The WD40 Repeat Protein Mutations: Genetics, Molecular Mechanisms and Therapeutic Implications

Caiguo Zhang*

Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO, USA

Abstract

Eukaryotes contain numerous WD40 repeat proteins which perform diverse functions in signal transduction, gene transcriptional regulation, vesicular trafficking, cytoskeleton assembly, genome stability maintenance and cell cycle control. Importantly, mutations of WD40 repeat proteins are genetically associated with diseases such as Lissencephaly, Cockayne syndrome, cancer, Parkinson’s disease, Chediak-Higashi syndrome, amyotrophic lateral sclerosis, cardiovascular diseases and colorectal neoplasms. Over the past years, great progresses have been made in understanding the molecular mechanisms and therapeutic strategies of these diseases. This review briefly provides current perspectives on WD40 repeat protein functions, and mainly summarizes the most recent understandings on molecular basis of WD40 repeat proteins involved in diseases and their therapeutic implications.

Keywords: WD40 repeat protein; Lissencephaly; Cockayne syndrome; Cancer; Parkinson’s disease

Introduction

The main functions of WD40 repeat proteins

WD40 repeat proteins are characterized by the presence of repeating units with 44–60 residues that ended with tryptophan (W) and aspartate (D) [1]. Structurally, the WD40 domain exhibits a β-propeller architecture, comprising seven blades in most cases [2]. WD40 repeat-containing proteins are widely distributed in eukaryotic organisms, for instance, there exist 349 predicted WD40 repeat-containing proteins in humans [3], and more than 200 putative WD40 repeat-containing proteins in plants such as Arabidopsis thaliana, Oryza sativa and Setaria italica [2,4,5]. In general, the basic function of WD40 repeat-containing proteins is to serve as a rigid platform for protein interactions [6]. Importantly, they can form E3 ligases to coordinate downstream events, such as ubiquitination and histone methylation [6,7]. WD40 repeat-containing proteins are involved in many biological processes such as apoptosis, cell death, DNA damage and repair, cell cycle control [2,6]. As such, mutations of WD40 proteins always lead to severe health problems and are genetically associated with a variety of diseases such as Lissencephaly [8], Cockayne syndrome [9], different types of cancer [10] and Parkinson’s disease [11] (Figure 1). However, the molecular mechanisms of these diseases caused by WD40 repeat protein mutations are extremely different. I’ll emphasize these WD40 proteins and their molecular functions in the following contents.

WD40 protein and lissencephaly

Lissencephaly is a rare brain formation disorder caused by defective neuronal migration [12,13]. It has been identified that malformations in five proteins, including LIS1 (lissencephaly-1) [14], 14-3-3ε [15], DCX (double cortin) [16], RELN (reelin) [17], and ARX (aristaless-related homeobox protein) [18], can lead to Lissencephaly. Of which LIS1 is a WD40 repeat-containing protein and its amino acid sequences share significant homology to βsubunits of heterotrimeric G proteins (Figure 2) [14]. The expression of LIS1 is mainly in fetal and adult brain, implying its critical roles in brain development [13]. Interestingly, LIS1 is a conserved protein with only one amino acid difference between the human and mouse versions, and 42% homology to NudF (nuclear distribution F), an ortholog found in Aspergillus nidulans [19-21]. Mice deleted mLIS1 (mLIS1+/−) die soon after implantation, whereas heterozygous mLIS1 (mLIS1+/−) mice exhibit hippocampal, cortical and olfactory bulb disorganization [22,23]. The human LIS1 has been evidenced to interact with Platelet-Activating Factor (PAF) acetyl hydrolases, cytoskeletal proteins and homologues of nuclear distribution proteins such as NudC and NudE [23], involving in a variety of biological processes including nucleokinesis, somal translocation, cell motility, mitosis and chromosome segregation [13]. Studies have demonstrated that LIS1 functions via a signaling pathway that includes NudE, NudL, cytoplasmic dynein, dynactin, and cytoplasmic linker protein CLIP-170 in mammalian cells [13].

Generally, the defects of Lissencephaly cannot be reversed [24]. The therapeutic implications aim to support and comfort affected patients. For instance, gastrostomy tubes may be provided to patients who have difficulty feeding and swallowing [25]. For patients with an excessive accumulation of cerebrospinal fluid, or experiences hydrocephalus, a surgical procedure that shunts fluid away from the brain may be necessary.

WD40 protein and cockayne syndrome

Cockayne Syndrome (CS) is a genetic recessive disorder characterized by growth retardation, impairment of nervous system development, hypersensitivity to sunlight and premature aging [26,27]. CS can be resulted from mutations in either CSA gene (also named as ERCC8 gene) or CSB gene (also named as ERCC6 gene) [28]. In which, CSA gene encodes a WD40 protein (Figure 2) [29]. Different mutations with CSA gene have been identified to cause CS to date [28]. Generally, many of these mutations result in a malfunctioning CSA protein. Studies have demonstrated that mutated CSA protein results in disruption of DNA repair, which allows abnormalities to accumulate.
Figure 1: The main diseases associated with WD40 repeat protein mutations. WD40 repeat-containing protein mutations can cause numerous diseases such as Lissencephaly (mutated by LIS1), Cockayne syndrome (mutated by CSA), cancer (mutated by β-TRCP1, β-TRCP2, FBXW7, CIAO1 and PALB2), Parkinson's disease (mutated by LRRK2), Chediak-Higashi syndrome (mutated by CHS1), amyotrophic lateral sclerosis (mutated by APAF1), cardiovascular diseases (mutated by APAF1), colorectal neoplasms (mutated by APAF1), depressive disorder (mutated by APAF1) and tuberous sclerosis (mutated by FBXW5).

Figure 2: Illustration of functional domains in the WD40 repeat proteins. LH, LIS-homology; CC, coiled-coil; ARM, armadillo domain; ANK, ankyrin repeat domain; ARM, armadillo domain; LRR, leucine-rich repeat; ROC, Ras of complex proteins; COR, C-terminal of ROC; KIN, kinase; DD, dimerization domain; DNA, DNA-binding domain; ChaM, chromatin-association motif; MRG15, MORF4-related gene on chromosome 15; CARD, caspase recruitment domain; NB-ARC, nucleotide-binding adaptor shared by APAF-1, R proteins, and CED-4.
in DNA [30]. These abnormalities impair cell function and eventually lead to cell death in many organs and tissues, which possibly contributes to features of CS, such as growth failure and premature aging [31]. In Arabidopsis, two CSA-like proteins, namely, CSAat1A and CSAat1B can form an E3 ligase with CUL4 (CULLIN4) and DDB1A (UV-Damage DNA Binding protein 1A) to regulate substrates in response to DNA damage through transcription coupled repair (TCR) [32]. These results may provide insights into the molecular mechanisms of CS caused by CSA gene mutations.

At present, there is no cure for CS. The classic method can only prevent the progression of symptoms via a mixture of symptom management and treatments [33]. Patients with CS generally require multi-disciplinary care including genetics, neurology, pediatrics, dentistry, ophthalmology, dermatology, audiology and physical therapy [34,35].

**WD40 protein and parkinson’s disease**

Parkinson’s Disease (PD) is a neurodegenerative disease clinically characterized by movement impairments, bradykinesia, rigidity, and resting tremor [36,37]. In recent years, several proteins, including leucine-rich repeat kinase 2 (LRRK2 or dardarin), alpha-synuclein (SNCA), parkin (PRKN), PTEN-induced putative kinase 1 (PINK1), parkinson protein 7 (PARK7, also named as DJ-1), and ATPas type 13A2 (ATP13A2), have been identified to directly associate with PD [38,39]. The LRRK2 gene encodes a WD40 protein and its biological functions have been extensively studied (Figure 2) [40]. The human LRRK2 protein consists of multiple domains, including an armadillo domain, an ankyrin repeat (ANK) region, a Leucine-Rich Repeat (LRR) domain, an ROC (Ras of complex proteins) GTPase domain, a COR (C-terminal of ROC) dimerization domain, a kinase domain, and a WD40 domain (Figure 2) [41]. It has been shown that the WD40 domain is a determinant for LRRK2 physiological and pathological activities [42,43], and is required to stabilize the LRRK2 dimer and to execute LRRK2-associated kinase activity as well as neurotoxicity [42]. Strong genetic association indicates that the substitution of G2385R variant is a common risk factor for sporadic PD in Chinese Han and Korean ethnicity [44]. It has been demonstrated that the G2385R variant correlates with a reduced binding affinity of LRRK2 WD40 to synaptic vesicles [45].

Moreover, the deletion of the WD40 domain of zebrafish LRRK2 (zLRRK2) protein can cause Parkinsonism-like phenotypes, including loss of dopaminergic neurons in diencephalon and locomotion defects [46].

Parkinson’s disease treatment is still beset with difficulties, but medications, surgery and multidisciplinary management can relieve the symptoms. The main drugs used for treating PD are levodopa, dopamine agonists and MAO-B inhibitors [47,48]. Surgery and deep brain stimulation are also used for cure when medications are not enough to control symptoms [49]. Further, palliative care is provided to improve quality of life in the end stages of PD patients [50].

**WD40 proteins and cancer**

Multiple WD40 proteins such as β-TRCP1 (β-transducin repeat-containing protein 1, also known as FBXW1A, F-box/WD repeat-containing protein 1A) [51,52], β-TRCP2 (also known as FBXW11) [51,52], FBXW7 (also known as FBW7 and CDC4) [53,54], CIAO1 (cytosolic iron-sulfur protein assembly 1) [55] and PALB2 (partner and localizer of BRCA2) (Figure 2) [56], have been identified to involve in carcinogenesis. These WD40 proteins generally function as tumour suppressors, or show emerging roles in suppressing tumorigenesis [10].

The β-TRCP1 and β-TRCP2 proteins have cell type-dependent and context-dependent roles in governing tumorigenesis [10]. The expression of β-TRCP1 is highly induced in pancreatic cancer and hepatoblastoma biopsy samples [57]. Similarly, elevated β-TRCP2 gene expression has also been observed in prostate, breast and gastric cancers [58]. These results indicate that the induction of β-TRCP1 or β-TRCP2 could be a common trend in human cancers [10]. Biochemically, β-TRCP1 and β-TRCP2 are considered to have functions in the recognition and degradation of their substrates [10,58]. A variety of β-TRCP1 and β-TRCP2 ubiquitin substrates have been identified, including β-catenin, CDC25A (cell division cycle 25 homolog A), FBXOS (F-box only protein 5), VEGFR2 (vascular endothelial growth factor receptor 2), IκB (inhibitor of nuclear factor-κB), PDCD4 (programmed cell death protein 4) and DEPTOR (DEP domain-containing mTOR-interacting protein) [10].

FBXW7 protein is a well-established tumour suppressor and responsible for substrate recognition in an SCF (Skp1-Cullin-F-box protein)-type ubiquitin ligase complex [10,59]. Numerous cancer-associated mutations of FBXW7 are detected in ovarian, breast and colorectal cancer cell lines, implicating its potential role in tumorigenesis [59]. Importantly, approximately 6% of all primary human cancers contain FBXW7 mutations [59]. The most frequent FBXW7 mutations were identified in cholangiocarcinoma (35%) and cell acute lymphoblastic leukaemia (T-ALL; 30%). The WD40 repeats in FBXW7 are essential for binding substrates such as cyclin E, MYC (myelocytomatosis oncogene), JUN (jun proto-oncogene) and Notch [59], which are also involved in a wide range of human cancers. Generally, tumors caused by FBXW7 mutations associate with abnormal cyclin E abundance or activity, and excess cyclin E activity can result in genomic instability [59]. Both MYC and Notch have also been demonstrated to involve human haematopoietic cancers associated with FBXW7 mutations [59], although their respective contribution to tumorigenesis remains to be determined. Moreover, some studies have indicated that the losses of FBXW7 and p53 cooperatively cause genetic instability and carcinogenesis, which is possibly resulted from the fact that induction of cyclin E and MYC by FBXW7 loss triggers p53 activation [59].

The CIAO1 gene encodes a novel WD40 protein, which can specifically interact with the Wilms tumor suppressor protein 1 (WT1) both in vitro and in vivo [55]. Inactivation of WT1 causes Wilms’ tumor and the WT1 protein has been found to stabilizep53 and inhibit p53-mediated apoptosis [60]. Notably, the interaction of CIAO1 with WT1 results in a change in the mobility of WT1-DNA complex in vitro [55]. CIAO1 possibly functions to modulate the transactivation activity of WT1 and regulate the physiological functions of WT1 [55], which further affect p53 and results in tumorigenesis.

PALB2 is a WD40 protein that functions in genome maintenance and can bind to BRCA2 (breast cancer 2) in nuclear, permitting the stable intranuclear localization and accumulation of BRAC2 [56]. Mutations in PALB2 gene have an increased lifetime risk of developing breast, ovarian and other cancers [61]. The WD40 domain in PALB2 can directly and independently bind RAD51C (RAD51 homolog C) and BRCA2 [61]. It has been demonstrated that the large truncations of PALB2 abolish interactions with RAD51C and BRAC2, whereas the L939W, T1030I and L1143P mutants within WD40 domain are associated with altered patterns of direct binding to the RAD51C.
RAD51 and BRCA2 proteins in vitro [61]. Clinically, PALB2 has been developed to test gene families where pancreatic cancer occurs in multiple family members [62].

**WD40 proteins and chediak-higashi syndrome**

Chediak-Higashi Syndrome (CHS) is a recessive disease characterized by immune system deficiency, ocoulcutaneous albinism, blood clotting, tremors, difficulty with movement and balance [63]. Mutations in the CHS1 gene (also known as LYST) have been found to associate with CHS [64].CHS1 contains three recognized domains: an ARM/HEAT domain, a BEACH (Beige and Chediak) domain and seven WD40 repeats at the carboxyl terminus [63]. The current understanding of CHS1 function is mainly focused on regulating vesicle trafficking and determining lysosome-related organelle size [65]. Several proteins involved in vesicle trafficking have been identified, including hepatocyte growth factor-regulated tyrosine kinase substrate (HRS), 14-3-3 protein, casein kinase II and lysosomal trafficking regulator–interacting protein 5 (LIP5) [65,66].

There is no specific and efficient treatment for CHS. Patients with CHS are commonly treated with antibiotics to control recurrent infections [65]. Anti-viral drugs such as acyclovir have been tried during the end phase of the disease [67]. Bone marrow transplants may help treat the defects of immune system and appear to have been successful in some patients [68].

**WD40 proteins and other genetic diseases**

Additionally, some WD40 proteins such as APAF1 (apoptotic peptidase activating factor 1) and FBXW5 (F-box/WD repeat-containing protein 5) are also genetically linked to a variety of diseases [69-71]. APAF1 is a cytoplasmic protein and plays critical roles in the apoptosis regulatory network [72]. This protein contains a Caspase Recruitment Domain (CARD), an ATPase domain (NB-ARC), and a WD40 repeat domain [72,73]. The WD40 repeat domain normally keeps APAF1 in an auto inhibitory state [74]. APAF1 is activated in the presence of dATP or ATP when cytochrome c binds to the WD40 repeat, forming an oligomeric apoptosome [74]. The apoptosome can activate caspase-9 and induce its dimerization and subsequent autocatalysis [75]. The APAF1/caspase-9 complex then functions as a resistance-gain mechanism as well as amyotrophic lateral sclerosis [77], cardiovascular diseases [78], colorectal neoplasms [79], depressive disorder [80], kidney neoplasms [81], tuberous sclerosis [82] and multiple myeloma [83].

Tuberous sclerosis (TSC) is an autosomal dominant disease characterized by hamartoma formation in various organs [84]. Generally, TSC is caused by mutations of either TSC1 or TSC2 genes [85]. It has been reported that a WD40 protein, namely, FBXW5, can recruit TSC2 to the DDB1-CUL4-ROC1 E3 ubiquitin ligase, resulting in the subsequent polyubiquitination and degradation of TSC2 [85]. Overexpression of FBXW5 or CUL4A gene can promote TSC2 protein degradation, whereas depletion of FBXW5, DDB1, or CUL4A/B protein stabilizes TSC2 [85]. Moreover, the CUL4A-DDB1-FBXW5 E3 ligase is suggested to be responsible for loss of DLC (deleted in liver cancer 1) protein [86], which encodes a GTPase-activating protein. Suppression of FBXW5 expression can restore DLC1 protein level, resulting in DLC1-dependent lung cancer cell growth suppression [86].

**Conclusion**

The clinical significance of WD40 repeat proteins has become increasingly clear. However, the molecular mechanisms of these proteins linked to diseases remain to be determined. The WD40 repeat protein mutations are commonly associated with genetic diseases, suggesting that there are no available treatments to stop or reverse the progression of these diseases. Current treatments only help with the symptoms of the diseases.


