation domain) (Figure 1) (e.g., Komuro et al., 2003; Hao et al., 2008). Since many studies have reported the use of a single cDNA of YAP1 in functional over-expression studies, but the YAP1 cDNAs used differed from each other either by the number of WW domains or by the variations within the TAD domain (e.g., Komuro et al., 2003; Oka et al., 2008; Hao et al., 2008; Muramatsu et al., 2011), we decided to define the spectrum of isoforms of the human YAP1 gene and propose a standardized nomenclature for their reference. In our study, we identified four additional, potential isoforms of human YAP1, which differ from the four isoforms reported recently (Muramatsu et al., 2011), proposing the existence of eight isoforms of YAP1 in total. We also investigated the exon-intron structure of the human YAP1 gene in detail and quantitatively measured the relative level of expression of the longest isoform of YAP1 (YAP1-2 δ) in human tissues, using real time PCR. In addition, we analyzed the evolutionary changes in the structure of the metazoan YAP1 orthologs (Hilman and Gat, 2011), with emphasis on exons 4, 5 and 6, the splicing of which generates the different isoforms of YAP1. We also suggest that all isoforms generated by additional sequences within the transcriptional activation domain would have their putative leucine zipper region disrupted in the predicted protein product, compared to the intact leucine zipper found in two YAP1 (α) isoforms.

2. Materials and methods

2.1 Panel of human tissues

First strand cDNA preparations from polyA⁺ mRNA, that was free of genomic DNA, and isolated from various human tissues (Human MTC Panels 1 and 2; Clontech, Mountain View, CA) were normalized for concentration against the following genes: α -tubulin, β -actin, G3PDH and phospholipase A2 and used as templates for PCR. cDNA for YAP1-1 and YAP1-2 were prepared from mRNAs isolated from HEK293 cells and these served as positive controls. Primers synthesized against the human YAP sequence were manufactured at Integrated DNA Technologies, Coralville, IA. The sequence of the Forward Primer 1 was: 5' AGAAGAGGTACCATGCTGTCCCAGATGAACGTCACA 3'. The sequence of the Reverse Primer 1 was: 5' ACAACATCTAGAATCCCGGGAGAAGACACTGGATTT 3'. Amplification was done using AccuPrime Taq DNA Polymerase System, (Invitrogen, Carlsbad, CA), following the protocol for cDNA using the following PCR program, denaturation at 94°C for 2 min, followed by 50 cycles of denaturation at 94°C for 15 s, annealing at 60°C for 30 s and elongation at 68 degrees for 30 s. This was followed by an extension at 68° for 2min. Samples were then incubated for 45 min at 72°C. PCR products were resolved by electrophoresis on a 3% gene composed of a mixture of agarose and Synergel clarifier gel (Diversified Biotech, Boston MA.), prepared according to the manufacturer's directions. A 50bp DNA Step Ladder (Promega, Madison, WI) was used as a molecular size standard.

2.2. YAP1-1 and YAP1-2 isoform analysis

The first strand human pancreas cDNA preparation was used as a template for amplification of YAP1 isoforms. The amplified product was ligated into the pCR 2.1-TOPO vector and transformed into chemically competent *E. coli* (TOPO TA Cloning, Invitrogen, Carlsbad, CA), with an incubation time of 30 minutes at room temperature. The pCR 2.1 TOPO library was then transformed into 'One Shot Top 10' competent cells via the One Shot Chemical Transformation Protocol, as outlined by the manufacturer, and plated on beta-galactose indicator plates. White colonies were selected from plates and plasmid DNA was prepared from each by use of the QIAprep Spin Miniprep Kit (Qiagen, Maryland). Clones were then digested with EcoR1 (Promega, Madison, WI) and resolved on a 3% agarose-Synergel, along with a 50bp DNA Step Ladder (Promega, Madison, WI), YAP1 and YAP2 positive controls isolated from HEK293 cells, and human pancreas amplified as stated above. Inserts showing slight size differences were chosen and the 4 YAP1-1 and 4 YAP1-2 isoform sequences were confirmed by direct sequence analysis at Genewiz, (South Plainfield, NJ).

2.3. Expression of YAP1-2 δ isoform in human tissues

YAP1-2- δ TaqMan primers (Applied Biosystems, Foster City, CA) were as follows, Forward, 6FAMAAGAACAAGACCACCTCTTGGCTAGACCCAAMGBNFQ and Reverse, VICGATTGCCTGTGGCC TCACCTMGBNFQ. The forward primer was located within the second WW domain of YAP, whereas the reverse primer corresponded to the sequences in exons 5A, 5B (also labeled as 5 extension) and 6, with the DNA corresponding to the coding region for QVRPQAMR (see Figure 2B). Quantitative Real-Time PCR was performed using the Applied Biosystems 7500 Fast Real-Time PCR System. Our TaqMan gene expression assay was custom designed to measure the expression level of the YAP1 isoform in 16 human tissue types. Comparative Ct quantitation experiments were run under standard cycling conditions using 10ul reactions containing 3ul of cDNA. GAPDH was used as the endogenous control. All reactions were run in triplicate.

2.4. Evolution of YAP1 gene

YAP1 and its paralog TAZ (WWTR1) were queried in the ENTREZ GENE database using the web-interface. For non-bilaterians, sequences were first verified with blastp against Refseq database. The sequences of the exons relevant to the different human isoforms were visualized on the genomic locus of different metazoans using the ENTREZ GENE graphic interface. Exon positions were manually assigned and corrected when needed.

2.5. Sequence and structural analysis of leucine zippers of YAP1 isoforms

The amino acid sequence alignment of the putative leucine zippers of various YAP1 isoforms was performed using ClustalW (Thompson et al., 1994). Structural models of putative leucine zippers of YAP1-1 α isoform were built using the MODELLER software based on homology modeling (Marti-Renom et al., 2000). Briefly, the solution structure of the leucine zippers of Jun transcription factor (PDB# 1JUN) was used as a template. It should be noted that the putative leucine zippers of YAP1-1 α share greater than 25% sequence identity with those from Jun. Importantly, the putative leucine zippers of YAP1-1 α contain a total of five highly-conserved signature leucine residues, designated herein L1-L5 in one chain and L1'-L5' in the other chain, at every seventh position in a manner akin to the leucine zippers of Jun. A total of 100 atomic models were calculated and the structure with the lowest energy, as judged by the MODELLER Objective Function, was selected for rendering using RIBBONS (Carson, 1991).

3. Results

3.1. Identification of YAP1 splice isoforms

Forward and reverse primers of human YAP1 were designed to identify not only both YAP1-1 and YAP1-2 isoforms, the latter of which differs from the former by the presence of an additional WW domain (Sudol et al., 1995), but also to identify suspected splicing isoforms at amino acid position 290 for YAP1-1 and 328 for YAP1-2. Surprisingly, gel analysis of RT-PCR products from the amplification reaction of human pancreas cDNA, which was performed under stringent conditions, revealed eight products, which we provisionally named as YAP1-1 α , β , γ , δ and YAP1-2 α , β , γ , δ (Figure 1). In a pilot experiment, the pancreas cDNA gave us the best yields and quality of the resolved bands, therefore we continued to use this cDNA as template for PCR amplifications, cloning and sequence analysis. The novelty of our data is that four additional isoforms of YAP, namely YAP1-1 α and γ plus YAP1-2 β and δ , were detected in addition to four other isoforms that were recently reported in normal esophagus and in an esophageal cancer cell line (Muramatsu et al., 2011). The PCR amplified DNA fragments were subcloned into pCR2.1 TOPO vector and submitted to direct DNA sequence analysis. The DNA sequences and the predicted amino acid sequences are in Supplementary Figure 1 and the schematic of the YAP1 isoforms is presented in Figure 2A and B.

Partial cDNA sequences of all four new YAP1 isoforms plus one isoform that was reported previously (Muramatsu et al., 2011) but was not deposited in the sequence data banks were submitted to the EMBL-EBI Nucleotide Sequence Database and assigned the following accession numbers: HE864159, HE864160, HE864161, HE864162, HE864163.

3.2. Genomic structure of YAP1 locus

The structure of the human YAP1 gene comprises 9 exons (Figure 3). Interestingly, the first WW domain resides within exon 2 and 3, which are flanked by the longest intron. The second WW domain is located within a single exon, exon 4. Exon 5 contains a short extension, which we named 5B*. More precisely, an alternative splice donor site 12 nucleotides into intron 5 generates some alternative transcripts containing the amino acids

VRPQ. Note two 5' splice donors in intron 5, (5A, CAGGTGAGG) (5B*, CAGGTTAGA). The 5A splice donor is fairly canonical, only the last G appears less often than T. In the 5B* splice donor, the first 5 nucleotides are conserved, including the obligatory GT, the 6th nucleotide is a T (least preferred), the 7th and 8th nucleotides are conserved, and the 9th nucleotide is the least preferred base. From the presence of least preferred bases in two sites, the 5B* could be a weaker 5' splicing site than 5A.

The exon 4, (which encodes the second WW domain), the exon 5 extension, 5B*, (which codes for amino acids VRPQ), and the exon 6, (which codes for amino acids AMRNINPSTANSPKCCQ), are alternatively spliced to generate the 8 primary transcripts of YAP1 gene (Figures 1–3).

3.3. Detection of 8 transcripts in human organs using non-quantitative PCR

We next sought to determine the relative expression profiles of all 8 YAP1 isoforms across numerous tissues. Commercial preparations of cDNA generated from polyA⁺ mRNA from human organs, which were normalized for concentration against 4 genes (see Methods) were used as templates for PCR. We assayed cDNAs from colon, small intestine, leukocytes, spleen, kidney, liver, skeletal muscle, brain, placenta, lung, heart, ovary and testes. With the exception of leukocytes, all other tissues or organs expressed all known isoforms of YAP1 (Supp. Figures 2 and 3). We also noted that in most of the organs the two shortest isoforms, namely YAP1-1 α and β were expressed at significantly lower levels, compared to other isoforms (Supp. Figure 2.)

3.4. Expression of the YAP1-2 δ isoform using TaqMan primers and quantitative PCR

To provide an example of the quantitative expression profile of one of the isoforms of YAP1, we selected YAP1-2 δ isoform because of the possibility of designing PCR primers that would not amplify any other isoforms (Figure 2B). The forward primer was placed within the second WW domain, whereas the reverse primer was at the very junction of spliced transcripts that included exon 5B* extension and exon 6 (Figure 2B and Materials and Methods). Indeed, the amplification reaction resulted in a single DNA fragment of the correct size (not shown).

From the TaqMan-RT-PCR results, we calculated a fold change using as a reference the lowest expression level found in leukocytes (Figure 4). Ovary and placenta expressed the highest level of the YAP1-2 δ isoform, whereas skeletal muscle and brain expressed lower levels of the isoform, but still more than 2.5 fold higher level than leukocytes (Figure 4). Equal amounts of cDNAs from human organs, which were normalized for the expression of 4 house keeping genes (see Material and Methods) were used in the TaqMan based Real Time PCR, run in triplicate. Although the normalization process is reassuring, we do assume that the cDNA synthesis efficiency was similar for all the RNA templates isolated from different organs.

3.5. Evolution of the splicing isoforms of YAP1 gene

We decided to analyze the evolutionary changes in the structure of the genomic region that includes the exons 5 and 6 coding for the carboxy-terminal isoforms of YAP1. In the basal metazoan, sea anemone, (*Nematostella vectensis*) the *YAP1* gene is encoded by two exons only (Figure 5). In this animal, the exon 2 harbors the sequence equivalent to the human exon 5, ending with a codon for the residue Q, which is followed by the sequence equivalent to the start of the human exon 7, initiating with the codon for residue E. In the basal chordate, sea squirt, (*Ciona intestinalis*) the exon equivalent to human exon 5 ends in a similar sequence, with Q as the last residue, and is followed by the exon 7 equivalent that starts at a more carboxy-terminal sequence of YAP1 protein.

An extension of exon 5, shown in green (Figure 5) can be first found in the vertebrate, Zebrafish (*Danio rerio*) and in another fish, pufferfish (*Tetraodon nigroviridis*), in the form of a sequence coding for three additional amino acids (RPQ). The amphibian representative, the pipid frog (*Xenopus tropicalis*) demonstrates an extension of 4 amino acids (VRPQ), which are identical to those added to the human isoform β when compared to isoform α . Importantly, the addition of the codon for V (GTG) before the RPQ coding sequence creates a novel consensus site for an intron-donor start site, which gives rise to alternative splicing products, isoforms β and δ , which include the sequence we have termed “Exon 5-ext”, while in the other isoforms it is spliced out. We can assume that VRPQ-containing-isoforms can be found in most or all tetrapods.

The first ortholog of human exon 6, (marked in blue, Figure 5) was identified in the amniote, an arboreal lizard (*Anolis carolinensis*). When compared to humans, this lizard has a very similar exon 6-like genomic sequence coding for 16 amino acids, and remarkably conserved exon-intron donor and acceptor sites. The genomic structure of all further amniotes up to man is also highly conserved, and we can thus expect the existence of these isoforms in them as well. The evolutionary appearance of exon 6 may be due to adjustments of amniotes to life on land or to the formation of amniote egg and its specific structures.

Interestingly, the YAP1 paralog gene, TAZ (WWTR1), (Kanai et al., 1999) resembles the primordial state of YAP1 in this region and shows no sign of either one of the evolutionary novel exons (Figure 5).

In conclusion, we can see an elaboration of the YAP1 gene structure along evolution, which led to the appearance of these isoforms. The basal form of YAP1 has no introns 3' to the WW domains and lacks both the exon 5-extension and the exon 6-like sequences. In early chordates, an intron was first introduced into that region, allowing further complexity. We first see a possible extension of the primordial exon 5, which may have occurred in the early vertebrate lineage after the formation of the TAZ gene. This was followed by a formation of an alternative intron start site, which allowed the existence of the extended isoform as well as the primordial form in the same organism. The next stage, allowing even more complexity, is first detected in amniotes with the “birth” of a new exon, in what was previously intron 5. This event doubled the number of YAP1 isoforms that could be expressed.

It is worthy to note that the exon 5-ext (VRPQ) is not identified in the NCBI annotations of genomic sequences. The lack of this exon extension is due to the low representation of this sequence in the consensus CDS protein set (CCDS) database (CCDS ID CCDS44716.1). CCDS has only one partial representation of this extension in GenBank, AK304485.1 (*). In contrast, there are more than 10 better quality EST's that lack this extension. This can be seen in the evidence viewer for YAP1 in ENTREZ GENE, geneID, 10413 (**).

3.6. Sequence and structural analysis of coiled coil region of YAP1 isoforms

To further understand the differences between the various isoforms of YAP1 identified here and how these differences may affect their biological function, we next performed detailed sequence and structural analysis (Figure 6). Our amino acid sequence alignment suggests that the various isoforms of YAP1 differ by the insertion of between 4–20 residues within the putative leucine zipper located within the TAD domain (Figure 6a). It is noteworthy that the leucine zipper is a highly conserved protein module involved in mediating protein-protein interactions pertinent to a plethora of cellular activities (Landschulz et al., 1988).

* (url, <http://www.ncbi.nlm.nih.gov/nucleotide/AK304485>).

** (url, http://www.ncbi.nlm.nih.gov/sutils/evv.cgi?taxid=9606&contig=NT_033899.8&gene=YAP1)

The leucine zipper is characterized by the presence of a signature leucine at every seventh position within the minimum four successive heptads of amino acid residues. Importantly, leucine zippers are thermodynamically unstable in isolation, but adopt continuous amphipathic α -helices in the context of a dimer, either through homo-association or hetero-association with other cellular partners, by virtue of their ability to wrap around each other into a parallel dimeric coiled coil with a slight left-handed twist (O'Shea et al., 1989a, b, 1992). Ironically, while the YAP1 α isoforms fulfill the above criteria of a stretch of amino acid sequence capable of generating natively folded leucine zippers (Figure 6b) primarily through inter-digitation of signature leucines within each chain, the insertion of amino acids (4–20 residues) between the third and fourth heptad in YAP1 β , YAP1 γ and YAP1 δ isoforms would likely disrupt the amino acid sequence requirement of leucine zippers, thereby leading to their thermodynamic destabilization. Accordingly, the inability of YAP1 β , YAP1 γ and YAP1 δ isoforms to adopt leucine zippers may have serious consequences on protein-protein interactions required for their hetero-association with other cellular partners. In short, our sequence and structural analysis suggests that the YAP1 β , YAP1 γ and YAP1 δ isoforms would not be able to participate in a wide variety of protein-protein interactions driven by the leucine zippers of YAP1 α isoforms.

Interestingly in TAZ protein, the YAP1 paralog, in the region that corresponds to the leucine zipper of YAP1 neither a coiled coil region could be detected (Lupas et al., 1991; Lutenic et al., 2012) nor TAZ gene undergoes splicing in this region (Figure 5).

4. Discussion

We identified 4 novel potential mRNA splicing isoforms of the human YAP1 gene, bringing the total number of isoforms to 8 (Sudol et al., 1995; Komuro et al., 2003; Muramatsu et al., 2011). We investigated the exon-intron structure of the human YAP1 gene and showed that differential use of exon 4, exon 5 extension and exon 6 gives rise to new isoforms of YAP1. We evaluated the relative levels of expression of the longest isoform of YAP1 (YAP1-26) in human tissues, using Real Time PCR, and showed that it is widely expressed, except in leukocytes. We also traced the evolutionary appearance of exon 6 to amniotes. Detailed analysis of the predicted protein region where the splicing of exon 5 extension and exon 6 occur, indicates that a leucine zipper would be disrupted in these isoforms compared to the unspliced α isoforms of YAP1. Finally, we proposed a simple nomenclature to distinguish the eight isoforms of YAP1 unequivocally.

The following aspects of this report deserve further comment: (i) the existence of the actual protein products for the 8 isoforms of YAP1; (ii) the known signaling differences among YAP1 isoforms; (iii) the relatively low expression of YAP1 isoforms in leukocytes; (iv) the disruption of the predicted leucine zipper in those YAP1 isoforms that contain exon 5 extension and exon 6; (v) the importance of the normalization of the nomenclature for YAP1 isoforms in the rapidly evolving field of Hippo pathway signaling.

We can easily distinguish YAP1-1 and YAP1-2 isoforms because they differ by 38 amino acids of the WW domain when YAP1 is precipitated with polyclonal antibodies. We cannot, however, resolve individual α - δ isoforms on a regular or gradient SDS-PAGE (e.g. Oka et al., 2008). Although various YAP1 cDNA clones have been generated and used by investigators in published reports, each encoding a distinct form of YAP1, the clones that we report here are novel. However, we cannot be sure that all 8 isoforms described here are expressed as proteins in cells. Moreover, various post-synthetic modifications, including serine (Sudol, 1994) and tyrosine phosphorylation (Levy et al., 2008) plus sumoylation (Lapi et al., 2008) have been shown for YAP1 protein and could add further to the heterogeneity of products resolved on SDS-PAGE gels, so that even the two dimensional gel

chromatography may not be the right tool to identify the protein products of YAP1. Either mass spectrometry analysis or a generation of monoclonal antibodies against spliced-in peptide sequences could provide answers as to whether all 8 isoforms of YAP1 are expressed as proteins.

Most of the studies published on YAP1 employ one of the YAP1-2 isoforms that contain two WW domains. Expression clones of YAP1-1 with one WW domain are used less frequently. However, YAP1-1 clearly differs in signaling from YAP1-2. For example, our group has shown that YAP1-1 does not bind p73 factor and cannot induce apoptosis when HEK293 cells are stressed by low serum conditions (Oka et al., 2008). In addition, YAP1-1 does not interact with angiomin while YAP1-2 does, and angiomin has the ability to inhibit the pro-apoptotic function of YAP1-2 by preventing its nuclear localization (Oka et al., 2012).

We observed a relatively low level of YAP1 expression in leukocytes compared to other tissues. This finding could be interpreted as following: YAP1 and perhaps other components of the Hippo pathway do not need to be operative in free floating cells such as leukocytes. It is known that one of the main routes by which the Hippo pathway controls YAP1 activity is via clues from cell-to-cell junctions, which help determine the final size of organs and the body (Zhao and Guan, 2010). Leukocytes and white blood cells do not form junctions. Three items of information further support the finding that YAP1 does not play a major role in the biology of leukocytes. Analyzing polyA⁺ mRNA isolated from human peripheral blood leukocytes on a Northern blot, we could not detect the YAP1 transcript, which is widely expressed in other human tissues and is 5 kb long. (Sudol et al., 1995). In mice, whose hematopoietic cells lack two kinases, Mst1 and Mst2, which act upstream of YAP1 in the canonical Hippo pathway, the number of mature T cells is significantly reduced (Mou et al., 2012). However, this defect in T cell proliferation seems unrelated to YAP1 (Mou et al., 2012). The ectopic expression of YAP1 in hematopoietic stem cells does not affect the quantity and function of these cells, neither during the steady state nor under stress conditions (Jansson and Larsson, 2012). This is in contrast to the effects of YAP1 on stem and progenitor cells in other organs.

Using individual mRNA splice isoforms of a gene for functional studies without specifying which isoform was used created controversy in the past, as exemplified by various reports using one of the four splice isoforms of ErbB4 receptor kinase (Sundvall et al., 2007; Veikkolainen 2011). The main message of our report is to propose a uniform nomenclature for the human YAP1 splice isoforms to avoid a similar scenario. This is important because the field of Hippo pathway signaling is progressing rapidly. A number of studies reported over-expression of individual and different isoforms of YAP1 in cells and animals (e.g., Komuro et al., 2003; Hao et al., 2008; Oka et al., 2008; Muramatsu et al., 2011; Schlegelmilch et al., 2011). Moreover, it is likely that exon 5 extension and exon 6 splicing variants of YAP1, which vary within the transcriptional activation domain (TAD), may select different repertoires of proteins in transcriptional complexes and affect the gene expression program in the YAP1 isoform- and tissue-specific manner. The disruption of the putative leucine zipper in exon 5 extension and exon 6 splicing variants of YAP1 may have functional consequences and we suggest that the two α isoforms of YAP1 will signal differently from β , γ and δ isoforms. Most likely the signaling differences between α and β , γ , δ isoforms of YAP1 will be subtle and a simple overexpression of these isoforms in cells followed by colony formation assay, for example, may not be the best way to characterize isoform-specific functions (Muramatsu et al., 2011).

Naming each of the isoforms of YAP1 either as suggested here (Figures 2A and 2B) or by the number of amino acids in the coding region of YAP1 isoform (e.g., YAP1-450 aa =

YAP1-1 α ; whereas YAP1-508 aa = YAP1-2 δ) will avoid miscommunications and should be helpful in the dissection of functional differences among various isoforms.

5. Conclusions

Four new mRNA splice isoforms of the human YAP1 gene were identified, making the total eight. Expression of all 8 isoforms of YAP1 is significantly lower in leukocytes compared to other tissues and organs. A new exon, exon 6 of the YAP1 gene, which is differentially spliced in various isoforms, appeared in evolution during the emergence of amniotes. All six isoforms of YAP1 with exon 5 extension, exon 6 or both, have their predicted leucine zipper regions disrupted. A new nomenclature for all 8 isoforms of YAP1 is proposed.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank our colleagues Kristin Berger, Kieran Harvey, Nikos Tapinos and Greg Yochum for valuable comments on the first version of the manuscript.

This work was supported by PA Breast Cancer Coalition Grants (#60707 and #920093) plus the Geisinger Clinic (to MS) and by funds from the National Institutes of Health (Grant# R01-GM083897) and the US Sylvester Berman Family Breast Cancer Institute to AF.

Abbreviations

bp	base pair
cDNA	DNA complementary to RNA
PCR	polymerase chain reaction
WW domain	tryptophan-tryptophan domain

References

- Bertini E, Oka T, Sudol M, Blandino G. YAP at the crossroad between transformation and tumor suppression. *Cell Cycle*. 2009; 8:49–57. [PubMed: 19106601]
- Bork P, Sudol M. The WW domain, a signalling site in dystrophin? *Trends Biochem Sci*. 1994; 19:531–533. [PubMed: 7846762]
- Carson M. Ribbons 2.0. *J Appl Crystallogr*. 1991; 24:958–961.
- Chen HI, Sudol M. The WW domain of Yes-associated protein binds a proline-rich ligand that differs from the consensus established for Src homology 3-binding modules. *Proc Natl Acad Sci U S A*. 1995; 92:7819–7823. [PubMed: 7644498]
- Hao Y, Chun A, Cheung K, Rashidi B, Yang X. Tumor suppressor LATS1 is a negative regulator of oncogene YAP. *J Biol Chem*. 2008; 283:5496–5509. [PubMed: 18158288]
- Harvey K, Tapon N. The Salvador-Warts-Hippo pathway – an emerging tumour-suppressor network. *Nat Rev Cancer*. 2007; 7:182–191. [PubMed: 17318211]
- Hilman D, Gat U. The evolutionary history of YAP and the hippo/YAP pathway. *Mol Biol Evol*. 2011; 28:2403–2417. [PubMed: 21415026]
- Jansson L, Larsson J. Normal hematopoietic stem cells function in mice with enforced expression of the Hippo signaling effector of YAP1. *PLoS One*. 2012:e32013. [PubMed: 22363786]
- Kanai F, Marignani PA, Sarbassova D, Yagi R, Hall RA, Donowitz M, Hisaminato A, Fujiwara T, Ito Y, Cantley LC, Yaffe MB. TAZ, a novel transcriptional co-activator regulated by interactions with 14-3-3 and PDZ domain proteins. *EMBO J*. 2000; 19:6778–6791. [PubMed: 11118213]

- Komuro A, Nagai M, Navin NE, Sudol M. WW domain-containing protein YAP associates with ErbB-4 and acts as a co-transcriptional activator for the carboxyl-terminal fragment of ErbB-4 that translocates to the nucleus. *J Biol Chem*. 2003; 278:33334–33341. [PubMed: 12807903]
- Landschulz WH, Johnson PF, McKnight SL. The leucine zipper: a hypothetical structure common to a new class of DNA-binding proteins. *Science*. 1988; 240:1759–1764. [PubMed: 3289117]
- Lapi E, Di Agostino S, Donzelli S, Gal H, Domany E, Rechavi G, Pandolfi PP, Givol D, Strano S, Lu X, Blandino G. PML, YAP, and p73 are components of a proapoptotic autoregulatory feedback loop. *Mol Cell*. 2008; 32:803–814. [PubMed: 19111660]
- Levy D, Adamovich Y, Reuven N, Shaul Y. Yap1 phosphorylation by c-Abl is a critical step in selective activation of proapoptotic genes in response to DNA damage. *Mol Cell*. 2008; 29:350–361. [PubMed: 18280240]
- Lupas A, VanDyke M, Stock J. Predicting coiled coils from protein sequences. *Science*. 1991; 252:1162–1164.
- Lutenic I, Doerks T, Bork P. SMART7: recent updates to the protein domain annotation resource. *Nucl Acid Res*. 2012:D302–5.
- Marti-Renom MA, Stuart AC, Fiser A, Sanchez R, Melo F, Sali A. Comparative Protein Structure Modeling of Genes and Genomes. *Annu Rev Biophys Biomol Struct*. 2000; 29:291–325. [PubMed: 10940251]
- Muramatsu T, Imoto I, Matsui T, Kozaki K, Haruki S, Sudol M, Shimada Y, Tsuda H, Kawano T, Inazawa J. YAP is a candidate oncogene for esophageal squamous cell carcinoma. *Carcinogenesis*. 2011; 32:389–398. [PubMed: 21112960]
- Oka T, Mazack V, Sudol M. Mst2 and Lats kinases regulate apoptotic function of Yes kinase-associated protein (YAP). *J Biol Chem*. 2008; 283:27534–27546. [PubMed: 18640976]
- Oka T, Schmitt AP, Sudol M. Opposing roles of angiomin-like-1 and zona occludens-2 on pro-apoptotic function of YAP. *Oncogene*. 2012; 31:128–134. [PubMed: 21685940]
- O'Shea EK, Rutkowski R, Stafford WF 3rd, Kim PS. Preferential heterodimer formation by isolated leucine zippers from fos and jun. *Science*. 1989a; 245:646–648.
- O'Shea EK, Rutkowski R, Kim PS. Evidence that the leucine zipper is a coiled coil. *Science*. 1989b; 243:538–542.
- O'Shea EK, Rutkowski R, Kim PS. Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. *Cell*. 1992; 68:699–708. [PubMed: 1739975]
- Pan D. The hippo signaling pathway in development and cancer. *Dev Cell*. 2010; 19:491–505. [PubMed: 20951342]
- Schlegelmilch K, Mohseni M, Kirak O, Pruszk J, Rodriguez JR, Zhou D, Kreger BT, Vasioukhin V, Avruch J, Brummelkamp TR, Camargo FD. Yap1 acts downstream of alpha-catenin to control epidermal proliferation. *Cell*. 2011; 144:782–795. [PubMed: 21376238]
- Sudol M. Yes-associated protein (YAP65) is a proline-rich phosphoprotein that binds to the SH3 domain of the Yes proto-oncogene product. *Oncogene*. 1994; 9:2145–2152. [PubMed: 8035999]
- Sudol M, Bork P, Einbond A, Kastury K, Druck T, Negrini M, Huebner K, Lehman D. Characterization of the mammalian YAP (Yes-associated protein) gene and its role in defining a novel protein module, the WW domain. *J Biol Chem*. 1995; 270:14733–14741. [PubMed: 7782338]
- Sudol M, Harvey KF. Modularity in the Hippo signaling pathway. *Trends Biochem Sci*. 2010; 35:627–633. [PubMed: 20598891]
- Sundvall M, Peri L, Maatta JA, Tvorogov D, Paatero I, Savisalo M, Silvennoinen O, Yarden Y, Elenius K. Differential nuclear localization and kinase activity of alternative ErbB4 intracellular domains. *Oncogene*. 2007; 26:6905–6914. [PubMed: 17486069]
- Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; 22:4673–4680. [PubMed: 7984417]
- Veikkolainen V, Vaparanta K, Halkilahti K, Iljin K, Sundvall M, Elenius K. Function of ERBB4 is determined by alternative splicing. *Cell Cycle*. 2011; 10:2647–2657. [PubMed: 21811097]

- Yagi R, Chen LF, Shigesada K, Murakami Y, Ito Y. A WW domain-containing yes-associated protein (YAP) is a novel transcriptional co-activator. *EMBO J.* 1999; 18:2551–2562. [PubMed: 10228168]
- Zhao B, Li L, Guan KL. Hippo signaling at a glance. *J Cell Sci.* 2010; 123:4001–4006. [PubMed: 21084559]

Highlights

Four new mRNA splice isoforms of human YAP1 gene are identified.

Exons 4, 5 and 6 of the human YAP1 gene are differentially spliced.

Exon 6 of the YAP1 gene appeared in evolution during emergence of amniotes.

Six out of eight YAP1 isoforms have disrupted leucine zippers.

Figure 1. Identification of eight transcripts that code for potential splice isoforms of YAP1 gene
Left Panel, Ethidium bromide stained agarose DNA gel with the 50 bp DNA ladder markers (left) and the products that were amplified from cDNA of human pancreas using Forward and Reverse primers shown in the *Right Panel*. The existence of YAP1-1 (with one WW domain) and YAP1-2 (with two WW domains) isoforms was documented previously (Sudol et al., 1995).

Figure 2. Schematic representation of four isoforms of YAP1-1 (2A) and four isoforms of YAP1-2 (2B)

The amino acid sequence encoded by exon 5-extension (in green) and exon 6 (in turquoise) are shown at the bottom. p-S indicates a major site of serine phosphorylation, p-Y, a site of tyrosine phosphorylation; TAD, transcriptional activation domain; PDZ-BM, PDZ domain binding motif (dark blue). The amino acid sequence of the coiled-coil region predicted by the SMART resource (Lupas et al., 1991; Luentic et al., 2012) is shown in orange. The location of sequences from exon 5 extension and exon 6, which disrupt the leucine zipper, is shown by green and turquoise arrowheads. The location of TaqMan primers to analyze the expression of YAP1-2δ isoform is indicated.

Figure 3. Structure of the YAP1 gene and the 8 alternatively spliced transcript isoforms
Top Panel, structure of the gene and its primary transcript drawn to scale. *Middle panel*, structure of primary transcript with introns of arbitrary length, but exons drawn close to scale with respect to one another. Selected protein features are indicated in exons by colors. Sequence context of alternative splicing at 3'-end of exon 5 is shown below primary transcript. An alternative splice donor site 12 nt into intron 5 generates some alternative transcripts containing the amino acids VRPQ. Note the two 5' splice donors in intron 5, 5A, CAGGTGAGG. This is fairly canonical, only the last G is less common than a T. 5B, CAGGTTAGA. First 5 nts are conserved, including the obligatory GT, the 6th nt is a T (least preferred), 7th and 8th nts are conserved, 9th nt is least preferred base. This is from the use of least preferred bases in two sites, a much weaker 5' SS than 5A. *Bottom Panel*, structure of eight mature (spliced) RNA transcripts. Dark green indicates transcribed regions and un-translated regions of mature transcripts; dark blue, no identified protein domain; red, WW domain; yellow with orange border, beginning of the TAD domain; light blue with black border, alternatively spliced exon 6 sequence, AMRNINPSTANSPKCQ.

Figure 4. The relative levels of YAP1-26 isoform expression in human tissues and organs

The data are represented in two forms. On the left, the distribution is summarized with a box-and-whisker plot: the median is indicated by the bold line, the range between the first and third quartiles is indicated by the box, the whiskers extend to 1.5 times the inter-quartile range (IQR), and circles indicate outliers (outside the 1.5 IQR). On the right, individual relative expression values are shown labeled with the source tissue. The expression levels were measured by TaqMan Real-Time PCR using human cDNAs normalized for expression to four house-keeping genes (see Methods). Note that the expression of YAP1-26 isoform in leukocytes is significantly lower than that in other tissues.

Figure 5. The structures of the genomic region that gives rise to alternative mRNA splicing products of the YAP1 gene in different metazoans

The exon structure in the region generating the different isoforms is shown and is compared to the longest human isoform δ . The sequences are marked in the same color used to depict the different human isoforms. Man, *H. sapiens* (*Homo sapiens*); Zebra Finch bird, *T. guttata* (*Taeniopygia guttata*); domestic chicken, *G. Gallus* (*Gallus gallus*); an arboreal lizard, *A. carolinesis* (*Anolis Caroliensis*); the pipid frog, *X. tropicalis* (*Xenopus tropicalis*); Zebrafish, *D. Rerio*, (*Danio Rerio*); sea squirt, *C. intestinalis* (*Ciona intestinalis*); sea anemone, *N. vectensis* (*Nematostella vectensis*).

Figure 6. Sequence and structural analysis of leucine zippers of YAP1 isoforms

(a) Amino acid sequence alignment of four isoforms of YAP1. Note that YAP1-1α isoform consists of five successive heptads of amino acid residues, designated herein H1-H5, with a signature leucine at every seventh position, designated herein L1-L5. Each of the five heptads is alternately colored blue and green, while the five signature leucines are shown in red. The YAP1-1β, YAP1-1γ and YAP1-1δ isoforms have amino acid inserts (4-20 residues) between heptads H3 and H4. (b) Structural model of the putative leucine zippers of YAP1-1α. One chain is labeled N-C, while the other N'-C'. Within each chain, each of the five heptads (H1-H5) is alternately colored blue and green, while the side chain moieties of five signature leucines (L1-L5) are shown in red. The prime (') distinguishes signature leucines and heptads of one chain from the other.