
©1996-2008 All Rights Reserved. *Online Journal of Bioinformatics* . You may not store these pages in any form except for your own personal use. All other usage or distribution is illegal under international copyright treaties. Permission to use any of these pages in any other way besides the before mentioned must be gained in writing from the publisher. This article is exclusively copyrighted in its entirety to OJB publications. This article may be copied once but may not be, reproduced or re-transmitted without the express permission of the editors. [This journal satisfies the refereeing requirements \(DEST\) for the Higher Education Research Data Collection \(Australia\)](#). Linking: To link to this page or any pages linking to this page you must link directly to this page only here rather than put up your own page.

OJBTM
*Online Journal of
Bioinformatics*®

9 (1):60-83, 2008

*Prediction of the putative function of
mouse WDR13 protein*

Murthy BSN, Pandit MW, Singh L.

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad – 500 007; India

ABSTRACT

Murthy BSN, Pandit MW, Singh L, Prediction of the putative function of mouse WDR13 protein, *Online Journal of Bioinformatics*, 9 (1): 60-83, 2008

Literature reports large number of WD-repeat proteins involved in variety of functions, although only few of them are characterized for their true physiological roles. As these proteins function in multi-protein complexes, it is rather difficult to individually express, purify, fold and functionally characterize these proteins. Thus, novel approaches are imperative in assessing their putative physiological functions. Evolutionary conservation of the protein sequence-structure and function is well known. Realizing that WDR13 is an unstable protein and expecting that protein instability could be an evolutionarily conserved property, we have analyzed a set of WD-repeat proteins for their Protein Instability Index and sequence motifs present in them, employing bioinformatics, and using this information predicted the physiological functions of WDR13. Nearly 65% of WD-repeat proteins are unstable and they apparently form functional clusters according to PII. N-terminal of WDR13 possesses novel nuclear localization signal and SOCS-homologous sequence, which suggested that WDR13 putatively participates in E3 ubiquitin ligase ECS complex known to associate with the Ubiquitin Proteasome System. BLAST of the Database of Interacting Proteins revealing sequence homology of WDR13 with Fbw7 (product of hCDC4), another component of E3 ubiquitin ligase (SCF) complex, on one side and

Eukaryotic Linear Motif analysis identifying presence of functional domains necessary for participation in ubiquitin ligase on the other, strongly suggests that WDR13 is the ECS component, functionally complimentary to Fbw7, which participated in the SCF complex. Above results suggest a new dimension of relation between primary sequences of proteins influencing their own fate as well as of the cells. Supplementary Data (Table S): Stable and unstable WD-repeat proteins available in the literature are shown in Table SA and SB along with their calculated Protein Instability Index (PII) values, identified / reported function, accession number and the organism.

Key words: WD-repeat protein; Memory related protein; Suppressor of cytokine signaling; SOCS; Elongin A; VHL protein; ECS Complex; SCF Complex; Ubiquitin ligase; Ubiquitin Proteasome System

INTRODUCTION

While looking out for genes involved in testicular differentiation and development, Suresh et al. [1] identified a novel and highly conserved WDR13 gene localized on the mammalian X chromosome, mapping to mouse XA1.1 locus and predominantly expressed in the testis. WDR13 contains 9 exons and 8 introns, its expression begins at an early stage of gonadal development and is maintained throughout adult life with predominant expression in germ cells of adult testis [1]. RNA in situ hybridization on testis and brain sections revealed cytoplasmic expression of WDR13. Alternatively spliced and significantly varied expression of WDR13 (transcripts) observed in several tissues suggested functional diversity of the gene [2]. Expression in unfertilized eggs and neuronal stem cells indicated functional significance of WDR13 during early stages of mouse development [1]. Analysis of EST clones from various tissues identified two splice variants: a full-length 53-kDa 485-amino acid (aa) and a 43-kDa 393-aa protein, devoid of N-terminal 92 aa region [2]. Significant levels of expression of WDR13 in the early embryo, testis and ovary suggested the involvement of WDR13 in differentiation/development/maintenance of gonads. However, its expression even in brain, heart and kidney on one hand and several normal as well as carcinoma tissues on the other, would suggest that it might be involved in a universal function in mammalian cells, the assumption which would perhaps fail to justify its localization on the X (one of the sex-) chromosome. This puzzling problem led us to set out for determining the true physiological function of WDR13.

Although, a large number of WD (WD40) repeat proteins is reported in the literature, only a few of these proteins are completely characterized for their physiological functions. Proteins forming WD-repeat family are characterized by the presence of repeated sequences of about 44-60 aa with conserved spacing and core, typically bordering GH (glycine-histidine) at the amino and WD (tryptophan-aspartic acid) at the carboxy termini of the repeat units, and the proteins contain 4-16 such repeat units [3 & 4]. Several WD-repeat proteins of varied and unknown functions from a variety of organisms have been reported in the literature [4, 5 & 6]. They are identified as regulatory components in multi-protein complexes that govern several processes such as transcriptional activation/repression, signal transduction, ubiquitination, cell-cycle entry/exit, pre-mRNA/rRNA processing, chromatin modification, RNA export, cell death, microtubule assembly and protein transport to/from and across membranes [4]. First discovered in the beta sub-unit of retinal GTP-binding protein, WD-repeat motifs are found in over 100 proteins with diverse cellular functions [4]. WD-repeats are thought to form 3D-propeller structure that provides a flexible surface for specific protein-protein interactions. Thus, whatever

may be the cellular process for which they are responsible, the proteins with this motif often participate in complexes with other proteins and sometimes coordinate sequential and/or simultaneous interactions involving several sets of proteins [4]; With little structural similarity and very diverse cellular roles, it is not clear as to why these proteins contain conserved GH and WD repeats and also what is the rationale behind categorizing them into one family? Recently, functional sub-families of WD-repeat proteins have been identified on the basis of the criterion of surface similarities in the propeller structure, which presumably reflects similarities in binding partners among the members of sub-families [4]. Further, several WD-repeat proteins have been reported to possess considerable homology between individual proteins across the eukaryotes, in addition to the conservation and the divergence of structure and function in this family [7].

Primary sequence guides proteins through the successive order of secondary, tertiary and quaternary levels of folding to subsequently acquire a stable and an active conformation. Invariably, amino acids involved in the active site present in different domains are brought into close proximity through various orders of folding. Therefore, protein sequence or its quantifiable parameters should be related to the activity determinants (the structure and hence the function) of the protein. Changes occurring in protein sequence through evolution (due to mutations) alter protein folding, activity and, consequently the physiological function. Thus, factors such as primary sequence, its quantifiable properties (or parameters), protein structures at various levels of folding and physiological function must be inter-related and conserved through evolution. Therefore, analysis of quantifiable sequence parameters and function of a specific group of proteins reported in the literature could in principle be used as an approach for predicting physiological function of unknown proteins of that group.

Upon cloning, over expression, purification to homogeneity and folding the 43 kDa WDR13 protein (see Methods & Results) we realized that the protein was insoluble, did not fold and underwent autolysis. As the protein preparation had no protease activity and autolysis produced large number of peptides, we suspected that it could be an unstable protein. We analyzed the protein sequence and concluded that it was an unstable protein by virtue of the presence of large number of unstable peptide bonds in the sequence. In view of the above, we anticipated evolutionary conservation of sequence elements contributing to protein instability and function in the WD-repeat family, reported in the literature. Earlier, it was reported that *in vivo* stability (half-life) of proteins, quantified as Protein Instability Index (PII) is related to its dipeptide composition [8]. Since PII is calculated on the basis of dipeptides (both quantity and type) present in proteins, similar functional domains and dipeptide composition would have similar PII and are likely to support similar/related functions, as each protein must have evolved from its ancestor, gradually, in terms of both sequence and function. This rationale encouraged us to evaluate PII values of a large number of WD-repeat proteins of various physiological functions reported in the literature, to assess if they form functional clusters according to PII, and use this data for comparing and predicting probable physiological function of WDR13.

METHODS

Cloning, expression, purification and folding of WDR13 protein: The cDNA of 393-aa-WDR13 was PCR amplified employing 5'ATTCATATGCGAATGGAGACTTTGAA3' as the forward and 5'ATTGAATCCCCTACTTTTGCTCTCGT3' as the reverse primers and resulting cDNA digested with NdeI and Bam HI restriction enzymes and cloned into the pET 28 vector. WDR13 containing plasmid was transformed into BL21DE3 bacteria and protein overexpressed by induction with IPTG, following standard laboratory protocols. Overexpressed WDR13, which was insoluble and accumulated into inclusion bodies was purified employing NiNTA Agarose under denaturing conditions employing the Quiagen protocols. Purified WDR13 was subjected to *in vitro* folding employing "foldit kit" manufactured by Hampton Research (<http://www.hamptonresearch.com>). WDR13 protein was diluted, by periodically releasing concentrated protein (1-2 mg/ml in 6M Gu.HCl) in small aliquots (40-80 µl/min) into 200 volumes of the buffer cocktails under constant swirling, and the protein was allowed to fold overnight at 4°C. Soluble protein was gradually dialyzed against the same buffer cocktail devoid of protein-folding agents. However, WDR13 protein, which was apparently dissolved, was in fact found degraded; the fragments obtained were characterized by ESI MS (Electron Spray Ionization Mass Spectrometry).

Characterization of WDR13 fragments: Dialyzed WDR13 protein obtained from *in vitro* folding was resolved employing C18 hydrophobic FPLC column, and eluted peptide fractions were pooled and freeze-dried and subjected to analysis by ESI MS. We identified a large number of well-resolved peptides of WDR13 observed in the ESI MS spectra, which were generated due to auto-degradation of the large number of unstable peptide bonds characteristic of WDR13 protein (see [Fig.1B next page](#)), and concluded that WDR13 is an unstable protein.

Collection of WD-repeat protein sequences: WD-repeat proteins were identified as regulatory components in multi-protein complexes that govern several cellular processes [4]. We took into account these functions and collected all the references cited in PubMed, which in turn were used for collecting WD-repeat proteins with their names as referred therein. These WD-repeat protein names were submitted to NCBI under protein sub-head for collecting as many unique WD-repeat proteins as possible. However, we retained the sequences having 300 to 600 aa, so as to achieve a collection of proteins ranging within $\pm 20\%$ of the size of WDR13 (393 and 485-aa).

PII analysis of WD-repeat proteins: Each of the WD-repeat protein sequences was submitted for computation of its PII to Protparam program available at <http://www.ExPASy.ch.org>. Protparam analysis employs two independent programs for computing the instability index of proteins. N-end rule is based on the observation that N-terminal amino acid of proteins plays a crucial role in deciding *in vivo* half-life of proteins [9, 10 & 11]; while the other is based upon *in vivo* short-lived proteins which are found to contain certain di-peptides more preferentially than in those that live longer in cells [8], suggesting that protein instability is an evolutionary consequence. Protparam takes into account the values obtained from both the above methods and gives a composite PII value in arbitrary units, where values < 40 are indicative of stable proteins and values > 40 indicate unstable proteins. After PII analysis we organized WD-repeat proteins in the ascending order of their computed PII.

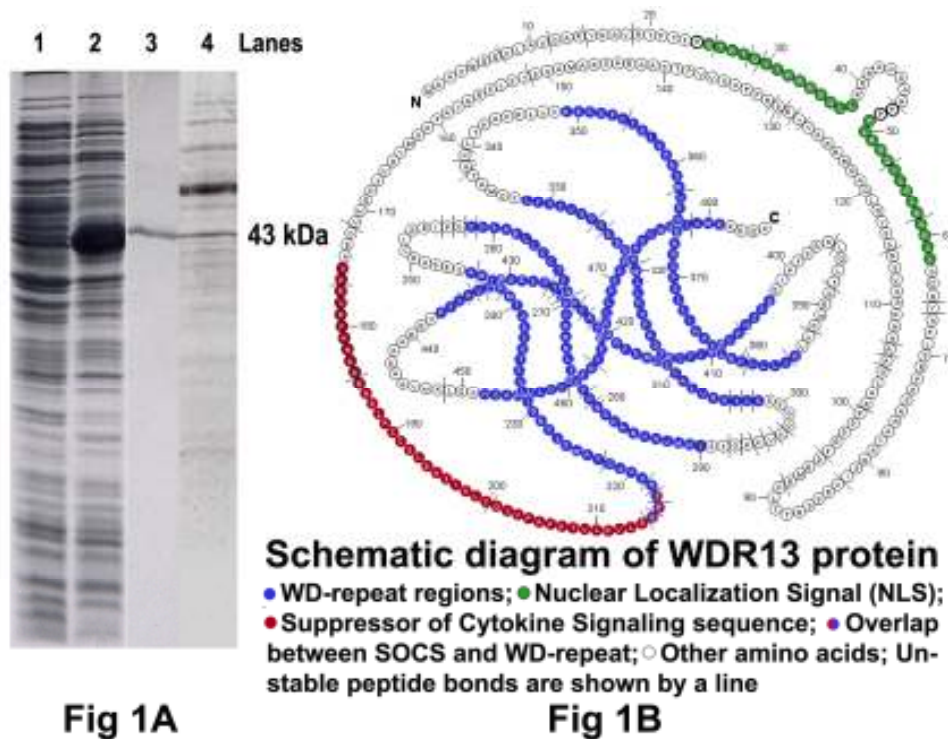


Fig. 1: (A) SDS PAGE profiles of overexpressed WDR13 protein: Lanes 1: un-induced control protein from bacteria containing WDR13 plasmids; Lane 2: overexpressed WDR13 (43 kDa protein can be seen as a thick band); (Bacterial culture of 1 ml with 0.80 OD grown/induced for 3 hours and electrophoresed; Lane 3: purified 43 kDa protein (50 μ g); lane 4: purified WDR13 after the protein folding attempt (200 μ g). (B) Schematic diagram of the sequence of full-length WDR13 protein depicting specific sequence domains identified by bioinformatics tools

Reporting protein function: We included in our analysis all the identified/predicted physiological functions of WD-repeat proteins reported in the literature. It is important to note the limitations of these reported functions. Only very few of the individual WD-repeat proteins are known for their precise physiological function. Therefore, protein similar to other known functionally identified proteins does not necessarily mean that the function of the concerned protein is clearly identified. In view of these limitations, the exercise of computing PII values allowed us to catalog, according to their PII, the known WD-repeat proteins hitherto reported, and look for apparent functional clusters in these proteins.

Analysis of proteins for WD-repeats/sequence motifs: We submitted each of the protein sequences to the SMART (Simple Molecular Architecture Research Tool) program available from EMBL at <http://www.ExPASy.ch.org>, under the head 'Similarity searches', and arrived at the WD-repeat and other sequence motifs present in these sequences. This program analyses sequence motifs present in proteins and displays in the priority order of SMART > PFAM > PROSPERO repeats > Signal peptide > Transmembrane > Coiled coil > Low complexity. We considered both predicted motifs with higher confidence as well as those predicted with a lower level of threshold. This is because the ability of a protein to perform a certain function is based on the strong functional sequence motifs present in it, and not because it could or could not be detected by EMBL SMART program.

Analysis for homology between SOCS-regions present in proteins: We collected SOCS sequence motifs present in WD-repeat-containing and other SOCS-box proteins reported in the literature and subjected these sequences to several multiple sequence alignment programs available at ExpASY Molecular Biology Server.

Bioinformatics analyses for comparison of sequence domains present in WDR13 with those present in known ubiquitin ligase proteins: Earlier bioinformatics analyses revealed that WDR13 possesses a SOCS-homologous sequence, which is likely to participate in ubiquitin ligase complex that ubiquitinates proteins in their degradation pathway (see Results and Discussion). Protein degradation, a universal process plays the pivotal role in cells and tissues in almost all physiological conditions like division (mitosis and meiosis), disease, stress, apoptosis, etc. Thus we decided to conduct a comparison of the sequence motifs present in WDR13 with those present in literature-reported well-known ubiquitin ligase components, for assessing the probability of participation of WDR13 in a novel ubiquitin ligase. We employed both SMART and Eukaryotic Linear Motif (ELM) analysis program as available at ExpASY proteomics server for detecting sequence domains present in known ubiquitin ligase proteins, human CDC4 [Fbw7a (gi|109659002)] and mouse Cdc20 (gi|8885513) and compared them with those present in full-length mouse WDR13 (gi|123283851).

RESULTS AND DISCUSSION

Attempts of folding WDR13 protein: We cloned, overexpressed and purified 393-aa-WDR13 as described in "Methods". Lanes 1 and 2 in [Fig 1A](#) show the patterns of the total proteins of un-induced control and induced bacteria, respectively, both containing the WDR13 cloned pET 28 Plasmids. Upon purification, we recovered homogenous WDR13 as shown in lane 3, of [Fig 1A](#). We attempted *in vitro* folding of WDR13 by diluting it into large volumes of different buffer cocktails (of 'fold it' kit) and succeeded in dissolving the protein in a cocktail of 50 mM Tris-HCl, pH 8.20, containing 550 mM L-Arginine HCl and 440 mM sucrose as major protein folding agents. Gradual dialysis of dilute protein solution retained WDR13 in solution after the removal of the protein folding agents. Upon dialysis, although WDR13 was retained in solution, it was degraded into large number of fragments. The lane 4 in [Fig 1A](#) shows the pattern of WDR13 retained in solution after the dialysis and concentration. High molecular weight protein bands (> 43 kDa) seen in the lane 4 of [Fig 1A](#) perhaps arose due to intermolecular disulfide cross-linking between degradation peptides of WDR13 due to the presence of oxidizing components present in the reducing agent (β -mercaptoethanol) [12] employed in SDS PAGE. Assay of WDR13 retained in solution for protease activity exhibited no activity whatsoever (data not shown) suggesting that WDR13 could be a labile protein. Hence, we attempted to identify labile peptide bonds present in WDR13 employing the method reported earlier [8]. [Fig.1B](#) presents a schematic diagram of 485-aa-WDR13 protein, in which all the amino acids forming labile peptide bonds are marked. It can be seen that di-peptides (labile peptide bonds) contributing to higher values of PII do not have a bias for occurrence either in WD-repeat containing or other sequence regions. We purified and analyzed the post degradation fragments of WDR13 by ESI MS; the de-convoluted mass spectra are shown in [Figure 2](#). Masses of several peptides from the spectra matched well with those that could be calculated theoretically, presuming the cleavage of WDR13 at the labile peptide bonds (data not shown).

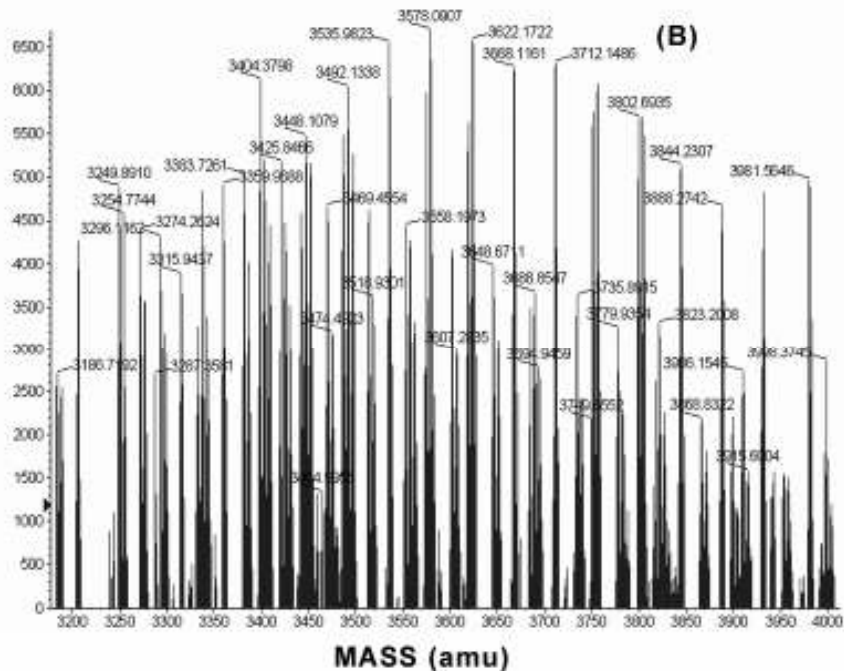
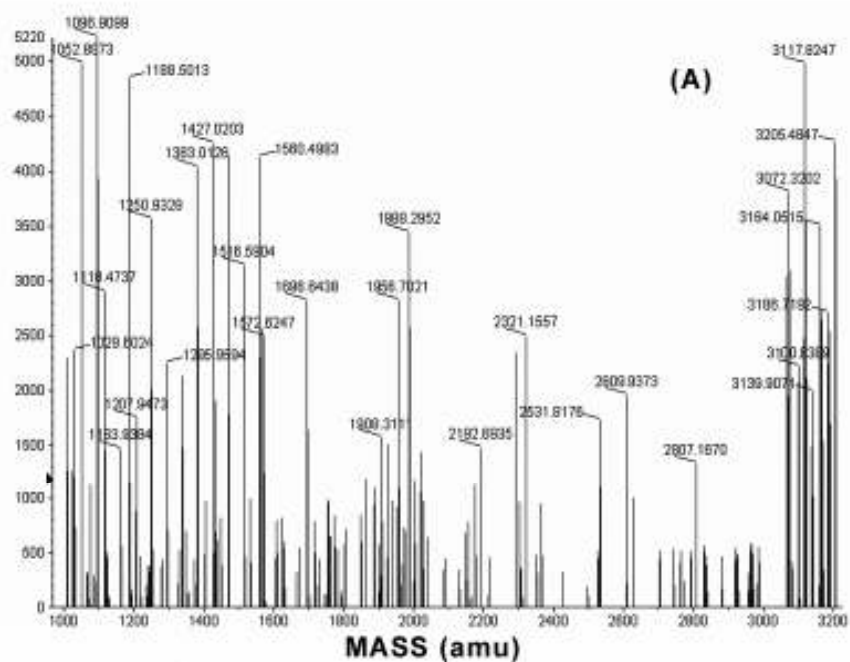


Fig 2 De-convoluted ESI MS spectra of peptides of degraded WDR13 protein. Mass spectra cover atomic mass units (amu) between 1000-3200 in (A) and between 3200-4000 in (B)

Fig 2

Instability of WD-repeat proteins and their functional clustering: We computed protein instability index (PII) of all the WD-repeat proteins collected from the literature. A catalogue of collected stable and unstable WD-repeat proteins along with their computed PII values, identified/reported protein function and accession numbers is provided as a supplementary data ([Table SA/SB](#)). Apparently, two thirds of these proteins were found to be unstable (PII > 40). When these were arranged in the ascending order of PII values, they appeared to form functional clusters, although there were overlaps between different clusters. Functional clusters of unstable WD-repeat proteins are shown in [Table I](#). PII values for 393- and 485-aa-WDR13 proteins were calculated to be 47.49 and 49.71, respectively. These values predominantly overlap with the PII values of WD-repeat proteins in [Table I](#), involved in M-phase checkpoint and F-box ubiquitination and hence provide some idea about the functional prediction of WDR13 proteins. Other functional clusters, with which WDR13 does not cluster, are also shown in [Table 1](#), for comparison.

Functional assessment of WDR13 vis-à-vis functional clusters: It is rather premature to arrive at the functional involvement of WDR13 protein only on the basis of PII and functional clustering. Therefore, we examined the sequence motifs in WDR13 proteins, which are likely to throw some light on the possible function of WDR13 proteins. EMBL SMART analysis for detecting the sequence motifs present in WDR13 revealed neither global homology of WDR13 with any of the M-phase checkpoint group of proteins, nor the presence of an F-box motif in WDR13. Thus, WDR13 does not appear to fit with certainty into either of these groups. The analysis further detected several sequence motifs in WDR13 proteins, which include SMR (161-256), SOCS (175-217), ARM (182-225), CLIP (191-257), DM16 (262-302), LH2 (279-377), DWB (360-462), GRAM (393--445), XPGI (404-470), PQQ (415-447) and WAP (419-439) (corresponding aa regions from the N-terminal of 485-aa-WDR13, are shown in the parenthesis). However, all these motifs were detected at lower threshold and therefore cast doubt on their functional significance; thus these motifs had to be evaluated individually. We evaluated each of these motifs in terms of their contribution to their known function of WD-repeat proteins. We found that none of these, except Suppressor of Cytokine Signaling (SOCS) motif, appeared to plausibly contribute to any reported function of WDR13 proteins. To assess the significance of the presence and thus the involvement of SOCS-region in WDR13 and its physiological function, we analyzed several isoforms of various SOCS proteins to EMBL SMART analysis. When we submitted the sequence of full-length WDR13 for analysis by EMBL SMART, interestingly a sequence representing SOCS-motif was detected, although with a lower level of confidence (indicated in the schematic diagram Fig. 1B), suggesting the probable relevance of SOCS-motif to the function of WDR13. Out of the six WD-repeat proteins (excluding WDR13) of SOCS group collected in our data ([Table S](#)), five (NP_599026, NP_599027, NP_056441, NP_062627 and NP_067514) were found to be stable proteins as indicated by their respective PII values. The sixth protein (Q9W5Z5) appears to be an unstable protein (Sl. no. 13 in [Table SB](#)), with a PII of 42.41. This indicates that (i) there could be some more of unstable WD-repeat proteins containing SOCS-box, which were not represented in our data collection of WD-repeat proteins; and (ii) SOCS containing unstable WDR-repeat proteins may be relatively recent in evolution and thus are not part of our data collection.

Cellular localization of WDR13 proteins: In view of the above, we decided to take into account the cellular localization of WDR13 on one side and all known physiological functions of known SOCS proteins on the other, for arriving at the contribution of SOCS-motif to the function of WDR13. Our earlier, confocal microscopy of WDR13 employing GFP-tagged protein expression in mammalian cells

revealed the presence of 485-aa-WDR13 only in the nucleus, while 393-aa protein was detected both in the nucleus as well as in the cytoplasm [Fig. 6B, Ref. 2]. Fig. 3 presents the cellular localization of WDR13. Thus, we expect WDR13 proteins to support cellular functions both in the cytoplasm as well as in the nucleus. To understand more about the localization of WDR13 proteins, we subjected these sequences to various bioinformatics tools available at <http://www.ExPASy.ch.org>. The significant results of these analyses are shown below:

MITOPORT takes into account the N-terminal region of proteins that is able to support mitochondrial targeting of proteins and the cleavage site. The analysis revealed that 485-aa-WDR13 possessed a mitochondrial targeting sequence, while 393-aa-WDR13 was devoid of it. This was not expected because neither our localization studies detected presence of WDR13 in mitochondria, nor WDR13 protein showed homology with any of the WD-repeat proteins involved in mitochondrial function. PSORTII predicts sub-cellular localization of proteins employing two sub-programs. NNCN, cytoplasmic/Nuclear discrimination method predicted cytoplasmic localization of 393-aa-WDR13 with reliability score of 76.70 and nuclear localization of 485-aa-WDR13 protein with a reliability score of 55.50. This prediction is in agreement with the earlier experimental findings, discussed above. However, 393-aa-WDR13 is also localized in the nucleus according to the earlier results, and needs a rational explanation. Further k-NN method predicted localization of 393- and 485-aa-WDR13 proteins into cytoplasm and mitochondria respectively with corresponding probability scores of 52.20% and 69.60%. The prediction of localization of 485-aa-WDR13 into the mitochondria instead of nucleus of the cells requires a better understanding, as MITOPORT also detected presence of mitochondrial targeting sequence in the N-terminal of this protein, as already mentioned above.

Predictions made by PSORTII rely on the presence of known nuclear localization signal (NLS) in the protein. NLS signals are characterized by the presence of either mono-partite or bi-partite clusters of basic amino acids, possessing a helix-breaking amino acid preceding the clusters, which are separated by 9-12 amino acids [13]. We observed that 485-aa-WDR13 in fact, possesses two highly basic and fairly homologous amino acid stretches, each containing 14 amino acids (25-38, QFRTQYIRRRSQLL and 49-62, ALRRQYLRLRGQLL), which are 10 amino acids apart, and also possess preceding helix-breaking pralines as marked in the NLS (Fig. 1B). These stretches could serve as a novel bipartite NLS signal, and thus provides a rationale for the observed localization of the protein in the nucleus. Our analysis of MITOPORT as well as k-NN methods of PSORT II identified localization of 485-aa-WDR13 into mitochondria, probably because of the absence of the above novel NLS sequence in the respective knowledge bases. MITOPORT identified the N-terminal 78-aa of 485-aa-WDR13 in the category of mitochondrial targeting sequence, which is obviously not available in 393-aa-WDR13, as it does not possess the N-terminal 92 aa portion of 485-aa-WDR13. Hence it is natural to expect that it would localize in the cytoplasm, which corroborates its existence in cytoplasm, experimentally. However, its existence in the nucleus can be explained by the observation that proteins devoid of NLS are found to localize in the nucleus as a complex entity [14]. It is quite possible that the presence of WD-repeats as well as SOCS in 393-aa-WDR13 which are known to aid in protein-protein interactions, may promote complex formation with other proteins possessing NLS and help in its localization to the nucleus. This would perhaps explain experimentally observed nuclear localization of 393-aa-WDR13.

An overview of physiological functions of proteins containing the SOCS-box:

Analysis of the reported WD-repeat proteins containing SOCS-box by EMBL SMART revealed that all these proteins possess SOCS-motif at their C-termini. The presence of C-terminal SOCS is found in about 40 proteins, which are classified into several families, such as CIS, SOCS-1 to SOCS-7, WSB, SSB, ASB, GTPases, ras, ras-like proteins, etc. [15]. This classification is based on the presence of various domains such as SH2-, WD-, SPRY- and ankyrin- repeats, in addition to SOCS. These domains are present in the N-terminal side to the SOCS-box. SOCS proteins suppress cytokine signal responses by negative feedback, which is thought to be important in the regulation of wide varieties of cellular functions controlled by extra-cellular signals. There is a lone class of protein, MUF-1 reported in the literature [15], which contain SOCS-box at the N-terminal unlike other families and at the same time is devoid of WD-repeats. We found that WDR13 is the first protein, which possess a SOCS-box on the N-terminal side of WD-repeats.

(i) Consequences of Suppression of Cytokine Signaling and related regulatory processes: Cytokines are secreted proteins that regulate a broad spectrum of important cellular responses such as proliferation and differentiation. Despite the functional diversity, they all use similar signal transduction pathways to elicit their responses [16]. Key events in cytokine signal transduction are well defined: Cytokines bind to the members of hemopoietin or cytokine receptor super family, which are expressed on the surfaces of responsive cells and induce dimerisation or aggregation [17 & 18]. Upon receptor dimerisation, associated JAK (Janus kinase) gets activated by cross-phosphorylation, and in turn phosphorylates tyrosines present in the cytoplasmic domains of these receptors, thus creating sites for docking of a host of cytoplasmic proteins such as signal transducers and activators of transcription (STAT) and other transcription factors. STATs and other transcription factors then get phosphorylated by JAKs, dissociate from the receptors, assemble into homo/hetero dimers, migrate to the nucleus, interact with specific sequence elements present in the promoters of cytokine responsive genes and increase their rate of transcription. Very elegant experimental evidence supports involvement of SOCS in efficiently attenuating signal transduction of most of the cytokine-induced responses [19, 20 & 21]. SOCS-1 and CIS attenuate cytokine signals by inhibiting catalytic activity of JAK family of cytoplasmic kinases [22, 23 & 24], and competitively binding to phosphorylated cytoplasmic receptor-domains, respectively [25 & 26]. SOCS was identified simultaneously on the basis of its interaction with JAKs, its intrinsic cross-reactivity to STATs, and its ability to inhibit the cytokine signaling [27]. Proteins containing C-terminal SOCS-box in addition to other repeat sequence motifs such as WD, ankyrin, SPRY, etc., are suggested to possess independent biological properties like protein-protein interaction. This was suggested on the basis of the observation that a variety of proteins including SOCS-1, SOCS-3, WSB-1, ASB-2 and Rar, could associate with Elongin C and, via this interaction, with Elongin B [28 & 29]. In view of the above, WDR13 protein possessing SOCS-region and WD-repeats may be expected to support other functions, in addition to suppression of cytokine signaling.

(ii) Involvement of SOCS proteins in polyubiquitination of proteins in their degradation pathway: Elongins B and C are found in mammalian transcription activator complex, bound to Elongin A (EA) through the SOCS-box of EA. Elongin B-C complex also interacts with VHL (von Hippel-Lindau), a tumor suppressor protein, possessing SOCS-homologous C-terminal α -domain. VHL is also involved in down-regulation of hypoxia-inducible genes, suggesting that Elongin B-C by itself does not regulate transcription [29 & 30]. Elongin B (ubiquitin-like protein) and Elongin C (skp1-like protein) form a complex Elongin B-C that binds to the motif (known as B-

C-box motif) present within the SOCS-regions of EA or VHL. The *skp-1* sequence motif present in both EA and VHL proteins recognize F-box motif present in F-box proteins, which are involved in polyubiquitination and degradation of proteins. In addition to its presence within the SOCS-region, B-C-box motif is present in several other proteins such as ras, WD-, SPRY- and ankyrin-repeat families. Elongin B-C complex is also reported to perform additional functions such as attenuation of JAK/STAT signaling, increasing half-life and concentration of SOCS proteins by interaction with these proteins through their SOCS-box, and inhibition of their degradation, etc. Above instances suggest that Elongin B-C is a multifunctional regulatory complex capable of modulating multiple pathways in the cell through interaction with a short degenerate (B-C) motif found in many different proteins [29]. typical example which illustrates this suggestion follows:

VHL participates in a multi-protein complex consisting of activating protein (E1) and conjugating protein (E2) that ubiquitinate hypoxia-inducible factor-1 (HIF1 α), resulting in the regulation of its degradation by proteasomes [30 & 31]. VHL and Elongins B and C are found in complex with cullin family member Cul-2 and RING-finger-protein Roc1, possessing ubiquitin ligase activity [31, 32 & 33]. Similar to other SOCS-box proteins, VHL also possesses N-terminal β -domain involved in protein-protein interaction that associates with a 'crucial proline' present in the transactivation domain of HIF1 α , only when this proline is hydroxylated. Under normoxic conditions, processes such as hydroxylation of 'crucial proline' in HIF1 α , VHL binding, polyubiquitination of HIF1 α and its proteasome-involved degradation take place. However, under hypoxic conditions HIF1 α fails to associate with VHL due to the absence of hydroxylation of the 'crucial proline', resulting in the accumulation of HIF-complex and transcriptional activation of its target genes. Thus, VHL is the key protein-interacting module involved in both transcription of HIF-responsive genes and their down-regulation. VHL participates in both tumor suppression as well as regulation of hypoxia inducible genes, while EA is involved in transcriptional activation of specific genes [29]. Therefore, similar functional properties may be expected for other proteins, including WDR13, possessing SOCS-box and protein-protein interaction domains.

Several SOCS proteins, Rar, EA and VHL possess either SOCS or SOCS-homologous domains in their C-terminal; and they specifically interact with Elongin C (B-C complex) and their substrates in a highly regulated manner, only after the substrates are modified post-translationally. Thus, it is speculated that all SOCS-box proteins participate in E3 ubiquitin ligase ECS (Elongin C-cullin-SOCS-box) complex [15] analogous to the best-studied SCF (*skp1*-cullin-F-box) ubiquitin ligase complex [34]. However, there are still open questions about such a speculation. In SCF complexes, F-box has been shown to play a role comparable to that proposed for SOCS in the speculated ECS complexes [28, 29, 30 & 34]. Similar to VHL/SOCS-1/SOCS-3 proteins, substrate recognizing domains of F-box also require post-translational modification (phosphorylation of serine or threonine) of target proteins for the interaction to occur. Thus, ECS and SCF pathways appear to have parallel mechanisms for polyubiquitination and proteasome-involved degradation of proteins. Thus, there is every likelihood that SOCS-domain present in WDR13 enables it to participate in a novel ubiquitin ligase and contributes to polyubiquitination of proteins in their degradation pathway; however this needs to be testified with appropriate evidence.

Evolutionary relation of SOCS domain present in WDR13: We subjected SOCS-regions present in WDR13 and several other WD-repeat containing and other SOCS proteins to multiple sequence alignment programs described in "Methods". Data of TCOFFEE multiple sequence analysis presented in [Fig. 4A](#) revealed identical sequence alignment scores for the SOCS-regions present in WDR13 and VHL proteins. Dendrogram of TCOFFEE analysis ([Fig. 4B](#)) revealed closest sequence relation between SOCS-regions of WDR13 and EA, while SOCS-region in VHL and SOCS1 proteins appear to possess relatively distant relationship with that of WDR13. Although sequence relationships between these SOCS-regions are somewhat varied between the outputs from different programs, all the analyses showed close relation of SOCS-region of WDR13 with either EA or VHL protein (shown by * in Fig. 4B, next page). In line with the SOCS sequence alignment analyses, we propose functional similarity between WDR13, EA and/or VHL. These results indeed reinforce our putative functional prediction of WDR13 arrived at from protein localization (experimental as well as bioinformatics analysis). Therefore, it appears that 393- and 485-aa-WDR13 proteins might perform suppression of cytokine signaling, transcriptional activation of certain genes and polyubiquitination of specific substrates in their degradation path way. By analogy and comparison with VHL and EA, mentioned above, we predict that 393-aa-WDR13 functions at the cytoplasmic level by binding to JAK/STAT-like factors and participates in the suppression of cytokine signaling. It might also participate in multi-protein (Elongin B-C) complexes, bind to the post-translationally modified substrates, migrate to the nucleus and involve in transcriptional activation of either targeted or a specific group of genes, similar to EA and VHL. Further, 393-aa-WDR13 protein might also take part in the E3 ubiquitin ligase ECS complex and involve in polyubiquitination and degradation of specific proteins in their degradation pathway. Presence of SOCS-motif in 393-aa-WDR13 protein and its localization, both in cytoplasm and nucleus, support these functional predictions.

BLAST analysis of WDR13 sequence in the database of interacting proteins: When we submitted amino acid sequence of 485-aa-WDR13 to BLAST search in the database of interacting proteins, very interestingly the analyses revealed that WDR13 is homologous to hCdc4 (gi|109659002), a protein known to participate in SCF complex, and is involved in ubiquitination of proteins in their degradation pathway. This is again in perfect agreement with the presence of functional SOCS-motif in WDR13 (as discussed earlier) and the instability index based functional clustering of WDR13, with F-box ubiquitination proteins, observed herein. Human Cdc4 expresses Fbw7 (F-box and 7 WD-repeat-containing) protein in three isoforms, which localize into the cytoplasm, nucleus and nucleolus. However, WDR13 has been reported to be expressed only in two isoforms [[1 & 2](#)]. Interestingly, careful re-look into the earlier report [[1](#)] revealed the presence of an EST clone of WDR13 expressing a 363-aa-WDR13, which was devoid of the N-terminal 122-aa portion of 485-aa-WDR13.

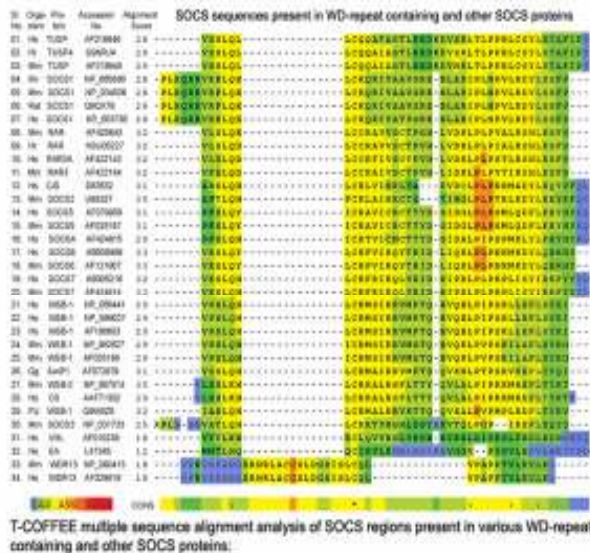


Fig 4A

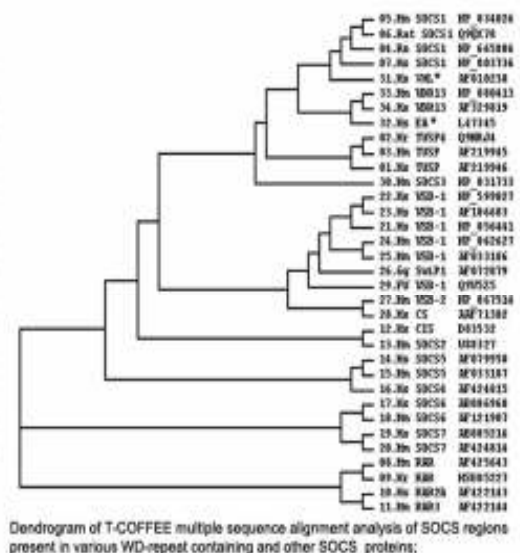


Fig 4B

Click on image and enlarge

This implies that analogous to Fbw7, WDR13 is also expressed in three isoforms: 485-aa, 393-aa (devoid of N-terminal 92-aa) and 363-aa (devoid of N-terminal 122-aa) portions of full length protein. It is equally interestingly to note that (i) all three WDR13 isoforms possess initiator methionine, suggesting their expression from different ORFs and (ii) SOCS motif represented by amino acids 175-217, present in full-length WDR13 is available in all three isoforms of WDR13. These observations strongly suggest that WDR13 (consisting the SOCS domain) participates in ECS ubiquitin ligase, analogous to the known SCF (functional homolog of ECS) ubiquitin ligase consisting the F-box domain.

Bioinformatics analyses for comparison of sequence domains in WDR13 with those in known ubiquitin ligase proteins: Analysis of large number of WD-repeat of SOCS protein sequences for their PII revealed that these proteins form functional clusters according to PII, and WDR13 belongs to either M-phase checkpoint or F-box ubiquitination proteins. BLAST analyses of protein sequence in the database of interacting proteins along with cellular localization studies suggested that WDR13 is homologous to Fbw7, participating in SCF ubiquitin ligase. Additionally WDR13 possesses a SOCS-homologous sequence, which apparently enables it to participate in ECS complexes known to involve in polyubiquitination of proteins in their degradation pathway. Detection of SOCS sequence region in WDR13 with a score relatively less than the required threshold need not cast any doubts about its functional abilities, as similar analysis of reported [15] SOCS proteins (like EA and MUF-1), employing EMBL SMART failed to detect presence of SOCS/SOCS-homologous sequence regions even with lower threshold level scores. For arriving at more plausible evidence of the participation of WDR13 in ECS ubiquitin ligase complexes, we attempted to compare the functional domains present in WDR13, like substrate recognition, protein-protein interaction, phosphorylation/glycosylation sites etc., with those present in known proteins like Fbw7 and Cdc20, participating in SCF and ECS ubiquitin ligase complexes. We performed Eukaryotic Linear Motif (ELM) and EMBL SMART analyses on the full-length WDR13 (gi|123283851), hCDC4 [Fbw7a (gi|109659002)] and mouse Cdc20 (gi|8885513). The data of the analyses is shown in [Table II](#). It is interesting to note that SMART identified a SOCS-homologous sequence in WDR13, although with a score less significant than the required

threshold and an F-box domain in Fbw7, while Cdc20 possessed neither of these domains. Further, SMART identified WD-repeat domains in all the three proteins. ELM analysis showed the presence of an APC/C-interacting (D_box_1) domain in both WDR13 and Fbw7, while Cdc20 showed presence of an alternative APC/C-interacting 'KEN_box_2' domain. Essentially all the domains found in Fbw7 and Cdc20 are present in WDR13. Particularly there is a one-to-one correlation between the sequence domains present in WDR13 and those present in Fbw7 and Cdc20. In addition, all the three proteins showed several phosphorylation sites and protein-protein interaction motifs. In view of these similarities with Fbw7 and Cdc20 and particularly because ELM looks for short sequence domains as the presence of several of such domains is necessary for a protein to participate in such multi-protein complexes; WDR13 has a potential to participate in the ECS complex. Thus, the existence of two proteins possessing similar domain-architecture reinforces the confidence in predicting the physiological function of proteins. We, therefore, conclude that Fbw7 and WDR13 are F-box and SOCS respectively, containing SCF and ECS ubiquitin ligase complexes and apparently both associate with the ubiquitin-proteasome system (UPS), involved in degradation of protein in their degradation pathway.

Recently SCF and ECS complexes participating in the Anaphase Promoting Complex (APC/C, the proteasome complex) were reported to regulate each other and the cell-cycle [35]. This implies that both SCF and ECS ubiquitin ligases are involved in ubiquitination dependent protein degradation in eukaryotic cells and are functionally complimentary. As protein degradation plays a key regulatory function in cells under several physiological conditions like differentiation, development, stress response and disease, SCF and ECS appear to determine the fate of cells under these conditions. Knock down of Fbw7 was observed to cause enlarged cells/nuclei [36], polyploidy, multiple centrosomes, reduced cell division and apoptosis [37]. Thus, loss of Fbw7 function appears to cause profound effect on cells, although cells manage to survive [37]. In view of these observations we believe that during gene knock down, loss of Fbw7 might lead to up-regulation of functionally complementary WDR13 due to which the cells do survive, although regulation of their division is lost. Therefore, knock-down of Fbw7 and WDR13 together in cells would perhaps be lethal for cell survival and this might possess interesting application in cancer cell biology. Evolutionarily conserved instability index determines the fate of unstable proteins (like WDR13), which in turn determine the fate of eukaryotic cells. Thus, WDR13 appears to be one of the key molecules involved in vital functions in eukaryotic cells.

CONCLUSION

From the results and literature reports discussed above, analysis of the WDR13 protein sequence, functional motifs present in and localization of WDR13, we hypothesize that by virtue of the presence of a SOCS-homologous region, WDR13 participates in the ECS ubiquitin ligase, associated with Ubiquitin Proteasome System involved in the degradation of proteins in their degradation pathway. Further, WDR13 proteins might also be involved in the suppression of cytokine signaling and transcriptional activation of certain genes.

REFERENCES

1. Suresh A, Shah V, Rani, DS, Singh BN, Prasad, GU, Subramanian S, Kumar S, Singh L. A mouse gene encoding a novel member of the WD family of proteins is highly conserved and predominantly expressed in the testis (*Wdr13*). *Mol Reprod & Dev* 2005; 72: 299-310
2. Singh BN, Amritha Suresh, Umaprasad G, Subramanian S, Sultana M, Sandeep Goel, Satish Kumar, Lalji Singh. A highly conserved human gene encoding a novel member of WD-repeat family of proteins (WDR13). *Genomics* 2003; 81: 315-328
3. van der Voom L, Ploegh HL. The WD-40 repeat. *FEBS Lett* 1992; 307: 131-134
4. Smith TF, Gaitatzes C, Saxena K, Neer EJ. The WD repeat: a common architecture for diverse functions. *TIBS* 1999; 24: 181-185
5. Duronio RJ, Gordon JL, Boguski MS. Comparative analysis of the beta transducin family with identification of several new members including PWP1, a nonessential gene of *Saccharomyces cerevisiae* that is divergently transcribed from NMT1. *Proteins* 1992; 13: 41-56
6. Neer EJ, Schmidt CJ, Nambudripad R, Smith TF. The ancient regulatory-protein family of WD-repeat proteins. *Nature* 1994; 371: 297-300
7. van Nocker S, Ludwig P. The Wd-repeat protein super family in Arabidopsis: conservation and divergence in structure and function. *BMC Genomics* 2003; 4: 50
8. Guruprasad K, Reddy BVB, Pandit MW. Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence. *Protein Engineering* 1990; 4: 155-161
9. Bachmair A, Finley D, Varshavsky A. In vivo half-life of a protein is a function of its amino-terminal residue. *Science* 1986; 234: 179-186
10. Gonda DK, Bachmair A, Wunning I, Tobias JW, Lane WS, Varshavsky A. Universality and structure of the N-end rule. *J Biol Chem* 1989; 264: 16700-16712
11. Tobias JW, Shrader TE, Rocap G, Varshavsky A. The N-end rule in bacteria. *Science* 1991; 254: 1374-1377
12. Suresh Kumar TK, Prasad VVH, Pandit MW. Origin of multiple bands of proteins on sodium dodecyl sulfate-polyacrylamide gel electrophoresis – Intermolecular disulfide cross-linking due to the presence of oxidizing components in the reducing agent. *Journal of Biochemistry and Biophysics* 1994; 28: 243-247
13. Boulikas T. Nuclear localization signals (NLS). *Crit Rev Eukaryot Gene Expr* 1993; 3:193-227
14. Hicks GR, Raikhel NV. Nuclear localization signal binding proteins in higher plant nuclei. *Proc Natl Acad Sci USA* 1995; 92: 734-738

15. Kile BT, Schulman BA, Alexander WS, Nicola NA, Martin HM, Hilton DJ. The SOCS-box: a tale of destruction and degradation. *Trends Biochem Sci* 2002; 27: 235-241
16. Ihle JN, Witthuhn BA, Quelle FW, Yamamoto K, Thierfelder WE, Kreider B, Silvennoinen O. Signaling by the cytokine receptor superfamily: JAKs and STATs. *Trends Biochem Sci* 1994; 19: 222-227
17. Darnell JE Jr, Kerr IM, Stark GR. Jak-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science* 1994; 264: 1415-1421
18. Ihle JN, Kerr IM. Jaks and Stats in signaling by the cytokine receptor superfamily. *Trends Genet* 1995; 11: 69-74
19. Marine JC, Topham DJ, McKay C, Wang D, Parganas E, Stravopodis D, Yoshimura A, Ihle JN. SOCS1 Deficiency Causes a Lymphocyte-Dependent Perinatal Lethality. *Cell* 1999; 98: 609-616
20. Naka T, Tsutsui H, Fujimoto M, Kawazoe Y, Kohzaki H, Morita Y, Nakagawa R, Narazaki M, Adachi K, Yoshimoto T, Nakanishi K, Kishimoto T. SOCS-1/SSI-1-deficient NKT cells participate in severe hepatitis through dysregulated cross-talk inhibition of IFN-gamma and IL-4 signaling in vivo. *Immunity* 2001; 14: 535-545
21. Horvat S, Medrano JF. Lack of SOCS2 expression causes the high-growth phenotype in mice. *Genomics* 2001; 72: 209-212
22. Starr R, Willson TA, Viney EM, Murray LJ, Rayner JR, Jenkins BJ, Gonda TJ, Alexander WS, Metcalf D, Nicola NA, Hilton DJ. A family of cytokine-inducible inhibitors of signaling. *Nature* 1997; 387: 917-921
23. Endo TA, Masuhara M, Yokouchi M, Suzuki R, Sakamoto H, Mitsui K, Matsumoto A, Tanimura S, Ohtsubo M, Misawa H, Miyazaki T, Leonor N, Taniguchi T, Fujita T, Kanakura Y, Komiyama S, Yoshimura A. A new protein containing an SH2 domain that inhibits JAK kinases; *Nature* 1997; 387: 921-924
24. Naka T, Narazaki M, Hirata M, Matsumoto T, Minamoto S, Aono A, Nishimoto N, Kajita T, Taga T, Yoshizaki K, Akira S, Kishimoto T. Structure and function of a new STAT-induced STAT inhibitor; *Nature* 1997; 387: 924-929
25. Yoshimura A, Ohkubo T, Kiguchi T, Jenkins NA, Gilbert DJ, Copeland NG, Hara T, Miyajima A. (1995) A novel cytokine-inducible gene CIS encodes an SH2-containing protein that binds to tyrosine-phosphorylated interleukin 3 and erythropoietin receptors. *EMBO J* 1995; 14: 2816-2826
26. Minamoto S, Ikegame K, Ueno K, Narazaki M, Naka T, Yamamoto H, Matsumoto T, Saito H, Hosoe S, Kishimoto T. Cloning and functional analysis of new members of STAT induced STAT inhibitor (SSI) family: SSI-2 and SSI-3. *Biochem Biophys Res Commun* 1997; 237, 79-83
27. Hilton DJ. Negative regulators of cytokine signal transduction. *Cell Mol Life Sci* 1999; 55: 1568-1577

28. Zhang JG, Farley A, Nicholson SE, Willson TA, Zugaro LM, Simpson RJ, Moritz RL, Cary D, Richardson R, Hausmann G, Kile BJ, Kent SB, Alexander WS, Metcalf D, Hilton DJ, Nicola NA, Baca M. The conserved SOCS-box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. *Proc Natl Acad Sci U S A* 1999; 96: 2071-2076
29. Kamura T, Sato S, Haque D, Liu L, Kaelin WG Jr, Conaway RC, Conaway JW. The Elongin BC complex interacts with the conserved SOCS-box motif present in members of the SOCS, ras, WD-40 repeat, and ankyrin repeat families. *Genes Dev* 1998; 12: 3872-3881
30. Kibel A, Iliopoulos O, DeCaprio JA, Kaelin WG Jr. Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C. *Science* 1995; 269: 1444-1446
31. Iwai K, Yamanaka K, Kamura T, Minato N, Conaway RC, Conaway JW, Klausner RD, Pause A. Identification of the von Hippel-lindau tumor-suppressor protein as part of an active E3 ubiquitin ligase complex. *Proc Natl Acad Sci USA* 1999; 96: 12436-12441
32. Pause A, Lee S, Worrell RA, Chen DY, Burgess WH, Linehan WM, Klausner RD. The von Hippel-Lindau tumor-suppressor gene product forms a stable complex with human CUL-2, a member of the Cdc53 family of proteins. *Proc. Natl. Acad. Sci* 1997; 94: 2156-2161
33. Lisztwan J, Imbert G, Wirbelauer C, Gstaiger M, Krek W. (1999) The von Hippel-Lindau tumor suppressor protein is a component of an E3 ubiquitin-protein ligase activity. *Genes Dev* 1999; 13: 1822-1833
34. Patton EE, Willems AR, Tyers M. Combinatorial control in ubiquitin-dependent proteolysis: don't Skp the F-box hypothesis. *Trends Genet* 1998; 14: 236-243
35. Vodermaier, H.C. (2004) APC/C and SCF: Controlling each other and the cell-cycle; *Current Biology* 14, R787-R796;
36. Welcker, M., Orian, A., Grim, J.E., Eisenman, R.N., and Clurman, B.E. (2004) A nuclear isoform of the Fbw7 ubiquitin ligase regulates c-Myc and cell size; *Current Biology* 14, 1852-1857;
37. Nakayama, K., Nagahama, H., Minamishima, Y.A., Matsumoto, M., Nakamichi, I., Kitagawa, K., Shirane, M., Tsunematsu, R., Tsukiyama, T., Ishida, N., Kitagawa, M., Nakayama, Kei-ichi, and Hatakeyama, S. (2000) Targeted disruption of Skp² results in accumulation of cyclin E and p27^{kip1}, polyploidy and centrosome overduplication. *EMBO J.* 19, 2069-2081

TABLE I

Functional clusters of unstable WD-repeat proteins with PII values between 40 and 52

M Phase checkpoint WD-repeat proteins			
GenBank Acc. No.	Reported / predicted / protein function	PII	Reference
AAK12629	Kinetochores-associated checkpoint protein [<i>Xenopus laevis</i>]	41.57	6
CAA21958	Possible TAF, expressing genes involved in progression through M phase of cell cycle [<i>Schizosaccharomyces pombe</i>]	42.90	16
NP_504039	Checkpoint kinase, transcripts for the two genes have 5' and 3' overlap [<i>C. elegans</i>]	43.57	19
NP_593798	Possibly mitotic checkpoint protein [<i>S.pombe</i>]	43.74	22
CAA20489	May be a TAF involved in gene expression, expression, in progression through M phase of cell cycle <i>Schizosaccharomyces pombe</i>	44.04	24
NP_564965	Similar to mitotic checkpoint protein in yeast/Arabidopsis BUB3; [<i>Homo sapiens</i>]	44.11	25
XP_341944	Similar to yeast BUB3 budding uninhibited by benzimidazole-3 homolog; mitotic phase checkpoint protein [<i>Rattus norvegicus</i>]	44.49	28
NP_033904	Mitotic checkpoint protein; similar to yeast BUB3; Takes part in control of cell cycle [<i>Mus musculus</i>]	45.48	35
NP_004716	Mitotic checkpoint protein; similar to yeast BUB3; budding uninhibited by benzimidazole [<i>Homo sapiens</i>]	47.37	47
NP_196888	Unknown; cell cycle switch protein [<i>Arabidopsis thaliana</i>]	47.93	48
F-box-Ubiquitination WD-repeat proteins			
NP_839891	F-box & WD-40 domain Protein 5 isoform 3; FBW5 [<i>Homo sapiens</i>]	42.27	10
NP_061871	F-box & WD-40 domain protein 5 isoform 1; FBW5 [<i>Homo sapiens</i>]	43.88	23
NP_958467	F-box and WD-40 domain protein1B [<i>Daino rerio</i>]	44.61	29
AAD08702	Receptor component of IkappaB, ubiquitin ligase containing F-box and WD-repeats [<i>Homo sapiens</i>]	45.07	32
XP_342057	Similar to F-box/WD-repeat protein 1A of beta-TrCP; (pIkappaB alpha-E3 receptor subunit) [<i>Rattus norvegicus</i>]	48.24	50
XP_220281	Similar to F-box & WD-40 domain protein 1B; [<i>Rattus norvegicus</i>]	48.47	52
BAA97451	Cell division control protein 20 (Cdc20) [<i>Mus musculus</i>]	48.70	54
NP_387449	F-box & WD-40 domain protein, Fbw; functions in phosphorylation dependent ubiquitination; [<i>Homo sapiens</i>]	48.84	56
AAI17247	F-box containing WD-40 domain protein [<i>Homo sapiens</i>]	50.84	65
Actin / Myosin-binding WD-repeat proteins			
NP_035909	Coronin; Actin-binding protein	40.24	1
NP_195164	Unknown; similar to Myosin heavy chain Kinase B [<i>Dictyostelium discoideum</i>]	40.92	3
P78774	Part of complex implicated in control of actin Poly-merization in cells; Binds to profiling [<i>S. pombe</i>]	42.93	17
Pre-mRNA splicing WD-repeat proteins			
P40968	Pre-mRNA splicing factor-PRP17; Cell division control point 40; [<i>PRP17_YEAST</i>]	41.05	4
NP_004688	U4/U6 small nuclear RNP, PrP4; involved in RNA splicing; contains motifs present in PrP18 / PrP4 [<i>Homo sapiens</i>]	46.20	38
NP_056975	Found in splicingosome 17, may involve in pre-mRNA splicing [<i>Homo sapiens</i>]	46.39	41
Locomotion involved WD-repeat proteins			
NP_491932	Enhancer of split groucho, Uncoordinated locomotion UNC-37; [<i>C.elegans</i>]	41.71	7
NP_703453	Unknown; Similar to Chlamydomonas reinhardtii flagellar Pf20 SW: [<i>CHLRE</i>]	42.23	9
NP_493279	Mec-8 gene; affects muscle and sensory neuron function; interacts with unc-52, smu-1 & smu-2 genes [<i>C. elegans</i>]	44.12	26
NP_083436	Unknown; PF20; sperm-associated WD-repeat protein [<i>Mus musculus</i>]	47.07	45
Transcription Associated / Factor - WD-repeat proteins			
CAA20489	May be a TAF involved in gene expression, in M phase of cell cycle [<i>S. pombe</i>]	44.04	24
S33263	Unknown; Drosophila TFIIID associated protein 44 [<i>D. melanogaster</i>]	29	27
NP_492169	TBP-Associated transcription Factor family member [<i>C. elegans</i>]	47.09	46
Note: PII: Protein instability index; Reference No. herein, refers to the number cited in the Supplementary Table IB			

TABLE II

Sequence domains identified in mouse WDR13, Cdc20 and human Cdc4 by ELM

Identified functional / sequence domain	Amino acid sequence and number in the protein sequence of			Comments about reported domain and protein function
	WDR13 (Wdr13) (gi 123283851)	CDC20(Cdc20) (gi 8885513)	hCDC4 (Cdc4) (gi 109659002)	
LIG_14-3-3_2/3	RSSRTL: 84-90 RTPTFP: 19-24 RAYSNS: 76-81	KVQTTP: 66-71 RVSSLS: 270-275 HVATLS: 301-306	HTHTNS: 137-142 HTSTVR: 460-465 RVVSGS: 473-578 HVVSGS: 553-558	14-3-3 proteins interaction motif 2; Consensus derived from natural interactors which do not exactly match the mode1 and mode2 ligands
LIG APCC-Dbox 1	RRSQLL: 33-38 RRQYLR: 51-56 LRGQLL: 57-62 SRTTLD: 86-91	-----	RRGELK: 366-371	An RXXL motif binds to either to Cdh1 / Cdc20 components of APC/C thereby targeting protein for destruction in a cell cycle dependent manner
LIG APCC_KE Nbox_2	-----	SKENQ: 96-100	-----	
LIG_Calther_C altBox_1	-----	LLQLD: 13-17 LQME: 219-223	LLALD: 250-254	Calthrin box motif found on cargo adaptor proteins, interacts with β -propeller structure located at N-terminus of Calthrin heavy chain
LIG_BRCT_BR CA1_1	FSFLF: 329-333	-----	ESDDF: 121-125 LSSPF: 157-161 YSLQF: 545-549	Phosphopeptide motif directly interacts with BRCT (carboxy-terminal) domain of the Breast Cancer Gene BRCA1 with low affinity
LIG_PDZ_3	DEAV: 172-175 NDIL: 229-232 AELL: 261-264 DESL: 464-467	PDRI: 172-175 GDIL: 214-217 GDYI: 226-229 NDNI: 329-332 KELI: 412-415 DETL: 464-467	QELL: 3-6 GEV: 57-60 QEWL: 235-238 PEKL: 247-250 LDEL: 253-256 DELI: 254-257 RDFI: 278-281 EEGI: 327-330 DEPL: 331-334	Class III PDZ domains binding motif
LIG_PP1	AVPRVRF: 175-181	AVKAWAWC: 357-364	-----	LIG_PP1, conserved (protein phosphatase1 catalytic subunit)-binding motif
LIG CYCLIN_1	RVLAL: 306-310 KRLVV: 344-348 KLLLY: 365-374	KILRL: 129-133 KVLV: 149-152 KELI: 412-415 RVLGL: 445-449	KLLAL: 249-253 KELAL: 286-290 KDLL: 299-302 RTL: 415-418 RTLKV: 441-445 RNLV: 658-661 KLLVL: 696-700	Substrate recognition site, interacting with cyclin, thereby increasing phosphorylation by cyclin / cdk complexes. Predicted protein also should have the MOD_CDK site. Also used by cyclin inhibitors
LIG_FHA_1/2	FRTQYIR: 26-32 KLTGRVL: 302-308 MATGKLT: 335-341	TPTKKEH: 106-112 GHTARVL: 441-447 DETLRLW: 464-470	THTNSVT: 138-144 TSTTGLV: 194-200 STTGLVP: 195-201 PTTFGDL: 206-212	Phosphothreonine motif binding a subset of FHA domains that have preference for an acidic amino acid at the Pt+3 position
LIG_TRAF6	-----	SGPGESGWA: 338-346	-----	TRAF6 binding site. Members of tumor necrosis factor receptor (TNFR) superfamily initiate intracellular signaling by recruiting C-domain of TNFR-associated factors through their cytoplasmic tails
LIG_WW_4	AYRTPT: 17-22 HEGSPV: 349-354	PAPSPM: 38-43 AGRTPG: 52-57 PGRTPG: 56-61 VQTTPS: 67-72	QLSSPF: 156-161 CSATPT: 202-207 FIHSPW: 346-351	Class IV WW domains interaction motif; phosphorylation-dependent interaction
LIG MAPK 1	RRQYLRL: 51-57 KLTKAKRLVV: 339-348	-----	-----	Molecules interacting with MAPK (MAPKs, substrates, phosphatases) carry docking

	REARDPSLLI: 365-374			motif that helps to regulate specific interaction in MAPK cascade. The classic motif approximates (R/K)xxx#x#, where # is a hydrophobic residue
LIG_SH2_GRB2	YSNS: 78-81	YQNR: 144-147 YLNL: 186-189	-----	GRB2-like Src Homology 2 (SH2) domains binding motif
LIG_SH2_SRC	-----	SSRKTCRY: 160-167	-----	Src-family, Src Homology 2 (SH2) domains binding motif
LIG_SH2_STA T3	YQLQ: 117-120	-----	YIRQ: 355-358	YXXQ motif found in cytoplasmic region of cytokine receptors binds the STAT3 SH2 domain
LIG_SH2_STA T5	YRR: 30-33 YLRL: 54-57	YIPS: 167-170 YLNL: 186-189 YLWN: 207-210 YISS: 228-231 YLAV: 240-243	YTKT: 162-165 YTST: 193-196 YIRQ: 255-358	STAT5 Src Homology 2 (SH2) domain binding motif
LIG_SH3_1 (_3)	YRTPTFP: 18-24	APIPNAP: 18-24 QTTPSKP: 68-74 RYIPSLP: 166-172	SSIVDLP: 147-153 ITSVQPP: 225-231 MMQVIEP: 268-274 RRKVIKP: 338-344	Motifs recognized by class I SH3 domains or that with a non-canonical class I recognition specificity
LIG_NRBOX	CLNKLL: 377-383	DLHSLQ: 9-15	-----	Nuclear receptor box motif (LXXLL) confers binding to nuclear receptors
MOD CK1_1	SSRTTLD: 85-91 SRTTLDR: 86-02 SRGSYQL: 113-119 SKLTGRV: 301-307 SVFSFLF: 327-333 SPVTSIS: 352-358 SARSWVS: 358-364 SWVSREA: 361-367	SRKTCRY: 161-167 SSLSWNS: 272-278 SWNSYIL: 275-281 SSGSRSG: 282-288 SGHSQEV: 306-312	SVGSKRR: 7-13 SEYTSTT: 191-197 SDDSSRE: 128-134 SEYTSTT: 191-197 SAVTGKC: 407-413 SGSTDRT: 436-442 SLDTSIR: 558-564 SLTSGME: 582-588	Ck1 Phosphorylation Site
MOD CK2_1	VPTSVAV: 137-143 GDTSLSE: 150-156	RSGSQME: 83-89 GTPTKKE: 105-111 VGTSNAE: 243-249	RFISVDE: 83-89 SDDSSRE: 128-134 QSWSGPE: 242-248 VLSFLE: 291-297 SLTSGME: 582-588	CK2 phosphorylation site
MOD_CDK	-----	PAPSPMR: 38-44 AGRTPGR: 52-58 PGRTPGK: 56-62 VQTTPSK: 67-73 DRGTPTK: 103-109	FIHSPWK: 346-352	Substrate motif for phosphorylation by CDK
MOD_GSK3_1	GPLSEPGS: 67-74 YSNSIVRS: 78-85 IVRSSRTT: 82-89 PPGSVPT: 132-139 KGGSSKLT: 297-304 DMATGKLT: 334-341 HEGSPVTS: 349-356	RSHSAGRT: 48-55 AGRTPGRT: 52-59 PGRTPGKS: 56-63 KSSSKVQT: 62-69 EVASFLLS: 89-96 VLYSQKAT: 150-157 QKATPGSS: 154-161 TPGSSRKT: 157-164 TSHSARVS: 265-272 SWNSYILS: 275-282 ILSSGSRs: 280-287 ILATGGGT: 370-377 DVHSQVCS: 397-404	DEHTHTINS: 135-142 HTHTNSVT: 137-144 HTNSVTNS: 139-146 NSVTNSSS: 141-148 QLSSPFYT: 156-163 DHGSEVRS: 173-180 CKVSEYTS: 188-195 VPCSATPT: 200-207 LIDSCEPT: 256-263 FIHSPWKS; 346-353	GSK3 phosphorylation recognition site
MOD_PIKK_1	QFRYQYI: 25-31 RRRSQLL:	RSGSQME: 83-89 VLYSQKA:	PSSSQVD: 23-29 QNDSQQG: 69-75	(ST)Q motif which is phosphorylated by PIKK

	32-38	150-156 SGHSQEV: 306-312 QTFTQHQ: 349-355	CEPTQVK: 260-266 VWSSQMR: 424-430	family members
MOD_PKA_1/2	RRRSQ: 32-38 VRSSRTT: 83-89 SRTTLDR: 86-92 HRRSVSR: 108-114 SRGSYQL: 113-119 DRGSVFS: 324-330	DRGTPTK: 103-109 LRLSGKP: 131-137 SRKTCRY: 161-167	RRRTGGS: 12-18 RRITSVQ: 223-229 IRASNTK: 675-681	PKA is a protein kinase involved in cell signaling; PKA phosphorylation site
MOD_PLK	GDTSLSE: 150-156 RDPSELLI: 368-374	LDNSVYL: 202-208 WCPWQSN: 463-469	EEHTARNGEVVG-VEPRP: 50-66 GHVAAVR: 399-405 TDRTLKV: 439-445 RDATLRV: 479-485 LDTsirV: 559-565 ADSTVKI: 599-605 DDGTVKL: 642-648	Site phosphorylated by the Polo-like-kinase
MOD_ProDKin_1	AYRTPTF: 17-23 HEGSPVT: 349-355	PAPSPMR: 38-44 AGRTPGR: 52-58 PGRTPGK: 56-62 VQTPPSK: 67-73 DRGTPTK: 103-109 QKATPGS: 154-160 ILWSPHY: 405-411 LTMSPDG: 449-455	QLSSPFY: 156-162 CSATPTT: 202-208 FIHSPWK: 346-352 ELKSPKV: 369-375	Proline-Directed Kinase (e.g. MAPK) phosphorylation site in higher eukaryotes
MOD_TYR_ITM	LLYRVV: 382-387	-----	-----	Immuno-receptor tyrosine-based inhibitory motif. Phosphorylation of ITIM motif, found in cytoplasmic tail of some inhibitory receptors (KIRs) that bind MHC Class I, leading to recruitment and activation of a protein tyrosine phosphatase
TRG_ENDOCYTIC_2	YLRL: 54-57 YGPL: 66-69 YRVV: 384-387	YLNL: 186-189 YLAV: 240-243 YKEL: 411-414 YPTM: 430-433	YWRI: 310-313 YDFM: 519-522	Tyrosine-based sorting signal responsible for the interaction with mu subunit of AP (Adaptor Protein) complex
TRG_LysEnd-APsAcLL_1	RRSQLL: 33-38 RDPSELLI: 368-373 DPSLLI: 369-374	DLHSLL: 9-14 SVYLWN: 215-220	DFISLL: 279-284 EPKDLL: 297-302 EETKLL: 693-698 ETKLLV: 694-699	Sorting & internalisation signal found in cytoplasmic juxta-membrane region of type I transmembrane proteins. Targets them from the Trans Golgi Network to the lysosomal-endosomal-melanosomal compartments. Interacts with adaptor protein (AP) complexes

Supplementary Data: Table S

Protein Instability Index (PII) values of various WD-repeat proteins available in the literature evaluated employing the ProtParam program available at the ExpASY proteomics server, in the ascending order of PII:

Table SA: Stable WD-repeat proteins			
Sl. No.	Identified / predicted protein function	PII	Protein Acc. No:
1	Unknown; Kidney expressed WD-repeat protein; [<i>Danio rerio</i>]	19.48	AAQ94566
2	Unknown: Strain PCC6803 of Cyanobacterium Synechocystis [<i>Synechocystis sp. PCC 6803</i>]	23.47	P74598
3	Unknown; Translational machinery interacting WDR protein; Asc1p [<i>S. cerevisiae</i>]	25.36	NP_013834
4	Unknown; pts2-binding protein involved in import of 3-oxoacyl-CoA thiolase into peroxisomes. May direct thiolase into the peroxisomes by shuttling between cytosol and peroxisomal membranes [<i>PEX7_YEAST</i>]	28.96	P39108
5	Unknown; Translated from <i>Drosophila melanogaster</i> Sop2 mRNA [<i>Drosophila melanogaster</i>]	29.32	AAF44823
6	Will dye slowly protein [<i>Drosophila melanogaster</i>]	29.33	Q9V3J8
7	Unknown; WD repeat domain 48 [<i>Homo sapiens</i>]	31.80	NP_065890
8	Unknown; This variant (2) has an alternate 3' sequence, compared to variant 1. Shorter isoform with distinct C-terminus compared to isoform 1 [<i>Homo sapiens</i>]	32.22	NP_653175
9	Part of a complex implicated in control of actin polymerization in cells (By similarity) [<i>ARC1A_HUMAN</i>]	32.31	Q92747
10	Unknown; Required for pre-mRNA splicing in vivo; Homologous to Arabidopsis thaliana pleiotro-pic regulators PRL1 and PRL2 [<i>S. cerevisiae</i>]	33.13	NP_015174
11	Embryonic ectoderm development isoform b;Protein-1 associating with integrin cytoplasmic tails [<i>Homo sapiens</i>]	33.89	NP_694536
12	Unknown; WD repeat domain 20 isoform 1; Encodes longest isoform than other transcripts [<i>Homo sapiens</i>]	34.57	NP_851808
13	Platelet-activating factor 1B α ; acetylhydrolase isoform; Lissen-cephaly-1 protein [<i>Mus musculus</i>]	34.71	NP_038653
14	Unknown; SOCS box-containing WD repeat protein SWiP-1 isoform-3; WSB1 protein [<i>Homo sapiens</i>]	34.79	NP_599026
15	Unknown; Similar to YMR116C of [<i>S. cerevisia</i>]	36.02	CAB11079
16	Unknown; WAIT-1; specifically interacts with cytoplasmic tails of beta7-integrins [<i>Homo sapiens</i>]	36.22	AAC68675
17	Unknown; SOCS box-containing WD protein SWiP-1 isoform 2; WSB1 protein; [<i>Homo sapiens</i>]	36.36	NP_599027
18	Transducin (beta)-like 2 isoform 1; Williams-Beuren syndrome chromosome region 13 [<i>Homo sapiens</i>]	36.44	NP_036585
19	Transcriptional modulator with roles in meiosis; silences UME1; negative regulator of meiosis [<i>UME1_YEAST</i>]	36.97	Q03010
20	SOCS box WD-repeat protein SWiP-1 isoform 1; WSB1 protein; WD SOCS-box protein 1 [<i>Homo sapiens</i>]	37.30	NP_056441
21	Retinoblastoma associated p48, synthetic multivulva protein, LIN-35 binding NURF, abnormal cell Lineage LIN-53 [<i>C. elegans</i>]	37.41	NP_492552
22	UNR-interacting protein; WDR protein; MAP activator	37.63	Q9Y3F4
23	Unknown; Smu-1 Homolog of <i>C. elegans</i> ; Similar to brain-enriched WDR protein [<i>Rattus norvegicus</i>]	38.11	NP_476543
24	WDR-5; GTP-binding protein; Transducin beta chain 5 [<i>Mus musculus</i>]	38.14	P54314
25	CS box-containing WD protein [<i>Homo sapiens</i>]	38.27	AAF71302
26	PAK1 / PLC-interacting protein 1 [<i>Homo sapiens</i>]	38.32	NP_060376
27	Pleiotropic regulator-1; Necessary for spliceosome assembly in pre-mRNA splicing; Binds to CDC5L; interacts with spliceosome [<i>PLRG1_HUMAN</i>]	38.40	O43660
28	Unknown; Chromosome 5 genomic sequence, [<i>Arabidopsis thaliana</i>]	38.71	NP_201106
29	Unknown; WDR protein with SOCS-box1; WSB-1 [<i>Mus musculus</i>]	38.86	NP_062627
30	Unknown; WDR protein tup1-fission yeast(<i>S. pombe</i>)	39.23	T38992
31	Unknown; Retinoblastoma binding protein-7; RbAp46; histone acetyl-transferase type B subunit 2;Interact with BRCA1 tumor-suppressor gene; [<i>Homo sapiens</i>]	39.44	NP_002884
32	PF20; sperm-associated RNA-processing WDR protein-29; Pleiotropic regulator-1; [<i>Homo sapiens</i>]	39.83	NP_078808
33	Unknown; WSB-2; WD-repeat & SOCS-box-2 containing protein [<i>Mus musculus</i>]	39.88	NP_067514

Table SB: Unstable WD-repeat proteins			
Sl. No.	Identified / predicted WD-repeat protein function	PII	Protein Acc. No:
1	Coronin; Actin-binding protein [<i>Mus musculus</i>]	40.24	NP_035909
2	Function unknown; WD-repeat protein [<i>Homo sapiens</i>]	40.31	NP_005819
3	Unknown; similar to Myosin heavy chain kinase B; [<i>Dictyostelium discoideum</i>]	40.92	NP_195164
4	Pre-mRNA splicing factor-PRP17; Cell division control point 40; [PRP17_YEAST]	41.05	P40968
5	Putative PAK inhibitor SKB15 [<i>Mus musculus</i>]	41.12	AAK51599
6	Kinetochores-associated checkpoint protein [<i>Xenopus Laevis</i>]	41.57	AAK12629
7	Enhancer of split groucho, UnCoordinated locomotion UNC-37; [<i>C.Elegans</i>]	41.71	NP_491932
8	Regulates steps in Dnm1p dependent process of mitochondrial fission [<i>S. cerevisiae</i>]	42.01	NP_012423
9	Unknown; Similar to Chlamydomonas reinhardtii flagellar Pf20 SW [PF20_CHLRE]	42.23	NP_703453
10	F-box and WD-40 domain protein 5 isoform 3; FBW5 [<i>Homo sapiens</i>]	42.27	NP_839891
11	Wdr6 protein [<i>Mus musculus</i>]	42.28	AAH35951
12	hP1P1; Regulation of p21-activated kinase (PAK), [<i>Homo sapiens</i>]	42.30	AAK57477
13	Unknown; WD-repeat & SOCS-1 (WSB1) [WSB1_FUGRU]	42.41	Q9W5Z5
14	Unknown; WD repeat and FYVE domain containing 2 [<i>Danio rerio</i>]	42.70	AAQ94581
15	Chromatin assembly factor 1, P55 subunit (CAF-1 P55); NURF [<i>Drosophila melanogaster</i>]	42.72	Q24572
16	Possible TAF, expressing genes involved in progression through M phase of cell cycle [<i>S. pombe</i>]	42.90	CAA21958
17	Part of a complex implicated in control of actin polymerization in cells. Interacts with profilin [<i>S. pombe</i>]	42.93	P78774
18	Function unknown; WD-repeat protein homolog to RBAP/MSI [<i>Zea mays</i>]	42.97	AAF97517
19	CHECKPOINT KINASE, transcripts for two genes have a 5' 3' overlap [<i>C. elegans</i>]	43.57	NP_504039
20	Cockayne syndrome 1 protein [<i>Homo sapiens</i>]	43.58	NP_000073
21	Unknown; Triple A syndrome is caused by mutations in this protein, AAAS [<i>Homo sapiens</i>]	43.73	CAC17465
22	Possibly mitotic checkpoint protein [<i>Sizosaccharomyces pombe</i>]	43.74	NP_593798
23	F-box & WD-repeat protein 5 isoform 1; FBW5 [<i>Homo sapiens</i>]	43.88	NP_061871
24	May be TAF involved in gene expression, in progression through M phase of cell cycle [<i>Schizosaccharomyces pombe</i>]	44.04	CAA20489
25	Similar to mitotic checkpoint protein Arabidopsis; and yeast BUB3 [<i>Homo sapiens</i>]	44.11	NP_564965
26	mec-8 gene of <i>C. elegans</i> ; affects muscle & sensory neuron function; interacts with 3 other genes, unc-52, smu-1 & smu-2; contains an ER membrane domain; Predicted to localize in cytoplasm [<i>C. elegans</i>]	44.12	NP_493279
27	<i>Drosophila</i> TFIIID associated protein [<i>Drosophila melanogaster</i>]	44.29	S33263
28	Similar to yeast BUB3, uninhibited by benzimidazole3, mitotic checkpoint protein [<i>Rattus norvegicus</i>]	44.49	XP_341944
29	F-box and WD-40 domain protein 1B [<i>Danio rerio</i>]	44.61	NP_958467
30	Maternal Effect Sterile MES-6, WD repeat protein [<i>C. elegans</i>]	44.66	NP_872079
31	Binds to N-terminal PTS2-type peroxisomal targeting signal & plays an essential role in peroxisomal protein import. Interacts with PEX5 [<i>Cricetus griseus</i>]	44.85	Q8R537
32	Receptor component of IkappaB, ubiquitin ligase; contains F-box & WD-repeat [<i>Homo sapiens</i>]	45.07	AAD08702
33	p21-activated kinase-interacting protein-1 [<i>Mus musculus</i>]	45.28	AAL40652
34	Unknown [<i>Gossypium hirsutum</i>]	45.34	AAM95645
35	Mitotic checkpoint protein; similar to yeast BUB3; involved in control of Cell cycle [<i>Mus musculus</i>]	45.48	NP_033904
36	Function unknown; Wdr4-prov protein [<i>Xenopus laevis</i>]	45.94	AAH43907
37	Regulates activity of ubiquitin ligase; confers substrate specificity to anaphase promoting complex during anaphase and telophase until onset of next S phase; third step in ubiquitin conjugation [FZR_MOUSE]	46.19	Q9R1K5
38	U4/U6 small nuclear RNP, PrP4; involved in RNA processing / splicing; contains motifs present in PrP18 & PrP4 [<i>Homo sapiens</i>]	46.20	NP_004688
39	Regulates activity of ubiquitin ligase; confers substrate specificity to anaphase promoting complex during anaphase and telophase until onset of next S phase; third step in ubiquitin conjugation [FZR_HUMAN]	46.28	Q9UM11
40	Involved in peroxisomal targeting and genome stability; microtubule integrity; Predicted to involve in mRNA maturation and TOR signaling pathway [<i>S. cerevisiae</i>]	46.34	CAB57925
41	Found in splicingosome 17, may involve in pre-mRNA splicing [<i>Homo</i>	46.39	NP_056975

	<i>sapiens</i>]		
42	Unknown; similar to Fzr1 (GI: 6463679) { <i>Homo sapiens</i> } & Srw1 [<i>S pombe</i>]	46.69	NP_192929
43	Unknown; WDR13; Possesses low threshold SMART detected SOCS [<i>Homo sapiens</i>]	46.72	AAH02507
44	Unknown; similar to hypothetical yeast YN57 WD-repeat protein [<i>Encephalitozoon cuniculi GB-M1</i>]	46.77	CAD25169
45	PF20; sperm-associated WD-repeat protein [<i>Mus musculus</i>]	47.07	NP_083436
46	TBP-Associated transcription Factor family member [<i>C. elegans</i>]	47.09	NP_492169
47	Mitotic checkpoint protein; Similar to yeast BUB3; budding uninhibited by benzimidazoles, [<i>Homo sapiens</i>]	47.37	NP_004716
48	Unknown; Cell cycle switch protein [<i>Arabidopsis thaliana</i>]	47.93	NP_196888
49	Unknown; may regulate meiosis; Strong expression in brain, liver, and testis. Also expressed in kidney and spleen; may have a role in development of mental symptoms in severe cases of myotonic dystrophy	47.99	Q090191
50	Similar to F-box/WD-repeat protein 1A of beta-TrCP; (pIkappaBalpha-E3 receptor subunit) [<i>Rattus norvegicus</i>]	48.24	XP_342057
51	Damage and replication checkpoint control protein - in [<i>Schizosaccharomyces pombe</i>]	48.37	S67437
52	Similar to F-box & WD-40 domain protein 1B; [<i>Rattus norvegicus</i>]	48.47	XP_220281
53	Unknown [<i>Arabidopsis thaliana</i>]	48.69	NP_196473
54	Cell division control protein 20 (Cdc20) [<i>Mus musculus</i>]	48.70	BAA97451
55	Unknown; WDR13; Memory related protein; [<i>Mus musculus</i>]	48.74	NP_080413
56	Unknown	48.80	NP_387510
57	F-box & WD-repeat protein, Fbw; functions in phosphorylation dependent ubiquitination; [<i>Homo sapiens</i>]	48.84	NP_387449
58	Unknown; Dystrophia myotonica-containing WD repeat motif [<i>Mus musculus</i>]	49.17	NP_034188
59	Unknown; WDR26 protein [<i>Mus musculus</i>]	49.17	AAH20044
60	Neurogenesis factor. May involve in segrregation of Neuroectoderm. May interact with notch and delta [<i>Drosophila melanogaster</i>]	50.45	P16371
61	Unknown; DMWD protein [<i>Homo sapiens</i>]	51.54	AAH19266
62	Unknown; WD repeat domain 21 [<i>Danio rerio</i>]	52.34	AAQ98008
63	Unknown; Nucleoporin SEH1-like (SEC13-like protein) [<i>Homo sapiens</i>]	52.72	Q96EE3
64	Chromatin assembly, chromosome segregation [<i>Schizosaccharomyces pombe</i>]	56.29	O14021
65	F-box & WD repeat domain containing 7 [<i>Homo sapiens</i>]	50.84	AAI17247

Protein Instability Index (PII) values of various WD-repeat proteins available in the literature evaluated employing the ProtParam program available at the ExpASY proteomics server, in the ascending order of PII