BIOCHEMICAL AND STRUCTURAL ANALYSES OF TBL1: INSIGHTS INTO THE FUNCTION OF A TRANSCRIPTIONAL REGULATOR

By

Yoana Nantcheva Dimitrova

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

in

Biochemistry

December, 2010

Nashville, Tennessee

Approved:

Professor Walter J. Chazin

Professor Brandt Eichman

Professor Scott Hiebert, Ph.D.

Professor Jennifer Pietenpol, Ph.D.

Professor Charles Sanders, Ph.D.

Brenda Schulman, Ph.D.

To my loving parents,

Lidia and Nantcho Dimitrovi

ACKNOWLEDGEMENTS

I will always be grateful to my advisor, Dr. Walter J. Chazin, for allowing me to pursue my Ph.D. in his laboratory. He is a wonderful mentor and a great scientist. I am very grateful that he has allowed me to be independent, even though it was an intellectual challenge – and sometimes an emotional rollercoaster – as often science is, and provided me with a wonderful opportunity to grow as a scientist under his invaluable guidance. He will always be my example of true dedication and discipline. I will never forget his understanding and support through the most difficult years of my life, when I lost both my mother and my close friend to cancer. I am thankful I can call him my mentor and my friend.

I would like to also acknowledge the committee members of my thesis work, who have been instrumental in my success. I have been fortunate to have had Drs. Brandt Eichman, Scott Hiebert, Jennifer Pietenpol and Charles Sanders from Vanderbilt University and Brenda Schulman from St. Jude's Research Hospital on my committee. I would like to thank Brenda Schulman for taking the time to travel for many of my meetings. I am very grateful for their interest in my project and thoughtful scientific input, the exciting discussions about the direction of my project and for the great support and encouragement throughout.

The success of my project would not have been possible if not for the generous assistance of many collaborators. I would like to especially thank Dr. Bill Weis and the members of his lab at Stanford University for introducing me to the world of β-catenin and the challenges of x-ray crystallography. I am grateful to Dr. David Friedman for his

interest in my project and for providing his expertise in mass spectrometry. The excitement of Dr. Cun-Yu Wang (UCLA) about my project has always been contagious and I am very grateful for the critical functional assays that were performed in his laboratory. I would also like to thank Robert Rambo at the Lawrence Berkeley National Laboratory for his assistance in performing SAXS experiments. I am also very indebted to Dr. Melanie Ohi for the success of the many biochemical and structural aspects of my project. I am so grateful for her great scientific input and training, for her valuable advice and mentoring, and most of all for her encouragement and support.

I am very glad to have had the opportunity to work with the Chazin group, who provided such wonderful, friendly environment. I would like to thank Shibani Bhattacharya, who was my direct mentor at the beginning of my work and who introduced me to the world of ubiquitination processes and multi-protein complexes. My thanks also to Susan Mayen, who was a wonderful lab manager and who instilled in me many scientific skills. I will always cherish the many great moments with Chris Brosey, Sarah Soss, and Benjamin and Marie-Eve Chagot, who have generously shared science, laughter and tears with me.

One of the most exciting aspects of my graduate school experience was the opportunity to know wonderful friends. I am so grateful and fortunate to have found friends like Gulfem Guler, Mert Karkas, Dalyir Pretto, Anne Karpay, Will Bush, Cindy Wilson and Andrew Morin to share the best and the worst in our lives together. I had known Anne for only a few years before she passed away, but her passion for life and love for science will always remain with me. I will always be amazed at the determination and grateful for the friendship of Dalyir, who completed graduate school

while caring for three young children. And I will always be thankful for the strength and love I have received from Will and Cindy. Gulfem has been one of my dearest friends, who often brought encouragement in the most difficult of days, and lots of fun in the never-ending discussions of experimental design, data interpretation and life as a scientist. I have been very fortunate to have found the love and support of Andrew. He has often brought me back to reality and found ways to keep my sanity during the most difficult times of my thesis work. I have loved our many discussions and intellectual engagements over what my data means and how science works. I am glad to have enjoyed many dinners and discussions with Gulfem, Mert, Will and Andrew. I will always cherish their wonderful friendship and will dearly miss them all after graduate school.

And finally, I would like to express my gratitude to my loving and supporting parents, Lidia and Nantcho Dimitrovi and to my amazing brother, Stratsimir. I will never forget the sacrifice of my parents in granting my wish and sending me to the United States to pursue my studies. I will always be grateful for the faith they have instilled in me and for their unconditional trust and belief in my abilities to accomplish what I have set to do. I am so thankful to my dad for showing me how to love, understand and appreciate classical music. And though my mother passed away last year, the excitement and charisma that she carried in her smile, even during the last days of her life, will always remain in me as an inspiration in my life and work. Finally, there is no one that can bring more excitement for music, good fun and a positive view of life than my wonderful and supporting brother, Stratsimir. I will always be grateful for his unconditional friendship and love.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

LIST OF ABBREVIATIONS

APC	Adenomatous Polyposis Coli
ATP	Adenosine triphosphate
AUC	analytical ultracentrifugation
βME	β Mercaptoethanol
CBP	CREB binding protein
CD	circular dichroism
Cdc	cell division cycle protein
ChIP	Chromatin Immunoprecipitation
CK1	casein kinase I
CRL	Cullin-RING ligase
Cyclin C	Carbohydrate binding and sugar hydrolysis
DBD	DNA binding domain
EBNA	Ebstein-Barr virus nuclear antigen
EBV	Epstein Barr virus
EM	electron microscopy
ER	estrogen receptors
ERE	estrogen response element
GPS2	G protein pathway suppressor 2
GSK3β	glycogen synthase kinase 3 β
HECT	Homologous to E6AP C-terminus
HEK	Human Embryonic Kidney

HPV	human papillomavirus
Hsp	Heat shock protein
H2B	Histone 2B
IPTG	isopropyl thiogalactoside
LB	Luria-Broth
LEF	Lymphoid enhancer factors
LRR	Leucine Rich Repeat
NCoR	Nuclear receptor Co-repressor
NHR	nuclear hormone receptors
NMR	nuclear magnetic resonance
LisH	Lis1 Homologous
MALS	multi-angle light scattering detectors
MMP	matrix metalloproteinase
OASD	Ocular albinism with late-onset sensorineural deafness
PAF-AH	platelet activating factor acetylhydrolase
Phyl	Phyllopod
PTM	post-translational modification
QELS	quasielastic light scattering
R_G	radius of gyration
RI	refractive index
RING	Really Interesting New Gene
SAXS	small angle x-ray scattering
SCF	Skp-Cullin-FBox

Sina	Seven in absentia homologue
SIP	Siah-1 Interacting Protein
SMRT	Silencing Mediator for the Retinoid and Thyroid hormone receptors
SRC1	steroid receptor co-activator 1
TBL1	Transducin beta-like 1
TBLR1	TBL1 related 1
TCF4	T cell factor 4
Ttk88	Tramtrack 88
U-box	Ufd2-box
VEGF	vascular endothelial growth factor

CHAPTER I

INTRODUCTION

Overview

All cells are exposed to an array of environmental signals that initiate intracellular processes and ultimately orchestrate their behavior. The translation of external signals results in a specific molecular response in the cell, including post-translational modification of proteins in the form of phosphorylation, ubiquitination and sumoylation that stimulate complex macromolecular assemblies acting anywhere from the cell membrane to the nucleus. These signal transduction pathways lead to processes as diverse as cell proliferation, differentiation, motility and death. The precise execution of each process is highly dependent on the rate and specificity of the signaling.

TBL1 (now known as TBL1X[*]) and TBLR1 are highly homologous proteins that function in mediating the Wnt signaling/β-catenin cascade and modulating the transcriptional activity of nuclear hormone receptors (NHRs). TBL1 was first implicated in the poly-ubiquitination and proteasomal degradation of β-catenin under conditions of cell stress. Subsequently, TBL1 and TBLR1 have emerged as essential transcriptional regulators of NHRs, facilitating the exchange between their co-activator and co-repressor complexes also through their ability to modulate protein ubiquitination. The mechanistic basis for TBL1 family members as functioning regulators of β-catenin and NHRs is unknown and is the main focus of my Ph.D. dissertation research.

[*] In order to align with previous publication, TBL1 and TBL1X will be used interchangeably in this thesis. Other TBL1 homologues will be identified strictly by their formal names (TBLR1, TBL1Y).

The TBL1 family of proteins

The *TBL1* gene was first identified in 1999 from patients with Ocular albinism with late-onset sensorineural deafness (OASD), who have partial or full deletion of the gene linked to the X-chromosome [3]. Subsequently, other TBL1 homologues have been identified and included in the TBL1 family comprised of: Transducin β-like 1, located at the X chromosome (TBL1X), TBL1X-related 1 (TBLR1) linked to chromosome 3 and TBL1Y, the Y-linked homologue of TBL1X [4]. Additional isoforms of X-linked TBL1 and TBLR1 have also been identified, underlying the diversity of this novel family of proteins. Sequence alignment of the proteins demonstrates very high (>80%) sequence identity between the different TBL1 homologues [5]. All TBL1 family members consist of an almost identical N-terminal LisH dimerization domain followed by a putative FBox domain and a C-terminal WD40 repeat domain, the main sequence variability comes from the length and sequence of the linker between the FBox and WD40 domains (Fig. 1.1) [5].

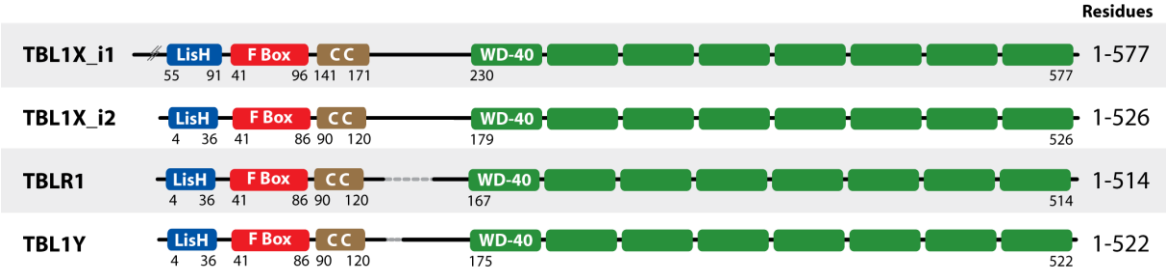


Figure 1.1. Schematic representation of human TBL1 homologues. TBL1X_i1- (isoform 1) and TBL1X_i2 – (isoforms 2).

TBL1 and TBLR1 are essential factors mediating the function of transcriptional co-activator β-catenin and nuclear hormone receptors [6-7]. TBL1Y is the least studied family member and its role has not been determined yet. Despite the high sequence identity between TBL1 and TBL1Y, the function of TBL1 cannot be compensated by TBL1Y. TBL1 is ubiquitously expressed in almost all tissues and bypasses X-chromosome inactivation [4]. TBL1Y has been proposed to have an essential distinct role in maleness.

TBL1 and TBLR1 share a diverse set of binding partners, yet most partners have preferentially higher affinity for one of the TBL1 homologues. Direct interaction of TBL1 family members have been demonstrated with β-catenin, TCF4, NCoR, SMRT, GPS2, CtBP1/2, Histone 2B (H2B), H4 and the 19S proteasome [6, 8-11]. Most of these interactions occur in the context of large multi-protein assemblies. Very little is known about the mechanism of assembly, specificity and release of these complexes. Recent studies have demonstrated that the function of TBL1 and TBLR1 is mediated by unique phosphorylation sites: Ser173, Thr334 and Ser420 for TBL1 and Ser123 and Ser199 for TBLR1 [10]. However, the basis of how TBL1 and TBLR1 facilitate this diverse set of interactions, as well as their regulation and functional outcome, is not known.

The first evidence for the role of TBL1 in mediating the formation of complexes that target proteins for ubiquitination and degradation comes from investigating the function of Ebi, the *Drosophila* homologue of TBL1. In *Drosophila*, the transcriptional repressor Tramtrack 88 (Ttk88) is recruited and poly-ubiquitinated by the Sina (Siah-1 homologue)/Phyllopod (Phyl)/Ebi complex for proteasomal degradation [12]. In a similar manner, mammalian TBL1 recruits non-phosphorylated β-catenin to Siah-1, which

mediates its poly-ubiquitination and degradation upon UV-induced DNA damage [13]. Despite the functional data from humans and *Drosophila* demonstrating the interplay between Siah-1 and TBL1 in substrate recruitment and ubiquitination, little is known about the specific mechanism of action in these functions.

Protein Ubiquitination

Post-translational modification of proteins by the covalent attachment of ubiquitin was initially discovered by Aaron Ciechanover, Avram Hershko and Irwin Rose in the early 1980s as a signal that targets substrates for degradation [14-16]. Since then, this finding was quickly expanded by an immense body of evidence demonstrating that post-translational modification by the attachment of ubiquitin serves many other purposes including chromatin remodeling, protein localization, cell cycle control, gene regulation and DNA repair [17]. Thus, this ubiquitous post-translational modification affects every aspect in the life, regulation and sustainability of the cell. Extensive in-depth reviews of various aspects of ubiquitination have been published [18-21]; a brief summary of the ubiquitination mechanism is presented below.

Mechanism of protein ubiquitination

Protein turn-over, the process of hydrolyzing the peptide backbone back to the basic amino acids for reuse, is essential to the maintenance of a stable and balanced cellular environment. The precise level of proteins in the cell is tightly regulated through coordination of the rate of protein synthesis and turnover. Post-translational modification of substrates by the covalent attachment of multiple ubiquitin moieties serves as a

tag/signal that is recognized by the 26S proteasome for degradation. The covalent attachment of ubiquitin to targets involves a cascade of enzymatic reactions carried out by the ubiquitin activating enzyme, E1, ubiquitin conjugating enzyme (E2), and ubiquitin ligating enzyme (E3) [16]. Ubiquitin is a small, 76 residue protein. It is activated by the E1 in an ATP-dependent manner to form a thioester bond between the E1 catalytic Cys and the C-terminal Gly residue of ubiquitin (Fig. 1.2). The ubiquitin~E1 complex recruits the E2 and transfers the covalently attached ubiquitin from the E1 to the catalytic Cys of E2. The final step of the ubiquitin transfer is catalyzed by the E3 ligase, which serves to recruit and catalyze the E2~ubiquitin (Ub) complex and the substrate as well as to activate the E2~Ub conjugate, resulting in transfer of a ubiquitin to lysine on the substrate. The process can be repeated multiple times, resulting in ubiquitination of other lysines on the substrate, as well as on the ubiquitin moieties of the ubiquitinated substrate [19].

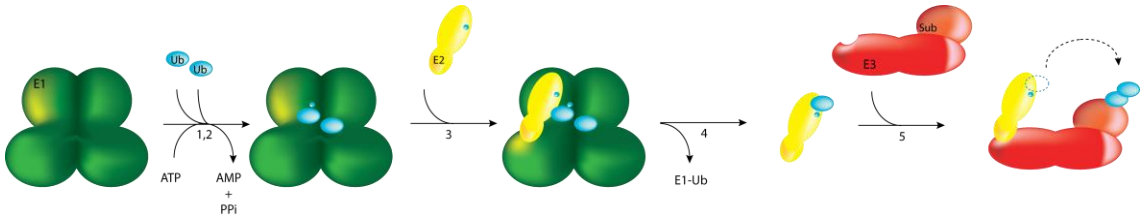


Figure 1.2. Schematic overview of the ubiquitin conjugation cascade. 1) The ubiquitin activating enzyme, E1 reacts with ubiquitin to form an activated ubiquitin~E1 complex by forming a thioester bond, in an ATP-driven reaction. 2) This step is repeated for the non-covalent attachment of a second ubiquitin to E1. 3) The E1 loaded with two ubiquitins binds the E2, ubiquitin conjugating enzyme. 4) The first covalently attached ubiquitin from the E1 is transferred to the E2 forming a thioester bond. 5) The E3 ligating enzyme, recruits an E2~Ubiquitin complex and a substrate to facilitate the transfer of ubiquitin to the substrate. Multiple cycles of E2~Ub complex recruitment are necessary for the formation of a poly-ubiquitin chain.

E3 ubiquitin ligases: recruiting substrates for ubiquitination

Substrate specificity is determined largely by the E3 ubiquitin ligases, based on the nature of their substrate-recruiting domains. The E3 enzymes can be broadly classified into two general subgroups: (i) simple E3 ligases, for which a single molecule performs the recruitment, activation and transfer of ubiquitin from E2 enzyme to the substrate; (ii) complex E3 ligases, represented by multi-protein assemblies that mediate the recruitment of the substrate, activation of the E2~Ub complex and transfer of ubiquitin. E3s are classified on the basis of their E2-binding domain as HECT (homologous to E6AP C-terminus), RING (really interesting new gene), and U-box (Ufd2-box) E3 ligases. The Cullin-RING ligases (CRLs) are the most well studied E3 ligases; they utilize a RING-module for the recruitment of E2~Ub [18].

HECT family of E3 ligases: HECT E3 ligases have at least 2 distinct domains, an E2- and a protein-protein binding domain. The transfer of ubiquitin to a substrate by HECT enzymes is a two step process that involves simultaneous recruitment of the substrate and the E2~Ub complex (Fig. 1.3A). First, ubiquitin from the E2 is transferred to a catalytic Cys on the HECT domain itself by forming an E3~Ub thioester, followed by transfer of ubiquitin from the E3~Ub intermediate complex to a Lys residue on the substrate.

RING and U-box family of simple E3 ligases: The RING and U-box domains of E3 ligases are structurally similar and both recognize diverse E2 enzymes. The general architecture and mechanism of RING/U-box ligases is similar to HECT proteins. They have a distinct substrate-binding domain and a RING/U-box domain that recruits specific E2 enzymes. The main difference is that the ubiquitin molecule is transferred directly

from the RING- or U-box-bound E2 to the targeted Lys on the substrate (Fig.1.3B). It was initially proposed that U-box ligases function as E4 ligases by extending the ubiquitin chains on an ubiquitin primed substrate [22]. However, U-box enzymes, such as CHIP and E4B can form poly-ubiquitin chains independent of other E3 ligases [23-24].

Complex E3 ligases, CRLs: The complex Cullin-RING ligases (CRLs) are the most abundant and well studied of the E3 ubiquitin ligases [18]. These include SCF (Skp-Cullin-FBox) E3 ligases, which are assembled using one of the Cullin family members as a molecular scaffold for bringing together the other components (Fig. 1.3C). In the paradigm SCF complex, $SCF^{\beta TrCP}$, the C-terminal domain of Cullin serves to recruit the RING-E3 protein Rbx1. The adaptor protein Skp1 is bound to the N-terminal region of Cullin and to the FBox domain of the substrate-recruiting protein. Once the complex is formed, SCF E3 ligases follow a direct mechanism of ubiquitin transfer from the E2~Ub complex to the substrate, similar to simple RING E3 ligases (Fig. 1.3A, C).

The substrate of the SCF E3 ligase is specified by one of the large family of substrate receptors. Substrates are bound to either a WD40 repeat, leucine-rich repeat (LRR), Kelch, carbohydrate binding and sugar hydrolysis (CASH) or Zinc finger domain [21]. The most abundant of these substrate binding domains are WD40 and LRR domains.

A specific binding motif for substrate interaction with SCF ligases has been identified, which is termed a phosphodegron, because it requires specific serines and threonines of the substrate to be phosphorylated. The number of phosphorylated residues in the phosphodegron has been associated with the rate of substrate poly-ubiquitination and degradation [25]. Other post-translational modifications including the addition of

oligosaccharides or a hydroxyproline epitope are also recognized by substrate receptors of E3 ligases, thereby expanding the diversity and specificity of protein ubiquitination [18, 21, 25-26].

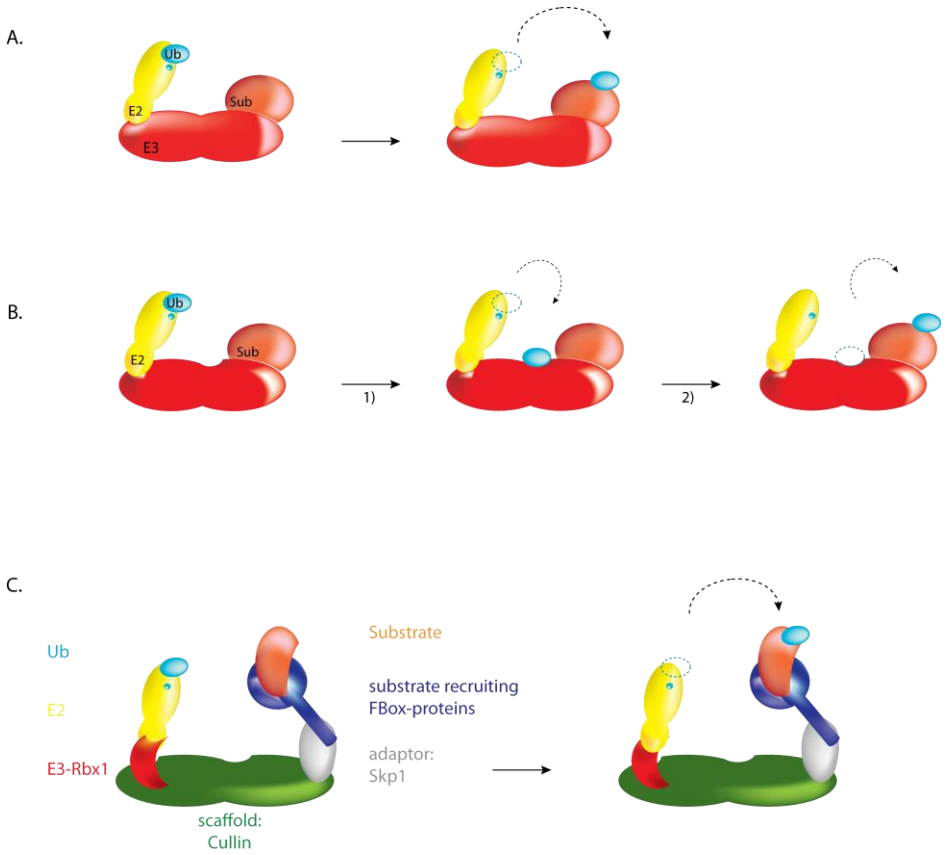


Figure 1.3. Representation of the E3 ubiquitin ligase subfamilies. A) RING/U-box ubiquitin ligases. Direct transfer of ubiquitin from the E2~Ub complex to the substrate; B) HECT ubiquitin ligases. Indirect transfer of ubiquitin from the E2~Ub first to the E3 (step 1), then to the substrate (step 2); C) Cullin-RING ubiquitin ligases. Formation of multi-protein complex that facilitates the direct transfer of ubiquitin from the E2~Ub to the substrate.

Physiological role of ubiquitin linkages

The functional readout of substrate ubiquitination depends on the number of ubiquitins added and the type of linkages formed. Mono-ubiquitination is the attachment of a single ubiquitin molecule to a Lys residue on the substrate, whereas ligation of a single ubiquitin to more than one of the Lys residues of the substrate is termed multi-monoubiquitination (Fig. 1.4A). Ubiquitin itself has seven Lys residues that are able to accept another ubiquitin moiety, which can lead to formation of either homogeneous linear: Lys6, Lys11, Lys27, Lys29, Lys33, Lys48 and Lys63, or heterogeneous or branched poly-ubiquitin chains (Fig. 1.4B). The functional effect of poly-ubiquitination is dictated by the type of chain. This can be understood in part by differences in structure of the poly-ubiquitin chain. For example, tightly packed Lys48 chains are structurally different from extended Lys63-linked poly-ubiquitin chains [27-28].

Recent advances in mass-spectrometry have been very instrumental in determining the identity of ubiquitin chains on specific substrates, although mapping the functional outcome associated with a specific chain type remains a challenge. In general, proteins with Lys48 and Lys11 linked ubiquitin chains are recognized by the 26S proteasome for degradation, whereas Lys63 chains are identified on substrates involved in DNA damage or cellular trafficking (Fig. 1.4C) [29-30]. The functional relevance of linkages through Lys27, Lys29 and Lys33 is not well understood, in part due to technical difficulties in identifying them in mass spectrometry experiments.

As noted above, besides homogeneous “linear” ubiquitin linkages, where only a single ubiquitin is added to a specific ubiquitin moiety on a substrate, branched or forked poly-ubiquitin chains can be formed. In these chains at least one ubiquitin moiety is

covalently bound to two or more ubiquitins (Fig 1.4C) [31]. The exact functional relevance of the branched and forked chains is not firmly established. However, evidence has accumulated suggesting that forked poly-ubiquitin chains prevent degradation of the substrate by the 26S proteasome [30-31].

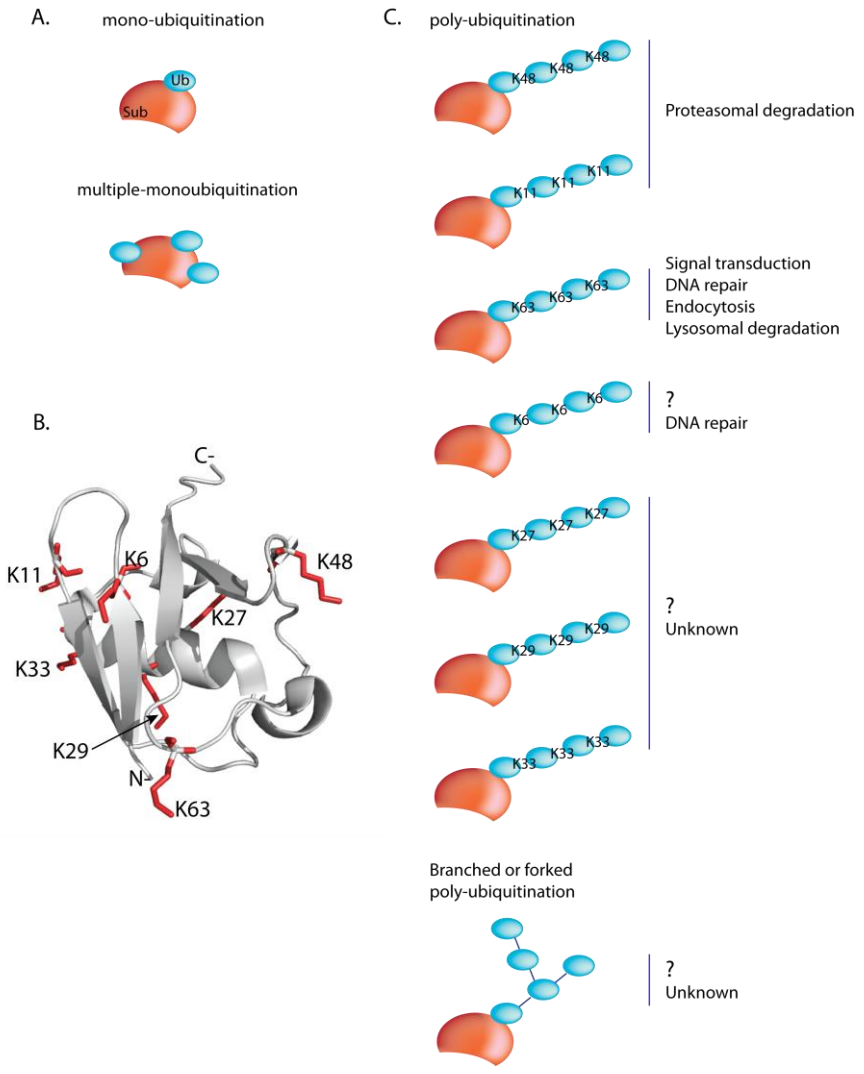


Figure 1.4. The ubiquitin signal. A) Ubiquitination forms from the attachment of a single ubiquitin to one (mono-ubiquitination) and many (multi-monoubiquitination) target site(s) on the substrate. B) The structure of ubiquitin where the seven lysine residues that function as acceptors for another ubiquitin molecule are colored in red. C) Types and functional relevance of poly-ubiquitin linkages.

Protein deubiquitination

Post-translational modification by ubiquitin is a reversible process. The covalent bond between ubiquitin and substrate can be hydrolyzed by a diverse family of proteases in a process called deubiquitination. Deubiquitinating enzymes (DUBs) are essential in determining the fate of ubiquitinated substrates by either removing the ubiquitin chain from the substrate or by editing the assembled linkage on the substrate [32-33]. Since the proteasome recognizes ubiquitin chains with a minimum of four ubiquitin molecules, editing already assembled linkages provides a DUB-dependent rate of substrate recruitment and degradation by the proteasome. Thus, DUBs introduce an additional means of regulation to control the levels of proteins in cells. DUBs have another role once substrates are degraded: their activity enables recycling of poly-ubiquitin chains released from the proteasome and free ubiquitin chains that are synthesized in the cell [33].

Notably, the proper function of deubiquitinating enzymes is also crucial for the ubiquitination process itself. Ubiquitin is actively expressed as a pro-ubiquitin molecule, a free linear poly-ubiquitin as a fusion to a ribosomal protein. DUBs provide the proteolytic activity required to generate single ubiquitin moieties Activation by DUBs includes cleavage of C-terminal residues of pro-ubiquitin, which must be trimmed to leave the Gly-Gly end required for function in ubiquitin ligation reactions.

Ubiquitination and disease

Since the ubiquitination machinery plays an essential role in keeping the levels and activity of many cellular proteins in check, defects in the ubiquitin cascade or

mutations in the substrate that reduce or abrogate binding to the E3 can lead to aberrations in the ubiquitination process resulting in a variety of diseases from cancer to neurodegenerative disorders. For example, mutations in the BRCA1 and BRCA2 E3 ligases leads to increased susceptibility to breast and ovarian cancer, and mutations in the Parkin E3 ligases are associated with Parkinson's disease [18, 34].

Interestingly, the ubiquitin-proteasome machinery has also been identified as a direct target of many oncogenic viruses such as human papillomavirus (HPV), Epstein Barr virus (EBV) and the adenovirus [35-36]. A common mechanism of these tumor viruses is to either modulate host E3 ligases or to produce viral ubiquitination and deubiquitination machinery in the host cell. Major signaling pathways such as the Wnt/β-catenin pathway are directly targeted by tumor viruses as are oncogenes including p53, c-Myc and cyclin D1 proteins [35]. Therefore, understanding the mechanisms of host and viral mediated ubiquitination and deubiquitination processes offers promise as a means to develop therapeutics, targeting key cancer proteins.

The role of TBL1 in mediating the function of β-catenin

β-Catenin is a ubiquitous transcriptional activator in the canonical Wnt signaling pathway. Wnt signaling is integral to many cellular processes including cell proliferation, cell fate and survival and adult stem cell differentiation and oncogenesis [37-38]. Regulation of β-catenin by the Wnt pathway is highly conserved throughout all animal species. The ability to study β-catenin in an array of model organisms has enabled numerous crucial discoveries. Importantly, this basic developmental pathway is directly associated with cancer.

Tight control of β-catenin levels and function is critical for the timely expression of Wnt target genes during embryogenesis and in adult homeostasis. Not surprisingly, mutations in β-catenin and regulatory factors from the Wnt pathway are associated with many cancers in addition to cardiovascular diseases, osteoarthritis and neurodegenerative diseases such as Alzheimer's and schizophrenia [37-39].

Defects and inappropriate stabilization of β-catenin are often associated with the initiation and progression of cancers such as breast, colorectal and ovarian [40-41]. One of the first discoveries that linked β-catenin to cancer was the identification of direct interaction with Adenomatous Polyposis Coli (APC) [42-43]. Loss of function mutations in APC that prevent its interaction with β-catenin or the $SCF^{\beta\text{-}TrCP}$ E3 ligase are found in many patients with colorectal cancers [39]. Furthermore, mutations of Ser33, Ser37, Thr41 and Ser45, which are part of the β-catenin phosphodegron, have been identified in patients with melanoma, colorectal, hepatocellular and gastric cancers [37].

Considering that defects in the Wnt/β-catenin pathway are strongly implicated in the pathogenesis of proliferative diseases, many components of this pathway have become targets for anti-cancer drug discovery [38, 44]. The challenge of interrupting or stimulating specific β-catenin-protein complexes without effecting other interactions is being actively pursued. Understanding the mechanisms that lead to the stabilization of β-catenin in its oncogenically active state and activation of Wnt target genes is currently being investigated and constitutes one of the primary motivations for my research.

Regulation of β-catenin by the Wnt signaling pathway

The Wnt signaling pathway controls cellular levels and the activity of β-catenin (Fig. 1.5). Wnt ligands are small Cys-rich glycoproteins utilized for short-range signaling. In the absence of Wnt ligands, β-catenin is phosphorylated by casein kinase I (CK1) and glycogen synthase kinase 3 β (GSK3β) [38]. Phosphorylated β-catenin is recognized and poly-ubiquitinated by the $SCF^{\beta\text{-}TrCP}$ E3 ligase, which targets it for degradation by the 26S proteasome (Fig. 1.5, Left) [37-38, 45]. Low β-catenin levels in the cell lead to its nuclear depletion, which results in T cell factor/Lymphoid enhancer factors (Tcf)/(Lef) transcription factors remaining in a repressed state, thereby preventing the expression of the Wnt target genes.

When Wnt ligands are secreted, GSK3β and other factors are recruited to the Frizzled transmembrane receptor, reducing poly-ubiquitination and degradation of β-catenin (Fig. 1.5, Right). Increased local concentration of β-catenin in the cytoplasm leads to its nuclear translocation. TBL1 and TBLR1 directly bind β-catenin and recruit it to the promoter site of Tcf/Lef factors, where β-catenin forms an activation complex with Pygopus and Bcl-9 [6, 37]. The formation of an activation complex displaces transcriptional repressors TLE1, HDACi and Groucho from the Tcf/Lef factors and triggers the expression of Wnt induced genes. Notably, TBL1 and TBLR1 are always found present at the promoter site in both the "on" and "off" states of the Wnt target genes [6, 37]. The activation of Wnt enhanced promoters by β-catenin contributes to the expression of many genes including: *cyclin D1, c-Myc, AXIN2, SOX9* and *vascular endothelial growth factor* (*VEGF).* The activation of these genes is important in processes ranging from cell proliferation (cyclin D1, c-Myc and SOX9) to development

(AXIN2) to signaling and cell motility (VEGF), which in turn highlights the importance of transcriptional activation by β-catenin in human physiology [37-38].

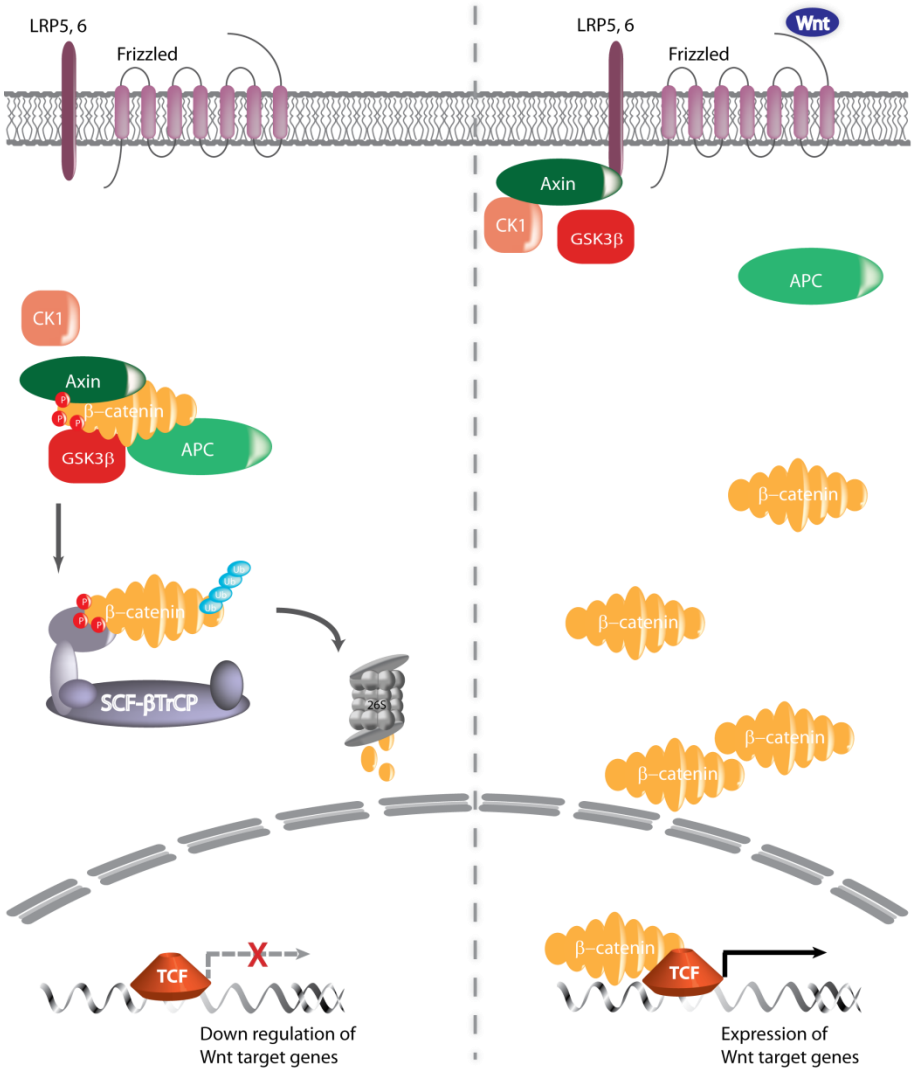


Figure 1.5. The Wnt signaling cascade. (Left) In the absence of a Wnt ligand, β-catenin is recruited for phosphorylation by GSK3β and CK1. Phosphorylated β-catenin is recognized for poly-ubiquitination by the $SCF^{\beta TrCP}$ complex, leading to its degradation by the 26S proteasome. (Right) In the presence of a Wnt ligand, the protein complex that recruits β-catenin for phosphorylation is localized to the cellular membrane. This prevents β-catenin from phosphorylation, leading to its accumulation in the cytoplasm and translocation to the nucleus. β -Catenin is recruited to the promoter site of Wnt genes, targeting their expression.

Degradation of β-catenin by the SCF$^{β\text{-}TrCP}$ E3 ubiquitin ligase

The critical role in transcriptional activation of genes has prompted investigation of post-translational modification and turnover of β-catenin. In 1997, Aberle and colleagues showed for the first time that the level of β-catenin is decreased by ubiquitination-mediated degradation in the proteasome [46]. Down regulation of β-catenin occurs only in the absence of Wnt ligands; expression of Wnt ligands suppresses the degradation process [46]. Soon after this report, it was shown that APC and AXIN2 recruit CK1 and GSK3β kinases and stimulate phosphorylation of β-catenin at Ser29, Ser33, Ser37, Thr41 and Ser45 [45, 47]. Phosphorylated β-catenin is recognized by the WD40 domain of β-TrCP, leading to poly-ubiquitination by the SCF$^{β\text{-}TrCP}$ E3 ubiquitin ligase. The assembly of poly-ubiquitin chains on Lys19 and Lys49 targets β-catenin for proteasomal degradation [1, 48].

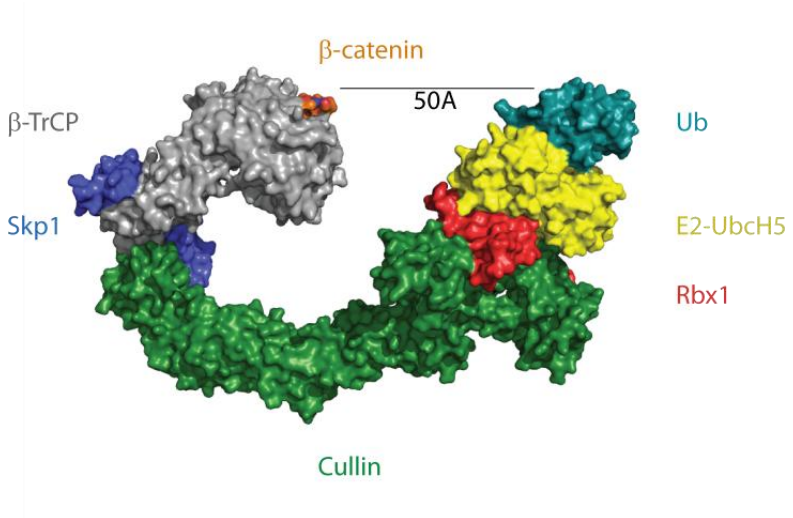


Figure 1.6. Structural model of SCF$^{β\text{-}TrCP}$ - β-catenin complex (reproduced based on the model in [1]). The extended modular architecture of Cullin facilitates the recruitment of phosphorylated β-catenin substrate through the Skp1-β-TrCP module and the RING-Rbx1 protein that binds the E2~Ub complex. Computation methods have been used to calculate an estimated distance from the active cystein of an E2 to the substrate to be ~50Å.

A structural model of $SCF^{\beta\text{-}TrCP}$ ligase in complex with phosphorylated β-catenin peptide has provided critical insights into the mechanism of substrate ubiquitination (Fig.1.6) [1]. The x-ray crystal structure of β-TrCP with β-catenin (residues 19-44) in combination with mutagenesis and *in vitro* ubiquitination studies have identified the sequence specificity of the doubly-phosphorylated phosphodegron, sufficient for the binding and ubiquitination of the β-catenin substrate. Asp32 and Gly34 in addition to the phosphate groups of pSer33 and pSer37 are critical for the strong binding with the WD40 β-propeller of β-TrCP. Based on the structural and biochemical experiments on β-catenin targeted ubiquitination by the $SCF^{\beta\text{-}TrCP}$ complex, the rate determining factor for ubiquitin chain assembly is proposed to be dependent on the position of a specific Lys residue on the β-catenin target sequence with respect to the residues in direct contact with the basic WD40 surface [1].

Siah-1-targeted degradation of β-catenin by an SCF^{TBL1}-like complex

Nearly a decade ago, a novel β-catenin ubiquitination pathway was discovered in cells stressed by UV-induced DNA damage [13]. Remarkably, this pathway targets the non-phosphorylated form of β-catenin. UV-induced DNA damage activates p53, which induces the overexpression of a RING E3 ligase, Siah-1. Increased levels of Siah-1 lead to the formation of an SCF-like complex comprised of Siah-1, Siah-1 Interacting Protein (SIP), Skp1, TBL1 and APC [13, 49]. Siah-1 binds the E2 ubiquitin conjugating enzyme through its RING domain and is linked to substrates through SIP, Skp1 and TBL1. TBL1 plays the key role in this complex by directly contacting non-phosphorylated β-catenin (Fig. 1.7). APC is associated with both β-catenin and Siah-1, and appears to play a

regulatory role in β-catenin localization, but its exact function has not yet been determined [50]. Even though Siah-1 does not form a prototypical SCF E3 ligase that contains a Cullin protein, it was named based on the established convention for SCF multi-protein E3 ligases: SCF^{TBL1}.

The substitution of SIP for the Cullin scaffolding protein has important implications for the SCF^{TBL1} complex because SIP and Cullin have very different architectures. Cullins are almost four times larger than SIP (~800 vs 229 residues). Since both Cullin and SIP serve as scaffolds in their respective complexes, the differences in their structures imply that existing SCF models would not be accurate representation of the SCF^{TBL1} complex. Moreover, both TBL1 and Siah-1 exist as oligomeric proteins [51-52]. Thus, despite the wealth of structural and biochemical information about SCF E3 ligases, at the start of this thesis work almost nothing was known about the mechanism of SCF^{TBL1}-mediated recruitment and poly-ubiquitination of non-phosphorylated β-catenin.

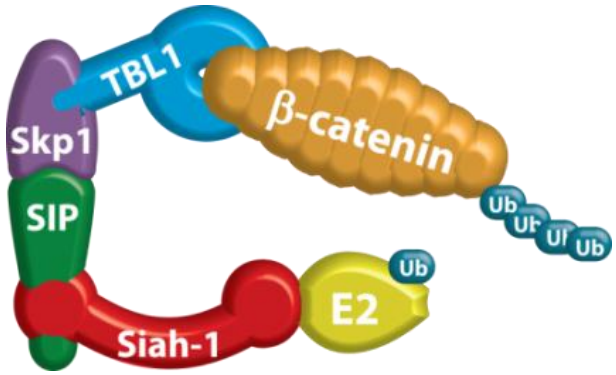


Figure 1.7. Schematic representation of the proposed SCF^{TBL1}-like E3 ubiquitin ligase. Overexpression of Siah-1-mediates the recruitment of SIP, which interacts with the adaptor protein Skp1. TBL1 binds the non-phosphorylated β-catenin substrate and recruits it to the SCF-like complex for poly-ubiquitination through cycles of E2~Ub complexes that bind RING-Siah-1 E3 ligase. TBL1 and Siah-1 are shown as monomers in this diagram for clarity of representing the complex.

The role of TBL1 in regulating the activity of nuclear hormone receptors

A wealth of knowledge about NHRs has accumulated over more than 100 years and has its roots from 1903, when Bayliss and Starling used the Greek work “hormone”, which means “I stimulate”, to characterize the chemical secretions from the pancreas. These early studies initiated the science of endocrinology and the idea of chemical messengers. Since then, science has unraveled the immense impact that mistakes in the regulatory pathways of NHRs have on a wide range of pathophysiologies from cancer to metabolic diseases such as diabetes [53].

NHRs are a large structurally related family of modular proteins that regulate many processes during development, differentiation and homeostasis [54-55]. They represent a superfamily of transcription factors activated by steroids and thyroid hormones as well as vitamins and numerous other small ligands. These modular proteins contain a conserved DNA binding domain (DBD) and a C-terminal ligand binding domain, both of which contribute to the formation of homo- and hetero-dimers. The N-terminal region is unstructured, has the highest degree of sequence variability within the NHR family members, and is responsible for the specificity of diverse protein interactions [54-55].

NHRs are divided into two major subfamilies: steroid receptors, such as the glucocorticoid receptors and estrogen receptors (ER); non-steroid receptors that include thyroid receptors, Vitamin D receptors and retinoic acid receptors [53, 56]. The activation pathways of the two receptor subfamilies differs mostly by the essential role that Heat shock proteins, Hsp70 and Hsp90, play in activating the steroid, but not the non-steroid receptors. Unliganded steroid receptors are localized predominantly in the cytoplasm and

upon ligand stimulation they form a complex with Hsp70 or Hsp90 chaperones, which mediate their nuclear translocation. Localization of steroid receptors to the nucleus leads to the dissociation from the heat shock proteins and formation of receptor homodimers at the promoter site of target genes. The association of ligand-bound receptors to the promoter facilitates the interaction with hormone elements and regulatory co-activator complexes that together trigger the expression of specific target genes.

Non-steroid receptors are activated and recruited to the promoter site of target genes independently of heat shock proteins [53, 56]. Ligand binding to the receptor leads to dissociation of the homodimers and formation of the more favorable heterodimer, which in turn is recruited to the response elements at the promoter site of target genes. In a similar manner to steroid receptors, the activation of gene expression requires the sum of many interactions from macromolecular assemblies of co-activators and signaling proteins, which at the end results in turning the transcriptional switch at the promoter "on" or "off".

TBL1-mediated transcriptional regulation of nuclear hormone receptors

Critical advances in understanding the dynamic mechanism of transcriptional regulation by NHRs have come from Chromatin Immunoprecipitation (ChIP) experiments using Estrogen Receptor α (ERα). Estradiol-bound ER is recruited to the promoter followed by activator complexes that initiates the expression of ER-driven genes [57-58]. Importantly, this event cycles with recruitment of the proteasome, which leads to subsequent degradation of the receptor and activator complexes, and replacement by co-repressor complex and the unliganded receptor [59]. The dynamic regulatory

exchange of these large protein machineries and the essential role of the proteasome in switching between activators and repressors have become the standard model for transcriptional regulation by NHRs.

In 2004, Perissi and colleagues demonstrated that TBL1 and TBLR1 are essential factors for the transcriptional activation of nuclear receptors, and absolutely necessary components in both the repressor and activator complexes of NHRs (Fig. 1.8) [60]. Notably, despite their high sequence identity, TBL1 and TBLR1 are not functionally redundant and show clear preference for activators or repressor of NHRs. Further analyses of the function of TBL1 proteins have shown that they recruit the 19S proteasome and mediate the ubiquitination and degradation of NHR and their co-factors at the promoter site (Fig. 1.8) [7, 60]. The mechanism that dictates the role of TBL1 scaffolding proteins in the precise exchange of these factors is still unknown and an active area of research in the field.

Therapeutic potential

Given the enormous impact that NHRs have on diverse physiological processes and their function via central signaling pathways and dynamic multi-protein machineries, it is not surprising that functional defects would result in disease. Defects in the function of NHRs have a physiological impact that ranges from proliferative, reproductive and metabolic conditions such as type II diabetes, obesity, cardiovascular diseases, infertility and many types of cancer [61-62].

Structural characterization of NHRs has promoted their consideration as drug targets and a number of therapeutics directed to the receptor superfamily have already

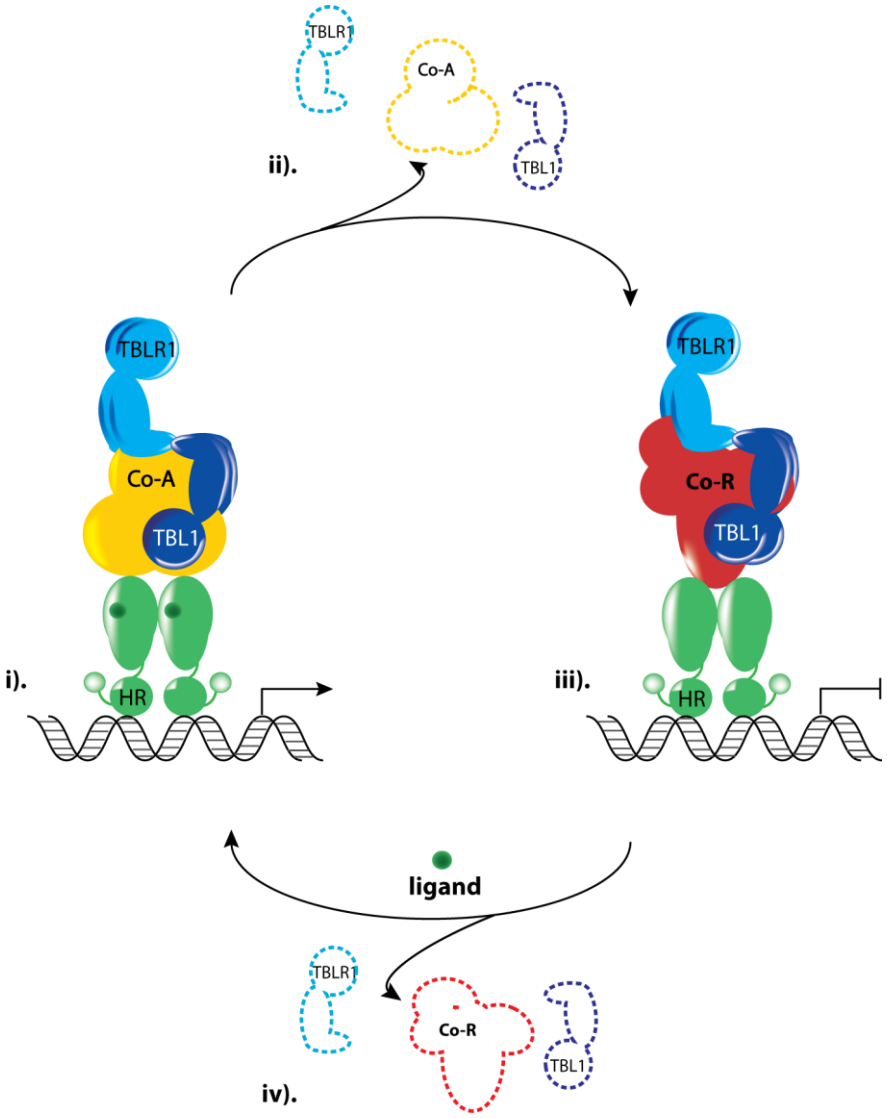


Figure 1.8. Mechanism of exchange between NHR co-activator (Co-A) and co-repressor (Co-R) complexes. (i) The ligand bound NHR recruits a Co-A complex in the presence of TBL1 and TBLR1. The formation of the multi-protein assembly leads to the transcriptional activation of HR target genes. (ii) The exchange of Co-A for Co-R complex, requires TBL1 mediated poly-ubiquitination and recruitment of the proteasome that leads to the degradation of Co-A components. (iii) TBL1 proteins replace the Co-A complex with the Co-R complex, which results in transcriptional inhibition of NHR target genes. (iv) Ligand induced stimulus targets the recruitment of an E3 ligase and the 19S proteasome by TBL1, leading to the proteasomal degradation of the Co-R complex and its exchange of a Co-A complex.

been developed [54]. However, although therapeutics that can specifically target NHR have been designed, a single receptor can activate multiple genes, which ultimately results in a drug with significant side effects. The current focus is on new therapeutics that recognize specific post-translational modifications and mediate the function of multi-protein assemblies [54]. Since TBL1 and TBLR1 are essential scaffolding proteins for these complexes, understanding their mechanism of function and how they mediate the exchange of co-activator and co-repressor complexes in activating NHRs will provide new avenues to pursue for identifying therapeutics.

Experimental Methods for Structural Analysis

Analytical Ultracentrifugation

Analytical ultracentrifugation (AUC) is a biophysical method used to analyze the molar mass and shape of a molecule and the stoichiometry and association constants of molecular complexes [63]. Two types of AUC experiments are typically performed: sedimentation velocity (SV) and sedimentation equilibrium (SE) [64-65]. In an SV experiment, the sedimentation of a molecule away from the center of rotation is monitored at high centrifugal speed, 40,000 – 60,000 rpm. Based on hydrodynamic principles, the data provides information about the size and shape of the molecule in solution. SE runs are performed at lower centrifugal speed, 15,000 – 30,000 rpm, where the sedimentation and back diffusion of the molecule are at equilibrium. Based on thermodynamic principles, the data is used to calculate the molar mass, stoichiometry and association constant of proteins and protein complexes in solution [63].

Analysis of the SV data can provide estimates of biophysical parameters such as the frictional ratio and the Stokes radius [64-65]. The frictional ratio, f/f_o is the experimentally determined frictional coefficient of a hydrated molecule divided by the calculated, non-hydrated spherical molecule with the same molar mass and gives information directly related to the shape of the molecule. Highly globular proteins have a lower frictional ratio, compared to proteins with elongated structural architectures. The Stokes radius, R_s, also known as the hydrodynamic radius of a molecule, relates to the radius of a hard sphere that diffuses at the same rate as the molecule.

For this thesis work, AUC velocity experiments were acquired with an ultracentrifuge cell with two separate sectors. One sector is used for the protein sample and the other used as a reference for the solvent. The loaded cell is placed in a rotor and the sedimentation of the sample monitored in real time through an optical system while centrifuged at 40,000 rpm. Analysis of the concentration distribution of the sample over time is performed with the program sedfit [63-64]. This analysis provides a direct measurement of the sedimentation coefficient, *s* (S), which represents the rate of migration of a molecule per unit field, determined using $s = v / (\omega^2 r) = M_b / f$, where v is the volume of the particle, ω the rotor speed, r the distance from the center of the rotor, M_b the buoyant molar mass and f the frictional coefficient [63-64].

In this thesis work, data from the velocity AUC experiments of TBL1 have been used to determine the oligomerization state of the protein. Furthermore, calculating the frictional ratio and the Stokes radius of the protein have been essential parameters in generating the structural model of TBL1 using SAXS (chapter III).

Nuclear magnetic resonance

Nuclear Magnetic Resonance (NMR) spectroscopy is a very sensitive and powerful technique used primarily for determining molecular structure, characterizing molecular dynamics and for mapping inter-molecular interactions. The NMR phenomenon arises from the precession of nuclear spins when placed in the presence of a magnetic field. The NMR frequency serves as a signature of the electronic structure surrounding the nucleus. NMR analysis is largely based on the detection of scalar (through-bond) and dipolar (through-space) couplings between nuclei. Over the past 30 years, the development of multi-dimensional, multi-nuclear experiments has enabled the study of large, complex molecules such as proteins.

NMR experiments on proteins are typically performed on samples with enrichment in one or more stable NMR-active isotopes such as ^{15}N, ^{13}C and ^{2}H. In a 2D ^{15}N-^{1}H HSQC (Heteronuclear Single Quantum Correlation) spectrum, each of the peaks represents the correlation of the ^{1}H and ^{15}N resonance frequencies of each amide in the protein. Since each peak arises from a different amino acid, this spectrum provides a powerful fingerprint of the protein structure.

Transverse relaxation optimized spectroscopy (TROSY)–HSQC experiments are often utilized to investigate large proteins or protein complexes, where the slow tumbling of the molecules leads to faster relaxation and correspondingly, larger line width [66]. It turns out that the resonances of even a simple, two-spin ^{15}N-^{1}H system, is composed of four components, which relax at different rates. TROSY filters out the fastest relaxing components leaving only the slow relaxing components. This leads to the ability to observe signals for larger molecules and much narrower resonances. Maximum TROSY

sensitivity requires acquisition of spectra at magnetic field strengths corresponding to 700 MHz or greater and deuteration of proteins.

Based on the peak distribution in an ^{15}N-^{1}H HSQC spectrum, essential information can be obtained in regards to the extent of ordered structure and stability of a protein. The chemical shifts of a globular protein are well dispersed in the ^{1}H dimension, whereas for an unfolded protein the range is significantly collapsed. Once the conditions for collecting a good HSQC spectrum of the protein have been established, additional experiments can be performed to assign each peak to the specific amino acid of the protein, which is the necessary prerequisite for any detailed NMR analysis such as mapping the interaction site of a binding partner or pursue structure determination.

The main application of NMR in this thesis work has been to confirm a direct physical interaction between two proteins using chemical shift perturbation assays. In these experiments, ^{15}N-labeled protein is used to record ^{15}N-^{1}H HSQC spectra alone and after titration of unlabeled binding partner. The unlabeled protein produces no signal, but the NMR chemical shifts of residues at the binding interface will be perturbed as the ligand is titrated. The chemical shifts are highly sensitive to the electronic environment of the nucleus therefore, direct physical contact between the two proteins leads to alterations in the chemical shifts at the binding interface. Ligand-induced conformational change of the labeled protein would also result in chemical shift perturbations at sites that may be remote from the binding site. Hence, careful evaluation of the data is required. The reciprocal experiment is often performed, switching which protein is isotopically labeled, in order to confirm the direct physical interaction between the two proteins and to map the binding site on the binding partner.

The protein-protein interactions investigated by NMR assays in this thesis work involved large proteins (between 30 and 85 kDa). In such cases, the larger protein complex can lead to increased relaxation rates and broadening of the signals below the detection limit in an HSQC spectrum. In these instances, definitive evidence for a direct protein interaction can be obtained by evaluating spectra under sub-stoichiometric conditions, i.e. with less than a molar equivalent of the unlabeled titrant [67-68]. When very small amounts of titrant are added, in favorable cases, it is possible to observe chemical shift perturbations and/or increases in line width of a select number of NMR signals that arise from residues at the binding interface [68].

In this work we showed the direct physical interaction between Siah-1 (44 kDa) and β-catenin (64 kDa) using TROSY-based NMR chemical shift perturbation assay (chapter II).

Small angle x-ray scattering

Small angle x-ray scattering (SAXS) is a powerful technique for obtaining low resolution structural information about the size, shape and the conformational dynamics of a molecule or molecular complex in solution. In a small angle scattering experiment, the sample is exposed to a monochromatic flux of highly-focused x-rays. The scattered x-rays from the sample are collected on a 2D-detector placed orthogonal to the incident beam. The raw scattering profile is transformed into a 1D curve of scattering intensity I(q) as a function of the scattering vector q, where $q = 4\pi/\lambda$ (sin θ), 2θ is the scattering angle between the incident and the scattered vectors and λ is the wavelength (Fig. 1.9A). The data are normalized based on the scattering intensity and the collection time. Finally,

the scattering of a buffer identical to the sample solvent is subtracted from the total scattering of the sample, providing the final scattering curve of the molecule in reciprocal space (Fig. 1.9B).

Guinier analysis of the scattering curve at low scattering angle q is performed to determine the quality of the scattering data and if exposure of the sample to x-rays has induced radiation damage [69-70]. The linearity of a Guinier plot in the range of $q\mathrm{Rg} \leq 1.3$ for spherical proteins is indicative of a monodisperse sample that has not been affected by radiation damage. A non-linear dependence of the scattering curve suggests the presence of protein aggregation, improper background subtraction or that the sample may be elongated in shape.

Additional structural parameters such as the radius of gyration (Rg), which is the mass distribution of the molecule around its center of gravity, and the molecular weight of the protein can be calculated based on the Guinier approximation. A linear fit of the scattering curve ln [I(q)] vs. q^2 at the lowest q range, is used to extrapolate the Rg and the scattering intensity at q=0, or I(0), used to calculate the molecular weight of the protein.

Direct information about the shape and the degree of conformational dynamics of the molecule in solution can be obtained by calculating the Kratky plot (q^2I(Iq) vs. q) from the scattering curve (Fig. 1.9D). Well folded spherical domains have a curve shaped like a parabola, whereas unstructured molecules lead to a profile that becomes linear at the high q values and does not return back to zero [69-70]. The presence of well folded domains linked by flexible linkers would produce a profile in between the two mentioned above.

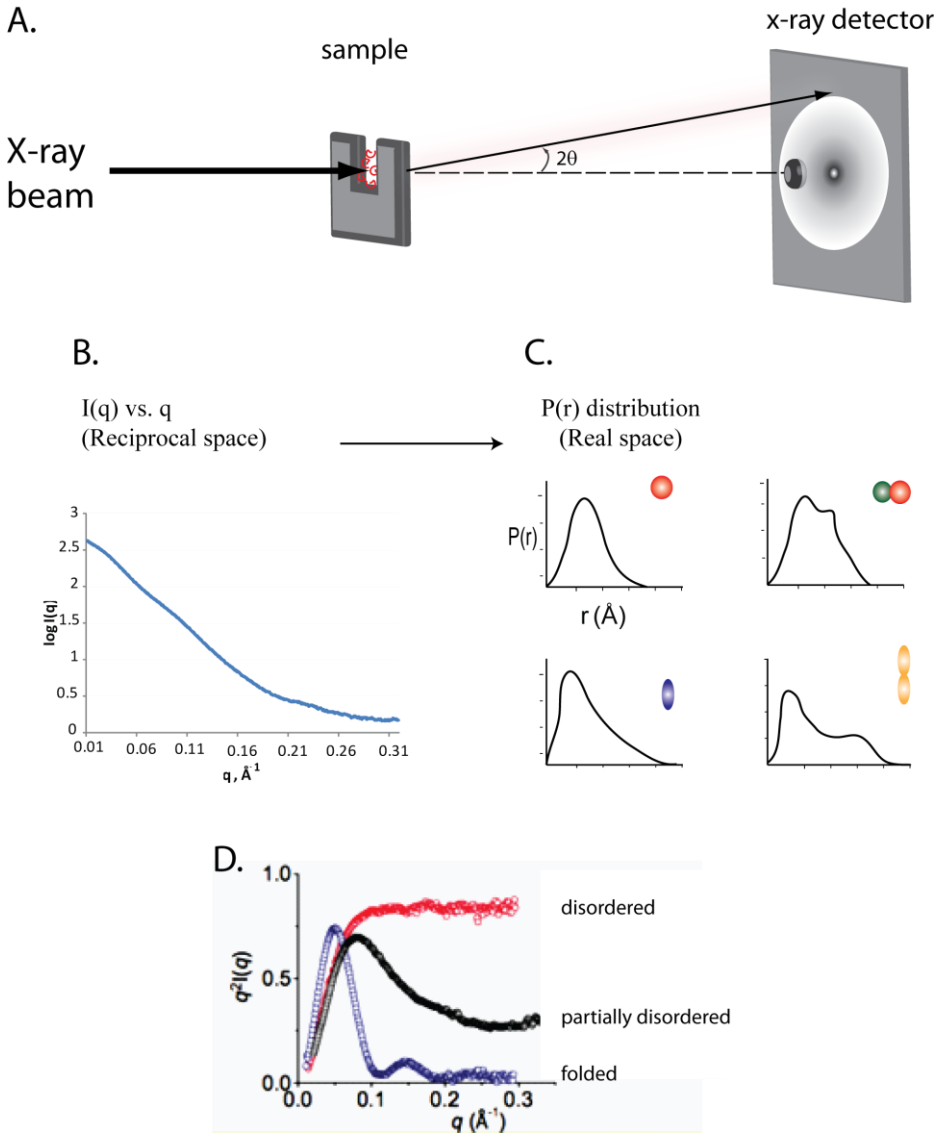


Figure 1.9. SAXS data collection and analysis. A) Schematic of SAXS data collection. B). An example of scattering curve from TBL1(1-90). C) Examples of the correlation between the P(r) function and the shape of the molecules. D) Kratky plot comparison of well folded and disordered proteins (adopted from Putnam, *et. al.*, 2007).

Fourier transform of the scattering data generates the pair distribution function, P(r), which represents the distance distribution between the electrons in the molecule plotted as a function of the radius, r (Å), in real space. Structural parameters such as the maximum dimension of the molecule (Dmax) can be estimated from the P(r) function (Fig. 1.9C) [71]. The process of calculating the P(r) function is iterative and involves choosing multiple Dmax values, calculating the P(r) function and evaluating the curve for a smooth convergence to zero at the Dmax value [69-70].

The P(r) function provides an estimate of the Rg in real space, which can be compared to the Rg from the Guinier analysis and used as a cross reference to confirm the validity of the analysis. There is a direct correlation between the shape of the P(r) curve and the shape of the molecule that can be very useful even in the absence of additional structural information. A domain with a spherical fold would have a bell shaped curve, whereas an elongated, rod-like domain would have a narrow peak with an extended shoulder (Fig. 1.9C). If the high resolution molecular structure is available, the Rg and Dmax values can be calculated based on the structure and used for direct comparison with the SAXS data. This analysis can provide important information about possible conformational changes of the molecule under different conditions.

Critical advances in SAXS data processing have established new algorithms for generating low resolution *ab initio* models of the molecule from the scattering data. DAMMIN and GASBOR are the most commonly used programs for *ab initio* 3D modeling [72-73]. DAMMIN utilizes simulated annealing procedure to search for the best bead model that satisfies the low resolution scattering data. The *ab initio* model generated in GASBOR takes into account the higher resolution SAXS data as well. The

program uses a simulated annealing algorithm in fitting the model, where the protein is represented as dummy residues, into the scattering data from reciprocal space or the P(r) function in real space. The models from GASBOR and DAMMIN are finally superimposed and averaged using the program DAMAVER in order to obtain the most probable low resolution shape reconstruction represented as ensembles of identical dummy residues or beads [74].

The most powerful application of SAXS data is when combined with high resolution structural information from x-ray crystallography, NMR or electron microscopy (EM) [70]. In the case where the structure of the protein has already been determine, the scattering profiles can be back-calculated. The goodness of the fit between the experimental and calculated scattering curves can be used for evaluating the models and for evaluating the presence of any conformational changes of the molecule.

The study presented in this thesis work utilizes SAXS as a technique to determine the overall structural architecture of TBL1, which is an oligomeric, multi-domain scaffolding protein. The approach used was to perform a series of SAXS experiments on a number of different TBL1 fragments and the full-length protein. These results provided information about the spatial organization of the TBL1 domains. Since high resolution structures of TBL1 or any of the distinct domains are not available, the studies presented in chapter III uses crystal structures of homologous TBL1 domains to generate homology models for each of the TBL1 domains. These models have been fit into the final *ab initio* envelopes including a compendium model of the full length protein.

Experimental Overview

The studies presented in this thesis work are composed of two distinct TBL1-focused research areas, which characterize (i) the mechanism of β-catenin recruitment and ubiquitination by the Siah-mediated SCF^{TBL1}-like complex and (ii) the structural architecture of TBL1 that provides insight into its role as a major scaffolding protein in the transcriptional regulation of NHRs. These studies have been performed with the scientific input of many generous collaborators.

Biochemical and functional approaches have been utilized to gain a deeper understanding of the mechanism of SCF^{TBL1}-like-targeted ubiquitination of non-phosphorylated β-catenin. Given the complexity of the multi-protein complex involved and the novelty of this E3 ligase, the molecular basis for the Siah-1-mediated ubiquitination of β-catenin was first investigated using an *in vitro* ubiquitination assay. Over-expression and purification protocols were developed for each of the SCF^{TBL1} proteins including Skp1, Siah-1, SIP and the E2 conjugating enzyme UbcH5a. Insightful details for working on full length β-catenin and different truncation mutants were provided through the collaboration with Bill Weis (Stanford University). A transient expression system using mammalian 293-6E cells was used for the expression of full length TBL1 (FL-TBL1) with the support of the Vanderbilt protein expression core facility.

Establishing an *in vitro* ubiquitination assay with purified proteins enabled a systematic analysis of β-catenin ubiquitination by the SCF^{TBL1}-like complex (chapter II). Surprisingly, the *in vitro* experiments showed that Siah-1/UbcH5a alone is sufficient to poly-ubiquitinate β-catenin and TBL1 inhibits Siah-1-mediated ubiquitination of β-

catenin. The direct physical interaction between Siah-1 and β-catenin was confirmed by nuclear magnetic resonance (NMR). The NMR chemical shift perturbation experiments proved especially powerful for studying dynamic interactions even in the case of these large, oligomeric proteins. To further characterize this novel pathway of targeting β-catenin for ubiquitination we established a collaboration with David Friedman (Vanderbilt University), who utilized the mass spectrometry-based proteomics tools to elucidate the ubiquitin modified sites and the chain linkages assembled on β-catenin by Siah-1.

The functional relevance of the biochemical data obtained in my studies was confirmed by experiments in the laboratory of Cun-Yu Wang (UCLA), using siRNA to knock down TBL1/TBLR1 and monitor their effect in cells during Wnt signaling. The data from the cell-based experiments demonstrates that TBL1 protects β-catenin from Siah-1-induced degradation during Wnt signaling. A model for the complex interplay between Siah-1-targeted degradation and TBL1-mediated protection of β-catenin was proposed (chapter II).

The second part of my thesis work is focused on characterizing the structural architecture of TBL1 in order to gain insight into its role as a major scaffolding protein in transcriptional regulation. To begin our studies, a library of TBL1 truncation mutants that contain different tandem domains of the protein was established. We used size exclusion chromatography coupled to multi-angle light scattering detectors (SEC-MALS) to obtain a direct readout for the molar mass of the proteins that is independent of their shape. The molar mass of FL-TBL1 corresponds well to a stable tetramer (chapter III). Biophysical experiments such as sedimentation velocity AUC in addition to SEC-MALS were

performed using TBL1 truncation mutants and identified the LisH as the core tetramerization domain of FL-TBL1. Site-directed mutagenesis with biochemical analyses were utilized to identify the precise residues at the LisH dimer of dimers interface.

Computational tools, Modeller and Rosetta, were used to generate homology models for each of the TBL1 domains and SAXS was utilized to generate a model for the FL-TBL1 tetramer (chapter III). SAXS is a powerful structural tool for characterizing the structure of large macromolecules and dynamic protein assemblies [70, 75]. The data were processed to obtain a low resolution *ab initio* model for the global architecture of FL-TBL1. The spatial organization of the TBL1 domains was deduced based on the *ab initio* models for the two shorter TBL1 constructs. The homology models for the domains were fitted in the SAXS model, bringing further insight into the binding region of different TBL1 substrates.

Finally, chapter IV provides a summary of the research and a proposed model for the role of TBL1 LisH-tetramer as the core unit of the scaffolding protein. Our results provide fundamental insights into the mechanism of TBL1 family members as major scaffolding proteins that serve in the transcriptional regulation of β-catenin and NHRs by mediating the precise switch between the multi-protein assemblies of activators and repressors.

CHAPTER II

DIRECT UBIQUITINATION OF β-CATENIN BY SIAH-1 AND REGULATION BY THE EXCHANGE FACTOR TBL1[1]

Introduction

β-Catenin is a ubiquitously expressed transcriptional activator in the canonical Wnt signaling pathway involved in cellular processes ranging from embryogenesis, cell proliferation, cell fate and survival to adult stem cell self-renewal and oncogenesis [40, 76]. Upon Wnt stimulation, β-catenin binds to T cell factor/Lymphoid enhancer factors (Tcf)/(Lef) initiating the expression of many genes including *cyclin D1, c-Myc, Axin2* and *vascular endothelial growth factor (VEGF)* [40]. Mutations in β-catenin and its regulatory factors, such as Adenomatous Polyposis Coli (APC), are associated with increased levels of nuclear β-catenin and in turn, to breast, colorectal, ovarian and other cancers [38, 40, 77].

Not surprisingly, the level and cellular localization of β-catenin are tightly regulated by a finely tuned balance of post-translational modifications and protein turnover [40, 77]. The most efficient means to lower the cellular levels of β-catenin is by poly-ubiquitination, leading to degradation in the 26S proteasome. Defects in this protein degradation machinery or mutations in β-catenin that prevent the recognition or processing by this machinery often lead to the stabilization of β-catenin in its oncogenically active state [77].

[1] The bulk of this chapter was published in Dimitrova YN, Li J, Lee YT, Rios-Esteves J, Friedman DB, Choi HJ, Weis WI, Wang CY, Chazin WJ. (2010) Direct ubiquitination of β-catenin by Siah-1 and regulation by the exchange factor TBL1. *J Biol.Chem*. 285(18):13507-16.

Poly-ubiquitination involves the serial action of E1, E2 and E3 enzymes, of which the substrate-recruiting E3 ligating enzymes are the most diverse. β-Catenin is recognized and ubiquitinated by a growing number of E3 ligases. Of these, the most well-studied is $SCF^{\beta\text{-}TrCP}$, a multi-protein complex that itself is regulated through the canonical Wnt signaling pathway [78-79]. In the absence of a Wnt ligand, β-catenin is phosphorylated by Glycogen synthase kinase-3β (GSK3β) and it is the phosphorylated state of the protein that is recognized by $SCF^{\beta\text{-}TrCP}$. Under conditions of genotoxic stress, activation of p53 occurs and an additional pathway for β-catenin degradation is initiated. p53 directly induces the expression of Siah-1 and in turn formation of a unique SCF-like complex SCF^{TBL1} comprised of Siah-1, Siah-1-interacting protein (SIP), Skp1, Transducin β-like 1 (TBL1) and APC [13, 49]. The physiological significance of Siah-1-targeted degradation of β-catenin is underscored by the discovery that this pathway is directly targeted by the viral oncoprotein latent membrane protein 1 (LMP1) [80]. In addition, recent studies identified two drugs, hexachlorophene and isoreserpine, which attenuate the function of β-catenin through activation of Siah-1 and subsequent proteasomal degradation [81-82].

In addition to functioning as part of the SCF^{TBL1} complex, Siah-1 functions as an E3 ligase. The Siah-1 RING E2-binding domain is linked to a substrate binding domain that directly recruits, and mediates poly-ubiquitination of many substrates including Deleted in Colorectal Cancer (DCC), Nuclear Co-Repressor (NCoR), c-Myb and synphilin-1 [83-86]. The ability of Siah-1 to serve as a simple E3 ligase as well as a component of an SCF-like complex raises the possibility of redundancy in the poly-ubiquitination pathways that lead to degradation of β-catenin.

The involvement of TBL1 in the SCF^{TBL1} E3 ligase complex is intriguing. Both TBL1 and its close isoform, TBL1-related protein (TBLR1) are exchange factors between nuclear co-activator and co-repressor complexes in the regulation of nuclear receptors and transcription factors [60]. Moreover, recent evidence shows that TBL1 acts as a co-activator of the Wnt signaling pathway by recruiting β-catenin to the promoter of Wnt target genes and stimulating their expression [6]. Thus, TBL1 appears to play a role in both activation and repression of β-catenin activity.

Previous studies by Matsuzawa and coworkers extensively characterized the ubiquitination of non-phosphorylated β-catenin through the action of SCF^{TBL1} in cells [13]. Here we report a combination of *in vitro* and cell based assays of β-catenin ubiquitination by SCF^{TBL1} and Siah-1 alone. Additionally, we mapped the physical interaction between Siah-1 and β-catenin and analyzed the effect of TBL1 on the Siah-1-mediated ubiquitination of β-catenin. These results highlight the role of TBL1 as a protector of β-catenin activity during Wnt signaling.

Results

In vitro ubiquitination of β-catenin by the SCF^{TBL1} complex

In order to further investigate the molecular basis for β-catenin degradation by the SCF^{TBL1} complex, we developed an *in vitro* ubiquitination assay using purified recombinant proteins. The E2 conjugating enzyme UbcH5a, Skp1, Siah-1, SIP, and the β-catenin substrate were overexpressed and purified from *E. coli* [87-89]. A transient expression system was used for the expression of full length TBL1 in mammalian 293-6E cells [90]. The protocols for the *in vitro* ubiquitination reaction were developed based on

previous reports [91], using proteins each purified to >95% homogeneity (Fig. 2.1A). In order to ensure that the full length β-catenin substrate is properly folded, a circular dichroism (CD) experiment was performed (Fig. A2.1). The analysis of this spectrum correlates well with the α-helical secondary structure elements observed in the x-ray crystal structure of the protein [92].

The ubiquitination reaction was initiated by addition of ATP after mixing all components and incubated for 60 minutes at 30 ^{o}C. Lane 5 in Figure 2.1B shows the

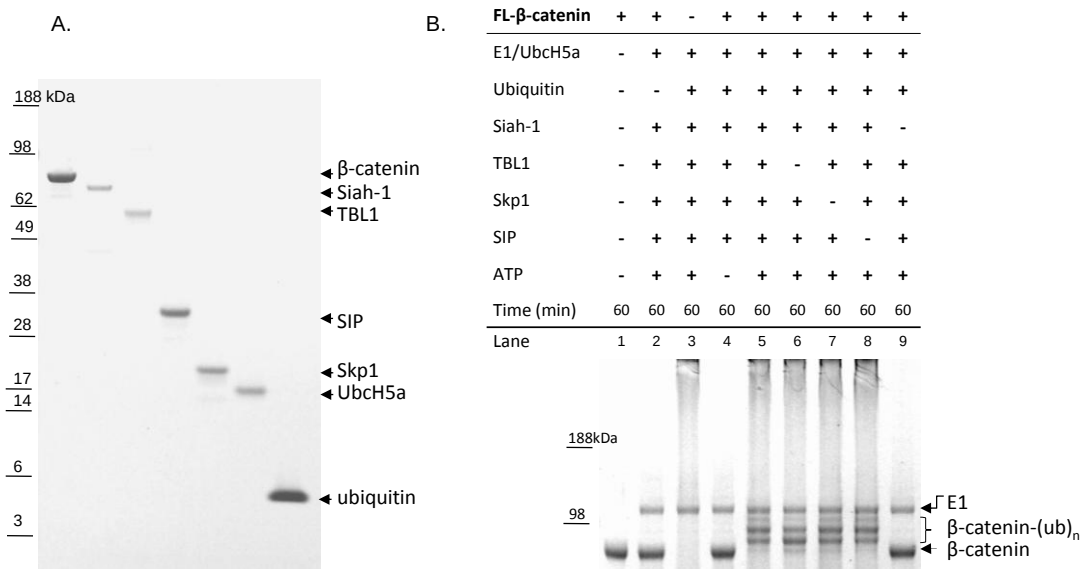


Figure 2.1. Reconstitution of the SCF^{TBL1}-like E3 ligase *in vitro*. (**A**) Recombinantly expressed and purified proteins: β-catenin, His_6-MBP-Siah-1, His_6-TBL1, His_6-SIP, His_6-Skp1, His_6-UbcH5a and ubiquitin (BostonBiochem) visualized by a Coomassie-stained SDS-PAGE gel. (**B**) *in vitro* ubiquitination of FL-β-catenin with E1, E2-UbcH5a, ubiquitin, Siah-1, SIP, Skp1, and TBL1. Upon activation with ATP the samples were incubated for 60 minutes at 30 ^{o}C. The reactions were stopped by the addition of SDS-Loading buffer and boiling for 15 minutes. The samples were resolved on a 4-12% gradient SDS-PAGE gel and visualized by Coomassie staining. Full scale Coomassie stained SDS-PAGE gel is shown in Supplementary materials Figure A2.2.

complete reaction with all components of the SCF^{TBL1} present. In this reaction, all free β-catenin has been consumed, resulting in formation of a characteristic ubiquitin ladder. Control experiments show no ubiquitination of β-catenin when ubiquitin (lane 2), β-catenin substrate (lane 3) or ATP (lane 4) are excluded from the reactions (Fig. 2.1B). The absence of Siah-1 also results in no discernable ubiquitination of β-catenin, presumably due to the inability to co-localize the E2-conjugated ubiquitin and the substrate (Fig. 2.1B, lane 9). In contrast, β-catenin poly-ubiquitination is observed even in the absence of TBL1, Skp1 or SIP (Fig. 2.1B, lanes 6-8). Together, these results show that the poly-ubiquitination of β-catenin by SCF^{TBL1} can be reconstituted *in vitro* and only Siah-1 is required for β-catenin poly-ubiquitination.

Siah-1 directly ubiquitinates β-catenin

The results from the *in vitro* ubiquitination of β-catenin by the SCF^{TBL1} complex were somewhat puzzling given the prevailing view of multi-protein SCF E3 complexes. Ideally, in the absence of any one component, the complex could not assemble and would lose its ubiquitinating activity. Our observation of β-catenin poly-ubiquitination in the absence of TBL1, Skp1 and SIP implied that non-phosphorylated β-catenin can be poly-ubiquitinated by an alternate Siah-1 mediated mechanism. Motivated by reports that Siah-1 serves as a simple E3 ligase and the observation that only Siah-1 is required for the ubiquitination of β-catenin (Fig. 2.1B, lane 9), a series of *in vitro* assays were performed to investigate if Siah-1 alone could serve as an E3 ligase for the β-catenin substrate and if so, was the reaction modulated by any of the other SCF^{TBL1} proteins.

A time-course of the ubiquitination reaction with Siah-1 alone as an E3 ligase is shown in Figure 2.2. Substantial mono-ubiquitination of β-catenin was observed within 10 minutes of activation of the reaction and both the extent of ubiquitination and the number of added ubiquitins increase over time (lanes 2-6). No ubiquitination was

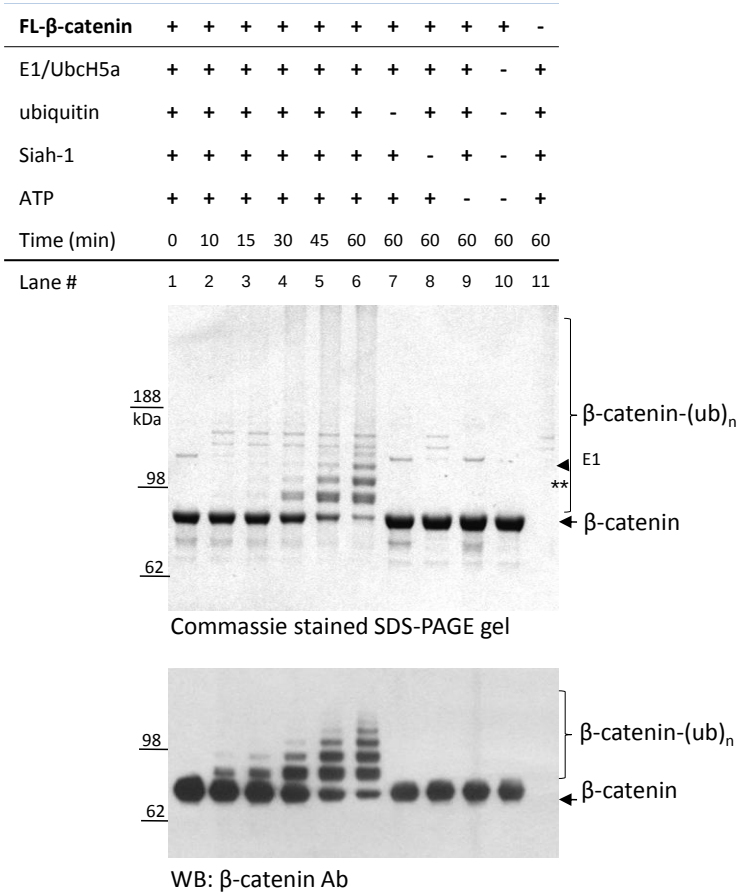


Figure 2.2. *In vitro* ubiquitination of β-catenin by Siah-1 E3 ubiquitin ligase. The *in vitro* ubiquitination of β-catenin by Siah-1 was reconstituted in a 20 µl reaction using E1 (BostonBiochem), ubiquitin, E2-UbcH5a, His_6-MBP-Siah-1 (0.1 µM) and β-catenin-FL (0.3 µM). After the addition of ATP the samples were incubated at 30 ^{o}C for different time periods. The reactions were terminated with SDS-Loading buffer and boiling for 15 minutes. ¾ of the sample volume from each of the reactions was resolved on a 4-12% gradient SDS-PAGE gel and visualized by Coomassie staining (top panel). The same ubiquitination reaction was detected on a Western blot with β-catenin antibody (Cell Signaling) using the remaining ¼ of the reaction (bottom panel).

observed in reactions lacking ubiquitin, Siah-1 or ATP (lanes 7-9). Siah-1 robustly auto-ubiquitinates but as shown in lane 11, the auto-ubiquitination reaction in the absence of β-catenin does not result in the distinctive ladder that is observed in the presence of the substrate (cf. lanes 6 vs. 11). Immunoblotting with β-catenin antibody confirms that the species observed by coomassie staining correspond to modified β-catenin substrate (Fig. 2.2, bottom panel). The specific pattern of two strong bands of ubiquitinated β-catenin seen in Figure 2.1 can arise from multiple mono-ubiquitination events. In order to confirm that β-catenin is poly-ubiquitinated we performed the *in vitro* ubiquitination reaction using K(0)-ubiquitin (Fig. A2.2B). This reaction shows a substantial overall decrease in the ladder of K(0)-ubiquitinated β-catenin compared to the ladder observed with wt-ubiquitin, confirming that β-catenin is indeed poly-ubiquitinated.

The specificity of Siah-1 for β-catenin in the *in vitro* ubiquitination assay was confirmed using a substrate Skp1 that does not bind to Siah-1. Figure A2.2A shows that Skp1 is not ubiquitinated. To further demonstrate that Siah-1 ubiquitination requires physical interaction, the assay was performed using SIP, which is known to directly interact with Siah-1 [87]. As anticipated, SIP is efficiently poly-ubiquitinated by Siah-1 (Fig. A2.2B). Thus Siah-1, acting on its own as an E3 ligase, specifically poly-ubiquitinates β-catenin *in vitro*, in a phosphorylation-independent manner.

Siah-1 forms K11 ubiquitin chains on β-catenin

In order to determine the specific ubiquitin chain assembled on β-catenin by Siah-1, the ubiquitinated substrate was characterized by tandem mass spectrometry. The products from the standard *in vitro* ubiquitination reactions were resolved on an SDS-

PAGE gel and the second band, labeled ** (Fig. 2.2, lane 6) was excised, subjected to in-gel trypsin digest, followed by liquid chromatography-mass spectrometry (LC-MS/MS). The strategy involved identifying ubiquitin peptides from the trypsin proteolysis that contain lysine residues with the mass addition of 114.04 Da from the covalently attached ubiquitin signature peptide (-GG) that remains attached after trypsin cleavage. A single modified peptide sequence, K_6TLTGK$_{11\text{-}\mathbf{GG}}$TITLEVEPSDTIENVK$_{27}$A, from ubiquitin was identified indicating a Lys11 linked ubiquitin chain on β-catenin. A number of studies have already shown that formation of Lys11 ubiquitin chains can target the substrates for degradation in the cell [93-94]. These data indicate that Siah-1/UbcH5a alone can assemble Lys11-linked ubiquitin chains on β-catenin that could lead to the proteasomal degradation of β-catenin in cells.

TBL1 inhibits in vitro poly-ubiquitination of β-catenin by Siah-1

Direct physical and functional interaction between TBL1 and β-catenin is well established [6, 13]. TBL1 and its isoform, TBLR1, are required for the transcriptional activity of β-catenin in cells by mediating recruitment of β-catenin to the promoter of Tcf/Lef target genes during Wnt signaling [6]. Moreover, in the SCF^{TBL1} complex, TBL1 plays the role of directly binding and recruiting non-phosphorylated β-catenin to the complex [13]. However, our *in vitro* ubiquitination data indicate that the presence of TBL1 is not required for the recruitment and poly-ubiquitination of β-catenin by Siah-1.

In order to obtain further insights into these observations, we compared Siah-1 *in vitro* ubiquitination assays performed under identical conditions for β-catenin alone and pre-incubated with SCF^{TBL1} (Fig. 2.3). In order to specifically monitor ubiquitination of β-catenin, the reactions were detected by Western blot with β-catenin antibody [95]. We noted that although species with long poly-ubiquitin chains are indeed generated in the reaction (Fig. 2.2), the amount of highly poly-ubiquitinated β-catenin molecules is below the detection limit and only the first few bands are observed using this approach.

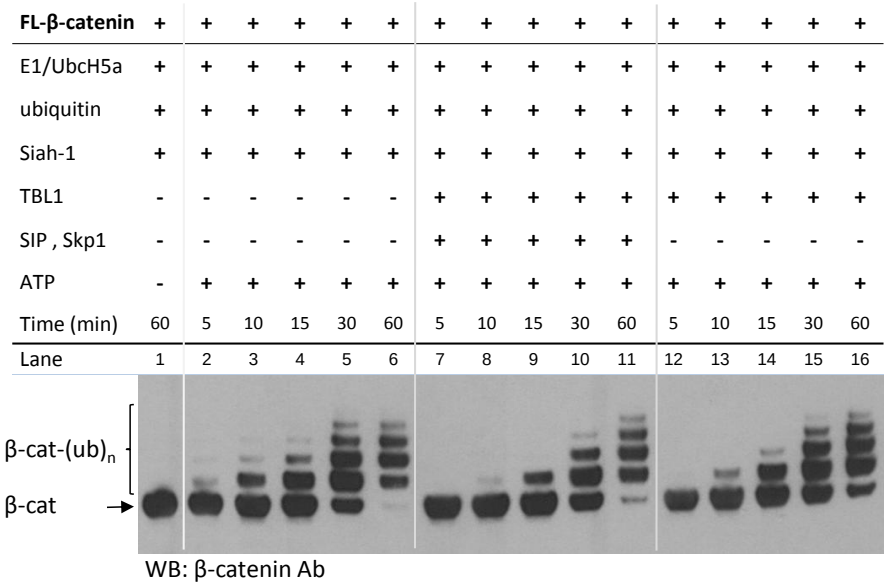


Figure 2.3. TBL1 inhibits Siah-1-mediated poly-ubiquitination of β-catenin. The *in vitro* ubiquitination assay was performed with E1, E2, ubiquitin and Siah-1. The substrate β-catenin was pre-incubated with SCF^{TBL1} or TBL1 for 30 minutes at RT before the rest of the ubiquitination components were added. The reactions was activated by the addition of ATP and incubated at 30 ^{o}C for different time periods. The reactions were resolved on an SDS-PAGE gel and detected on a Western blot with β-catenin antibody (Cell Signaling). Full scale Coomassie stained SDS-PAGE gel is shown in Figure A2.4.

Interestingly, Siah-1 alone appeared to ubiquitinate β-catenin at a faster rate than the SCF^{TBL1} complex. Figure 2.3 presents a time-course of β-catenin ubiquitination by Siah-1 in the absence (lanes 2-6) and presence (lanes 7-11) of SCF^{TBL1}. Free β-catenin was depleted by 60 min in the absence of SCF^{TBL1}, whereas a significant amount of non-ubiquitinated β-catenin still remained when TBL1 was present in the reaction (Fig. 2.3, lanes 6 vs. lane 11). Furthermore, mono-ubiquitination of β-catenin by Siah-1 alone was observed at the 5 minute point, and two additional bands of ubiquitinated β-catenin also appear at 10 and 15 min (lanes 2-4). In contrast, only one band of ubiquitinated β-catenin was observed at 15 min after pre-incubation with SCF^{TBL1} (lane 9). This clearly shows that the formation of ubiquitin chains on β-catenin is delayed in the presence of all SCF^{TBL1} proteins. Since the differences between the time courses are modest, the experiments were repeated several times, including using proteins from different purification stocks. A high reproducibility was observed between experiments, which validates the observations.

To further investigate if a single component from the SCF^{TBL1} is able to attenuate Siah-1-dependent ubiquitination of β-catenin, the effect of pre-incubating SIP, Skp1 or TBL1 with β-catenin was tested. Lanes 12-16 in Figure 2.3 show that the addition of TBL1 alone had an effect similar to the SCF^{TBL1} complex (lanes 7-11). Together, these results show that TBL1 inhibits the *in vitro* poly-ubiquitination of β-catenin by Siah-1.

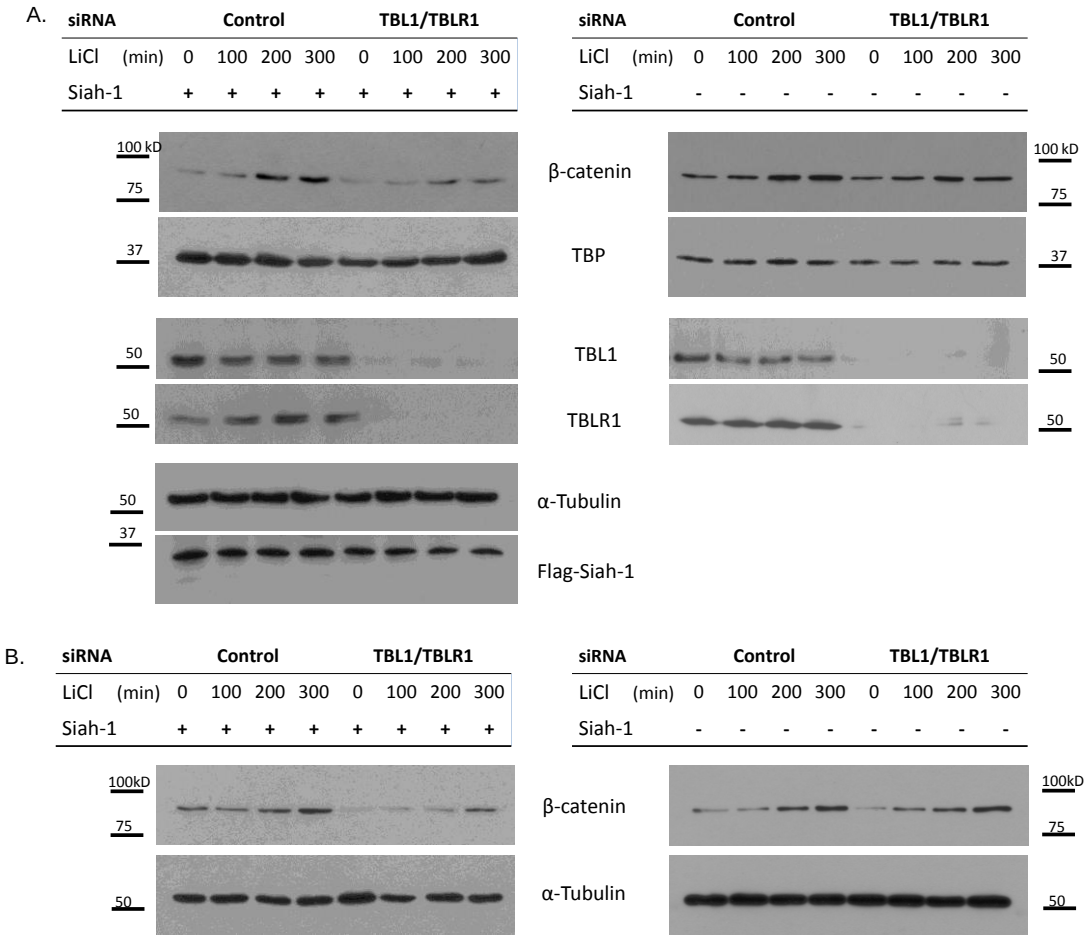


Figure 2.4. TBL1/R1 protects β-catenin from Siah-1-mediated poly-ubiquitination and subsequent degradation *in vivo*. TBL1 and TBLR1 were knocked down in 293T cells with siRNA. 24 hours later the cells were transfected with Siah-1 and treated with 20 mM LiCl. (**A**) β-catenin levels were detected in the nuclear fraction by immunoblotting. The panels on the left are with transient expression of Siah-1 and on the right are without expression of Siah-1. The levels of transfected Siah-1 in the whole cell extract is shown in the bottom two panels. (**B**) Detection of β-catenin levels in the cytoplasmic fraction of the cell.

TBL1 inhibits Siah-1-mediated β-catenin degradation in cells

Expression of Siah-1 had been correlated with reduced levels of β-catenin in cells [13, 49, 80, 96]. Overexpression of Siah-1 in the absence of a Wnt ligand reduced levels of β-catenin and lowered the induction of Tcf/Lef-target genes [13, 49]. Wnt activation also stimulated the interaction of TBL1 and TBLR1 with β-catenin [6]. In the presence of a Wnt ligand, expression of Siah-1 also decreased the amount of expressed Tcf/Lef-target genes, promoting the degradation of β-catenin, but the effect was not as substantial as in the absence of a Wnt ligand [49].

To determine the effect of TBL1/TBLR1 on β-catenin degradation during Wnt signaling we examined the levels of β-catenin in cells. It is important to note that the endogenous level of Siah-1 in cells is extremely low and does not rise to appreciable levels until p53 is activated. Thus, endogenous Siah-1 cannot ordinarily be observed on western blots. In essence, because Siah-1 is not prevalent in quiescent cells, the "normal" condition when Siah-1 is actively involved in regulating β-catenin corresponds only when there is stress to the cells so that Siah-1 is upregulated through activation of p53. Consequently, the standard and now well-accepted approach to study the effects of Siah-1 is via overexpression [13, 49, 83-86]. In order to study the roles of non-canonical pathways for β-catenin degradation, the basal phosphorylation-dependent poly-ubiquitination of β-catenin by the $SCF^{\beta TrCP}$ complex must be down regulated. The standard protocol for this is to use LiCl to inhibit GSK-3β kinase, which suppresses the phosphorylation of β-catenin and therefore its ability to be recognized and degraded by the $SCF^{\beta TrCP}$ complex [6, 38].

In our experiments, small interfering RNA (siRNA) was used to knock down the expression of TBL1 and TBLR1 in HEK293T cells, and LiCl was added after transfection of Siah-1 to monitor the effect on degradation of β-catenin. The amount of nuclear and cytoplasmic β-catenin was detected over a period of time by Western blot analysis. More than 80% reduction of TBL1 and TBLR1 expression was achieved in 293T cells by the siRNA (Fig. 2.4A, bottom panels). Expression of Siah-1 led to a significant decrease in the level of nuclear β-catenin over the time course of Wnt induction when TBL1 and TBLR1 were knocked down. In contrast, the presence of TBL1 and TBLR1 was seen to protect β-catenin from Siah-1-mediated degradation in the nucleus (Fig. 2.4A). The same effect was observed with cytoplasmic β-catenin, where a significant decrease in the level of β-catenin was seen when TBL1 and TBLR1 are knocked down (Fig. 2.4B).

In order to ensure that the level of transfected Siah-1 was comparable to previously published data and to the amount of protein expressed upon activation of p53 during DNA damage, we measured the level of Siah-1 after transfection and upon addition of adriamycin [13, 49, 97]. Our transfection protocol did not result in gross overexpression of Siah-1, and in fact the level of Siah-1 was similar to the level observed upon activation of p53 (Fig. A2.5). Taken together, our results indicate that upon Wnt signaling, TBL1 and TBLR1 serve to protect β-catenin from Siah-1-induced degradation.

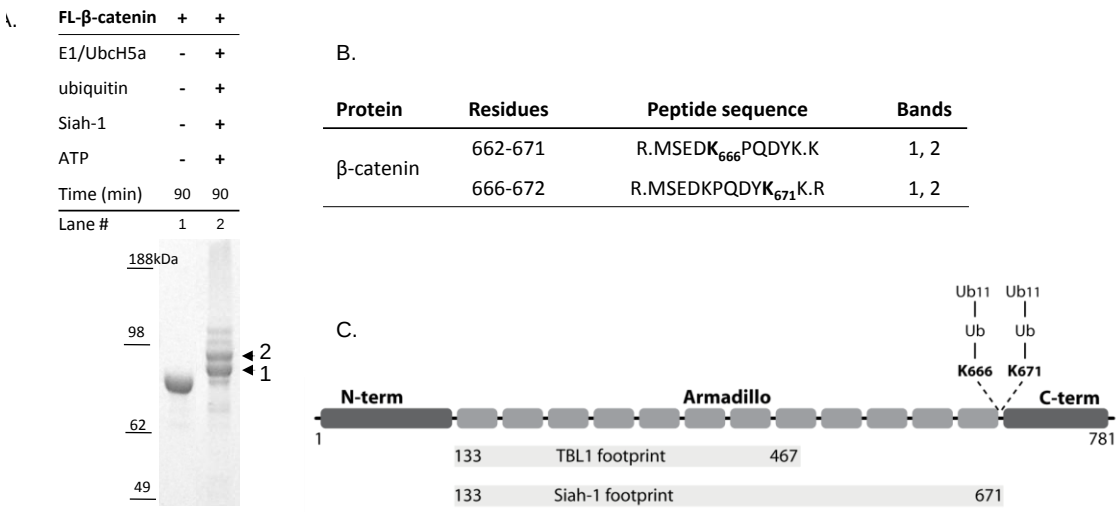


Figure 2.5. Siah-1 poly-ubiquitinates β-catenin on Lysines 666 and 671. (**A**) The *in vitro* ubiquitination assay was performed using recombinantly produced proteins as explained above. The reactions were resolved on an SDS-PAGE and stained with Coomassie blue. Bands 1 and 2 from lane 2 were excised from the gel, subjected to trypsin digestion and analyzed by liquid chromatography-mass spectrometry (LC-MS/MS). (**B**) Summary of the peptides identified from β-catenin and ubiquitin in each of the analyzed bands shown in A. The specific Lysine residues modified with the –GG signature peptide from ubiquitin is shown in bold. (**C**) Schematic presentation of β-catenin domains indicating the position of Lys666 and Lys671 ubiquitinated by Siah-1 and previously mapped TBL1 and Siah-1 binding sites.

Siah-1 ubiquitinates β-catenin at lysines outside the TBL1 binding site

The core of β-catenin is an armadillo (arm) repeat domain, which has an elongated α-helical structure that facilitates the interaction of the majority of β-catenin substrates [89, 92]. The interaction site of TBL1 with β-catenin was previously mapped to the N-terminal region of this (arm) domain, specifically residues 134-467 [6]. In order to test if TBL1 directly blocks access to the lysine residues targeted for ubiquitination by Siah-1, mass spectrometry was used to identify the ubiquitination sites on β-catenin [29, 98]. The standard *in vitro* ubiquitination reaction was carried out and the reaction

products were run on an SDS-PAGE gel (Fig. 2.5A). Bands 1 and 2 were excised, digested by trypsin and analyzed by LC-MS/MS. Lys666 or Lys671 at the C-terminal (Ct) domain of β-catenin (Ct: residues 665-781) were identified as the predominant sites of ubiquitination in both bands 1 and 2 (Fig. 2.5B). Notably, in all of the MS/MS experiments performed, no peptides with both Lys residues ubiquitinated were ever observed. Together, these data show the lysine residues on β-catenin targeted for poly-ubiquitination by Siah-1 are well outside the TBL1 binding site (Fig. 2.5C).

Siah-1 binds at the β-catenin armadillo domain

The ability of TBL1 to protect β-catenin from ubiquitination by Siah-1 suggests there may be direct competition between TBL1 and Siah-1 for β-catenin. Since the TBL1 binding site on β-catenin has been mapped to (arm) domain residues 134-467 [6], we set out to characterize the β-catenin binding site for Siah-1 using the *in vitro* ubiquitination assay. A series of experiments were performed on β-catenin-full length (FL) and truncation mutants containing the: N-terminal (Nt:1-133), Nt + armadillo repeat (Nt+arm:1-671), arm + C-terminal (arm+Ct:134-781) and (arm:134-671) domain. Poly-ubiquitination of β-catenin was observed only for (Nt+arm), (arm+Ct) and the isolated (arm) domain. These observations indicate the (arm) domain is necessary and sufficient for Siah-1 recruitment and ubiquitination (Fig. 2.6A, B).

An NMR chemical shift perturbation experiment was performed to confirm the direct physical interaction between Siah-1 and β-catenin. An ^{15}N-enriched sample of Siah-1 Substrate Binding Domain (SBD) was purified and used to acquire ^{15}N- ^{1}H TROSY-HSQC spectra in the absence and presence of unlabeled β-catenin-FL protein.

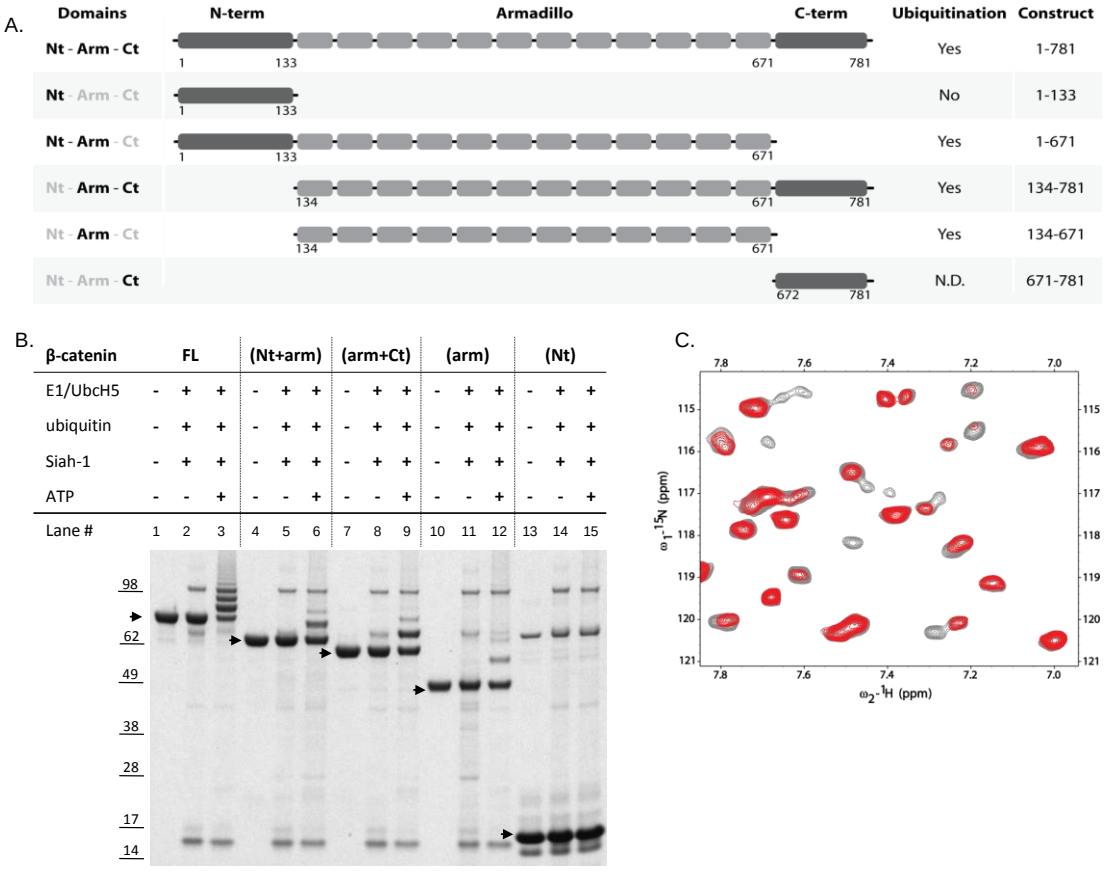


Figure 2.6. The armadillo (arm) repeat domain of β-catenin is necessary and sufficient for binding and poly-ubiquitination by Siah-1. (A) Schematic diagram of β-catenin single and tandem domains tested for Siah-1 poly-ubiquitination in (B) by *in vitro* ubiquitination assay and a summary of the ubiquitination results for each of the constructs tested. (**B**) The *in vitro* ubiquitination assay was performed with purified recombinant β-catenin FL, (Nt+arm), (arm+Ct), (Ct) and (Nt) domain as a substrate in addition to the ubiquitination machinery: E1, E2, Siah-1, ubiquitin and ATP. The ubiquitination reaction is visualized by Coomassie blue. (**C**) ^{15}N-^{1}H NMR chemical shift perturbation assay confirming the interaction between Siah-1 and β-catenin. Overlay of TROSY-HSQC spectra of ^{15}N-labeled Siah-SBD (90-282) acquired in the absence (gray) and presence (red) of β-catenin-FL. The experiment was collected at pH 8 and 30 ^{o}C with ^{15}N-Siah-SBD at 122 μM, and β-catenin added to 91.5 μM.

Binding of the ^{15}N-Siah-SBD as a dimer of 44 kDa to the 85.5 kDa β-catenin would result in a complex larger than 100 kDa, which ordinarily would lead to slow tumbling times in solution and near complete loss of signal intensity. In such solutions, useful information on binding can be obtained when examining the early part of the titration curve for systems that exhibit binding in the μM range [67-68]. Upon addition of a sub-stoichiometric amount of β-catenin to ^{15}N-Siah-SBD, spectral changes were observed including loss of intensity and perturbation of chemical shifts for a subset of peaks in the spectrum (Fig. 2.6C and for full spectrum: Fig. A2.6). These observations confirm a relatively weak, but specific interaction between Siah-1 and β-catenin.

Discussion

Efficient poly-ubiquitination and degradation of β-catenin during genotoxic stress is critical to preventing constitutive cell proliferation and preserving genomic stability. Upon UV-induced DNA damage, β-catenin is targeted for poly-ubiquitination and subsequent proteasomal degradation by a p53-induced mechanism that does not require phosphorylation of the substrate [13, 49]. The critical protein in this pathway is Siah-1, which mediates an efficient depletion of β-catenin, thereby down regulating transcription of Wnt target genes. Siah-1-mediated degradation of β-catenin was initially demonstrated through the formation of an SCF-like complex (SCF^{TBL1}), comprised of Siah-1, SIP, Skp1 and TBL1 [13]. We established an *in vitro* ubiquitination assay with reconstituted SCF^{TBL1} to investigate the mechanism of action of this complex, but interestingly we found that Siah-1 alone functions as an E3 ligase that is able to directly bind and poly-ubiquitinate β-catenin *in vitro* (Fig. 2.2 and 2.6C).

In addition to demonstrating that poly-ubiquitination of β-catenin can occur through a direct interaction with Siah-1, we established that Siah-1/UbcH5 assembles Lys11 ubiquitin chains at a novel β-catenin ubiquitination site. We have also found that during Wnt signaling TBL1, a transcriptional co-activator of β-catenin, plays a role in protecting β-catenin from Siah-1-mediated poly-ubiquitination and proteasomal degradation. TBL1 and Siah-1 were found to bind β-catenin at the (arm) domain. The absence of poly-ubiquitination of TBL1 by Siah-1 in the *in vitro* ubiquitination assay (Fig. A2.2A) rules out the possibility of TBL1 competing with β-catenin as a substrate for Siah-1. While inhibition of β-catenin ubiquitination via an allosteric binding effect cannot be ruled out, our results support a model in which TBL1 protection of β-catenin from poly-ubiquitination by Siah-1 is due to direct competition for the Siah-1 binding site.

The function of TBL1 as transcriptional co-activator and co-repressor of β-catenin

TBL1 appears to serve two roles in regulating the activity of β-catenin. Besides the initially identified role of TBL1 in recruiting β-catenin to the SCF^{TBL1} complex, it has also been shown to function as a transcriptional co-activator of β-catenin in recruiting it to the promoter site of Wnt-target genes (left side Fig. 2.7) [6]. Our results indicated that TBL1 can inhibit the poly-ubiquitination of β-catenin by Siah-1 *in vitro* (Fig. 2.3) and stabilize β-catenin in cells by protecting it from Siah-1-mediated ubiquitination and proteasomal degradation (Fig. 2.4).

We note that the *in vitro* ubiquitination of β-catenin by Siah-1 is not an extremely efficient reaction. APC is required to observe Siah-1-mediated degradation of β-catenin

in cells and that APC directly interacts with Siah-1 and with β-catenin [49, 99]. Our data indicates that Siah-1 binding to β-catenin is very dynamic and weak (Fig. 2.6C). It is likely that APC serves as a molecular bridge for Siah-1 and β-catenin, thus stabilizing the complex to increase the efficiency of ubiquitin chain formation in cells. The situation in cells would also be different because TBL1 recruits β-catenin to the Tcf/Lef transcription factors, which in effect protects β-catenin from access by Siah-1. This interpretation is supported by the inhibitory effect of TBL1 in the cell based experiments (Fig. 2.4). Furthermore, our data indicates that TBL1 (TBLR1) forms a complex with β-catenin in the cytoplasm and nucleus, since β-catenin can be protected from Siah-1 induced degradation in both cellular compartments (Fig. 2.4). Together, our results suggest a model in which down regulation of Wnt target genes during DNA damage involves binding of β-catenin by Siah-1 and poly-ubiquitination of TBL1-unboud, non-phosphorylated β-catenin (Fig. 2.7). Once a complex between β-catenin and TBL1 (TBR1) is formed upon Wnt stimulation, it binds the Wnt-target gene promoter and transcription factors such as Tcf, therefore limiting the ability of Siah-1 to access and poly-ubiquitinate β-catenin. We propose that the activation of Wnt-target genes during genotoxic stress and the regulation of cellular processes such as cell proliferation and apoptosis depend on the balance of Siah-1-mediated degradation of β-catenin versus TBL1 (TBLR1)-facilitated protection and activation of β-catenin targeted genes.

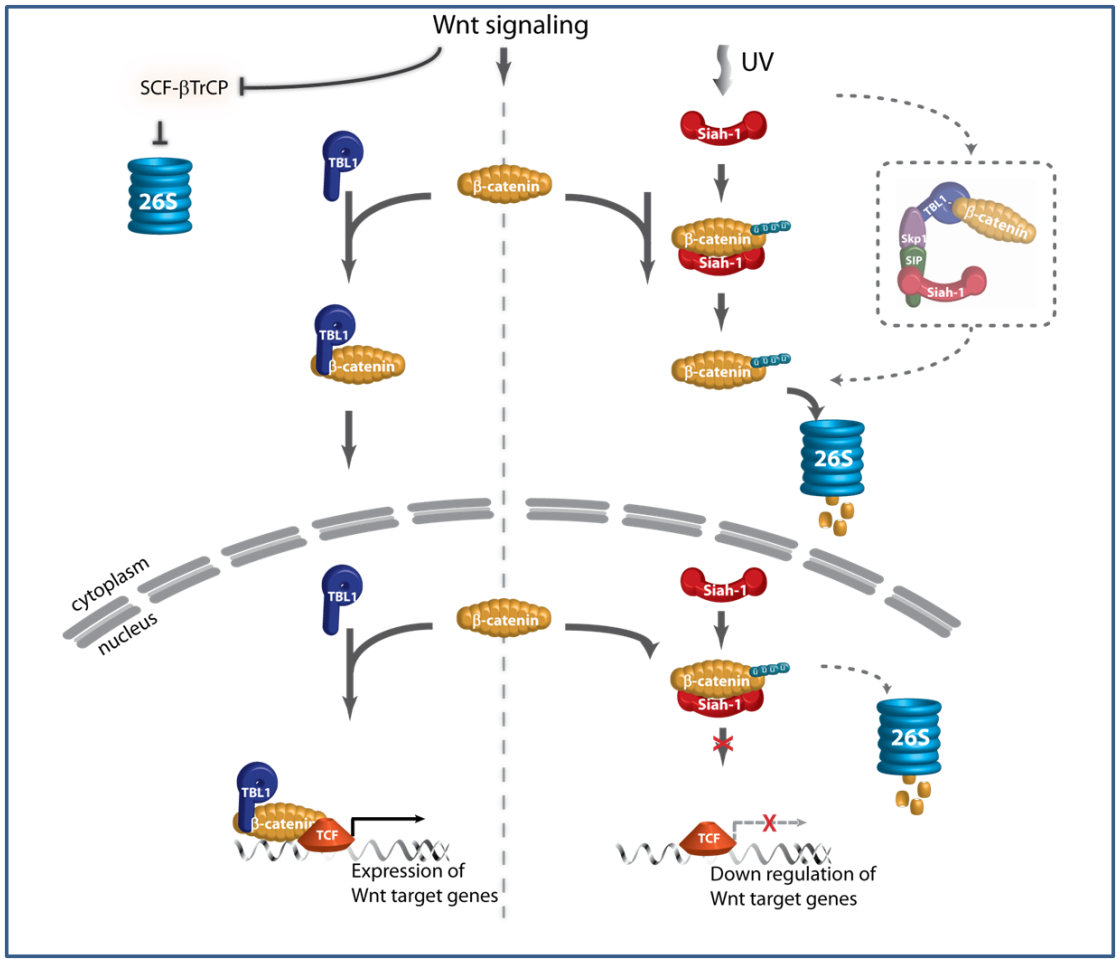


Figure 2.7. Model of the mechanism of TBL1-mediated activation and Siah-1-induced poly-ubiquitination of β-catenin. (**A**) During Wnt activation, the degradation of β-catenin is inhibited leading to its accumulation in the cytoplasm and translocation to the nucleus. TBL1/TBLR1 form a complex with β-catenin and mutually co-localize to the promoter site of Wnt-target genes, stimulating their transcription. (**B**) Upon DNA damage during Wnt signaling, Siah-1 targets free β-catenin in the cytoplasm and nucleus for poly-ubiquitination and proteasomal degradation, resulting in down regulation of Wnt-target gene transcription.

The role of SCF^{TBL1} components SIP and Skp1 on β-catenin poly-ubiquitination and degradation

Our results leave open the question as to whether Siah-1 is sufficient for the recruitment and poly-ubiquitination of β-catenin in cells or additional components from the SCF^{TBL1} complex are necessary. Siah-1 has a substrate binding domain for recruiting substrates and a RING domain for binding the E2-ubiquitin complex. It has been shown to act alone in poly-ubiquitinating many substrates such as DCC, NCoR and synphilin-1 [83-84, 86]. Our data also indicates that Siah-1 can function as a simple E3 ligase for poly-ubiquitination of β-catenin. If Siah-1 functions on its own, what is the role of SIP? One possibility is that SIP, like DCC and NCoR, is simply another target of Siah-1 for poly-ubiquitination and degradation. *In vitro* ubiquitination reactions show an efficient poly-ubiquitination of SIP in the presence or absence of Skp1, TBL1 and β-catenin (Fig. A2.2B). Over-expression of SIP mutants that cannot interact with Siah-1 prevented the efficient degradation of β-catenin in HEK293T cells and the level of SIP was essential for β-catenin degradation in gastric and renal cancer cells [100-102]. However, overexpression of only Siah-1 was sufficient for β-catenin degradation in cells and the *in vitro* ubiquitination of β-catenin was not affected by the addition of SIP (Fig. 2.4 and A2.3B) [49, 80]. It is possible that the interaction between SIP and Siah-1 is important for the proper localization of Siah-1 or for protection of Siah-1 from auto-ubiquitination and degradation.

The role of Skp1 as an adaptor is to provide an anchor in SCF ligases between Cullins, in our case SIP, and the F Box–protein TBL1. Interaction between SIP and Skp1 has been demonstrated by affinity chromatography and NMR chemical shift perturbation

experiments [87]. However, evidence for an effective interaction between the F Box domain of TBL1 and Skp1 is lacking. A wealth of biochemical and structural data have established that the binding between Skp1 and F Box domains from Skp2 or β-TrCP is very strong with a half-life of the Skp1-Skp2 complex longer than 9 hours and the surface area of the core interface of more than 2000Å^2 [103-104]. TBL1 has a putative F Box domain that appears to have a very low affinity for Skp1. Previous studies using the *Drosophila* homologue of TBL1, Ebi have been unsuccessful to detect an interaction between Ebi and Skp1 [105]. We have also been unable to detect this interaction by pull down experiments with purified proteins, co-expression of Skp1 and TBL1 followed by pull downs and negative gel shift assays. Using the very sensitive NMR chemical shift perturbation assay, we found an extremely weak (K_D ~mM) interaction between Skp1 and TBL1 (1-170) (Fig. A2.7). Thus, it is unclear how Skp1 is able to play the role of the adaptor in the SCF^{TBL1} complex. Although additional experiments are required to systematically evaluate if the SCF^{TBL1} does indeed form *in vivo* as postulated, the role of Siah-1 in degradation of β-catenin and the modulation of this activity by TBL1 are clear.

The SCF^{TBL1} complex closely resembles the *Drosophila* Sina/Phyl/Ebi complex. In *Drosophila*, the transcriptional repressor Ttk88 is recruited and poly-ubiquitinated by the Sina (Siah-1 homologue)/Phyllopod (Phyl) complex for proteasomal degradation [12]. Phyl functions as an adaptor of Sina in recruiting different substrates for poly-ubiquitination during *Drosophila* neurogenesis [106]. Both Phyl and mammalian SIP contain a Siah-binding motif or a degron sequence that has been crystallized in complex with Siah-SBD, demonstrating an almost identical interaction [107]. Ttk88 can directly bind to Sina and Phyl, but it also requires Ebi (a homologue of TBL1) for an efficient

poly-ubiquitination and degradation of Ttk88. In contrast to TBL1, which directly binds β-catenin, Ebi has a strong affinity for Sina and Phyl, but weak and indirect interaction with the substrate Ttk88 [105]. The similarity between mammalian and *Drosophila* Siah/Sina-mediated ubiquitin complexes suggests the formation of a Siah-1-mediated E3 ligase that does not resemble the conventional SCF complex. Siah-1 and Sina can directly bind and recruit the substrate for poly-ubiquitination, but they also require the assistance of the adaptor protein SIP/Phyl. Skp1 has been demonstrated to be dispensable to the function of the complex by us and others [105]. Most interestingly Ebi does not interact directly with the Ttk88 substrate, but with Sina and Phyl, whereas TBL1 appears to play a dual role in facilitating both activation and degradation of protein substrates.

Tight regulation of β-catenin by multiple E3 ligases

The existence of multiple ubiquitination pathways leading to degradation implies the physiological importance of tightly regulating β-catenin. Besides the major E3 ligases $SCF^{\beta\text{-}TrCP}$ and the p53-induced Siah-1, poly-ubiquitination of β-catenin is also initiated by a recently identified single subunit E3 ligase, Jade-1 [108]. Interestingly, membrane-associated β-catenin is regulated by a different set of E3 ligases, such as Hakai and the muscle specific Ozz E3 ubiquitin ligase [109-110]. Thus the mechanism of recognition, recruitment and poly-ubiquitination of β-catenin by different E3 ligases as well as the signaling event initiating the degradation of the substrate appears to vary substantially. The activation of the $SCF^{\beta\text{-}TrCP}$ depends on the presence or absence of a Wnt ligand, whereas Siah-1 mediated β-catenin degradation is induced by activation of p53 during genotoxic stress. In the $SCF^{\beta\text{-}TrCP}$ complex, β-TrCP binds the phosphorylated N-terminal

domain of β-catenin leading to the attachment of ubiquitin chains on Lys19 and Lys49 [1, 48]. Here, we show that Siah-1 binds the non-phosphorylated β-catenin, leading to the attachment of ubiquitin chains at the very C-terminal region of the (arm) domain on Lys666 and Lys671. Notably, this coincides with the binding surface of most transcription co-activator complexes such as p300/CBP and TRRAP/TIP60 histone acetyltransferaces (HATs) and the PAF1 complex [39]. Thus poly-ubiquitination of β-catenin by Siah-1 could prevent the transcription of many Wnt target genes. Mutations in β-catenin can lead to inefficient degradation and therefore greater accumulation in the nucleus. The corresponding over activation of Tcf/Lef genes by β-catenin is believed to be responsible for the initiation and progression of many types of cancer [38]. Understanding the uniqueness and precise mechanism of action of each E3 ubiquitin ligase targeting β-catenin for degradation and the molecular basis for defective activity of mutants is an important goal for understanding the accumulation of β-catenin in the cell and its relationship to oncogenesis.

Experimental Procedures

Bacterial protein expression and purification

Full length Siah-1 (residues 1-282) was expressed as a His_6-maltose binding protein (MBP)-fusion protein. The human cDNA was sub-cloned in pLM302 plasmid (Laura Mizoue, Center for Structural Biology, Vanderbilt University), which contained a 3C-precision protease cleavage site after the MBP tag. The SBD of Siah-1, (residues 90-282) was sub-cloned in a pET28a vector (Novagen) with an N-terminal His_6-tag containing a thrombin cleavage site. SIP and Skp1 were expressed as His_6-constructs

from pET28a vectors as previously described [87]. UbcH5a was expressed as a His_6-fusion construct in a pET15 vector. Full length murine β-catenin (residues 1-781) and four truncation mutant constructs (Nt: residues 1-133), (Nt+arm: residues 1-671), (arm: residues 134-671) and (arm+Ct: residues 1-781) were expressed as glutathione S-transferase (GST)-fusion proteins from pGEX vectors as previously described [89].

The proteins were overexpressed in the *Escherichia Coli* BL21(DE3) strain. Cells were grown at 37 ^{o}C until they reached A_{600} of 0.6-0.8 and were then induced with 0.5 mM isopropyl thiogalactoside (IPTG) for 3-4 hours at 25 ^{o}C for Siah-1 constructs, 30 ^{o}C for β-catenin constructs and 37 ^{o}C for SIP, Skp1 and UbcH5a. Purification of Siah-1, SIP, Skp1 and UbcH5a was performed by Ni-NTA (Qiagen) followed by a Source Q chromatography [87, 111]. Expression of ^{15}N-labeled Siah-SBD was carried in minimal medium supplemented with $^{15}NH_4Cl$ as the sole nitrogen source following the same protocol. β-catenin constructs were affinity purified on Glutathione Sepharose 4B (Amersham) followed by cleavage of the GST-fusion tag, Source Q and size exclusion chromatography, as previously described [89].

Recombinant expression of TBL1 in mammalian 293-6E cells

Full length murine TBL1 (residues 1-526) was sub-cloned into pTT5 vector used for intracellular expression of proteins in mammalian 293 cells [90]. N-terminal His_6-tag followed by a 3C-precision protease cleavage site was cloned into the XbaI/EcoRI site of pTT5. Then the TBL1 cDNA was inserted into the EcoRI/NotI site of pTT5. Human Embryonic Kidney (HEK) 293-6E cells (Invitrogen) with stably expressing Ebstein-Barr virus nuclear antigen 1 (EBNA1) were grown in suspension in low-calcium-hybridoma

serum-free medium (HSFM), under standard conditions at 37 ^{o}C and 5% CO_2, supplemented with 1% bovine calf serum (BCS), 50 µg ml^{-1} Genericin, 0.1% Pluronic F-68 (Sigma) and 10 mM HEPES [90]. The cells were resuspended in fresh HSFM medium with 1% BCS at a density of 1.0 x 10^6 cells ml^{-1} three hours before transfection. TBL1 was transfected using linear Polyethylenimine (PEI) (Aldrich) into 300 ml of 293-6E cells. 48 hours post transfection, the cell pellet is harvested and stored at -20 ^{o}C.

Recombinant His_6-TBL1 was affinity purified on a Ni-NTA resin using 25 mM Tris-Cl, pH 7.5, 300 mM NaCl, 5 mM Imidazole, 5 mM β-Mercapto Ethanol (βME) as NiA buffer and 300 mM Imidazole in the NiB buffer. The final yield of purified TBL1 from 293-6E cells is between 4-6 mg/L.

In vitro ubiquitination assay of FL-β-catenin

All ubiquitination experiments were carried out at a final volume of 20 µl including: E1(BostonBiochem) at 52 nM, His_6-E2-UbcH5a at 0.6 µM, ubiquitin (BostonBiochem) at 50 µM, His_6-MBP-Siah-1 at 0.18 µM, SIP at 1.5 µM, Skp1 at 1.5 µM, and His_6-TBL1 at 1.5 µM. The assay was performed in ubiquitination buffer containing 100 mM NaCl, 1 mM DTT, 5 mM $MgCl_2$ and 25 mM Tris-Cl, pH 7.5 for 0.25 µM β-catenin-FL and (Nt) domain or at pH 8, for β-catenin (arm) and (arm+Ct) constructs. The reactions were activated with 5 mM ATP and incubated for different time periods, as indicated on the specific experiments, at 30 ^{o}C. To stop the ubiquitination reaction, the samples were incubated for 15 minutes at 90 ^{o}C after the addition of 5 µl SDS-Loading buffer. Reactions were resolved on a NuPAGE 4-12% Bis-Tris gradient gel (Invitrogen) and detected by a SimplyBlue SafeStain (Invitrogen). Ubiquitination

reactions analyzed by Western blotting using C-terminal β-catenin primary antibody (Cell Signaling) and visualized with goat anti-rabbit-horseradish peroxidase (HRP) by SuperSignal West Pico Chemiluminescent Substrate (Thermo Scientific).

NMR samples and chemical shift perturbation assay

^{15}N- ^{1}H TROSY HSQC spectra were acquired for ^{15}N-enriched Siah-SBD at 122 μM in 25 mM Tris-Cl, pH 8, 100 mM NaCl and 10 mM βME in 90% H_2O/ 10% D_2O. One sample was free Siah-SBD and the second also contained 91.5 μM unlabeled β-catenin-FL. The NMR experiments were performed using a Bruker DRX 800 MHz spectrometer equipped with a CryoProbe. The ^{15}N-^{1}H TROSY HSQC spectra were acquired at 303 K with 128 scans. Data were processed using Topspin 2.0b (Bruker) and analyzed with Sparky [81].

Mapping ubiquitination sites on β-catenin and ubiquitin chain formation by mass spectrometry analysis

Proteins were separated by SDS-PAGE and the gel was stained with SimplyBlue SafeStain (Invitrogen). Individual protein bands were excised, equilibrated in 50 mM NH_4HCO_3, reduced with DTT (3 mM in 100 mM NH_4HCO_3, 37 °C for 15 min) and alkylated with iodoacetamide (6 mM in 50 mM NH_4HCO_3 for 15 min). The gel slice was then dehydrated with acetonitrile and rehydrated with 15 μL 12.5 mM NH_4HCO_3 containing 0.01 μg/μL modified trypsin (Promega), and trypsin digestion was carried out for >2 hours at 37 °C. Peptides were extracted with 60% acetonitrile, 0.1% trifluoroacetic acid, and dried by vacuum centrifugation and resuspended in 10 μL 0.1%

formic acid. LC-MS/MS analysis of the peptides was performed using a Thermo LTQ ion trap mass spectrometer equipped with a Thermo MicroAS autosampler and Thermo Surveyor HPLC pump, Nanospray source, and Xcalibur 2.0 SR2 instrument control. The peptides were resolved on a fused silica capillary column, 100 µm × 15 cm, packed with C18 resin (Jupiter C_{18}, 5 µm, 300 Å, Phenomonex, Torrance, CA) using a 95 min gradient of increasing acetonitrile with 0.1% formic acid. MS/MS scans were acquired using an isolation width of 2 *m*/*z*, and activation time of 30 ms, activation Q of 0.250, and 35% normalized collision energy using 1 microscan and maximum injection time of 100 for each scan. The MS/MS spectra of the peptides were acquired using data-dependent scanning (top-five) with dynamic exclusion (60 sec exclusion, list size = 50, repeat count = 1). Individual MS/MS fragmentation spectra were then searched against the IPI_mouse database (Feb 2008) allowing for complete carbamidomethylation of cysteine, and partial modification by oxidation of methionine. Partial modification of lysine by +114 Da was also selected to detect ubiquitinated peptides (mass shift due to GG ubiquitin sequence that remains after trypsin digestion). All candidate spectra were manually inspected for verification of ubiquitination.

Cell culture, siRNA transfection and Western blot analysis

Human embryonic kidney 293T cells were cultured in DMEM media containing 10% FBS at 37°C in a 5% CO_2/ 95% air atmosphere. For siRNA transfections, 1×10^6 293T cells were seeded into 6-well plates for 12 hours and then transfected with various amount of siRNA using Lipofectamine 2000 according to the manufacturer's protocol (Invitrogen). Expression vector of Flag-Siah-1 or control empty vector were transfected

using Fugene6 (Roche) 24 hours after the transfection with siRNA. Cells were treated, unless otherwise indicated, with 20 mM LiCl. Nuclear and cytosolic extracts were prepared and 5 µg of extract was loaded on SDS-PAGE and followed by Western blot analysis using an enhanced chemiluminescence reagent (Thermo Scientific). siRNAs were synthesized by Dharmacon Research and sequences are used as previously described [6]. All antibodies used in the Western blot are commercial available: β-catenin (BD), TBL1 (Abcam), TBLR1 (Bethyl Laboratories), and α-Tubulin (Sata Cruz).

CHAPTER III

TAMING THE BEAST: THE STRUCTURAL BASIS FOR THE FUNCTION OF TBL1

Introduction

The mechanism controlling the switch between gene activation and repression is critically important for understanding transcriptional regulation. Recent models for regulated gene expression in higher eukaryotes involve highly controlled, dynamic exchange between co-activators and co-repressors that are recruited to DNA bound transcription factors. TBL1 and TBLR1 are nuclear exchange factors initially identified as part of the nuclear receptor co-repressor (N-CoR) and silencing mediator for the retinoid and thyroid hormone receptors (SMRT) complexes [9, 60]. TBL1 and TBLR1 have very high sequence identity, and are expressed from two different chromosome locations, *TBL1* from chromosome X and *TBLR1* from chromosome 3 [4]. Upon ligand activation of nuclear hormone receptors, TBL1 and TBLR1 mediate the release of the NCoR/SMRT co-repressor complex, and its poly-ubiquitination and subsequent proteasomal degradation. The discharge of the co-repressor complexes occurs in parallel with the recruitment of co-activators to the DNA promoter site. Interestingly, TBL1 and/or TBLR1 were found present at the promoter site of hormone receptors in both the activated and the repressed state of the genes.

In a similar manner, TBL1 and TBLR1 are also found to facilitate the exchange of co-activators with co-repressors from the Tcf/Lef transcription factors at Wnt promoter sites [6]. β-Catenin is a transcriptional co-activator in the Wnt signaling pathway that

mediates the expression of target genes such as *c-Myc*, *cyclin D1* and *matrix metalloproteinases* (*MMPs*). Previous studies, as explained in more detail in chapter II, have demonstrated that in the absence of a Wnt ligand, the promoters of Wnt-targeted genes is occupied by the co-repressors TLE and HDAC1 that inhibit the firing of Wnt target genes [6]. The release of Wnt signals prevents the proteasomal degradation of β-catenin, leading to the accumulation and TBL1-mediated recruitment of β-catenin to the promoter site of Wnt inducible genes. The localization of TBL1- β-catenin complex at the promoter displaces the co-repressors and induces the expression of the Wnt genes. ChIP analysis of Wnt target promoters identified TBL1 at the repressed Wnt promoter with TCF4, TLE and HDAC1, as well as in a complex with co-activators, β-catenin and TCF4, similar to the regulatory mechanism of nuclear hormone receptors [6]. These results demonstrate the function of TBL1 proteins as nuclear exchange factors that function in transcriptional control of gene expression. TBL1 and TBLR1 also play a unique role, not only participants in different multi-protein complexes, but also as facilitators of protein-protein interactions. Despite the wealth of functional data, the molecular mechanism of TBL1-orchestrated recruitment of co-activators and co-repressors is not yet understood. Obtaining information about the structural architecture and the domain organization of TBL1 will provide valuable insights how it can facilitate the recruitment and formation of protein assemblies.

TBL1 is a multi-domain protein that based on sequence homology is predicted to have an N-terminal LisH domain followed by FBox, coiled coil (CC) and WD40 domains (Fig. 3.1C). The LisH region of TBL1 is a known dimerization motif identified in more than 100 eukaryotic proteins [112]. The importance of the small LisH domain was

demonstrated in a mutational analysis of the residues conserved in TBL1-LisH [51]. Mutants with perturbed dimer subunit interface are not able to translocate to the nucleus and have a significant decrease in protein stability and half-life in the cell.

LisH domains are often found in concert with WD40 repeat domains, as is the case for TBL1. WD40 domains are well recognized protein-binding modules in signaling proteins with a characteristic β-propeller structure. The ability of TBL1 to facilitate many protein interactions and to play a major role as a nuclear exchange factor is expected to be largely facilitated by the WD40 domain and the potentiality of TBL1 to oligomerize. Previous cell based experiments have shown that TBL1 forms homo- and TBL1/TBLR1 hetero-oligomers, but what remains unknown is the exact oligomeric state of the complexes [51, 113].

In order to gain insight into the function of TBL1, we have determined the oligomeric state of TBL1 and generated a structural model by combining small angle x-ray scattering (SAXS) with homology modeling, and biochemical experiments including analytical ultracentrifugation (AUC) and size-exclusion chromatography - multi-angle light scattering (SEC-MALS) approaches.

Results

Production and oligomerization state of full-length TBL1

In order to pursue biochemical and structural studies of TBL1, a method for the recombinant expression and purification of the full length protein was established. FL-TBL1 constructs with N-terminal His_6-, His_6-SUMO- and GST- tags were expressed in different *E.coli* strains at 18, 22, 25 and 37^{o}C. These expression and solubility tests

revealed that TBL1 was expressed predominantly in inclusion bodies, except for His_6-TBL1 and His_6-SUMO-TBL1 grown at 18^oC in Rosetta(DE3) cells. These conditions yielded soluble protein. Since the His_6-TBL1 protein has a smaller N-terminal tag it was selected for all further experiments.

To test the purity and homogeneity of the protein, we used negative stain electron microscopy (EM). The electron micrographs showed that these samples are highly contaminated, presumably with the *E.coli* chaperone GroEL. Further mass spectrometry analysis of the purified His_6-TBL1 confirmed that the sample was contaminated with a 60 kDa *E. coli* chaperone. The similarity of the molecular masses of His_6-TBL1 and the *E.coli* chaperone (58.8 and 60 kDa, respectively) made the differentiation between the two proteins very challenging by SDS-PAGE or native gels. Moreover, TBL1 bound tightly to the chaperone and efforts to separate the two proteins were not successful. Thus, two new expression systems for TBL1 were developed, one involving baculovirus Sf9 insect cells and the other Human Embryonic Kidney (HEK) 293-6E cells (Invitrogen) with stably expressing Ebstein-Barr virus nuclear antigen 1 (EBNA) [90]. His_6-TBL1 was successfully expressed in 293-6E cells in collaboration with Vito Quaranta's laboratory and then produced by the protein expression core at Vanderbilt University.

High yield of purified FL-TBL1 was produced following Co^{2+}-affinity, anion exchange and size exclusion chromatography (SEC) (Fig. A3.1). Notably, the elution profile of TBL1 from SEC suggested that it was an oligomeric protein. In order to determine the oligomeric state of purified His_6-tagged TBL1, SEC-MALS detection was undertaken (Fig. 3.1A, B). Analysis of the SEC-MALS profile indicated that the molecular mass of TBL1 is 245.7 kDa (1% error), which corresponds well to a tetramer.

These results are consistent with previous research demonstrating that Sif2p, the *S. cerevisiae* homologue of TBL1, also forms a tetramer [114].

TBL1 oligomerizes through the LisH domain

Disorder predictions indicate that TBL1 has two structured regions, the N-terminal LisH-FBox (1-121) and the C-terminal WD40 domain (170-527), connected by a 50 residue linker (Fig. 3.1C). The LisH-FBox region covers residues 2-86, leaving residues 87-121 structurally uncharacterized. According to COILS prediction algorithms,

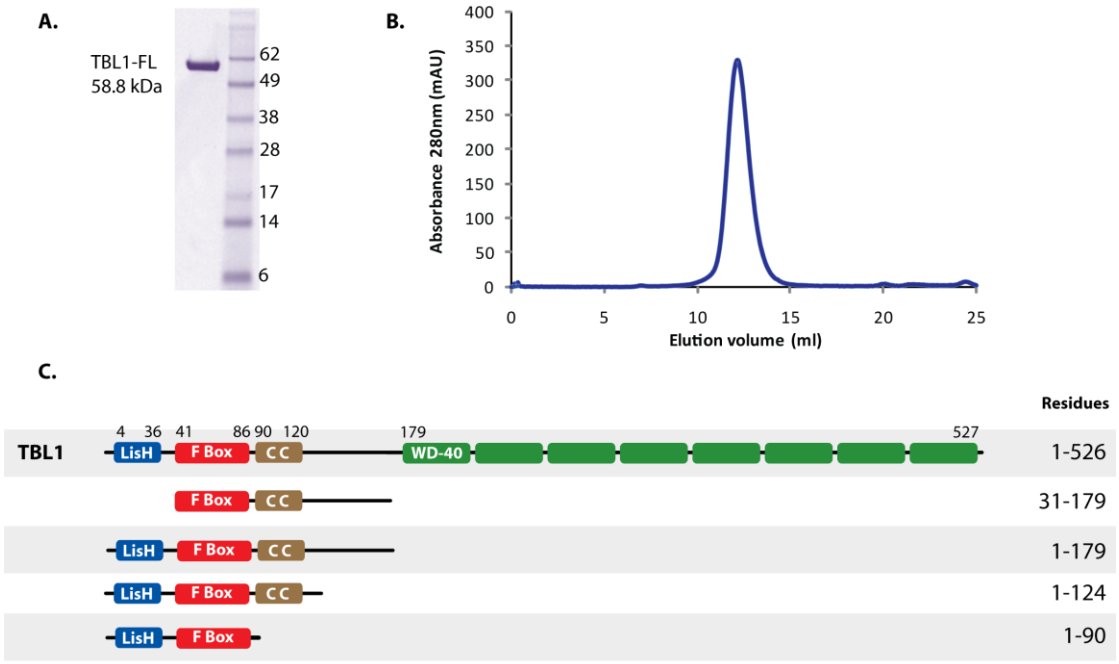
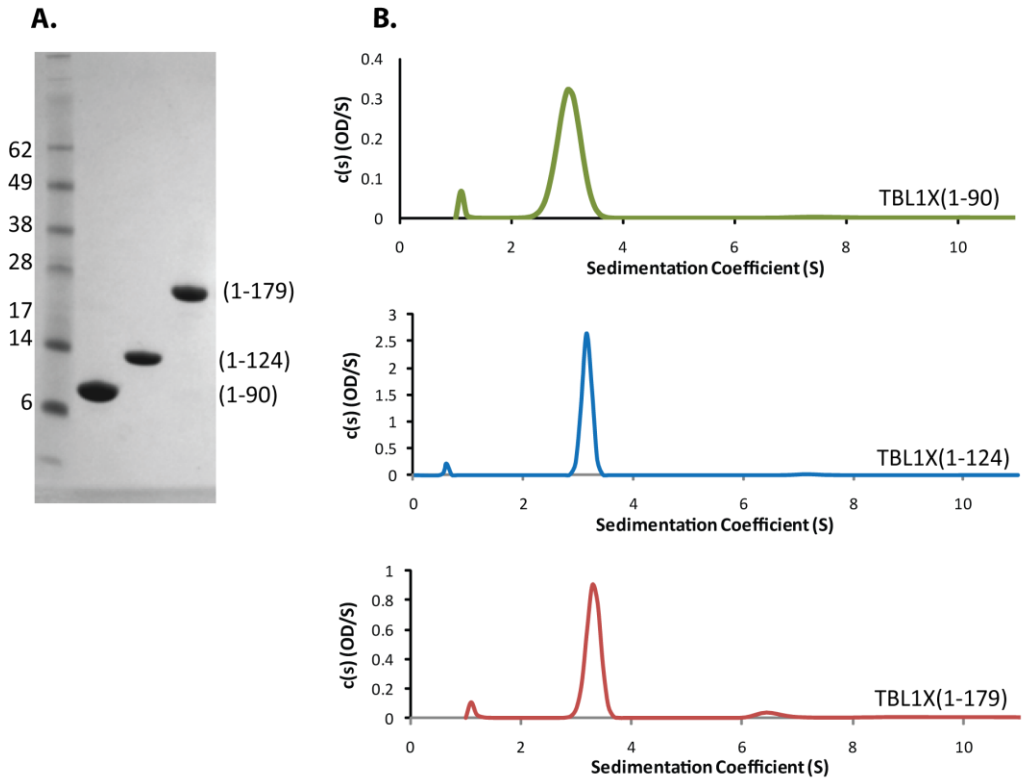


Figure 3.1. Biochemical analysis of FL-TBL1, domain representation and constructs. A) Coomassie stained SDS-PAGE of FL-TBL1. B) Elution profile of TBL1 from Superose 6 SEC-MALS in 25 mM sodium phosphate pH at 7.5, 200 mM NaCl, 5% Glycerol and 5 mM βME. C) Schematic representation of TBL1 domains and list of deletion constructs.

residues (90-121) should form a coiled coil [115]. In order to further characterize the oligomerization state of TBL1 and determine the domains contributing to oligomerization, sedimentation velocity AUC experiments were performed using three TBL1 constructs: (LisH-FBox, 1-90), (LisH-FBox-CC, 1-124), (LisH-FBox-CC-linker, 1-179) (Fig. 3.1C). A coomassie-stained gel of the pure recombinant proteins is shown in Fig. 3.2A. Analysis of the AUC velocity data for the largest construct, TBL1 (1-179, 21 kDa) shows that it has a sedimentation value (*s*) of 3.30 +/- 0.004 (S). This corresponds to a molecular mass of 73 kDa and suggests the formation of a tetramer (Fig. 3.2B, C). Truncation of the 50 residue linker results in a highly stable construct TBL1(1-124) with a mass of 15 kDa, which sediments with an s value of 3.16 +/- 0.006, consistent with the formation of a 61 kDa tetramer. Thus, the linker does not have a significant contribution to tetramerization. Next, to test if the LisH-FBox domains are sufficient for the formation of a tetramer, sedimentation velocity experiment using TBL1(1-90) were performed. Analysis of the data shows that TBL1 (1-90, 12 kDa) has sedimentation value (*s*) of 3.19 +/- 0.005, which is calculated to be 49 kDa and corresponds to a tetramer as well (Fig. 3.2). Based on these results we can conclude that the LisH domain is the primary mediator of tetramerization and that the linker and putative CC are not required.

The LisH is a known oligomerization domain that has been shown to induce protein dimerization in the Lis1 and FOP proteins, and tetramerization in Sif2p [114, 116-117]. However, in order to confirm that the LisH is sufficient for the tetramerization of TBL1, a deletion mutant of TBL1 that lacks the LisH domain, but includes the FBox-CC-linker (residues 31-179) was characterized. This mutant was not very stable and more



C.

protein	s(w)	rmsd	frictional ratio	experimental molar mass	calculated molar mass
TBL1 (1-90)	3.186	0.0051	1.51	11,999	48,815
TBL1 (1-124)	3.156	0.0056	1.76	15,397	60,991
TBL1 (1-179)	3.305	0.0044	1.90	21,010	73,257

Figure 3.2. Analysis of sedimentation velocity data for TBL1 deletion mutants. (A) Purified His_6-TBL1 constructs (1-90), (1-124) and (1-179) resolved on a coomassie-stained SDS-PAGE gel. (B) Sedimentation coefficient distribution for each of the TBL1 constructs, the calculated sedimentation value [c(s)] is plotted as a function of the sedimentation coefficient (s). (C) Summary of the results from the sedimentation velocity analysis.

prone to aggregation than constructs that contain the LisH domain. Thus, the LisH also plays an important structural role in stabilizing the N-terminal region of TBL1. Considering the instability of the construct, we selected SEC-MALS instead of AUC to characterize the oligomerization state of this construct. A comparison of the elution profiles of TBL1(31-179) with TBL1(1-124) and TBL1(1-170) is presented in a chromatogram in Figure 3.3. The analysis of the MALS data showed that TBL1(31-179) had molecular mass of 17.4 kDa (error of 0.6%), which correspond well to the monomer mass of 17.6 kDa (Fig. 3.3). Together, these data clearly show that the predicted coiled coil region does not induce oligomerization on its own. We conclude that the LisH domain drives the formation of a stable TBL1 tetramer.

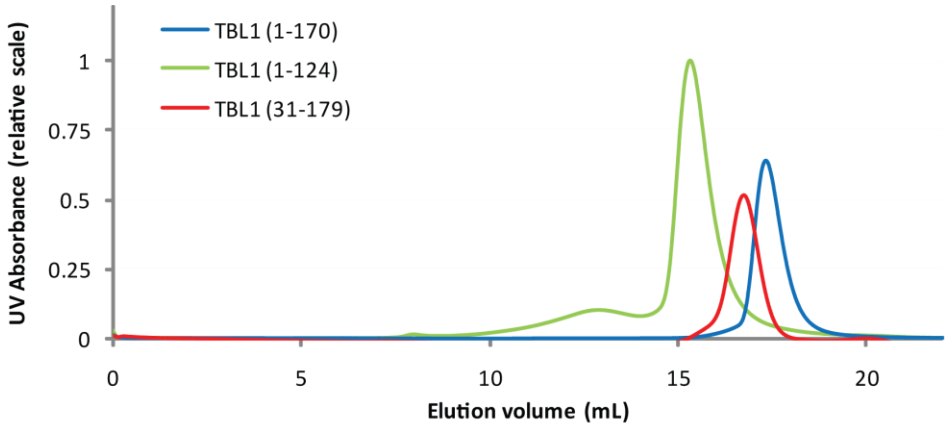


Figure 3.3. SEC-MALS analysis of TBL1 truncation constructs. Overlay of chromatograms for TBL1(1-124), (31-179) and (1-170). Purified proteins were resolved on a Superose 6 column, by injecting 50 µl of TBL1(1-170) and (31-179) at 5 mg/ml and TBL1(1-124) at 6 mg/ml. The molar mass was calculated to 54.1 kDa (0.2% error) for TBL1(1-124), 17.4 kDa (0.6% error) for TBL1(31-179) and 73.6 kDa (0.1% error) for TBL1(1-170).

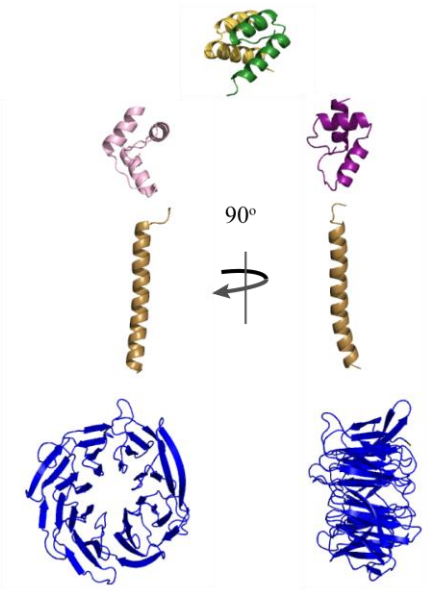
Modeling of TBL1 domains

LisH: The LisH (Lis1 homologous) is a known oligomerization domain found in more than 100 eukaryotic proteins with a variety of functions [112]. A model of the TBL1 LisH domain (LisH, 2-36) was generated by homology modeling using the program Modeller [118]. The 1.75 Å resolution x-ray structure of the Lis1-LisH dimer with sequence identity of 22 % was used as the template for the model of the TBL1-LisH

A.

Domain	aa	Modeling	Program	Template	% Identity	% Similarity
LisH	34	Homology	MODELLER	Lis1: 1uuj	22	42
FBox	40	*de novo*	ROSETTA	-	-	-
FBox*	40	Homology	MODELLER	1fqv, 2ovp, 1p22, 1nex, 2e31	2 - 12	17 - 30
CC	30	Homology	MODELLER	anti-freeze: 1wfb	32	27
WD40	352	Homology	MODELLER	Sif2p: 1r5m	25	37

B.



C.

```
                                        10        20        30
TBL1X    -      M. musculus     -SITSDEVNFLVYRYLQESGFSHSAFTFGIESHISQ
Lis1     1uuj   H. sapiens      SQRQRDELNRAIADYLRSNGYEEAYSVFKKEAELDM
```

Figure 3.4. Homology modeling of TBL1 domains. (A) Summary table of the details in generating the models of TBL1 domains. (B) Structures of the models arranged according to the table in A. Top: Homology model of the LisH dimer calculated by Modeller. The two protomers are colored in light and dark green; (next down) de novo fold of the putative FBox domain generated by Rosetta in purple; (next down) homology model of the α-helix from the putative coiled coil in brown; (bottom) homology model of WD40 repeat domain in blue. C) Structure based sequence alignment of TBL1-LisH to Lis1-LisH template generated by Modeller. Conserved residues within the dimer interface are marked in green.

domain (Fig. 3.4). Structure-based sequence alignment was generated and used as a basis to build the model (Fig. 3.4B, C). The LisH subunit formed a helical hairpin, which assembles into an anti-parallel LisH homodimer. Despite the limited overall sequence identity between TBL1-LisH and the Lis1-LisH template, the key anchor residues Val12, Tyr15, Leu16, Phe21 and Phe28 of the mostly hydrophobic interface between the subunits are well conserved (Fig. 3.4C). Furthermore, Glu31, which is crucial for the stabilization of the helices across the dimer interface, is highly conserved within all LisH domains including the TBL1-LisH domain [117].

FBox: Immediately C-terminal to the LisH domain is the predicted FBox-like motif. FBoxes are domains of approximately 40 residues that bind Skp1, the adaptor protein in SCF E3 ubiquitin ligases [21, 103]. However, biochemical and functional data in our laboratory indicates that the FBox of TBL1 has unusually low (mM range) affinity for Skp1 and functional data from our laboratory suggests that TBL1 does not function as a classical FBox-substrate recruiting protein (chapter II) [119]. Despite these observations, we initially proceeded to generate a homology model of the TBL1 FBox-like domain with Modeller. A template with multiple structure alignment of FBox domains from Skp2, Fbs1, Cdc4, β-TrCP, Fbw7 was created and used for the structure-based sequence alignment of TBL1-FBox (Fig. A3.2A). The sequence identity between TBL1 FBox and the other FBox domains is extremely low, between 2-12%, which suggests care in interpreting the results of a homology model.

Given the uncertainty of the FBox homology model, we utilized an orthogonal approach to generate additional structural models, i.e. *de novo* fold generated from sequence using Rosetta [120]. Based on the size (only 40 residues) and predicted α-

helical nature, the FBox is well within the range for reliable structure prediction by Rosetta. The Rosetta calculation generated a three helix bundle that highly resembles the fold of the FBox domain (Fig. 3.4A, B, Fig. A3.2).

Coiled coil: Between the LisH-FBox region (residues 2-86), residues 87-169 form a bridge/linker to the WD40 repeat domain (170-527). According to secondary structure prediction algorithms residues 90-121 will occupy α-helical conformation and based on COILS prediction algorithms, it should form a coiled coil [115]. Notably, 14 out of the 34 amino acids in that region are alanines, 10 of which are consecutive in the sequence (Ala110 - Ala119), which strongly predicts to helical secondary structure. To build a model of this region, we selected a template, using the BioInfoBank MetaServer that has poly-Ala content and forms a coiled coil [121]. The 1.5 Å resolution x-ray structure of the antifreeze protein from winter flounder was used as a template with sequence identity of 35%, for the model of TBL1(90-120) (Fig. 3.4). The homology model of TBL1(90-120) generated in this way, is an extended α-helix (Fig. 3.4B).

The results from the oligomerization studies and the model of the TBL1 (90-120) raise the question whether residues 90-120 form a coiled coil or is simply a region with an α-helical character. The propensity of isolated TBL1(90-120) to fold as a coiled coil was tested by circular dichroism (CD). The ratio of molar ellipticity at 222 nm and 208 nm ($[\theta_{222}/\theta_{208}]$) is used as a relative measure for the coiled coil nature of a protein, where a ratio ≥ 1 is indicative of coiled coil and a ratio of ≤ 1, of isolated helices. The predicted CC(90-120) was not produced alone, since the constructs that lack the LisH domain TBL1(31-179) and TBL1(31-103) are very unstable and prone to aggregation. Consequently, the molar ellipticity for that region was calculated by subtracting the CD

spectrum of TBL1(1-90) from that of TBL1(1-124) (Fig. A3.3). The helical content calculated for the two constructs using K2D2 is >80 % and the ratio of [$\theta_{222}/\theta_{208}$] is 0.68. This observation suggests TBL1(90-120) does not form a coiled coil, but exists as an isolated α-helix. We note that based on the biochemical data presented above, the predicted CC region of TBL1 is not sufficient to induce protein oligomerization, at least in the absence of the LisH domain (Fig. 3.3).

WD40: TBL1(179-527) is predicted to fold into commonly known interaction domain, a WD40 repeat domain. More than 20 WD40 repeat domain structures have already been determined, including the WD40 repeat domain of Sif2p, the homologue of TBL1 from *S. cerevisiae* [114]. The homology model of TBL1 WD40 was calculated using the structure based sequence alignment of Sif2p-WD40 template, which has 25% sequence identity with TBL1 (Fig. 3.4).

WD40 repeats have a β-propeller fold, generally with 6 to 7 blades, with a few exceptions of proteins such as Sif2p and Cdc4, which have 8 blades [114, 122]. Each of the blades is a structural motif of four anti-parallel beta-strands formed from ~40 amino acids that end in a Trp-Asp sequence. Based on the well conserved sequence similarity and length between TBL1 and Sif2p WD40 domains, it is highly likely that TBL1 also has an unusual 8-bladed β-propeller fold. There are four loop regions in the Sif2p WD40 structure: 191-196, 213-219, 303-317 and 477-490 [114]. In order to obtain the complete model of the TBL1-WD40 domain, the loops were built separately using Modeller. The final WD40 repeat domain model is shown in Figure 3.4B.

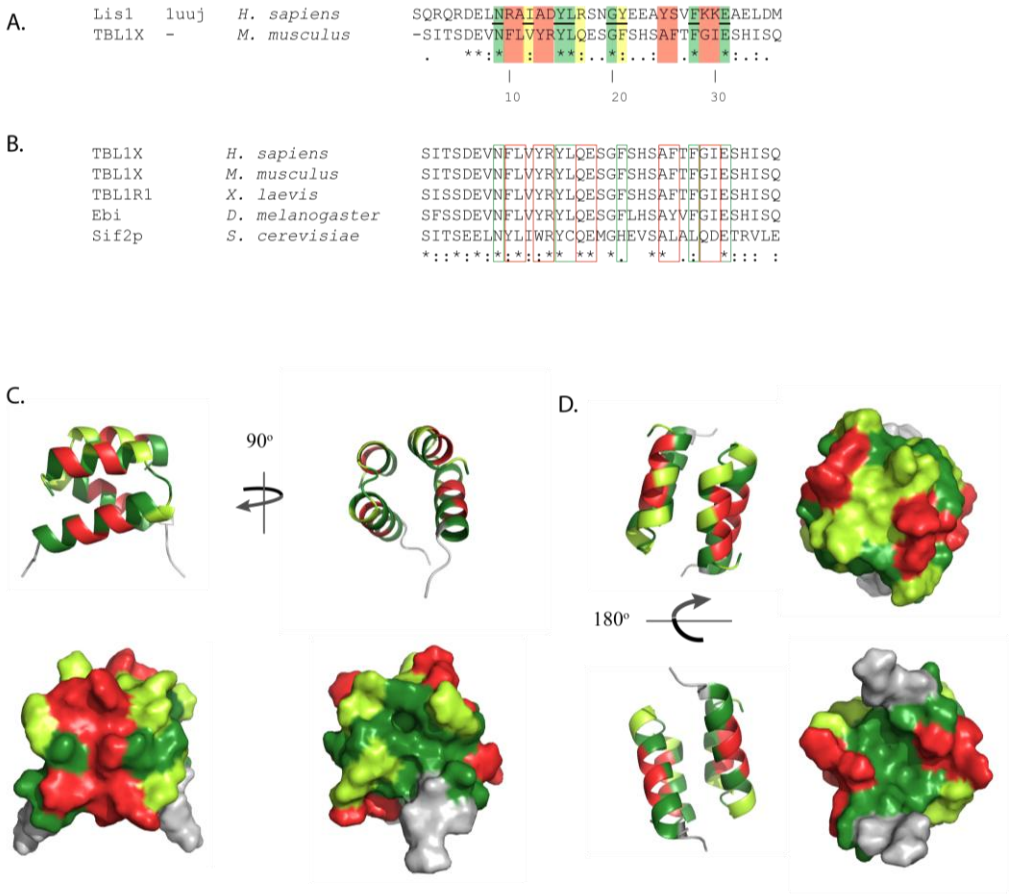


Figure 3.5. Sequence analysis of TBL1-LisH domain. A) Sequence comparison of TBL1-LisH and Lis1-LisH domains. Conserved residues within the dimer interface of the two LisH sequences are underlined. Residues highlighted in green and yellow are identical or similar, non-conserved residues are in red. B) Sequence alignment of TBL1-LisH orthologues. Resides in green are at the conserved dimer interface and in red non-conserved residues with the Lis1-LisH domain. C) The structure of the TBL1-LisH model and the surface representation below. The color scheme is consistent with the one in (A).

Design of LisH mutants that perturb TBL1 tetramerization

The ability of TBL1-LisH and its yeast homologue, Sif2p, to form stable tetramers is a novel feature compared to the well characterized LisH-containing proteins Lis1 and FOP. This raises interesting questions as to the functional basis of tetramerization and which residues within this highly conserved LisH domain mediate formation of the tetramer. Answers to these questions are highly relevant to generating the structural model of TBL1 as a tetramer. As a result, we decided to map the interface between the LisH dimers of TBL1 by using site-directed mutagenesis and use this information to refine the structural model of the tetramer.

The design of a mutant that can perturb the tetramerization of TBL1 was directed at the interface between the common LisH domain dimers. Our efforts were guided by the structural information of Lis1-LisH dimer and through sequence analysis. The conserved residues between these two LisH domains lie within the interface of the dimer, as described in detail above (Fig 3.4C). The sequence variability between the exposed residues of the LisH dimer from Lis1 and TBL1 is shown in a surface representation in Figure 3.5, with highly conserved residues in green and yellow, and non-conserved residues at the surface in red. As expected, the area with the highest conservation is at the dimer subunit interface and at the cavity formed on the bottom of the dimer (Fig. 3.5B and D). This generally hydrophobic cavity serves as the packing area of the domain positioned C-terminal to the LisH as illustrated in the Lis1 and FOP structures [116-117]. The non-conserved solvent-exposed residues of TBL1-LisH dimer include Phe10, Leu11, Tyr13, Arg14, Ser22, Phe26, Ile30 and His33, which represents a relatively hydrophobic platform and a possible oligomerization site (Fig. 3.5A). Sequence analysis of the LisH

motif of TBL1 homologues shows that Leu11, Tyr13, Arg14, Gln17, Glu18 are identical from human to yeast and Phe10 and Phe26 are also highly conserved as large hydrophobic residues (Fig. 3.5B). His33 is highly homologous throughout all species, whereas Ser22 and Ile30 are not conserved in yeast (Fig.3.5B). Based on the sequence conservation within the TBL1-LisH homologues and the difference of these residues compared to the LisH of Lis1 and FOP dimers, we hypothesize that Phe10, Leu11, Tyr13, Arg14, Ser22, Phe26, Ile30 and His33 are at the tetramer interface of the TBL1-LisH dimers.

Design of TBL1 dimer mutant

In order to test the functional role of TBL1 tetramerization, we set out to produce a TBL1 mutant that only forms dimers based on the hypothesis generated from sequence analysis. We began with the full seven point mutant of TBL1(1-124), where each residue was swapped to the corresponding amino acid from the Lis1-LisH domain: Phe10Arg, Arg14Asp, Gln17Arg, Glu18Ser, Ser22Glu, Phe26Ser and His33Glu (Fig. 3.6A). The TBL1(1-124) construct with these 7 mutations (TBL1(1-124)-7M) was expressed and purified using the protocol developed for the wild type (wt) construct. The oligomerization state of this protein was initially determined by size exclusion chromatography. Comparison of the elution profiles from a Superose 6 column of the wt TBL1 with the mutant shows a peak shift of 1.1 ml (Fig. 3.6B). This result is consistent with a change in molecular mass from approximately 60 kDa tetramer to a 30 kDa dimer.

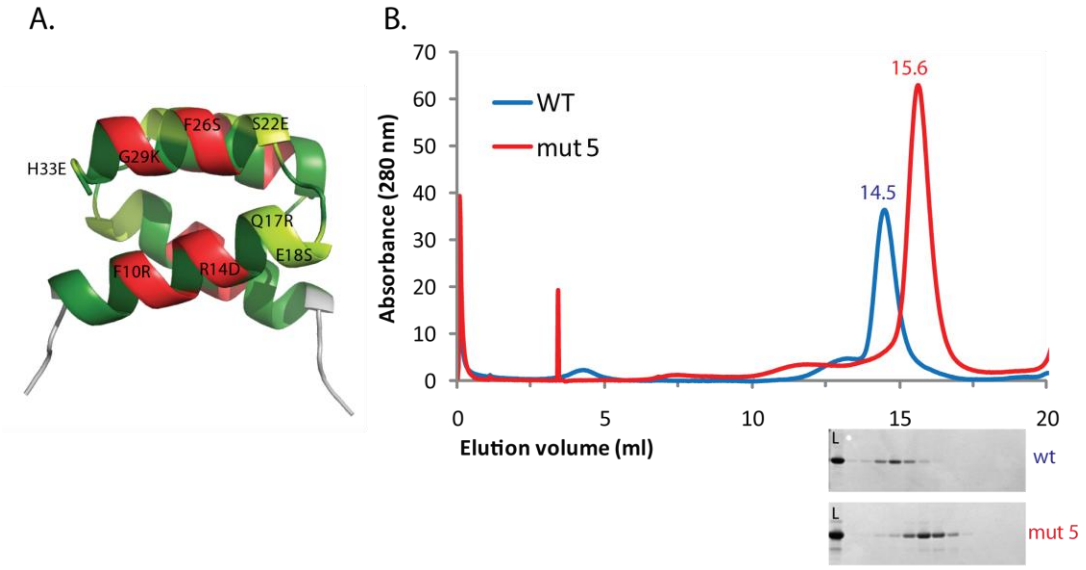


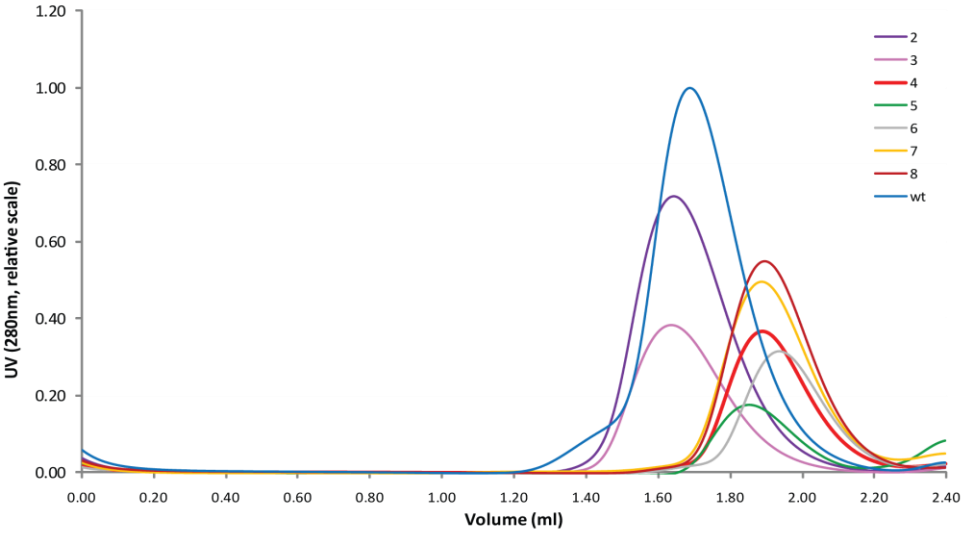
Figure 3.6. Design of TBL1-LisH mutants. A) Homology model of TBL1-LisH dimer. The marked residues are not conserved with the Lis1-LisH domain and used for designing TBL1-LisH mutants. B) Comparison of elution profiles for TBL1(1-124) wt at 80 µM and TBL1(1-124) mutant 5: Phe10Arg, Arg14Asp, Gln17Arg, Glu18Ser, Ser22Glu, Phe26Ser, His33Glu at 120 µM. The proteins were run on a Superose 6 column in buffer containing 25 mM Tris-Cl, pH 7.5, 200 mM NaCl, 5 mM βME. Coomassie stained SDS-PAGE of the eluted proteins.

Having been successful with the design of TBL1(1-124)-7M, we next sought to determine the minimum number of residues required for tetramerization. Seven constructs with different numbers and combinations of point mutations were designed, expressed and purified then analyzed by CD in order to confirm that α-helical secondary structure is retained (data not shown). The oligomeric state of these mutants was investigated by SEC-MALS (Fig. 3.7A, B). Mutants 4, 5, 6, 7 and 8 have a significant peak shift to a molecular mass of a dimer, compared to the wt protein and mutants 2 and 3, which form tetramers (Fig. 3.7A and B). Sequence comparison of the point mutants indicate that the change of Phe26 to Ser is present in all constructs that form a dimer

A.

mutant #	muations							oligomerization	stability
1	F10R	R14D						tetramer	good
2	F10R	R14D	Q17R	E18S				tetramer	good
3	F10R	R14D	Q17R	E18S	S22E			tetramer	good
4	F10R	R14D	Q17R	E18S	S22E	F26S		dimer	some aggregation
5	F10R	R14D	Q17R	E18S	S22E	F26S	H33E	dimer	good
6		R14D	Q17R	E18S		F26S		dimer	aggregation
7		R14D	Q17R	E18S		F26S	H33E	dimer	aggregation
8						F26S	H33E	dimer	good

B.



C.

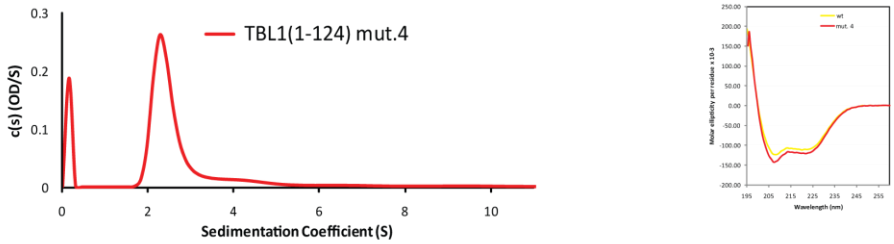


Figure 3.7. Biochemical analysis of TBL1-LisH(1-124) mutants. A) Summary table of mutants. B) SEC-MALS elution profile of mutants. The proteins were run on a Superdex 200 analytical column in a buffer containing 25 mM Tris-Cl, pH 7.5, 200 mM NaCl, 5 mM βME. B) Analysis of TBL1(1-124) mutant 4. Left: Sedimentation coefficient distribution, the calculated sedimentation value [c(s)] is plotted as a function of the sedimentation coefficient (s). Right: CD spectra of TBL1(1-124) wt and mutant 4.

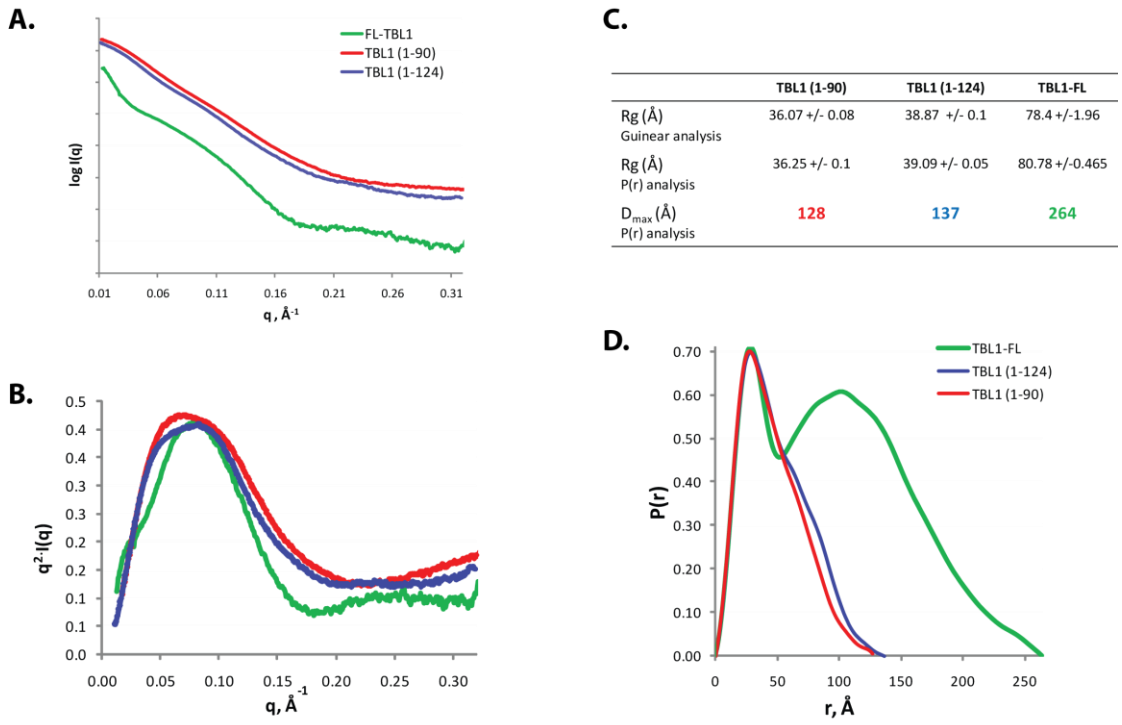
(mutants 4-8), which suggests that Phe26 alone is the critical hydrophobic residue for the formation of a stable tetramer. However, even though the elution profile of mutant 8 (Phe26Ser, His33Glu) is indicative of a dimer it has a polydispersity greater than 50%, which suggests that mutation of Phe26 is not sufficient to shift the equilibrium from a tetramer to a stable dimer. The construct also containing: Arg14Asp, Gln17Arg and Glu18Ser produces a dimer with much lower polydispersity. However, this mutant was very prone to aggregation during concentration of the protein. Addition of mutations Phe10Arg and Ser22Glu proved sufficient to prevent aggregation. Thus, the mutant containing Phe10Arg, Arg14Asp, Gln17Arg, Glu18Ser, Ser22Glu and Phe26Ser forms a stable dimer with molar mass of 28 kDa and polydispersity of less than 10%. The CD spectrum of this mutant shows the expected α-helical secondary structure (Fig. 3.7C). Furthermore, AUC sedimentation velocity experiments with this mutant produced a sedimentation value (*s*) of 2.6, which corresponds to a molar mass of 35 kDa, confirming the formation of a dimer.

Structural analysis of TBL1 by SAXS

Since we have directly established that TBL1 is a tetramer, the next step was to build a tetramer model. There are numerous potential tetramer architectures, even if only the LisH domain is considered. Thus, in order to directly characterize the molecular architecture of TBL1 tetramer in solution, we used SAXS. SAXS is a powerful low resolution technique that can provide information about the size, shape and conformational dynamics of a molecule or molecular assemblies. Our approach in studying TBL1 by SAXS was to collect data on a series of constructs including: TBL1

(LisH-FBox, 1-90), TBL1 (LisH-FBox-CC, 1-124)) and the full length protein, FL-TBL1(1-527) (Fig. 3.1C).

SAXS data was first acquired for TBL1(1-90) and TBL1(1-124) (Fig.3.8A). Since the scattering data represents the electron pair distribution of all molecules in solution, the quality of the sample is absolutely critical for proper data analysis. The monodispersity of TBL1(1-90) and TBL1(1-124) was confirmed by SEC-MALS and only the samples with correct molecular weight, ~50 µl from the top of the elution peak, were used for SAXS data acquisition (Fig. A3.4). SEC-MALS and Guinier analyses were used to verify that there was no sample aggregation (Fig. A3.4). Analysis of the Kratky plot (Fig. 3.8B) for the two constructs indicates that the proteins are well folded



	TBL1 (1-90)	TBL1 (1-124)	TBL1-FL
Rg (Å) Guinear analysis	36.07 +/- 0.08	38.87 +/- 0.1	78.4 +/-1.96
Rg (Å) P(r) analysis	36.25 +/- 0.1	39.09 +/- 0.05	80.78 +/-0.465
D_{max} (Å) P(r) analysis	128	137	264

Figure 3.8. SAXS analysis of TBL1(1-90), (1-124) and FL proteins. A) Scattring curves, B) Kratky plot, C) Table with SAXS data statistics and D) P(r) functions.

with small degree of flexibility. Based on the symmetry of the curves, there appear to be no major changes in the LisH-FBox construct upon the addition of the putative CC region (Fig. 3.8B). The radius of gyration for the two TBL1 constructs, based directly on the Guinier analysis, differ by only 2.8 Å: R_g = 36.1 +/- 0.08 for (1-90); R_g = 38.9 +/- 0.1 for (1-124) (Fig. 3.8C). This suggests that the entire TBL1(1-124) construct forms an integrated structural unit.

Distance distribution functions (P(r)) generated by GNOM for TBL1(1-90) and TBL1(1-124) have a characteristic bell shape, which is indicative of the presence of a near spherical globular domain (Fig. 3.8D). The presence of a slightly elongated shoulder at longer distances suggests an overall extended conformation of the protein. The maximum dimensions (Dmax) of TBL1(1-90) and TBL1(1-124) in the P(r) functions are 128 Å and 137 Å, respectively (Fig. 3.8C). In comparing to proteins with similar molecular weight, the Dmax value of ~130 Å suggested an overall extended conformation of the protein [123]. These results were further supported by velocity AUC experiments, which provided frictional ratios for TBL1(1-90) of 1.5 and TBL1(1-124) of 1.8 (Fig. 3.2C). Together, the SAXS and AUC results imply that the LisH-FBox(-CC) region of TBL1 has an elongated structural organization rather than a tightly packed spherical globular architecture.

SAXS data were collected for FL-TBL1 (Fig. 3.8A). Analysis by SEC-MALS showed the sample was a tetramer with a monodispersity of less than 2% (Fig. A3.4). The radius of gyration calculated by Guinier analysis was 78.4 +/- 1.96, almost two times larger than that of the N-terminal domain constructs. This was consistent with an extended conformation of the protein (Fig. 3.8). Kratky analysis of FL-TBL1 suggested a

relatively rigid overall organization compared to the two truncated constructs, evident from the parabolic shape of the curve and closer return to baseline at higher scattering angle (Fig. 3.8B). Therefore the addition of the WD40 domain appears to stabilize the overall structure.

Comparison of the pair distribution function of FL-TBL1 with TBL1(1-90) and TBL1(1-124) revealed two distinctive peaks at 25 Å and at 100 Å (Fig. 3.8D). The first peak correlates very well with the peak from the curves of TBL1(1-90) and TBL1(1-124), which reflects the scattering from the globular, nearly spherical LisH domain. The second distinctive peak in the curve of the FL protein corresponds to the large WD40 repeat domain, which contains 66 % of the total TBL1 sequence. The Rg value of 78.4 Å for the FL-TBL1 calculated from the P(r) function was consistent with the Rg value of 80.8 Å from the Guinier analysis, which supported the validity of the analysis (Fig. 3.8C). Comparison of the Dmax values for TBL1(1-90) and TBL1(1-124) of ~130 Å to the Dmax of 264 Å for the FL protein implies a spatial organization of the domains in which the WD40 is not folded onto the N-terminal region of TBL1, but extends away from the tetramerization core.

Structural model of full-length TBL1 tetramer

The biochemical and biophysical analysis presented above clearly demonstrates that TBL1 forms a stable tetramer through the LisH core with an extended organization of the domains. Formation of the LisH dimer was predicted with high confidence based on the homologous high resolution structures of Lis1-LisH and FOP-LisH, which formed stable dimers, but what remained unclear was whether the oligomeric TBL1 formed a

parallel dimer of dimers or an extended, anti-parallel dimer of dimers. The evidence from the SAXS data was an excellent tool for addressing this critical question. The estimated length based on the structural models for each of the TBL1 domains was ~25 Å for LisH, ~25 Å for FBox and 50 Å for the CC region. Therefore the two TBL1 constructs (LisH-FBox: 1-90) and (LisH-FBox-CC: 1-124) have expected Dmax of approximately 50 Å and 100 Å respectively, if the proteins form a parallel dimer of dimers. Comparison of these measurements to the experimental SAXS values of 128 Å and 137 Å for TBL1(1-90) and TBL1(1-124), respectively, strongly supported a model in which TBL1 formed an anti-parallel dimer of dimers. Furthermore, the estimated maximum dimension for a fully parallel tetramer of TBL1 would correspond to the Dmax of a single subunit, i.e. ~ 150 Å. In contrast, a tetramer formed from anti-parallel dimers would have a Dmax value substantially larger than 200 Å, consistent with the experimentally observed Dmax of 264 Å.

To obtain further insight into the structure of TBL1, *ab initio* molecular envelopes were generated for TBL1(1-90) and TBL1(1-124) using DAMMIN. The ten low resolution models obtained from the different runs were aligned, averaged and filtered based on occupancy with the program DAMAVER (Fig. 3.9A) [74]. The structural model for each of the truncation constructs has an elongated shape with a distinct dip or narrowing in the middle of the structure, which would correspond to the interface between the LisH dimer of dimers.

Ab inito shape reconstruction was then undertaken for FL-TBL1 using DAMMIN and GASBOR. The low-resolution models from each of the programs were filtered with DAMAVER [74]. Similar results were obtained whether or not information about the

symmetry of the tetramer was enforced. The FL-TBL1 model has an overall S-shape that resembles a two-arm windmill, each arm representing one TBL1 dimer. The rotational axis through the center of the molecule runs through the interface of the LisH dimers (Fig. 3.9A). The width of the base is ~170 Å, which fits the molecular envelopes of TBL1(1-90) and TBL1(1-124). Thus the LisH-FBox-CC domains are positioned at the tetramerization core of the model as expected (Fig. 3.9B). The two large globular domains at the end of each arm have a spherical shape and approximate dimensions of 50x65x70 Å, which corresponds well to the two WD40 domains per arm.

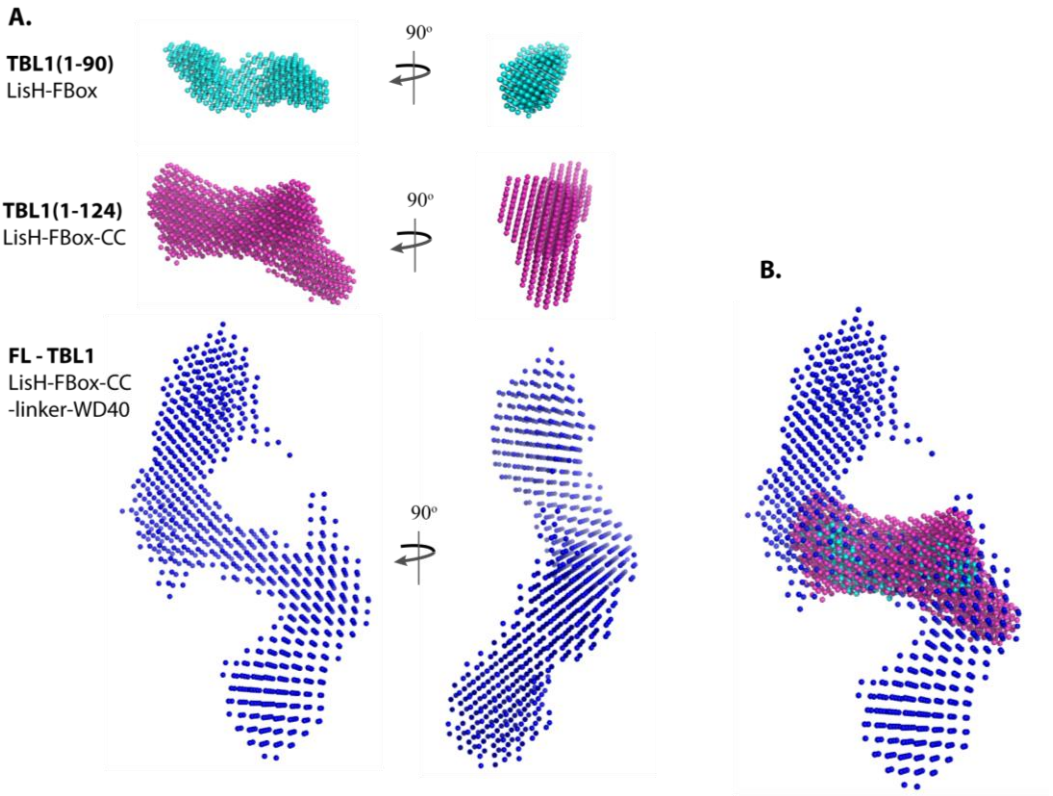


Figure 3.9. SAXS *ab initio* models. A) Averaged *ab initio* models for TBL1(1-90) in cyan, TBL1(1-124) in purple and FL-TBL1 in blue. B) Overlay of the three TBL1 models.

In order to further refine the spatial organization of each of the domains, we fitted the homology models for each domain into the molecular envelope. As noted above, the LisH domain provides the tetramerization core of the model and fits into the center of the molecular envelope. The structure of the LisH tetramer was manually fitted into the envelopes calculated for the truncated constructs and translated into the model for the FL-TBL1. Two FBox domains per dimer were also manually fitted into the low resolution envelope, although the exact orientation or packing between the FBoxes or onto the LisH domain was unclear (Fig. 3.10). The two WD40 domain structures fit well into the

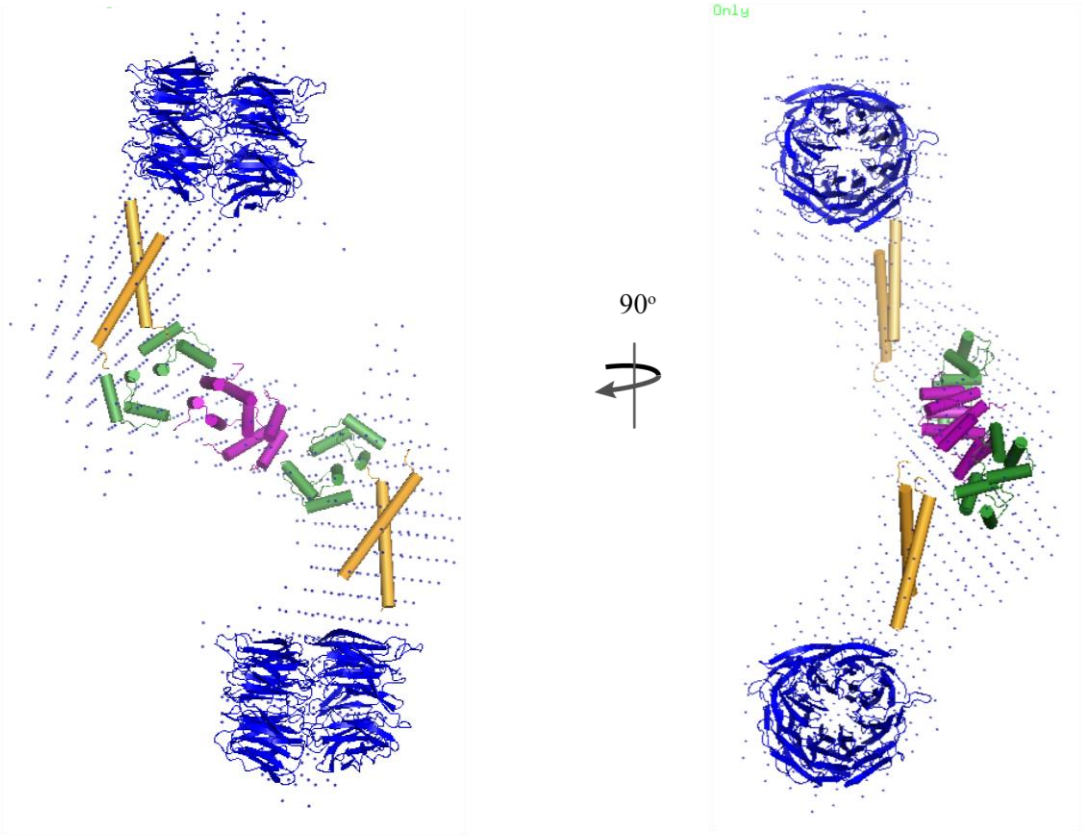


Figure 3.10. Structural model of FL-TBL1. The homology models for the TBL1 domains are manually fitted in the SAXS based *ab initio* envelope of FL-TBL1.

globular envelopes extended out from the core (Fig. 3.10). The exact placement of the putative CC (90-120) remained unclear. Comparison of the Dmax values for TBL1(1-90) and (1-124) showed a difference of only ~10 Å, which was further reflected in nearly complete overlapping of the models (Fig. 3.9). This suggested that in the absence of the WD40 domain, the predicted CC-like α-helical structure does not extend away from, but rather folds onto the LisH-FBox region. Between the CC and the WD40 repeat domains are 50 residue linkers, which according to disorder prediction algorithms are highly unstructured. However, the Kratky analysis of the FL protein implied there were no extraneous unstructured residues (Fig. 3.8B). Furthermore, limited proteolysis experiments of FL-TBL1 also implied that the tetramer is highly packed, since it did not provide flexible linkers easily accessible for digestion (Fig. A3.5). Therefore, it is likely that the CC and linker regions are structured and position the WD40 repeat domains in a specific orientation. Overall, we find highly extended structural organization of the tetramer, which provides a large surface area for the recruitment of binding partners and for the formation of multi-protein assemblies.

Discussion

TBL1 proteins are essential transcriptional regulators of the Wnt/β-catenin pathway and nuclear hormone receptors. The role of TBL1 and TBLR1 in recruitment and exchange of co-activators and co-repressors at DNA promoter sites has been demonstrated, but their mechanism of function remains unknown. To begin to address this challenge, we pursued structural and biochemical characterization of TBL1. Using sedimentation AUC and SEC-MALS analyses we demonstrated that TBL1 forms a stable

tetramer through the LisH oligomerization domain. Further insights into the structural architecture of full length TBL1 and the spatial organization of the domains were obtained by SAXS. TBL1 forms an anti-parallel dimer of dimers and an *ab initio* model, generated from SAXS data has an S-shape that resembles a two-arm windmill, each arm representing one TBL1 dimer (Fig. 3.9). A homology model of the LisH tetramerization core fits into the center of the molecular envelope and the substrate-binding WD40 repeat domains are at the end of each arm (Fig. 3.10).

Our structural model of the TBL1 tetramer provides insights into its ability to form multi-protein complexes. The extended global architecture of an anti-parallel dimer of dimers presents a large surface area that can accommodate multiple binding partners. In order to test the functional role of TBL1 tetramerization in facilitating the formation of multi-protein assemblies, we have designed and produced a TBL1 mutant that forms only a dimer, thus setting the stage for functional analysis in cell-based experiments.

Insights into the TBL1 model: comparison to LisH-containing proteins Lis1 and FOP

To gain further insights into the functional mechanism of TBL1, we compared our model to the structures of Lis1 and FOP proteins, which also contain LisH and/or WD40 repeat domain. The LisH motif was initially identified as a ubiquitous dimerization sequence, present in more than 100 eukaryotic proteins [112]. Interestingly, about 40% of these proteins are also found to contain a WD40 repeat domain, a major protein interaction unit. The Lis1 and FOP proteins provided the first insights into the structural organization of LisH-containing proteins, demonstrating that they exist as stable dimers [116-117, 124-125]. The domain composition of TBL1 is very similar to the 410-residue

Lis1 protein, with an N-terminal region containing LisH-CC domains and a C-terminal seven bladed WD40 repeat. The only identifiable region from the sequence analysis of the 399-residue FOP protein is the N-terminal LisH domain.

The crystal structure of Lis1(1-86) first revealed the LisH dimerization fold. This structure confirmed an anti-parallel four-helix bundle that extends into an asymmetric CC with two parallel helices [116]. The asymmetry of the LisH-CC structure comes from a distinct kink in one of the CC helices, which according to NMR analysis provides a high degree of flexibility allowing the exchange between open and closed conformations of the two CC helices [124]. The Lis1-CC region is not sufficient to induce protein dimerization, but contributes to the stability of the N-terminal Lis1 fragment, similar to what we have observed for the predicted CC region of TBL1. The second LisH structure is from the FOP protein [117]. It revealed the same anti-parallel dimerization fold, and a conserved hydrophobic network of residues at the interface of the dimer core, which is also conserved in TBL1.

Information about the domain organization of a full length oligomeric protein that has LisH and WD40 domains is available from the crystal structure of Lis1 in complex with platelet activating factor acetylhydrolase (PAF-AH) substrate [125]. The Lis1 homodimer binds the PAF-AH homodimer through the WD40 β-propeller domains. Even though the crystals contained full length Lis1, the N-terminal region of Lis1(1-90) was not observed in the electron density. Based on the distance between the two WD40 propellers that sandwiched the PAF-AH dimer, a large conformational change in the CC from a closed to an open state, must have occured in order to facilitate the binding of the substrate.

The *ab initio* model of TBL1 was calculated in the absence of a substrate. The overall S-shape of the tetramer indicates that the domains within a TBL1 dimer are not aligned orthogonally; instead, there is a kink in the structure of the FBox, CC or the linker region (Fig 3.9, 3.10). This structural organization is seen in the asymmetric LisH-CC fragment of Lis1 that is induced by a significant kink in one of the CC helices [116]. Thus, the domain organization of Lis1 in the absence of a substrate may resemble that of an asymmetric TBL1 dimer. The large conformational change in the Lis1 CC that accommodates the binding of a substrate may also be necessary in order for the TBL1 WD40 repeat domains to recruit a substrate. Performing SAXS experiments with TBL1 in complex with a substrate would provide a direct confirmation if there is a common substrate-mediated switch between open and closed conformations in these LisH-WD40-containing proteins.

TBL1 and TBLR1 function as scaffolding proteins

TBL1 and its close homologue, TBLR1 are transcriptional regulators essential for the proper function of nuclear hormone receptors and Tcf/Lef families of transcription factors. While TBL1 and TBLR1 have been identified as “nuclear exchange factors”, they share many common features with scaffolding proteins: (i) interactions with different proteins within a specific pathway or process, (ii) recruitment of substrates to a designated cellular location, (iii) coordination of the function of positive and negative signaling factors, and (iv) facilitation of the formation of multi-protein complexes [126-127]. TBL1 and TBLR1 tightly regulate the activation of many genes in the Wnt signaling pathway. In addition to interacting with the transcription factor TCF4,

TBL1/TBLR1 also bind and recruit co-activator β-catenin to the promoter site upon activation of Wnt signaling. In the absence of a Wnt ligand, TBL1/TBLR1 are associated with co-repressors TLE and HDAC1 at Wnt promoters and are required for the exchange of these repressors with the β-catenin activator [6]. Taken together, these observations suggest the involvement of TBL1/TBLR1 in negative and positive transcriptional control of Wnt signaling and that they function as scaffolding proteins.

In a similar manner, TBL1 and TBLR1 play an essential role as scaffolding proteins in regulating the transcriptional outcome of nuclear hormone receptors. TBL1 and TBLR1 are involved in coordinating the negative and positive signaling factors of NHRs. In the absence of a ligand, TBL1/TBLR1 recruits transcriptional co-repressors, such as CtBP or the NCoR/SMRT complex, therefore inhibiting the activation of the nuclear receptors. Upon ligand stimulation, TBL1/TBLR1 facilitates the release and proteasomal degradation of co-repressors by directly binding and recruiting the ubiquitination machinery. Further investigation of the mechanism of TBL1 and TBLR1 as scaffolding proteins will provide critical insights into the mechanism of regulating the transcriptional activity of NHRs.

Experimental methods

Subcloning

The subcloning of the expression plasmid for recombinant murine TBL1 has been previously described (chapter II and [119]). To obtain TBL1 truncation constructs (1-90), (1-124), (1-170) and (31-179), the DNA sequence was amplified from the full length cDNA of TBL1 and inserted into a pET28a vector (Novagen) with NdeI and EcoRI

restriction enzymes. The plasmid contains a thrombin-cleavable N-terminal His_6-tag. Point mutants of TBL1(1-124) in pET28a vector were generated using a Quikchange (Stratagene) mutagenesis protocol. The correct sequence of all constructs was confirmed by DNA sequencing (Vanderbilt University DNA sequencing core).

Purification of FL-TBL1

The expression and purification protocol for recombinant FL-TBL1 has been previously described (chapter II and [119]). This protocol has been further optimized to produce high purity FL-TBL1 for analytical biochemical and structural studies. 250 mL of FL-TBL1 growth from HEK 293-6E cells was resuspended in 20 mL lysis buffer: 50 mM Na_2HPO_4 at 7.5 pH, 500 mM NaCl, 10% Glycerol, 10 mM Imidazole, 5 mM β-Mercapto Ethanol (βME), 1mM PMSF, 0.1% Tween-20, 0.2 mg/ml lysozyme and one Complete Mini EDTA-Free protease inhibitor cocktail tablet (Roche). The cells were incubated on ice for 20 min. and lysed by sonication at 4 ^oC. Soluble protein was isolated by centrifugation at 21,000 rpm for 25 min. at 4 ^oC.

His tag purification (Co): His_6-tagged TBL1 was purified by gravity flow at 4 ^oC using TALON® Co^{2+}-affinity resin (Clontech). The equilibration and wash buffers included 50 mM Na_2HPO_4 at pH 7.5, 500 mM NaCl, 10% Glycerol, 5 mM βME and 10 mM or 20 mM Imidazole for the two buffers, respectively. The supernatant from the cell lysis was loaded onto the pre-equilibrated 2.5 mL Co-resin and incubated for 30 min at 4 ^oC. After 10 CV wash with the wash buffer, FL-TBL1 was eluted in 25 mL wash buffer containing 300 mM Imidazole.

Anion exchange chromatography: The purity of the sample was verified on a NuPAGE gel (Invitrogen) stained with SimplyBlue SafeStain (Invitrogen). TBL1 was purified by anion exchange chromatography using Source-Q 10/10 column (Amersham Pharmacia Biotech). The pure fractions were pooled and dialyzed against the QA buffer containing 50 mM Na_2HPO_4 at pH 7.5, 50 mM NaCl, 10% Glycerol and 5 mM βME for ~14 hrs at 4 ^oC. Protein was loaded onto the Source-Q column pre-equilibrated in a QA buffer and washed with 12 CV of QA buffer. A two step elution gradient was used to elute TBL1, 5 CV from 0 - 210 mM NaCl and 5 CV from 210 – 1000 mM NaCl.

Size exclusion chromatography: The last step of TBL1 purification and initial characterization of the oligomeric state of the protein was done by size exclusion chromatography. After equilibration of the Superose 6 10/300 GL (GE Healthcare) with 2 CV of running buffer (50 mM HEPES at pH 7.5, 200 mM NaCl, 5% Glycerol and 5 mM βME), 0.5 mL of TBL1 was loaded and run at 0.3 mL/min.

Protein Expression and purification of TBL1 constructs

TBL1 truncation constructs (1-90), (1-124), (1-170), (31-179) and TBL1(1-124) mutants were expressed in the *Escherichia Coli* Rosetta (DE3) strain. Starter cultures were grown overnight in Luria-Broth (LB) at 37 ^oC and used to inoculated 1 L LB cultures for protein expression. The cells were grown until they reached A_{600} of 0.8-1.0 and were then induced with 0.1 mM isopropyl thiogalactoside (IPTG) for 4 hrs at 37 ^oC. The cells were harvested by centrifugation at 8,000 x g for 15 min and the pellet stored at -80 ^oC.

His tag purification (Ni): The His_6-tagged proteins were affinity purified on a Ni^{2+}-NTA column (Qiagen) using 25 mM Tris-Cl at pH 7.5, 300 mM NaCl, 10 mM Imidazole, 5 mM βME as NiA buffer and 300 mM Imidazole in the NiB buffer. A 20 mL hand poured Ni^{2+}-column was equilibrated with NiA resin before loading the cell lysate. After washing with 10 CV of NiA containing 20 mM Imidazole, the protein was eluted in a 5 CV linear gradient to 100 % NiB buffer.

Anion exchange chromatography: TBL1 constructs were further purified by anion exchange chromatography as for FL-TBL1, except the QA buffer contained 30 mM Tris-Cl at pH 7.5, 50 mM NaCl and 5 mM βME. The protein was eluted in a 10 CV linear gradient of NaCl from 50 – 500 mM NaCl.

Size exclusion chromatography: The final step in the purification of the TBL1 constructs was size exclusion chromatography using Superdex 75 10/300 (Amersham Pharmacia) column. Between 0.1 – 1 mL of protein sample was loaded per run in a buffer containing 30 mM Tris-Cl at pH 7.5, 200 mM NaCl and 5 mM βME.

All of the TBL1(1-124) mutants were purified following the protocol for FL-TBL1.

Circular dichroism (CD)

TBL1 constructs were buffer exchanged into 10 mM Na_2HPO_4 at pH 7.5, 2.7 mM KCl, 1.8 mM KH_2PO_4 and 100 mM NaCl. The far-UV CD sectrum was collected using Jasco J-180 spectrophotometer (Easton, MD). The spectra were scanned form 195 – 260 nm at 22 ^oC using protein samples with final concentration between 1.5 – 10 µM. Percentage of secondary structure was estimated with the K2D2 web server [128].

Analytical ultracentrifugation (AUC)

Sedimentation velocity experiments were performed with a Beckman Optima XL-A analytical ultracentrifuge (Beckman-Coulter, Fullerton, CA) with absorbance optics. TBL1 samples were dialyzed into 25 mM Tris-Cl at pH 7.5 and 200 mM NaCl and tested at concentrations from 50 – 400 µM. Double sector cells with charcoal-filled Epon centerpiece (pathlength 1.2 cm) and quartz windows were filled with 395 µl protein sample and 400 µl of matching buffer. Using a 4-hole An60Ti rotor the samples were centrifuged at 40,000 rpm at 22 ^{o}C. The velocity experiments were analyzed with the program Sedfit (version 9.3) [64]. Approximately 300 scans collected 60 s apart were used for each analysis for a final confidence level of $p = 0.95$ and a resolution of n = 200.

Limited proteolysis

FL-TBL1 was subjected to limited proteolysis with chymotrypsin in buffer containing 25 mM Tris-Cl at pH 7.5, 200 mM NaCl, 5 mM βME and 10 mM $CaCl_2$. The digestion was carried out using 1:400 ratio of protease:protein at room temperature. Aliquots were taken at different time points and the reactions stopped with the addition of trypsin inhibitor. The samples were resolved on an SDS-PAGE and stained with SimplyBlue SafeStain (Invitrogen).

Size-exclusion chromatography - multi-angle light scattering (SEC-MALS)

SEC-MALS analyses were performed for mass determination and evaluation of sample polydispersity. MALS experiment were performed using a DAWN HELEOS 8+ light scattering detector connected to a refractive index (RI) concentration detector and

quasielastic light scattering (QELS) detector (Wyatt Technology). Superose 6 10/300 GL (GE Healthcare) and Superdex 200 (GE Healthcare) columns were used for the analysis of TBL1 constructs and full length protein, respectively. SEC running buffer for TBL1(1-90) and (1-124) contained 30 mM Tris-Cl at pH 7.5, 200 mM NaCl, 5% Glycerol and 5 mM βME and for FL-TBL1, 50 mM Na_2HPO at pH 7.5, 200 mM NaCl and 1% Glycerol. The system was equilibrated with 50 µl (Superose 6) and 25 µl (Superdex200) of BSA at a concentration of 10 mg/ml in SEC running buffer. Analysis of TBL1(1-90), (1-124) and full length protein were run on samples with concentrations ranging from 1 – 10 mg/ml in the specified SEC running buffer.

SAXS data collection and processing

SAXS data of TBL1 constructs were collected at the Advanced Light Source at the SYBILS beamline (12.3.1) at Lawrence Berkeley National Laboratory (Berkeley, CA). Scattering measurements were collected on 20 µl protein sample in serial dilution between 10 - 0.5 mg/ml and on the corresponding matching buffers using a 96-well plate (Nunc) covered with protective film. A Hamilton syringe robot was used for an automated loading of the protein samples into the sample cuvette, as previously described [75]. Sequential exposures (1, 10 and 1s) for all TBL1 samples were taken at room temperature. The first and last exposures for each sample were used to evaluate for radiation-damage. The scattering vector, q is defined as $q = 4\pi \sin(\theta/2)/\lambda$, where $\theta/2$ is the scattering angle and λ is the wavelength in Å. The data were collected over the q range of 0.007-0.31 $Å^{-1}$ with the sample-to-detector distance maintained at 1.486 m and

with a wavelength of 1.3 Å (12,000 eV). Analysis of the data was performed with the ATSAS software package [129].

The raw scattering data, I(q), were scaled and the buffer subtracted. The experimental data were first evaluated for aggregation based on the Guinier plots using the program PRIMUS. A linear correlation in the plot of log(I(q)) vs. q^2, confirms the quality of the sample. The scattering data from different concentrations were scaled and merged using the program PRIMUS [130]. The radii of gyration (R_g) was derived based by the Guinier approximation with limits of $qR_g<1.3$ for TBL1(1-90), TBL1(1-124) and $qR_g<1.5$ for FL-TBL1. The pair distribution function P(r) was calculated using the program GNOM. This program also directly extracts an estimate of the maximum dimension (Dmax) of the molecule [71].

Kratky plots were generated for each of the TBL1 constructs to evaluate the foldedness of proteins using PRIMUS. The samples of well-folded particles with no interparticle interference were selected for *ab initio* modeling. For each SAXS dataset, multiple independent DAMMIN and/or DAMMIF runs were performed, averaged and filtered with the program DAMAVER [74]. The molecular graphics were generated using PyMol (DeLano Scientific, Paulo Alto, CA).

Homology modeling

The modeling of TBL1 LisH, FBox, CC and WD40 domains were performed using the Modeller 9v8 software [118]. The templates were chosen based on the highest sequence identity of available 3D structures; the specific templates for the domains are listed in Fig. 3.4. The structure based sequence alignment was obtained using the

standard “align2d.py” script for LisH, CC and WD40 domains and the model was built using the “model- single.py” script. A template with five aligned FBox x-ray structures was created with “salign.py” (Fig. A3.2). The structure based alignment of the FBox from TBL1 with FBox 3D structures was obtained with “align2d_mult.py” and the FBox models built with “model_mult.py”.

The models were evaluated using the “evaluate_model.py” script and the final structure selected based on the molpdb score and the lowest discrete optimized protein energy (DOPE) potential.

Rosetta

De-novo structure prediction of the Fbox domain of TBL1 was performed using Rosetta software version 3.1 [131]. The 51-residue sequence [INGTLVPPAALISILQKGLQYVEAEISINEDGTVFDGRPIESLSLIDAVM] was initially folded at low resolution using Monte Carlo fragment insertion with the -abinitio and -fastrelax keywords and the residue-level centroid scoring function to generate 1000 candidate structures. The best-scoring 25% of structures were subjected to clustering with a per-cluster RMSD radius of 4Å using Rosetta's clustering function, to identify the putative most native-like fold [132]. The lowest energy member of the largest cluster was refined using Rosetta's relax mode, which performs cycles of local perturbation and gradient-based minimization to produce a high-resolution full-atom model. The final model was additionally refined with AMBER version 10 to remove any remaining clashes, using 20 cycles of steepest-descent and 480 cycles of conjugate-gradient minimization in the ff99SB force field [133-134].

CHAPTER IV

DISCUSSION AND FUTURE DIRECTIONS

Summary of this work

TBL1 in β-catenin activation of Wnt target genes

The studies presented in this work provide information about the structural architecture of TBL1 isoforms and insights into their function as transcriptional regulators. The origin of this work was an investigation of the biochemical and structural basis for phosphorylation–independent poly-ubiquitination of β-catenin by the SCF^{TBL1} complex. By establishing an *in vitro* ubiquitination system, we were able to characterize the proposed SCF^{TBL1} complex, testing separately the function of each component of the complex including Siah-1, SIP, Skp1 and TBL1 (chapter II).

We discovered that Siah-1 alone can function as the E3 ubiquitin for non-phosphorylated β-catenin and that it assembles Lys11 ubiquitin chains on the substrate. Substrates with Lys11 linkages are known targets for proteasomal degradation, which is in agreement with cell based studies performed by us and other groups demonstrating that overexpression of Siah-1 results in a decrease of cellular β-catenin levels [13, 49, 119]. Our results also lead directly to questioning the formation and physiological significance of SCF^{TBL1}. In addition, we mapped the ubiquitination sites on β-catenin to Lys666 and Lys671, which indicates that Siah-1 targets different lysine residues compared to the Wnt-regulated $SCF^{\beta\text{-}TrCP}$ complex that assembles ubiquitin chains on Lys19 and Lys49 of phosphorylated β-catenin substrate [48]. These data provided further evidence for the

existence of multiple pathways that are involved in precisely regulating the turnover and therefore the activity of such a critical transcriptional co-activator as β-catenin. Most significantly, our studies brought to light new ideas about the roles of TBL1 in regulating the activity of β-catenin. Evidence from the *in vitro* ubiquitination data suggests that TBL1 protects β-catenin from recruitment and ubiquitination by Siah-1 (chapter II). These experiments are supported by our cell based studies demonstrating that TBL1 also inhibits Siah-1 targeted poly-ubiquitination and degradation of β-catenin during Wnt signaling. This work established that both Siah-1 and TBL1 bind the armadillo repeat domain of β-catenin, suggesting a competitive mode of interaction. The ensemble of data was used to generate a model for the regulation of β-catenin by Siah-1 and TBL1 (Figure 2.7). Hence, what started out as an investigation of a complex SCF-like E3 ubiquitin ligase, turned into a study of the role of TBL1 as a scaffold for the exchange of nuclear transcription factors.

TBL1 and TBLR1 function as scaffolding proteins

We demonstrated that TBL1 forms a stable tetramer through the LisH oligomerization core (chapter III). Starting from multiple sequence alignment of LisH homologues, the binding interface between pairs of LisH dimers was mapped by site directed mutagenesis. A stable TBL1 LisH dimer mutant was produced and confirmed with biochemical and biophysical methods. Direct information about the global architecture of the TBL1 tetramer was obtained from SAXS experiments that were used to generate an *ab initio* model for the protein. We found that TBL1 has an extended S-shape conformation with two-fold symmetry through the LisH tetramer (Fig. 3.9). The

spatial organization of the TBL1 domains was deduced based on direct comparison of the *ab initio* models for TBL1(1-90), TBL1(1-124) and the full length protein. The homology models for each of the TBL1 domains were generated and fitted into the low resolution envelope, providing critical information about the placement of the N-terminal LisH tetramerization core and the C-terminal WD40 domain (Fig. 3.10). The structure of the FBox-CC-linker region has a significantly higher degree of uncertainty due to the lack of high resolution structural information or a high confidence homology model.

The model for the global architecture of TBL1 has provided a number of critical insights and set the stage for new experiments. The extended conformation of the protein indicates that the position of the WD40 domains in a TBL1 dimer is distant from the LisH core, an architecture that can facilitate the formation/stabilization of large protein complexes. Furthermore, the formation of the TBL1 dimer of dimers through the LisH core suggests a possible mechanism of action for TBL1 in which the LisH dimers serve as a hinge during the exchange of co-activator and co-repressor complexes at the promoter site of NHRs. Thus, the model has provided fundamental new insights into TBL1 function and a foundation for the design of biochemical experiments to test the role of TBL1 family members as nuclear exchange factors.

Implication of the results

The mechanism of transcriptional activation mediated by β-catenin and NHRs is of high interest because defects in their regulatory pathways are found to be a major cause for many cancers, cardiovascular diseases, and neurodegenerative disorders [37-38, 61-62]. In fact, components from these pathways have become an attractive target for

anti-cancer drugs [38, 44]. The challenge of interrupting or stimulating specific protein complexes that affect the transcriptional activities of β-catenin and NHRs without affecting other functions is being heavily pursued. TBL1 and TBLR1 have been only recently identified as transcriptional exchange factors of co-activators and co-repressors of β-catenin and NHRs, and their mechanism of function remains unknown. Determining that TBL1 forms a stable tetramer with an extended S-shape provides new insight into the basis for formation of regulatory complexes of β-catenin and NHRs. These findings may have a significant impact in advancing our understanding of the assembly/disassembly of the transcription machinery and of the mechanism of exchange between transcriptional co-activators and co-repressors.

The function of TBL1 isoforms

There are three known TBL1 homologues in humans, the X-linked TBL1 protein, chromosome 3-linked TBLR1 and Y-linked TBL1Y, in addition to a number of other isoforms identified for TBL1. X-linked TBL1 and TBLR1 have been implicated in mediating the function of the transcriptional co-activator β-catenin and of NHRs [6-7]. TBL1Y is the least studied family member and its role has not been determined yet. In general, many of the X-linked genes that are duplicated in the Y-chromosome perform the same functional role in the cell. However, despite the high sequence identity between X- and Y-chromosome linked TBL1 proteins, TBL1Y cannot compensate for the function of TBL1, which is highly ubiquitous and can bypass the X-chromosome inactivation [4].

TBL1 family members share >80 % sequence identity, but have two distinctive features that may be contributing to their functional differences and substrate specificity: unique phosphorylation sites and distinct lengths of the linker between the CC region and the WD40 repeat domain. Figure 4.1 represents a sequence alignment of TBL1 family members from human and mouse. The overall sequence identity between LisH, FBox-like, CC and WD40 domains is very high. Furthermore, sequence identity between human and mouse X-linked TBL1 orthologues is 95% and TBLR1 orthologues is 99% (Fig. 4.1).

One of the clear differences between X-linked TBL1 and TBLR1 are the unique Ser and Thr sites in the WD40 domain and the linker (Fig. 4.1). Studies using mouse TBL1 and TBLR1 demonstrated that the release of co-repressors from each of the proteins is dependent on the phosphorylation of Ser123 and Ser199 for mouse TBLR1 versus Ser175, Thr334 and Ser420 for mouse TBL1 [10]. The phosphorylation sites are also different for each of the TBL1 or TBLR1 homologues in human and mouse. Thus, phosphorylation appears to be a primary factor in determining the functional specificity of TBL1 family proteins.

The second distinctive feature within the TBL1 family members is the linker between the CC region and the WD40 repeat domain. TBL1 has a 45 amino acids long linker that is 12 residues longer than TBLR1 and notably six of the extra residues are either Ser or Thr. This suggests a link between the linker length/identity and phosphorylation. It is possible that recruitment of substrates to TBL1 proteins is dependent on the pattern of phosphorylation within the linker, which is coupled as well to the length of the linker.

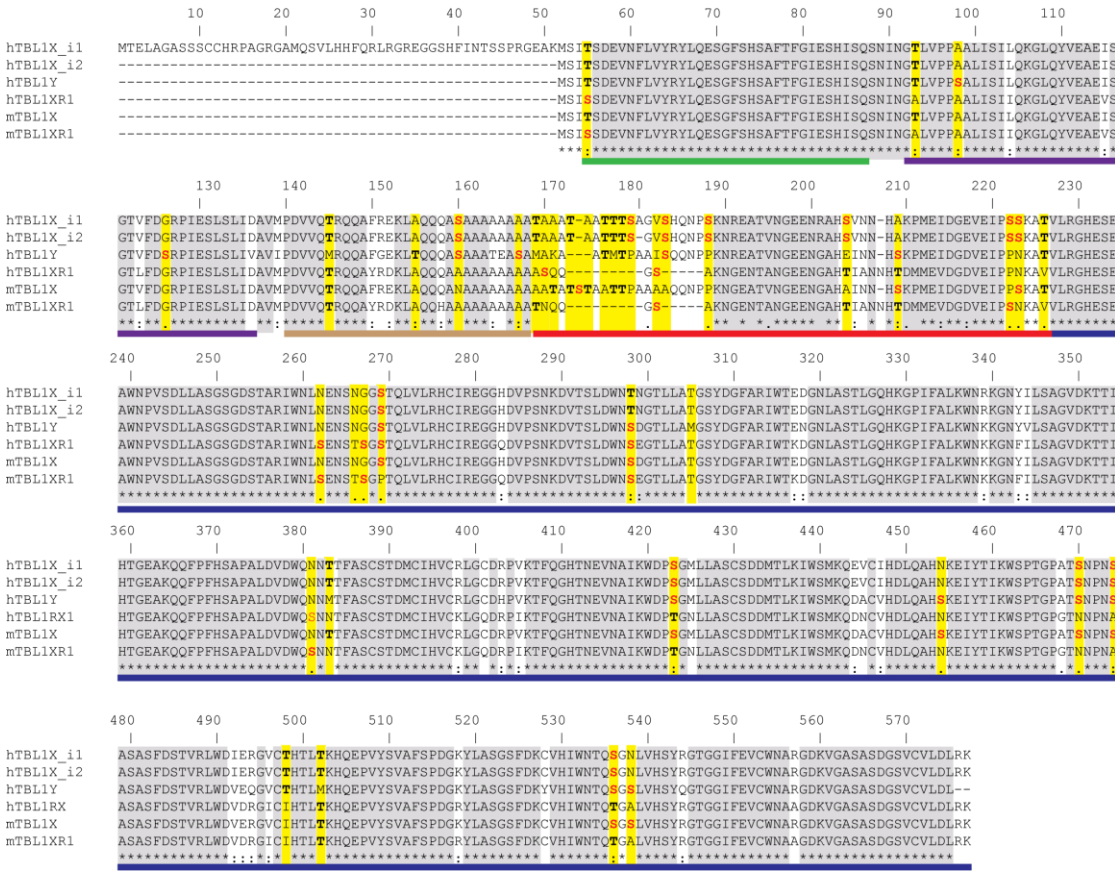


Figure 4.1. Sequence alignment of TBL1 homologues generated in CLUSTALW [2]. The proteins in the alignment are: Human TBL1, linked to chromosome X_isoform 1 (hTBL1X_i1) – ID: O60907; hTBL1X_i2 - O60907-2; hTBL1, chromosome Y (hTBL1Y) - Q9BQ87; hTBL1X related 1, chromosome 3-linked (hTBL1XR1) - Q9BZK7. Mouse TBL1 X –linked (mTBL1X) - Q9QXE7; mTBL1-related 1, chromosome 3 linked, (mTBL1XR1) - Q8BHJ5.
Residues marked with "*" are identical, ":" strongly similar, "." weakly similar.
The alignment highlighted in gray represents identical regions, in yellow - Ser and Thr residues. The color lines under the sequence represent TBL1 domains, green - LisH, purple – FBox-like, CC - beige, linker – red, WD40 – blue.

Our structural model of substrate-free TBL1 places two of the WD40 domains in relatively close spatial proximity to each other (Fig. 3.10). Therefore, in order for a substrate to bind at the surface of the WD40 repeats it is likely that the two WD40 repeat domains are separated. This may arise from lateral displacement of the domains or a conformational change in the FBox-CC-linker region. It is possible that the phosphorylation of specific residues in the loop can trigger such structural changes to increase access to the WD40 domains. Another level of substrate specificity could derive from the length, where the shorter linker of TBLR1 homologues could provide a geometric restraint on the ability to recruit large binding partners.

Interestingly, SAXS and limited proteolysis experiments suggest that mouse TBL1 (mTBL1), which has a 50 residue linker, has a fairly overall rigid architecture (chapter III, Fig. 3.8B, A3.5). This suggests that the length of the linker itself is not sufficient to induce high degrees of interdomain flexibility. Additional experiments would be necessary to determine if phosphorylation of mTBL1 would have an effect on the dynamics of the protein and its ability to recruit binding partners. In order to test this, phosphorylated TBL1 and TBLR1 proteins can be purified from *in vitro* kinase reactions and the flexibility of the proteins measured by SAXS and limited proteolysis. A similar approach can be used to test the role of the linker length by producing deletion mutants for SAXS and proteolysis experiments. Comparison between the wild type and TBL1 mutants will enable more specific statements to be made about the extent and importance of TBL1 structural dynamics.

Tight links between TBL1 and Siah-1, but not in an SCF complex

The role of TBL1 in mediating the activity of β-catenin by recruiting it to the promoter site of Wnt target genes has been established [6]. Our research also supports the role of TBL1 as a co-activator of β-catenin by protecting it from Siah-1-targeted poly-ubiquitination and proteasomal degradation during Wnt signaling (chapter II). Previous experiments also demonstrated that TBL1 forms a complex not only with the Wnt co-activators, β-catenin and TCF4, but with co-repressors such as CtBP1/2 [6, 10]. The obvious question is whether TBL1 is actively involved in exchanging β-catenin co-activators for co-repressors at the promoter site of Tcf/Lef transcription factors. The role for TBL1 family members in the exchange between co-activator/co-repressor complexes of NHRs has already been established and it is highly probable that a similar mechanism is in place for the regulation of Wnt target genes [60]. The experimental evidence from Perissi and colleagues also suggests that the exchange is facilitated by Siah-1-targeted poly-ubiquitination and degradation by the proteasome. Furthermore, the experimental evidence suggests that both Siah-1 and the 19S proteasome could be recruited to the co-activator and co-repressor complexes by TBL1 family members [10, 60].

The recently developed evidence for the correlation between Siah-1 and TBL1 provides some insights into the initial conclusions that TBL1 is involved in a Siah-1-mediated SCF^{TBL1} complex. The authors hypothesized that Siah-1 mediates the formation of an SCF complex that targets β-catenin for ubiquitination because (i) Siah-1 interacts with APC, a known binding partner of β-catenin that mediates its degradation through another SCF ligase, the $SCF^{\beta TrCP}$ complex, and (ii) transient transfection of Siah-1 leads to a drastic decrease of β-catenin levels in cells [13, 49, 135]. The role of TBL1 in this

Siah-1-mediated complex was pursued based on experimental evidence for the function of Ebi, the *Drosophila* homologue of TBL1, in recruiting Ttk88 for poly-ubiquitination by SINA (the *Drosophila* homologue of Siah-1) [12].

Matsuzawa and colleagues tested the hypothesis that Siah-1-mediates an SCF-like complex containing Siah-1, SIP, Skp1 and TBL1 that targets β-catenin for poly-ubiquitination. The evidence presented for the protein-protein interactions between E2-Siah-1, Siah-1-SIP, SIP-Skp1 and Skp1-TBL1 were based on yeast two-hybrid and three-hybrid assays further "confirmed" by transient transfections followed by co-immunoprecipitations [13]. The lack of strong experimental evidence demonstrating a direct physical interaction of the proteins in the putative SCF^{TBL1} complex and the possibility that a direct interaction between Siah-1 and TBL1 could have contributed to the misleading conclusions by Matsuzawa and colleagues. Furthermore, our biochemical and *in vitro* data demonstrating that the TBL1-FBox domain does not bind the adaptor protein, Skp1, thus precluding the formation of a stable SCF complex (chapter II), suggests that the proposed SCF^{TBL1} complex is not physiologically relevant.

So, what is the relationship between Siah-1 and TBL1? Notably, there is a growing number of binding partners that interact with both Siah-1 and TBL1 (Table 4.1). Interestingly, TBL1 is necessary for the Siah-1-mediated poly-ubiquitination of certain substrates such as CtBP1/2 and NCoR [10-11]. In addition, the degradation of these substrates depends on TBL1-mediated recruitment of the 19S proteasome to DNA promoter sites. Our experiments have shown that TBL1 does not function as an FBox-WD40-substrate recruiting receptor for SCF complexes and confirms the function of Siah-1 as a simple E3 ligase (chapter II). Thus, it is highly possible that TBL1 serves as a

chaperone in recruiting substrates to Siah-1 for poly-ubiquitination, and subsequently to the proteasome for degradation.

Table 4.1. Summary of TBL1 and Siah-1 binding partners.

TBL1 binding proteins	method	TBL1 interacting domain	References
β-catenin	IP, native gel shift assay	(LisH-Fbox-CC-linker, 1-142)	(Li, 2008)
NCoR	GST-pull down, Co-IP, Y2H	(LisH-Fbox-CC-linker, 1-142)	(Guenther, 2000; Rual, 2005; Zhang, 2002)
SMRT	GST-pull down	(LisH-Fbox-CC-linker, 1-142)	(Guenther, 2000; Zhang, 2006)
GPS2	GST-pull down	(LisH-Fbox-CC-linker, 1-142)	(Zhang, 2002)
Siah-1	Y2H, GST-Pull down	ND	(Rual, 2005)
HDAC5	Y2H, GST-Pull down	ND	(Rual, 2005)
CtBP1/2	GST-pull down	(WD40, 190-526)	(Perissi, 2008)
NCoR complex: NCoR, SMRT, HDAC3, GPS2	affinity column purification	ND	(Guenther, 2000; Zhang, 2002)
NCoR complex: NCoR, HDAC3, Emerin	affinity purification with Emirin-beads, Co-IP	ND	(Holaska, 2007)

Siah-1 ubiquitination targets	method	Siah-1 interacting domain	References
β-catenin	*in vitro* ubiquitination; cell-based assay*	HSQC experiment; (SBD, 90-282)	(Dimitrova, 2010; Liu, 2001; Matsuzawa, 2001)
NCoR	cell-based assay*	GST-pull down(SBD, 108-325)	(Zhang, 1998)
	cell-based assay*	ND	(Perissi, 2004)
	cell-based assay* (Siah-2)	ND	(Frasor, 2005)
HDAC3	cell-based assay*	ND	(Perissi, 2004)
CtBP1/2	cell-based assay*	ND	(Perissi, 2008)

*cell-based assay - The experiment consisted of transient transfection of Siah-1 followed by measuring the level of the substrate. In most cases proteasomal inhibitor MG132 is added to confirm that this is a proteasome dependent process.

A physical interaction between TBL1 and Siah-1 has been identified in a large yeast two-hybrid proteomics screen and reproduced by GST-pull down [136]. Of course these experiments need further confirmation, but a direct physical interaction between Siah-1 and TBL1 would be additional evidence for the chaperone function of TBL1 as a scaffolding protein in recruiting both Siah-1 and substrates to close proximity. Once a substrate is poly-ubiquitinated, the chaperone-like function of TBL1 could be extended to binding and recruiting the proteasome for efficient substrate degradation.

If TBL1 proteins function as chaperones of Wnt regulators, could TBL1 be also involved in the Siah-1-mediated poly-ubiquitination of β-catenin? Despite our finding that TBL1 inhibits Siah-1 targeted degradation of β-catenin during Wnt signaling, our

data does not provide evidence if additional signaling events favor the release of β-catenin for poly-ubiquitination. Since the phosphorylation of TBL1 leads to the release and degradation of CtBP, which is a co-repressor for NHRs and β-catenin, post-translational modifications of TBL1 could be a direct signal for stimulating the chaperone role of TBL1 proteins. In order to test this, phosphorylated TBL1 can be produced in an *in vitro* kinase assay and the effect of the post-translational modification investigated by performing an *in vitro* ubiquitination reaction of β-catenin with Siah-1 as the E3 ligase. If phosphorylated TBL1 releases β-catenin, and serves as a chaperone for the Siah-1 E3 ligase, we would expect to observe an efficient *in vitro* poly-ubiquitination of β-catenin.

The role of TBL1 oligomerization

One of the critical features characterized in our biochemical and structural analyses is the ability of TBL1 proteins to form a stable dimer of dimers through the N-terminal LisH domain. Even though we have been able to map the residues at the binding interface of the dimers, determining the exact binding affinity has been challenging. The methods that were used to determine the oligomerization state of TBL1 include SEC-MALS, interference AUC and SAXS, with the lowest detectable concentrations of TBL1 estimated to be at least 2 μM. The data from these experiments have captured TBL1 only at a tetrameric state, which is a strong indication that the affinity of the two dimers is *at least* in the nM range (chapter III).

However, this data does not provide direct information about the oligomeric state of TBL1 in cells. TBL1 is a predominantly nuclear protein and its concentration in the cytoplasm is very low, which could lead to dissociation into a TBL1 dimer. The high

abundance of TBL1 and TBLR1 in the nucleus and especially at DNA promoter sites would increase the local concentration of the proteins, leading to the formation of homo- and hetero-tetramers. Even more interesting is the effect of binding partners on the oligomerization state of TBL1 family members. One possibility is that TBL1 binds substrates as a dimer in the cytoplasm and after nuclear translocation it associates into homo- and hetero-tetramers to stabilize the formation of large protein assemblies. It is also possible that TBL1 forms tetramers only at promoter sites of target genes of Wnt or NHRs. However, if the affinity of the TBL1 dimer of dimers is very high, it is also likely that TBL1 and TBLR1 recruit specific substrates to the promoter sites as homo-tetramers. Then, the assembly of the multi-protein co-activator or co-repressor complexes may require the formation of a TBL1/TBLR1 hetero-tetramer through subunit exchange, with each of the homo-dimers bringing a specific substrate to the complex.

The formation of tetramers could be an essential function for the scaffolding mechanism of TBL1 proteins. For instance, β-catenin binds both TBL1 and TBLR1, whereas transcription factor TCF4, which directly interacts with β-catenin to induce transcriptional activation of Wnt target genes interacts with TBL1, but not TBLR1. However, it has not been demonstrated if the formation of a TBL1 hetero-tetramer is required for the simultaneous recruitment of TCF4 and β-catenin at the Wnt promoters. The mutants that we designed to perturb TBL1 tetramerization and form only a stable dimer would be an excellent tool for testing the effect of TBL1 oligomerization in cells. Considering the complexity of the system and the fact that cell-based experiments do not provide a reliable answer for the direct interaction of protein complexes, an *in vitro* reconstitution of the system would be necessary. Determining the stoichiometry and the

structural organization of Wnt regulatory complexes such as co-activators TBL1-β-catenin-TCF4 and co-repressors TBL1-CtBP, TLE-HDACi would provide important new insights into the mechanism of TBL1 as an essential transcriptional regulator of Wnt signaling.

Functional implications of TBL1 in regulating NHRs

It has been well established that transcriptional regulation of NHRs at a target promoter site involves a tightly controlled switch between the recruitment of co-activators (Co-A) and co-repressors (Co-R) (Fig. 1.7) [10, 60]. Using cell-based experiments, a number of groups have demonstrated the critical scaffolding role of TBL1 and TBLR1 in mediating the direct exchange of these regulatory complexes from the NHRs by recruiting them for poly-ubiquitination and proteasomal degradation. Our structural and biochemical studies revealed for the first time the extended global architecture of the mTBL1 tetramer (chapter III), setting the stage for pursuing further functional and structural studies.

Proposed model for the mechanism of TBL1 proteins as nuclear exchange factors

Even though the high affinity of the LisH dimers suggests the formation of a TBL1 tetramer *in vitro* and in cells, we cannot exclude the possibility that low protein concentration and the interaction with certain binding partners could alter the oligomerization state. We propose the following model for the recruitment and exchange of substrates by TBL1 and TBLR1 (Fig. 4.2). The low concentration of TBL1 and TBLR1 in the cytoplasm allows the formation of homo-dimers that bind and recruit substrates to the nucleus or NHR promoter sites (Fig 4.2, (1) and (4)). The increased local

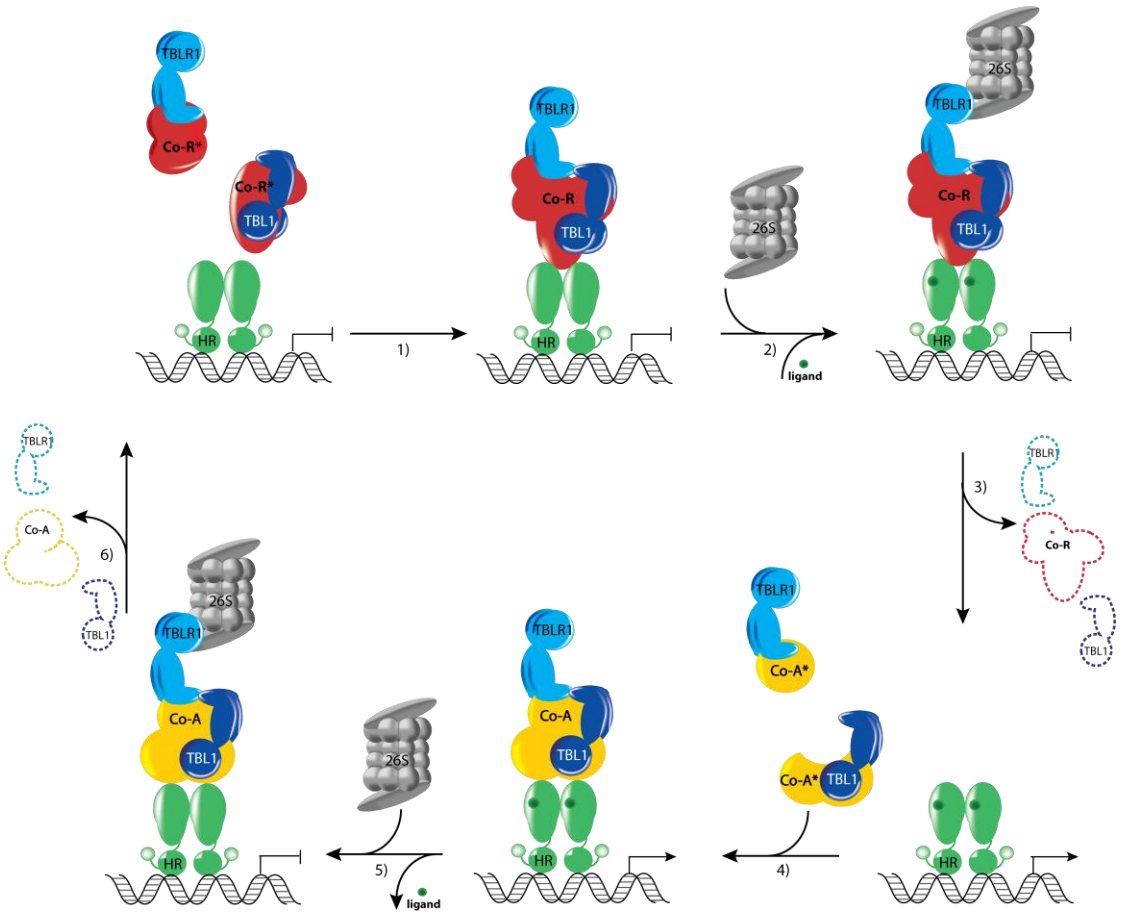


Figure 4.2. Proposed model for the mechanism of TBL1 homologues as scaffolding proteins of NHRs. 1) In the absence of a ligand, TBL1 and TBLR1 homo-dimers bind and recruit different co-repressor components (Co-R*) to the hormone receptor (HR). TBL1 proteins facilitate the assembly of the entire co-repressor (Co-R) complex at the HR promoter by forming a TBL1/TBLR1 hetero-tetramer through the N-terminal LisH domain. The Co-R complex inhibits HR-targeted gene transcription. 2) Ligand stimulation of the HRs leads to TBL1-mediated poly-ubiquitination and recruitment of the proteasome to the Co-R complex. 3) Release and degradation of the TBL1/TBLR1 Co-R complex, which leads to low transcriptional activation of HR genes. 4) Co-activator components (Co-A*) are recruited by TBL1/TBLR1 homo-dimers to ligand activated HRs. Assembly of the entire co-activator (Co-A) complex through the LisH hinge of the TBL1/TBLR1 hetero-tetramer leads to the complete activation of HR-targeted gene transcriptional. 5) Transcriptional repression of HRs is restored by TBL1-mediated poly-ubiquitination of the Co-A complex and recruitment of the proteasome. 6) The exchange of the Co-A complex for Co-R complex is allowed by the release and proteasomal degradation of the HR activator components.

concentration in the nucleus and at specific promoter sites would lead to the assembly of stable tetramers. Since the LisH is the core oligomerization domain, which is completely conserved between TBL1 family members, it is conceivable that TBL1 and TBLR1 hetero-oligomers can form (Fig. 4.2, (2) and (5)). In order to facilitate the exchange between the Co-A and Co-R complexes, TBL1 proteins have been suggested to directly recruit E3 ubiquitin ligases such as Siah-1, for poly-ubiquitination of the complex. Subsequently, TBLR1 binds the 19S proteasome and mediates the release and proteasomal degradation of the NHR bound complex (Fig. 4.2, (3) and (6)). Whether TBL1 and TBLR1 are also degraded during the exchange of the Co-A/Co-R complexes remains unclear. The two proteins are found at both the activated and repressed promoter sites, but the evidence was obtained from chromatin immunoprecipitation (ChIP) experiments and does not provide information about the dynamics of the exchange and the fluctuation of TBL1/TBLR1 levels [60].

Future directions

TBL1 and TBLR1 are essential factors for the transcriptional activation of NHRs, and absolutely necessary components in both the co-repressor and co-activator complexes [60]. Further analyses of the function of TBL1 proteins have shown that they recruit the 19S proteasome and mediate the ubiquitination and degradation of NHR and their co-factors at the promoter site [7, 60]. However, despite a wealth of functional data demonstrating the critical role that TBL1 and TBLR1 proteins play in mediating the transcription of NHRs, the mechanism that dictates the role of these scaffolding proteins in the precise exchange of the co-activator and co-repressor factors remains a critical gap

in knowledge. The research presented in my thesis work demonstrates that TBL1 forms a stable tetramer that has an extended S-shape conformation that may facilitate the recruitment of large protein assemblies. These data provided fundamental new insights into the mechanism of TBL1 and TBLR1 in mediating the activity of NHRs. Furthermore, this work can have a significant impact in the design of new experiments to precisely investigate this novel mechanism of TBL1 as an exchange factor of NHRs.

In order to test the model for the function of TBL1 (Fig. 4.2), an experimental approach going beyond characterizing single, isolated proteins is required, one which will undertake the challenge of studying the entire system both *in vitro* and in cells. Due to the number and complexity of TBL1-protein interactions, the goal of the experimental design is to establish a basic, but comprehensive *in vitro* system that will serve as a guide to test the mechanism of TBL1 proteins in cells. The discoveries will be confirmed and the mechanism tested in a cell-based luciferase reporter assay that will serve as a functional readout for the transcriptional activation of NHRs, such as the estrogen receptor (ER). The flowchart in Figure 4.3 is a representation of the proposed future goals that would provide an in-depth understanding of the mechanistic role of TBL1 proteins, and this strategy is summarized in the following sections.

I. In vitro reconstitution of the system

In order to compare the role of TBL1 family members in the exchange of co-activator and co-repressor factors from NHRs, the complexes need to be recombinantly expressed and purified (Fig. 4.3, I). These complexes include the ER co-activators: steroid receptor co-activator 1 (SRC1), p300 and CREB binding protein (CBP); and co-repressors: NCoR, SMRT, HDAC3 and GPS2 [11, 137]. Many of these factors have been

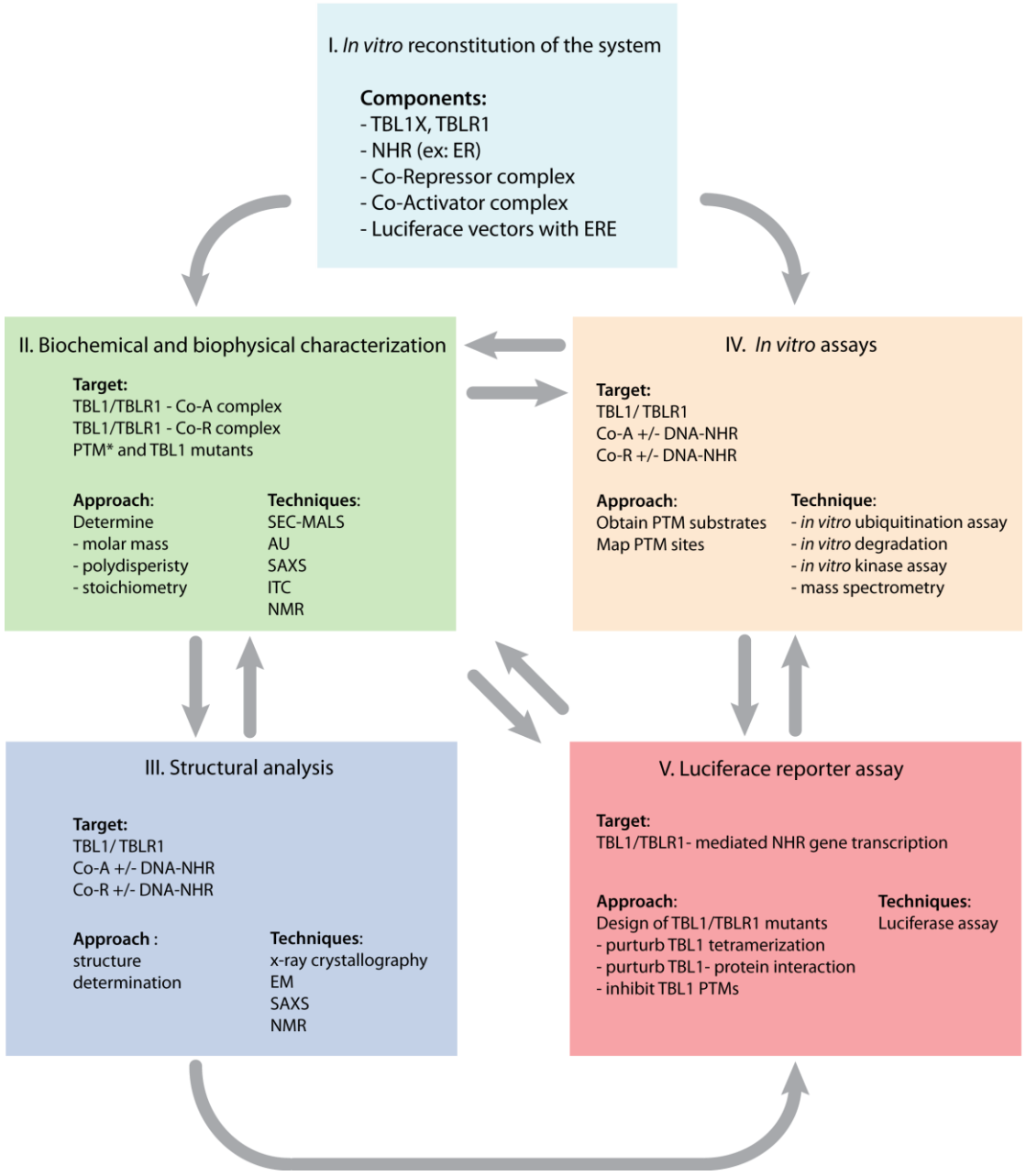


Figure 4.3. Experimental plan for studying the mechanism of TBL1-facilitated exchange between NHR co-activator and co-repressor complexes. Arrows show the information feedback of the different experimental approaches.

successfully produced before, but in order to make the functional full length proteins, different heterologous and mammalian expression systems should be pursued. Since the regulatory system of NHRs is mostly conserved in yeast and the function of Sif2p, the yeast homologue of TBL1, also resembles the one in humans, if necessary the experiments could be performed in a more basic and less complex system with the co-activator and co-repressor complexes from *S. cerevisiae*.

II. Biochemical and biophysical characterization

Our finding that TBL1 forms a stable dimer of dimers has lead to some very interesting questions regarding the formation of activator and repressor complexes and the mechanism that mediates their exchange from the NHRs. These questions include: what is the oligomerization state of TBL1 and TBLR1 alone and in the complexes with binding partners: are they in a homo- or hetero- oligomeric state? What is the stoichiometry of the various complexes and how does this relate to the oligomeric state of TBL1? And finally, as proposed in our model, are individual substrates recruited by TBL1 dimers that associate with the complex through the formation of a TBL1 tetramer?

In order to address these questions, a combination of biochemical and biophysical techniques such as AUC, SEC-MALS, SAXS will need to be used to determine the molar mass of the targeted complexes and the stoichiometry of the components (Fig. 4.3, II). NMR experiments will need to be performed to map the binding interface of specific protein interactions. ITC and AUC equilibrium experiments would be informative for measuring the affinity of dynamic protein interactions. TBL1 mutants that were designed to perturb the dimer of dimer interface (chapter III) would be very instrumental in

addressing many of the mechanistic questions and confirming the oligomerization state of TBL1 proteins in the co-activator and co-repressor complexes.

III. Structural analysis

The most powerful approach for characterizing the structural architecture of large protein assemblies is to collect complementary data with multiple techniques such as crystallography, NMR, EM and SAXS (Fig.4.3, III). X-ray crystallography is the main technique to determine the structures of large, conformationally-stable proteins and protein sub-complexes, and can provide valuable high resolution data. Considering the internal flexibility and the size of the co-activator and co-repressor complexes (>800 kDa), crystallization may not be possible and negative stain EM might provide a more general approach for elucidating the structural architecture of these macromolecules. The precise positioning of the proteins within the complex can be obtained by specifically labeling each of the components individually. Knowledge of the spacial organization of the proteins within the complex is crucial for reconstructing the entire assembly and for modeling the high resolution structures of the sub-complexes. Additional structural information about the sub-complexes in solution can be obtained by SAXS. Since the data collected by SAXS experiments is a summation of all molecules in solution, it provides low resolution structural information. However, when combined with high resolution NMR and crystallography data, it can bring critical knowledge about the shape, conformational dynamics and the spatial organization of different components in the complex.

IV. In vitro assays

The function of many regulatory factors, including TBL1 and TBLR1 is regulated by post-translational modifications (PTMs) with two of the most common ones being protein phosphorylation and ubiquitination. Furthermore, the exchange between the co-repressor and co-activator complexes has been found to depend on the TBL1-mediated poly-ubiquitination and proteasomal degradation of the components in each complex [10, 60]. In order to test the effect of these PTMs, (i) an *in vitro* kinase assay for TBL1 family members and (ii) an *in vitro* ubiquitination assay for co-repressor/co-activator components should be established (Fig. 4.3, IV).

In vitro kinase assays are well established and can be performed using GSK3 and Casein kinase 1 (CK1), which have been demonstrated to target specific phosphorylation sites of TBL1 and TBLR1 [10]. The precise PTMs will be verified by mass spectrometry analysis. Post-translationally modified substrates will be purified and the effect of these modifications on the formation and dynamics of the co-activator and co-repressor complexes can be further characterized using structural and biochemical approaches (Fig. 4.3, II and III).

Previous studies have shown that some of the TBL1-binding proteins from the co-activator and co-repressor complexes are poly-ubiquitinated by Siah-1, leading to their proteasomal degradation [60]. The *in vitro* ubiquitination assay we have established with Siah-1 (chapter II) can be used to test activity on TBL1-binding partners as well as the effect of TBL1 phosphorylation. The protocol for the *in vitro* ubiquitination assay can be optimized for different E3 ubiquitin ligases. It would also be interesting to determine the rate of poly-ubiquitination and proteasomal degradation of the entire complex by adding

purified 26S proteasome to the ubiquitination reaction. The ubiquitin linkage assembled on the substrates could be analyzed by mass spectrometry.

V. Establishing luciferase reporter assay

A luciferase assay for the gene activation of NHRs can be used to specifically test TBL1 mutants that provide insight into the proposed nuclear exchange mechanism. A luciferase reporter assay would provide the functional readout in testing the mechanistic role of TBL1 determined by the structural and biochemical analyses (Fig. 4.3, V). The luciferase assay is a well established method used to quantitate gene expression. The luciferase DNA vector can be designed to include an estrogen response element (ERE) [137]. The transcriptional activation of the ER leads to the expression of luciferase protein and a bioluminescent reaction when exposed to a luciferin substrate. The light emitted by the reaction is directly proportional to the amount of the expressed enzyme, and therefore the binding and activation of the ER.

To begin the analysis, the functional relevance of perturbing TBL1 tetramerization on the activation of NHR would need to be tested. Next, the effect of mutating specific TBL1-protein interaction sites and PTM sites should be measured with the assay, after knocking down the wt proteins and introducing the mutants. Finally, the functional data can be used to develop a better experimental design for future structural and biochemical research.

VI. Summary

The proposed future directions in this chapter are highly specialized in characterizing the mechanism of TBL1 proteins in mediating the transcriptional activity

of NHRs. The experimental approach includes reconstitution of a co-activator and co-repressor complex of NHR for an in-depth biochemical and structural characterization (Fig. 4.3). Details from the structural information will be used to formulate hypotheses that can be further examined in cells and in *in vitro* assays. The main priority in the experimental approach of investigating the mechanistic role of TBL1 is to pursue high resolution structure determination of TBL1 in a complex with co-repressors such as NCoR and SMRT. The hypotheses generated from the structural data can be tested by designing TBL1 mutants that can be used in the biochemical and functional experiments presented above. The cross-reference approach can provide critical insights into the complex system of the scaffolding TBL1 proteins.

Concluding remarks

The necessity of proper transcriptional activation by NHRs and β-catenin is demonstrated by the observation that dysregulation of these processes participate in the initiation and progression of many diseases, including cancer. Though undeniably critical, our knowledge of the role TBL1 proteins play in regulating gene transcription through NHRs and β-catenin is quite limited. Since it is an essential regulator of these transcription factors, in-depth understanding of the functional mechanism of TBL1-mediated protein assemblies is important. The extended structural architecture of tetrameric TBL1, discovered in my thesis work, can have important implications in understanding its mechanistic role as a scaffolding protein and nuclear exchange factor.

Another critical part of my thesis work is providing new understanding of the interplay between Siah-1, as an E3 ubiquitin ligase and TBL1, a mediator of its function

and a possible chaperone. The evidence presented in chapter II demonstrates that Siah-1 can directly target β-catenin for poly-ubiquitination and proteasomal degradation and previously published data demonstrates that Siah-1 also targets specific co-regulators of NHRs [138]. Notably, many of the Siah-1 targets are also binding partners of TBL1 (Table 4.1). My work provides evidence that Siah-1 and TBL1 do not work together in an SCF^{TBL1} complex, but possibly through a novel mechanism, where TBL1 functions only as a chaperone that modulates the E3 ligase activity of Siah-1.

Protein turnover is one of the major regulatory processes of protein function, therefore insights into the mechanism of Siah-1 and TBL1 in mediating the levels of β-catenin and co-regulators of NHR could have a significant impact in finding new ways to regulate their function and transcriptional activation. Considering the complexity of the system, the answers will not emerge from a single approach, whether structural or functional. Moreover, an interdisciplinary approach that utilizes the data generated from a combination of biochemical, structural, cellular and clinical research will enable the translation of our knowledge into a deeper understanding of diseases and effective therapeutic design.

APPENDIX A

Supplementary Figures

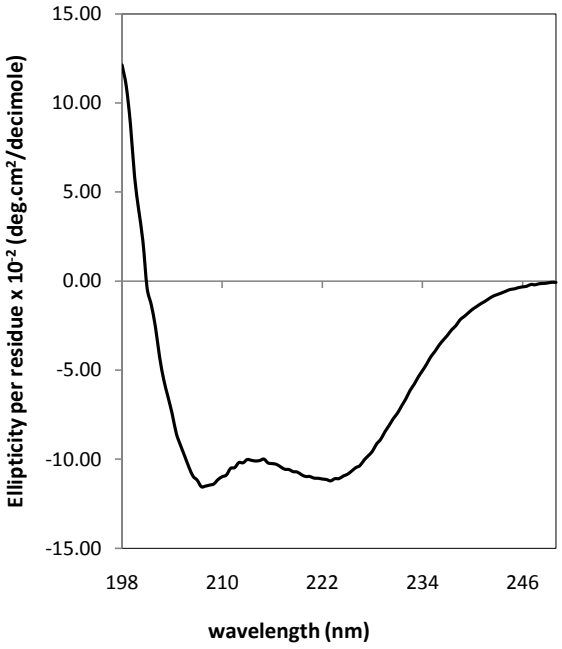


Figure A2.1. CD spectrum of β-catenin-FL.

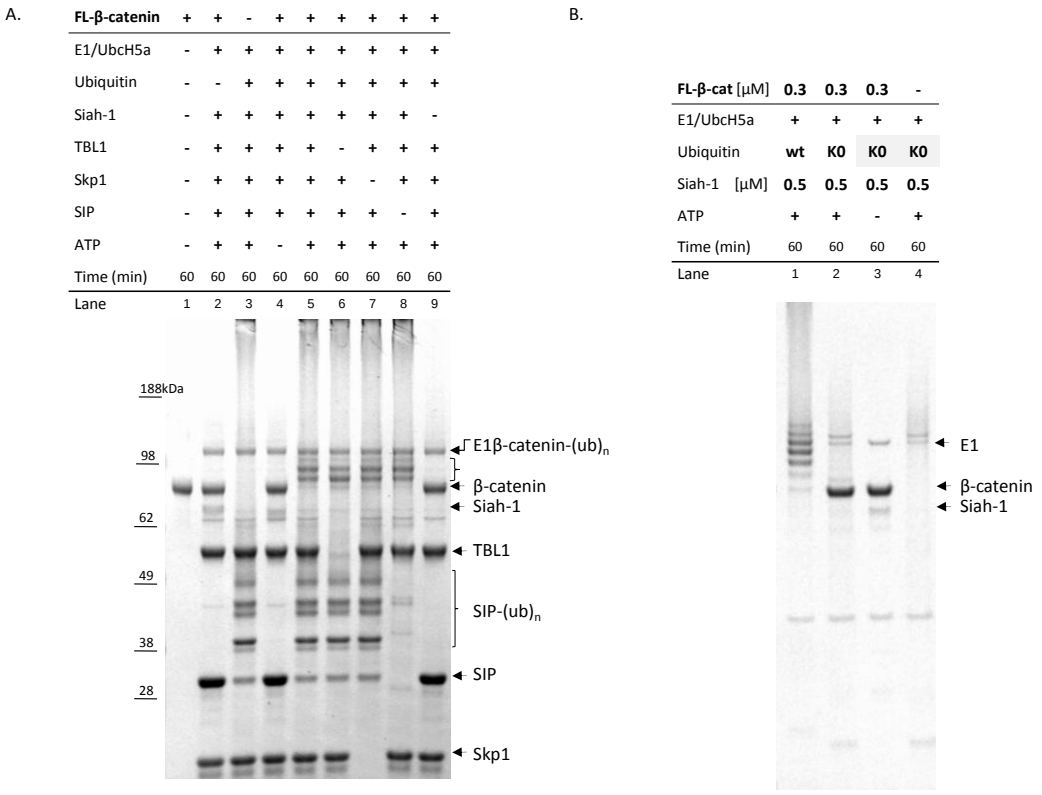


Figure A2.2. Full scale of the Coomassie stained SDS-PAGE gel from Figure .2.1B.(A) *In vitro* ubiquitination of β-catenin by Siah-1 with wt- and K(0)-ubiquitin(B).

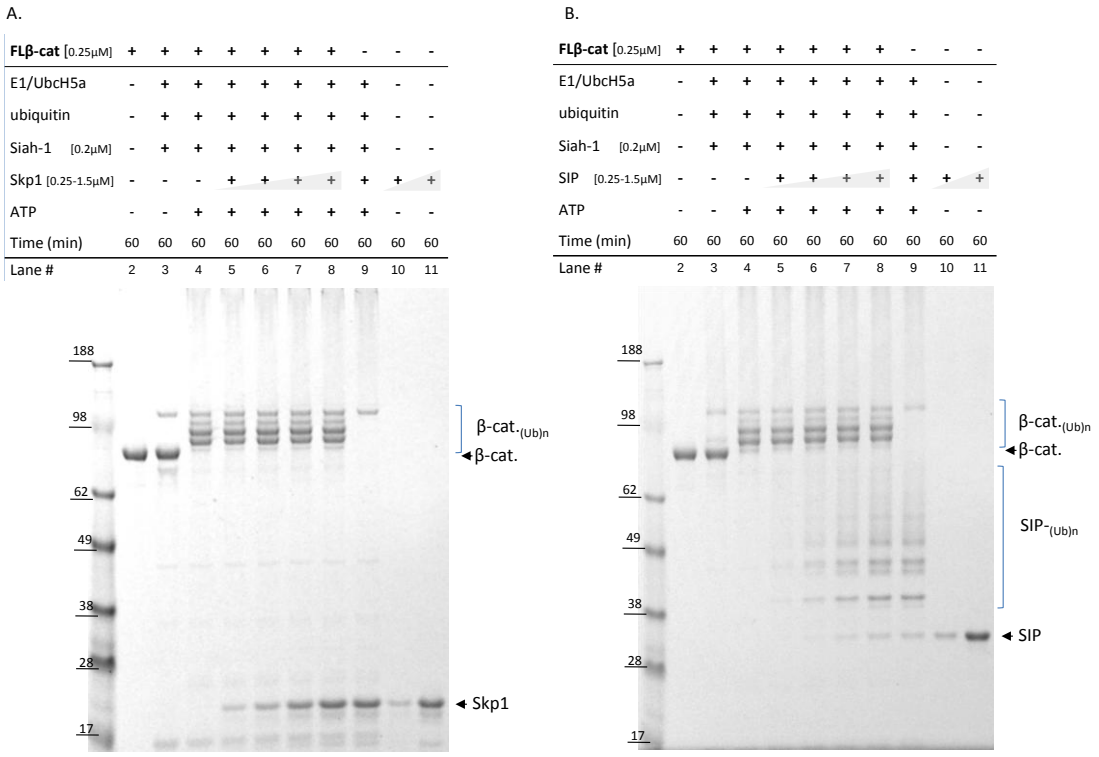


Figure A2.3. *In vitro* ubiquitination of β-catenin by Siah-1 in the presence of (A) Skp1 and (B) SIP.

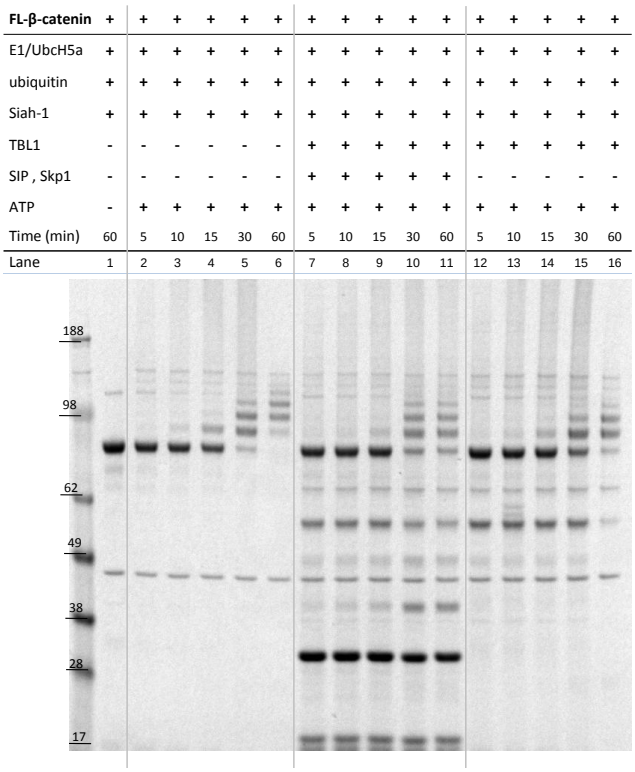


Figure A2.4. Full scale Coomassie stained SDS-PAGE gel of Figure. 2.3.

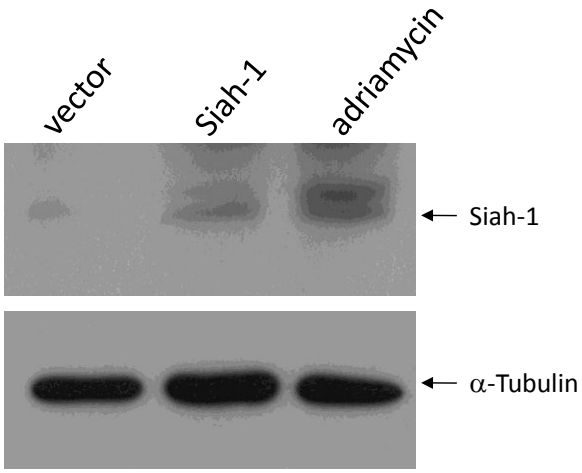


Figure A2.5. Comparison of Siah-1 expression levels in HEK293T cells.

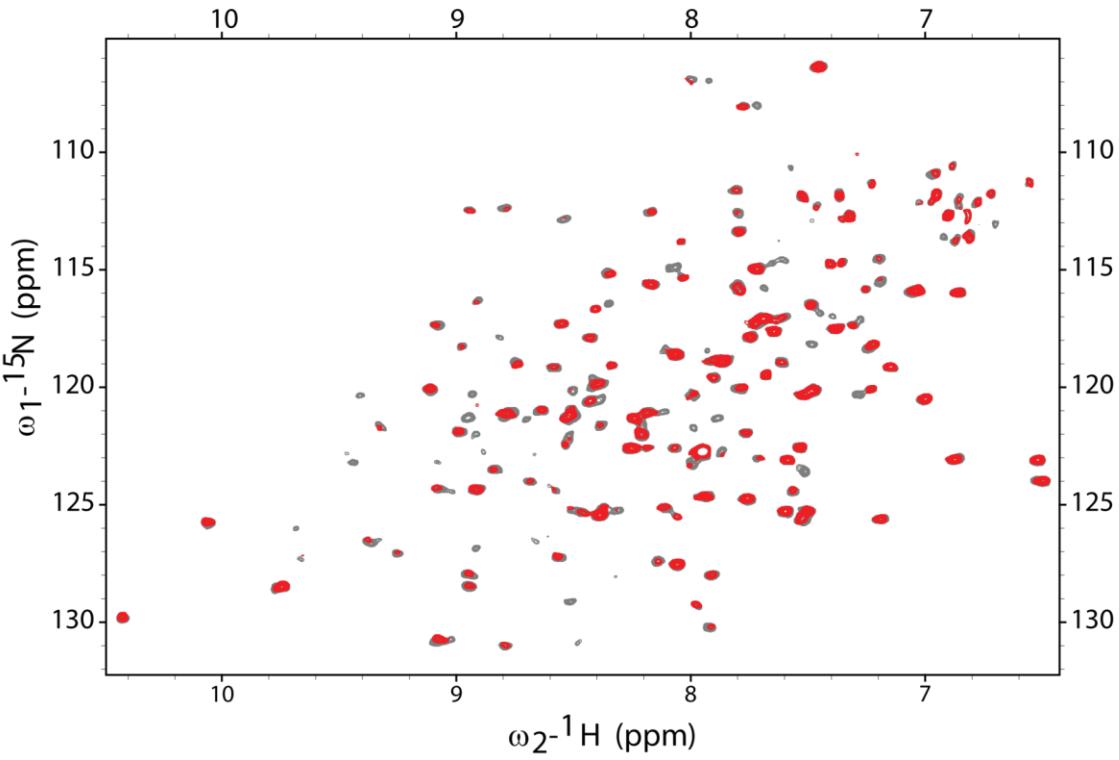


Figure A2.6. Overlay of full ^{15}N-^{1}H TROSY-HSQC spectra of ^{15}N-enriched Siah-SBD in the absence (blue) and presence (red) of β-catenin-FL.

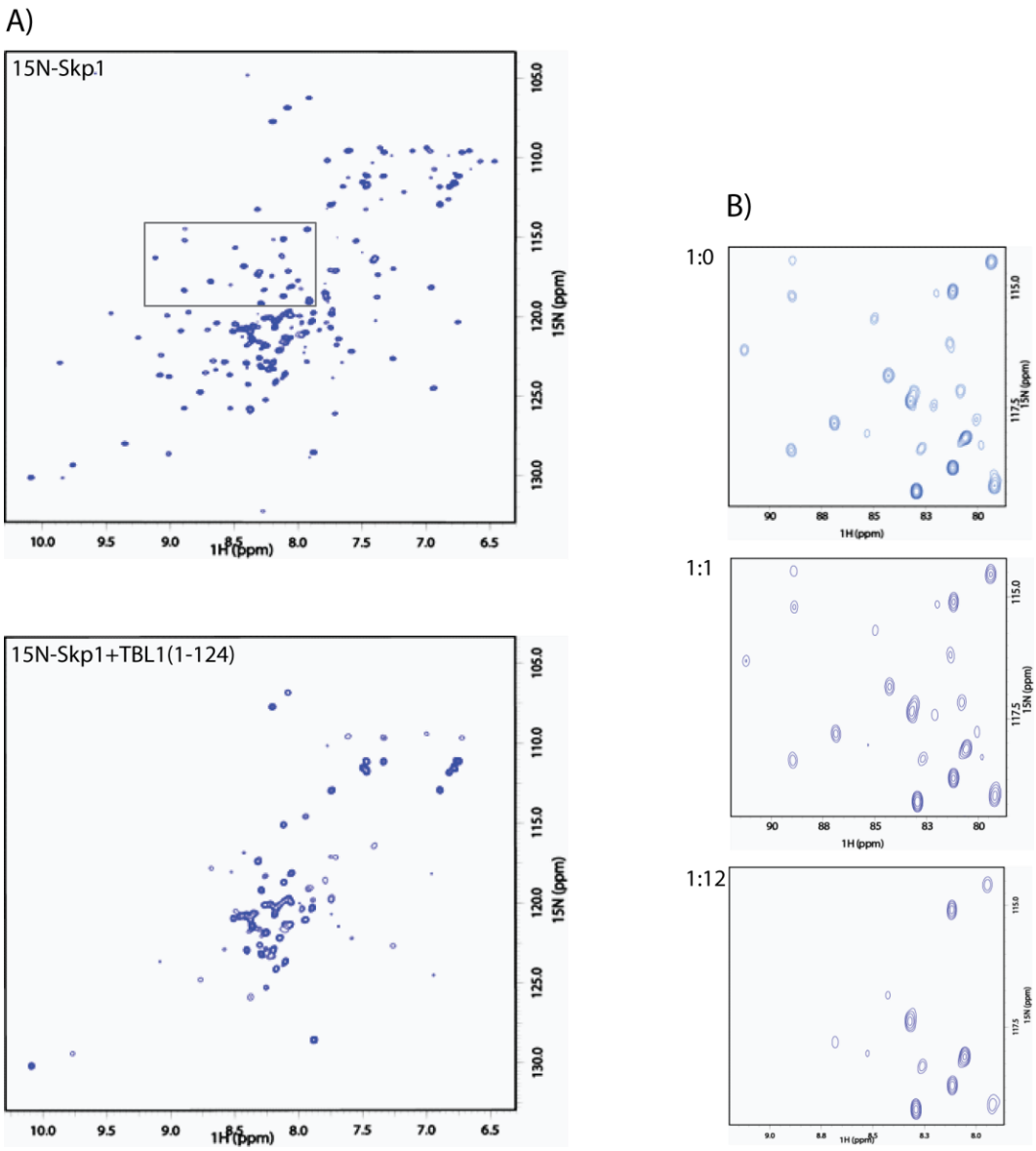


Figure A2.7. ^{15}N-^{1}H NMR chemical shift perturbation assay using ^{15}N- Skp1 and TBL1(31-179). A) ^{15}N-^{1}H HSQC spectra of ^{15}N-labeled Skp1 alone at 50 µM (top) and after titration of TBL1(31-179) (bottom) at 1:12 molar ratio. B) Zoom in view on the spectra in A. The spectra represent the titration of TBL1(31-179) at 1:1 and 1:12 molar ratio. The experiment was collected in 20 mM Na_2HPO_4 at pH 7, 20 mM NaCl and 20 mM βME.

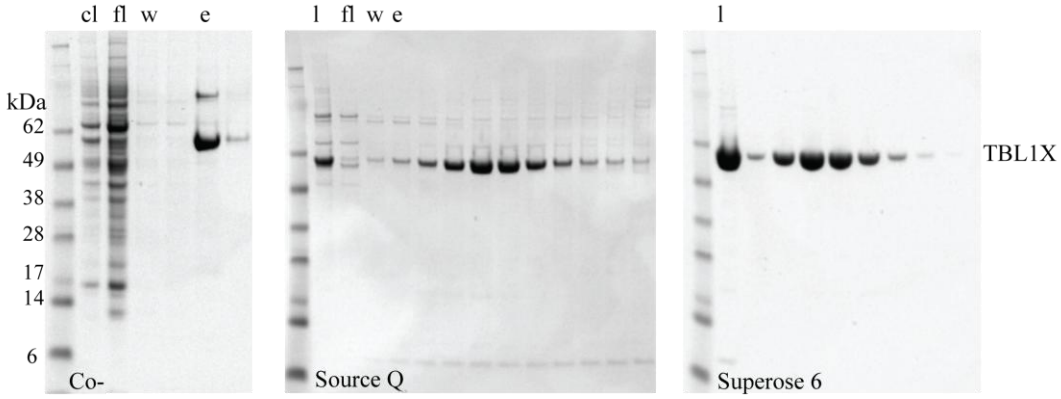


Figure A3.1. Purification of FL-TBL1X using Co-affinity, anion exchange and gel filtration chromatography. Lane abbreviations are cl- cell lysate, fl - flow through, w - wash, e - elution, l- load.

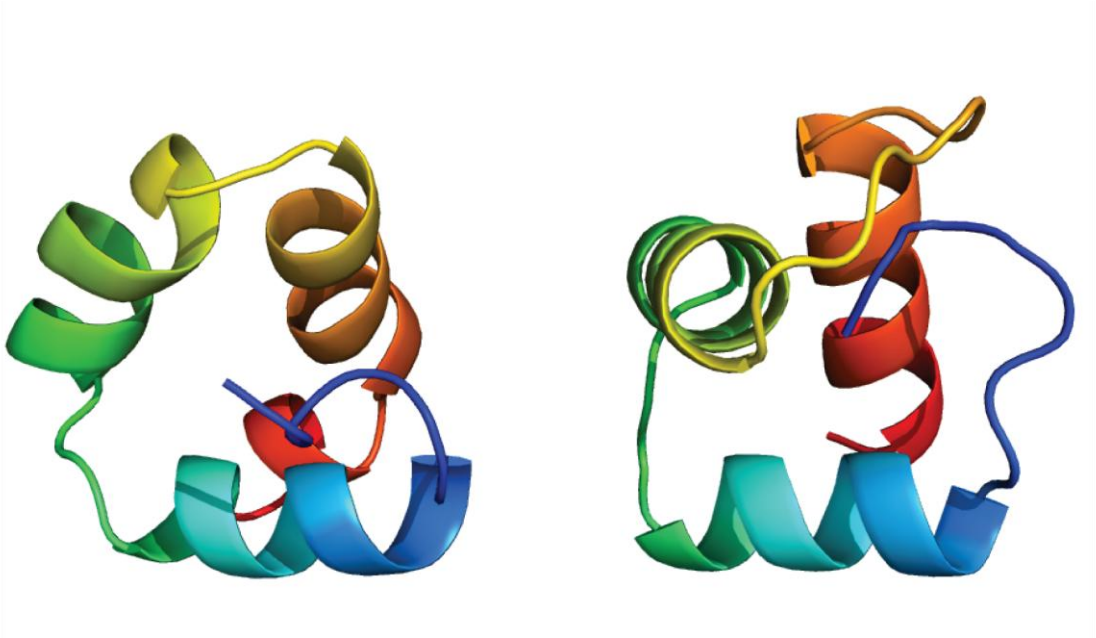
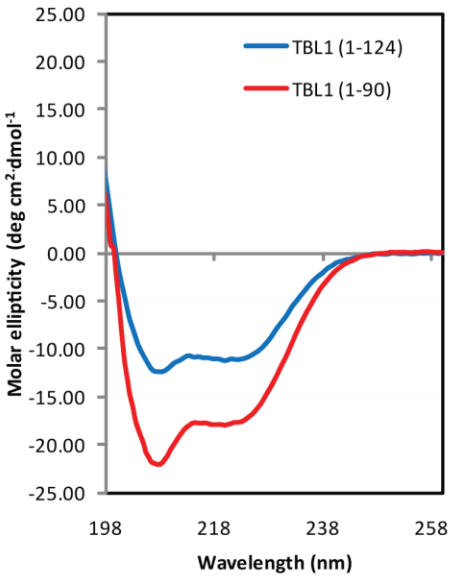


Figure A3.2. Comparison of TBL1-FBox domain models. Homology model generated with Modeller (left), de novo fold of the FBox with Rosetta (right).

A.



B.

```
TBL1 (90-120)   DVVQTRQQAFREKLAQQQANAAAAAAAAAAT
                --HHHHHHHHHHHHHHHHHHHHHHHHHHH--
```

Figure A3.3. Analysis of TBL1 coiled coil region. A) CD spectra of TBL1(1-124) and (1-90). B) Secondarys tructure prediction of TBL1(90-120) using the BioInfoBank MetaServer. “H” designates helical structure prediction.

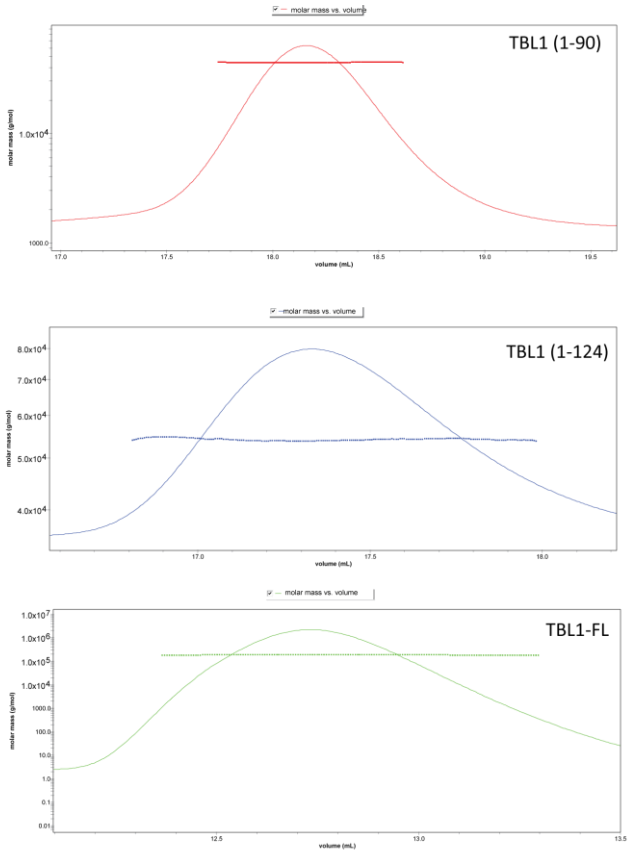


Figure A3.4. Molar mass distribution of TBL1(1-90), (1-124) and FL proteins from the SEC-MALS elution profile

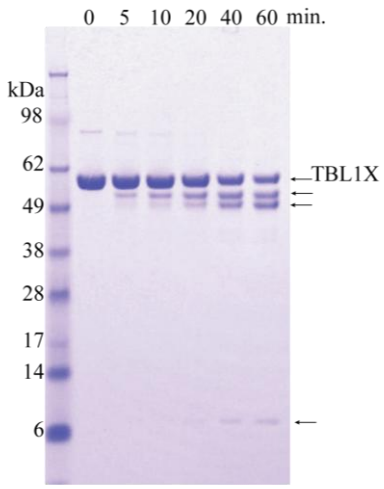


Figure A3.5. Limited proteolysis of FL-TBL1X with Chymotrypsin. The digestion reaction was carried out at 1:400 ratio of protease:protein at room temperature. Aliquots of the reaction were taken at the specified time points.

BIBLIOGRAPHY

1. Wu, G., et al., *Structure of a beta-TrCP1-Skp1-beta-catenin complex: destruction motif binding and lysine specificity of the SCF(beta-TrCP1) ubiquitin ligase.* Mol Cell, 2003. **11**(6): p. 1445-56.

2. Combet C., B.C., Geourjon C. and Deléage G., *NPS@: Network Protein Sequence Analysis*. 2000. p. 147-150.

3. Bassi, M.T., et al., *X-linked late-onset sensorineural deafness caused by a deletion involving OA1 and a novel gene containing WD-40 repeats.* Am J Hum Genet, 1999. **64**(6): p. 1604-16.

4. Yan, H.T., et al., *Molecular analysis of TBL1Y, a Y-linked homologue of TBL1X related with X-linked late-onset sensorineural deafness.* J Hum Genet, 2005. **50**(4): p. 175-81.

5. Zhang, X.M., et al., *TBLR1 regulates the expression of nuclear hormone receptor co-repressors.* BMC Cell Biol, 2006. **7**: p. 31.

6. Li, J. and C.Y. Wang, *TBL1-TBLR1 and beta-catenin recruit each other to Wnt target-gene promoter for transcription activation and oncogenesis.* Nat Cell Biol, 2008. **10**(2): p. 160-9.

7. Perissi, V., et al., *Deconstructing repression: evolving models of co-repressor action.* Nat Rev Genet. **11**(2): p. 109-23.

8. Choi, H.K., et al., *Function of Multiple LisH/WD-40 Repeat-Containing Proteins in Feed-Forward Transcriptional Repression by SMRT/N-CoR Corepressor Complexes.* Mol Endocrinol, 2008.

9. Guenther, M.G., et al., *A core SMRT corepressor complex containing HDAC3 and TBL1, a WD40-repeat protein linked to deafness.* Genes Dev, 2000. **14**(9): p. 1048-57.

10. Perissi, V., et al., *TBL1 and TBLR1 phosphorylation on regulated gene promoters overcomes dual CtBP and NCoR/SMRT transcriptional repression checkpoints.* Mol Cell, 2008. **29**(6): p. 755-66.

11. Zhang, J., et al., *The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2.* Mol Cell, 2002. **9**(3): p. 611-23.

12. Tang, A.H., et al., *PHYL acts to down-regulate TTK88, a transcriptional repressor of neuronal cell fates, by a SINA-dependent mechanism.* Cell, 1997. **90**(3): p. 459-67.

13. Matsuzawa, S.I. and J.C. Reed, *Siah-1, SIP, and Ebi collaborate in a novel pathway for beta-catenin degradation linked to p53 responses.* Mol Cell, 2001. **7**(5): p. 915-26.

14. Hershko, A., *Ubiquitin: roles in protein modification and breakdown.* Cell, 1983. **34**(1): p. 11-2.

15. Hershko, A., et al., *Proposed role of ATP in protein breakdown: conjugation of protein with multiple chains of the polypeptide of ATP-dependent proteolysis.* Proc Natl Acad Sci U S A, 1980. **77**(4): p. 1783-6.

16. Hershko, A., et al., *Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown.* J Biol Chem, 1983. **258**(13): p. 8206-14.

17. Glickman, M.H. and A. Ciechanover, *The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction.* Physiol Rev, 2002. **82**(2): p. 373-428.

18. Deshaies, R.J. and C.A. Joazeiro, *RING domain E3 ubiquitin ligases.* Annu Rev Biochem, 2009. **78**: p. 399-434.

19. Dye, B.T. and B.A. Schulman, *Structural mechanisms underlying posttranslational modification by ubiquitin-like proteins.* Annu Rev Biophys Biomol Struct, 2007. **36**: p. 131-50.

20. Harper, J.W. and B.A. Schulman, *Structural complexity in ubiquitin recognition.* Cell, 2006. **124**(6): p. 1133-6.

21. Petroski, M.D. and R.J. Deshaies, *Function and regulation of cullin-RING ubiquitin ligases.* Nat Rev Mol Cell Biol, 2005. **6**(1): p. 9-20.

22. Koegl, M., et al., *A novel ubiquitination factor, E4, is involved in multiubiquitin chain assembly.* Cell, 1999. **96**(5): p. 635-44.

23. Nordquist, K.A., et al., *Structural and functional characterization of the monomeric U-box domain from E4B.* Biochemistry. **49**(2): p. 347-55.

24. Jiang, J., et al., *CHIP is a U-box-dependent E3 ubiquitin ligase: identification of Hsc70 as a target for ubiquitylation.* J Biol Chem, 2001. **276**(46): p. 42938-44.

25. Ravid, T. and M. Hochstrasser, *Diversity of degradation signals in the ubiquitin-proteasome system.* Nat Rev Mol Cell Biol, 2008. **9**(9): p. 679-90.

26. Yoshida, Y., *F-box proteins that contain sugar-binding domains.* Biosci Biotechnol Biochem, 2007. **71**(11): p. 2623-31.

27. Eddins, M.J., et al., *Crystal structure and solution NMR studies of Lys48-linked tetraubiquitin at neutral pH.* J Mol Biol, 2007. **367**(1): p. 204-11.

28. Komander, D., et al., *Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains.* EMBO Rep, 2009. **10**(5): p. 466-73.

29. Kirkpatrick, D.S., C. Denison, and S.P. Gygi, *Weighing in on ubiquitin: the expanding role of mass-spectrometry-based proteomics.* Nat Cell Biol, 2005. **7**(8): p. 750-7.

30. Ye, Y. and M. Rape, *Building ubiquitin chains: E2 enzymes at work.* Nat Rev Mol Cell Biol, 2009. **10**(11): p. 755-64.

31. Kim, H.T., et al., *Certain pairs of ubiquitin-conjugating enzymes (E2s) and ubiquitin-protein ligases (E3s) synthesize nondegradable forked ubiquitin chains containing all possible isopeptide linkages.* J Biol Chem, 2007. **282**(24): p. 17375-86.

32. Amerik, A.Y. and M. Hochstrasser, *Mechanism and function of deubiquitinating enzymes.* Biochim Biophys Acta, 2004. **1695**(1-3): p. 189-207.

33. Reyes-Turcu, F.E., K.H. Ventii, and K.D. Wilkinson, *Regulation and cellular roles of ubiquitin-specific deubiquitinating enzymes.* Annu Rev Biochem, 2009. **78**: p. 363-97.

34. Nalepa, G., M. Rolfe, and J.W. Harper, *Drug discovery in the ubiquitin-proteasome system.* Nat Rev Drug Discov, 2006. **5**(7): p. 596-613.

35. Shackelford, J. and J.S. Pagano, *Tumor viruses and cell signaling pathways: deubiquitination versus ubiquitination.* Mol Cell Biol, 2004. **24**(12): p. 5089-93.

36. Shackelford, J. and J.S. Pagano, *Role of the ubiquitin system and tumor viruses in AIDS-related cancer.* BMC Biochem, 2007. **8 Suppl 1**: p. S8.

37. Klaus, A. and W. Birchmeier, *Wnt signalling and its impact on development and cancer.* Nat Rev Cancer, 2008. **8**(5): p. 387-98.

38. Moon, R.T., et al., *WNT and beta-catenin signalling: diseases and therapies.* Nat Rev Genet, 2004. **5**(9): p. 691-701.

39. MacDonald, B.T., K. Tamai, and X. He, *Wnt/beta-catenin signaling: components, mechanisms, and diseases.* Dev Cell, 2009. **17**(1): p. 9-26.

40. Chien, A.J., W.H. Conrad, and R.T. Moon, *A Wnt Survival Guide: From Flies to Human Disease.* J Invest Dermatol, 2009.

41. Lin, S.Y., et al., *Beta-catenin, a novel prognostic marker for breast cancer: its roles in cyclin D1 expression and cancer progression.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4262-6.

42. Rubinfeld, B., et al., *Association of the APC gene product with beta-catenin.* Science, 1993. **262**(5140): p. 1731-4.

43. Su, L.K., B. Vogelstein, and K.W. Kinzler, *Association of the APC tumor suppressor protein with catenins.* Science, 1993. **262**(5140): p. 1734-7.

44. Daniels, D.L., K. Eklof Spink, and W.I. Weis, *beta-catenin: molecular plasticity and drug design.* Trends Biochem Sci, 2001. **26**(11): p. 672-8.

45. Marikawa, Y. and R.P. Elinson, *beta-TrCP is a negative regulator of Wnt/beta-catenin signaling pathway and dorsal axis formation in Xenopus embryos.* Mech Dev, 1998. **77**(1): p. 75-80.

46. Aberle, H., et al., *beta-catenin is a target for the ubiquitin-proteasome pathway.* EMBO J, 1997. **16**(13): p. 3797-804.

47. Salomon, D., et al., *Regulation of beta-catenin levels and localization by overexpression of plakoglobin and inhibition of the ubiquitin-proteasome system.* J Cell Biol, 1997. **139**(5): p. 1325-35.

48. Winer, I.S., et al., *Lysine residues Lys-19 and Lys-49 of beta-catenin regulate its levels and function in T cell factor transcriptional activation and neoplastic transformation.* J Biol Chem, 2006. **281**(36): p. 26181-7.

49. Liu, J., et al., *Siah-1 mediates a novel beta-catenin degradation pathway linking p53 to the adenomatous polyposis coli protein.* Mol Cell, 2001. **7**(5): p. 927-36.

50. Henderson, B.R., *Nuclear-cytoplasmic shuttling of APC regulates beta-catenin subcellular localization and turnover.* Nat Cell Biol, 2000. **2**(9): p. 653-60.

51. Gerlitz, G., et al., *Novel functional features of the Lis-H domain: role in protein dimerization, half-life and cellular localization.* Cell Cycle, 2005. **4**(11): p. 1632-40.

52. Polekhina, G., et al., *Siah ubiquitin ligase is structurally related to TRAF and modulates TNF-alpha signaling.* Nat Struct Biol, 2002. **9**(1): p. 68-75.

53. Tata, J.R., *Signalling through nuclear receptors.* Nat Rev Mol Cell Biol, 2002. **3**(9): p. 702-10.

54. Huang, P., V. Chandra, and F. Rastinejad, *Structural overview of the nuclear receptor superfamily: insights into physiology and therapeutics.* Annu Rev Physiol. **72**: p. 247-72.

55. Ribeiro, R.C., P.J. Kushner, and J.D. Baxter, *The nuclear hormone receptor gene superfamily.* Annu Rev Med, 1995. **46**: p. 443-53.

56. Bain, D.L., et al., *Nuclear receptor structure: implications for function.* Annu Rev Physiol, 2007. **69**: p. 201-20.

57. Metivier, R., et al., *Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter.* Cell, 2003. **115**(6): p. 751-63.

58. Shang, Y., et al., *Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription.* Cell, 2000. **103**(6): p. 843-52.

59. Reid, G., et al., *Cyclic, proteasome-mediated turnover of unliganded and liganded ERalpha on responsive promoters is an integral feature of estrogen signaling.* Mol Cell, 2003. **11**(3): p. 695-707.

60. Perissi, V., et al., *A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors.* Cell, 2004. **116**(4): p. 511-26.

61. Schulman, I.G., *Nuclear receptors as drug targets for metabolic disease.* Adv Drug Deliv Rev.

62. Hotamisligil, G.S., *Inflammation and metabolic disorders.* Nature, 2006. **444**(7121): p. 860-7.

63. Cole, J.L., et al., *Analytical ultracentrifugation: sedimentation velocity and sedimentation equilibrium.* Methods Cell Biol, 2008. **84**: p. 143-79.

64. Brown, P.H. and P. Schuck, *Macromolecular size-and-shape distributions by sedimentation velocity analytical ultracentrifugation.* Biophys J, 2006. **90**(12): p. 4651-61.

65. Schuck, P., *Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling.* Biophys J, 2000. **78**(3): p. 1606-19.

66. Pervushin, K., et al., *Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution.* Proc Natl Acad Sci U S A, 1997. **94**(23): p. 12366-71.

67. Takeuchi, K. and G. Wagner, *NMR studies of protein interactions.* Curr Opin Struct Biol, 2006. **16**(1): p. 109-17.

68. Matsuo, H., Walters, K. J., Teruya, K., Tanaka, T., Gassner, G.T., Lippard, S. J., Kyogoku, Y., and Wagner G., *Identification by NMR Spectroscopy of Residues at Contact Surfaces in Large, Slowly Exchanging Macromolecular Complexes.* Journal of the American Chamical Society, 1999. **121**(42): p. 9903-9904.

69. Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution.* Q Rev Biophys, 2007. **40**(3): p. 191-285.

70. Rambo, R.P. and J.A. Tainer, *Bridging the solution divide: comprehensive structural analyses of dynamic RNA, DNA, and protein assemblies by small-angle X-ray scattering.* Curr Opin Struct Biol, 2010. **20**(1): p. 128-37.

71. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria.* J Appl Crystallogr, 1992. **25**: p. 495-503.

72. Svergun D., B.C., Koch M. H. J., *CRYSOL – a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates.* J Appl Crystallogr, 1995. **28**(6): p. 5.

73. Svergun, D.I., *Advanced solution scattering data analysis methods and their applications.* Journal of Applied Crystallography, 2000. **33**(3): p. 4.

74. Volkov, V.a.S., DI, *Uniqueness of ab initio shape determination in small-angle scattering.* J Appl Crystallogr, 2003. **36**: p. 860-864.

75. Hura, G.L., et al., *Robust, high-throughput solution structural analyses by small angle X-ray scattering (SAXS).* Nat Methods, 2009. **6**(8): p. 606-12.

76. Willert, K. and K.A. Jones, *Wnt signaling: is the party in the nucleus?* Genes Dev, 2006. **20**(11): p. 1394-404.

77. Polakis, P., *The oncogenic activation of beta-catenin.* Curr Opin Genet Dev, 1999. **9**(1): p. 15-21.

78. Latres, E., D.S. Chiaur, and M. Pagano, *The human F box protein beta-Trcp associates with the Cul1/Skp1 complex and regulates the stability of beta-catenin.* Oncogene, 1999. **18**(4): p. 849-54.

79. Yost, C., et al., *The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3.* Genes Dev, 1996. **10**(12): p. 1443-54.

80. Jang, K.L., et al., *Up-regulation of beta-catenin by a viral oncogene correlates with inhibition of the seven in absentia homolog 1 in B lymphoma cells.* Proc Natl Acad Sci U S A, 2005. **102**(51): p. 18431-6.

81. Goddard, T.D., and Kneller, D.G., *SPARKY 3*. 2006, University of California, San Francisco.

82. Gwak, J., et al., *Isoreserpine promotes beta-catenin degradation via Siah-1 up-regulation in HCT116 colon cancer cells.* Biochem Biophys Res Commun, 2009. **387**(3): p. 444-9.

83. Hu, G., et al., *Mammalian homologs of seven in absentia regulate DCC via the ubiquitin-proteasome pathway.* Genes Dev, 1997. **11**(20): p. 2701-14.

84. Liani, E., et al., *Ubiquitylation of synphilin-1 and alpha-synuclein by SIAH and its presence in cellular inclusions and Lewy bodies imply a role in Parkinson's disease.* Proc Natl Acad Sci U S A, 2004. **101**(15): p. 5500-5.

85. Tanikawa, J., et al., *p53 suppresses the c-Myb-induced activation of heat shock transcription factor 3.* J Biol Chem, 2000. **275**(20): p. 15578-85.

86. Zhang, J., et al., *Proteasomal regulation of nuclear receptor corepressor-mediated repression.* Genes Dev, 1998. **12**(12): p. 1775-80.

87. Bhattacharya, S., et al., *The modular structure of SIP facilitates its role in stabilizing multiprotein assemblies.* Biochemistry, 2005. **44**(27): p. 9462-71.

88. Brzovic, P.S., et al., *Binding and recognition in the assembly of an active BRCA1/BARD1 ubiquitin-ligase complex.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 5646-51.

89. Choi, H.J., A.H. Huber, and W.I. Weis, *Thermodynamics of beta-catenin-ligand interactions: the roles of the N- and C-terminal tails in modulating binding affinity.* J Biol Chem, 2006. **281**(2): p. 1027-38.

90. Durocher, Y., S. Perret, and A. Kamen, *High-level and high-throughput recombinant protein production by transient transfection of suspension-growing human 293-EBNA1 cells.* Nucleic Acids Res, 2002. **30**(2): p. E9.

91. Hao, B., et al., *Structural basis of the Cks1-dependent recognition of p27(Kip1) by the SCF(Skp2) ubiquitin ligase.* Mol Cell, 2005. **20**(1): p. 9-19.

92. Huber, A.H., W.J. Nelson, and W.I. Weis, *Three-dimensional structure of the armadillo repeat region of beta-catenin.* Cell, 1997. **90**(5): p. 871-82.

93. Jin, L., et al., *Mechanism of ubiquitin-chain formation by the human anaphase-promoting complex.* Cell, 2008. **133**(4): p. 653-65.

94. Xu, P., et al., *Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation.* Cell, 2009. **137**(1): p. 133-45.

95. Su, Y., et al., *APC is essential for targeting phosphorylated beta-catenin to the SCFbeta-TrCP ubiquitin ligase.* Mol Cell, 2008. **32**(5): p. 652-61.

96. Park, S., et al., *Hexachlorophene inhibits Wnt/beta-catenin pathway by promoting Siah-mediated beta-catenin degradation.* Mol Pharmacol, 2006. **70**(3): p. 960-6.

97. Kim, C.J., et al., *Inactivating mutations of the Siah-1 gene in gastric cancer.* Oncogene, 2004. **23**(53): p. 8591-6.

98. Peng, J., et al., *A proteomics approach to understanding protein ubiquitination.* Nat Biotechnol, 2003. **21**(8): p. 921-6.

99. Eklof Spink, K., S.G. Fridman, and W.I. Weis, *Molecular mechanisms of beta-catenin recognition by adenomatous polyposis coli revealed by the structure of an APC-beta-catenin complex.* EMBO J, 2001. **20**(22): p. 6203-12.

100. Ning, X., et al., *Calcyclin-binding protein inhibits proliferation, tumorigenicity, and invasion of gastric cancer.* Mol Cancer Res, 2007. **5**(12): p. 1254-62.

101. Santelli, E., et al., *Structural analysis of Siah1-Siah-interacting protein interactions and insights into the assembly of an E3 ligase multiprotein complex.* J Biol Chem, 2005. **280**(40): p. 34278-87.

102. Sun, S., et al., *Overexpressed CacyBP/SIP leads to the suppression of growth in renal cell carcinoma.* Biochem Biophys Res Commun, 2007. **356**(4): p. 864-71.

103. Schulman, B.A., et al., *Insights into SCF ubiquitin ligases from the structure of the Skp1-Skp2 complex.* Nature, 2000. **408**(6810): p. 381-6.

104. Zheng, N., et al., *Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex.* Nature, 2002. **416**(6882): p. 703-9.

105. Boulton, S.J., et al., *A role for Ebi in neuronal cell cycle control.* Embo J, 2000. **19**(20): p. 5376-86.

106. Chang, P.J., et al., *Negative-feedback regulation of proneural proteins controls the timing of neural precursor division.* Development, 2008. **135**(18): p. 3021-30.

107. House, C.M., et al., *Elucidation of the substrate binding site of Siah ubiquitin ligase.* Structure, 2006. **14**(4): p. 695-701.

108. Chitalia, V.C., et al., *Jade-1 inhibits Wnt signalling by ubiquitylating beta-catenin and mediates Wnt pathway inhibition by pVHL.* Nat Cell Biol, 2008. **10**(10): p. 1208-16.

109. Fujita, Y., et al., *Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex.* Nat Cell Biol, 2002. **4**(3): p. 222-31.

110. Nastasi, T., et al., *Ozz-E3, a muscle-specific ubiquitin ligase, regulates beta-catenin degradation during myogenesis.* Dev Cell, 2004. **6**(2): p. 269-82.

111. Christensen, D.E., P.S. Brzovic, and R.E. Klevit, *E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages.* Nat Struct Mol Biol, 2007. **14**(10): p. 941-8.

112. Emes, R.D. and C.P. Ponting, *A new sequence motif linking lissencephaly, Treacher Collins and oral-facial-digital type 1 syndromes, microtubule dynamics and cell migration.* Hum Mol Genet, 2001. **10**(24): p. 2813-20.

113. Yoon, H.G., et al., *Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1.* Embo J, 2003. **22**(6): p. 1336-46.

114. Cerna, D. and D.K. Wilson, *The structure of Sif2p, a WD repeat protein functioning in the SET3 corepressor complex.* J Mol Biol, 2005. **351**(4): p. 923-35.

115. Lupas, A., M. Van Dyke, and J. Stock, *Predicting coiled coils from protein sequences.* Science, 1991. **252**(5009): p. 1162-4.

116. Kim, M.H., et al., *The structure of the N-terminal domain of the product of the lissencephaly gene Lis1 and its functional implications.* Structure (Camb), 2004. **12**(6): p. 987-98.

117. Mikolajka, A., et al., *Structure of the N-terminal domain of the FOP (FGFR1OP) protein and implications for its dimerization and centrosomal localization.* J Mol Biol, 2006. **359**(4): p. 863-75.

118. Sali, A. and T.L. Blundell, *Comparative protein modelling by satisfaction of spatial restraints.* J Mol Biol, 1993. **234**(3): p. 779-815.

119. Dimitrova, Y.N., et al., *Direct ubiquitination of beta-catenin by Siah-1 and regulation by the exchange factor TBL1.* J Biol Chem, 2010. **285**(18): p. 13507-16.

120. Kaufmann, K.W., et al., *Practically useful: what the Rosetta protein modeling suite can do for you.* Biochemistry. **49**(14): p. 2987-98.

121. Ginalski, K., et al., *3D-Jury: a simple approach to improve protein structure predictions.* Bioinformatics, 2003. **19**(8): p. 1015-8.

122. Orlicky, S., et al., *Structural basis for phosphodependent substrate selection and orientation by the SCFCdc4 ubiquitin ligase.* Cell, 2003. **112**(2): p. 243-56.

123. Navarro, M.V., et al., *Structural analysis of the GGDEF-EAL domain-containing c-di-GMP receptor FimX.* Structure, 2009. **17**(8): p. 1104-16.

124. Mateja, A., et al., *The dimerization mechanism of LIS1 and its implication for proteins containing the LisH motif.* J Mol Biol, 2006. **357**(2): p. 621-31.

125. Tarricone, C., et al., *Coupling PAF signaling to dynein regulation: structure of LIS1 in complex with PAF-acetylhydrolase.* Neuron, 2004. **44**(5): p. 809-21.

126. Locasale, J.W., A.S. Shaw, and A.K. Chakraborty, *Scaffold proteins confer diverse regulatory properties to protein kinase cascades.* Proc Natl Acad Sci U S A, 2007. **104**(33): p. 13307-12.

127. Shaw, A.S. and E.L. Filbert, *Scaffold proteins and immune-cell signalling.* Nat Rev Immunol, 2009. **9**(1): p. 47-56.

128. Andrade, M.A., et al., *Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network.* Protein Eng, 1993. **6**(4): p. 383-90.

129. Maxim V. Petoukhov, a., b Peter V. Konarev,a,b Alexey G. Kikhneya and Dmitri I. and D. Svergun, *ATSAS 2.1 – towards automated and websupported small-angle scattering data analysis.* Journal of Applied Crystallography, 2007. **40**: p. 6.

130. Konarev PV, V.V., Sokolova AV, Koch MHJ, Svergun DI, *PRIMUS: A Windows PC-based system for small-angle scattering data analysis.* J Appl Crystallogr, 2003. **36**: p. 1277-1282.

131. Rohl, C.A., et al., *Protein structure prediction using Rosetta.* Methods Enzymol, 2004. **383**: p. 66-93.

132. Simons, K.T., et al., *Ab initio protein structure prediction of CASP III targets using ROSETTA.* Proteins, 1999. **Suppl 3**: p. 171-6.

133. D.A. Case, T.A.D., T.E. Cheatham, III, C.L. Simmerling, J., et al., *AMBER 10*. 2008, University of California, San Francisco.

134. Hornak, V., et al., *Comparison of multiple Amber force fields and development of improved protein backbone parameters.* Proteins, 2006. **65**(3): p. 712-25.

135. Matsuzawa, S., et al., *p53-inducible human homologue of Drosophila seven in absentia (Siah) inhibits cell growth: suppression by BAG-1*. Embo J, 1998. **17**(10): p. 2736-47.

136. Rual, J.F., et al., *Towards a proteome-scale map of the human protein-protein interaction network.* Nature, 2005. **437**(7062): p. 1173-8.

137. Liu, D., et al., *Estrogen-enhanced gene expression of lipoprotein lipase in heart is antagonized by progesterone.* Endocrinology, 2008. **149**(2): p. 711-6.

138. Frasor, J., et al., *Estrogen down-regulation of the corepressor N-CoR: mechanism and implications for estrogen derepression of N-CoR-regulated genes.* Proc Natl Acad Sci U S A, 2005. **102**(37): p. 13153-7.