



12-2014

# Strengths and Weaknesses of Hybrid TPR Technology for Obtaining Structural and Mechanistic Insights into TPR Proteins

Shanshan Yu

*University of Tennessee Health Science Center*

Follow this and additional works at: <https://dc.uthsc.edu/dissertations>

 Part of the [Amino Acids, Peptides, and Proteins Commons](#), [Investigative Techniques Commons](#), [Medical Biochemistry Commons](#), and the [Medical Microbiology Commons](#)

## Recommended Citation

Yu, Shanshan, "Strengths and Weaknesses of Hybrid TPR Technology for Obtaining Structural and Mechanistic Insights into TPR Proteins" (2014). *Theses and Dissertations (ETD)*. Paper 400. <http://dx.doi.org/10.21007/etd.cghs.2014.0368>.

This Dissertation is brought to you for free and open access by the College of Graduate Health Sciences at UTHSC Digital Commons. It has been accepted for inclusion in Theses and Dissertations (ETD) by an authorized administrator of UTHSC Digital Commons. For more information, please contact [jwelch30@uthsc.edu](mailto:jwelch30@uthsc.edu).

---

# Strengths and Weaknesses of Hybrid TPR Technology for Obtaining Structural and Mechanistic Insights into TPR Proteins

**Document Type**

Dissertation

**Degree Name**

Doctor of Philosophy (PhD)

**Program**

Biomedical Sciences

**Track**

Microbiology, Immunology, and Biochemistry

**Research Advisor**

Brenda A. Schulman, Ph.D.

**Committee**

John V. Cox, Ph.D. Eric J. Enemark, Ph.D. Roderick T. Hori, Ph.D. Stephen White, Ph.D.

**DOI**

10.21007/etd.cghs.2014.0368

**Strengths and Weaknesses of Hybrid TPR Technology for Obtaining Structural and  
Mechanistic Insights into TPR Proteins**

A Dissertation  
Presented for  
The Graduate Studies Council  
The University of Tennessee  
Health Science Center

In Partial Fulfillment  
Of the Requirements for the Degree  
Doctor of Philosophy  
From The University of Tennessee

By  
Shanshan Yu  
December 2014

Copyright © 2014 by Shanshan Yu.  
All rights reserved.

## **DEDICATION**

To my parents, Suozhen Chen, Muqing Yu  
my brother, Ge Yu  
and all my lovely friends of Memphis  
for their love and support

## 献 给

我的父母，陈锁珍，余木清  
我的弟弟，余戈  
和孟菲斯我可爱的朋友们  
谢谢你们的爱与支持

## ACKNOWLEDGEMENTS

Above all, I would like to thank my mentor, Brenda Schulman, for immersing me into the world of cutting-edge research and teaching me what it takes to keep doing a leading science. The training will serve me well always. I want to thank Brenda for creating a collaborative, informative and productive environment for doing sciences—both in terms of recruiting fantastic people and providing great resources. Also, I would like to thank committee members: John Cox, Eric Enemark, Roderick Hori and Stephen White for their effort and the advices they put into guiding me through graduate school.

I would like to thank the members of the Schulman lab for all that they have taught me and making this journey a fun one: Danny Scott for his insightful scientific questions and suggestions; Dave Duda for all the brilliant crystallography tips and tricks; David Miller for keeping the lab fully stocked and equipment always running; Hari Kamadurai for all the experiment guides over the years; Jing Wang and Masaya Yamaguchi for their help on projects we did together; Jeremy Frye for great advices towards experiments and structural softwares; Ryan VanderLinden for all the scientific discussion and suggestions; Nicholas Brown and Kuen-Phon Wu for the inspiring conversation and help; Randy Watson a fellow PhD candidate, for all the scientific (and nonscientific) discussion, advices sand being very helpful; Vladislav Sviderskiy for being helpful and for all the fun times. Manjeet Mukherjee and Yumei Zheng (a fellow graduate student), for sharing parts of the journey; Shelia Boezman for all the care, administrative support and activity organization (not forgetting the chocolates, candies and doughnuts); Jenny Olszewski, Shein Ei Cho and Omar Alsharif for being great lab mates and the fun time together.

Additionally, my thanks are for: Billy Dye for the guidance with cloning and all the smart jokes; Steve Kaiser for all the scientific discussions and the insights to the career; Mathew “Matt-man” Calabrese for the taking the time to teach; Julie Monda, Alan “DDR” Deng and Steven Seyedin for spreading the joy and for being ever helpful.

I would also like to thank the following people their tremendous help along the way: Igor Polyakov for data collection and phase calculation. Darcie Miller for her help with structure refinement; Luigi Iconaru, Hari Kamadurai, and Weixing Zhang for NMR experiment operation and interpretations; John Bollinger and John Rogers for the computer/IT support. David King for the Mass Spectrometry and Intact Mass data; Amanda Nourse for the analytical ultracentrifugation data; Randy Watson, Vladislav Sviderskiy and Jeremy Frye for the dissertation proofreading.

Finally, a big THANK YOU to my parents, my brother and all my friends who have made this possible!

## ABSTRACT

Tetratricopeptide (TPR) repeats are a 34-residue helix-turn-helix motif that when repeated pack into a superhelical structure. TPR domains are frequently found mediating protein-protein interactions, often through a central groove. One protein complex bearing numerous TPR repeats is the Anaphase Promoting Complex (APC). The anaphase-promoting complex (APC) is a multi-subunit complex, which orchestrates mitotic cell cycles. APC is an E3 ligase in the ubiquitin cascade, and directs the 26S proteasome degradation of cell-cycle regulators. Throughout mitotic progression, proteins that are key regulators of the cell cycle are assembled with polyubiquitin chains by APC.

One domain of the human APC is comprised of four related TPR proteins, APC8, APC6, APC3, and APC7, with each found in pairs. Crystal structures of some of these indicate that each has an N-terminal dimerization domain and a C-terminal domain that APC3 extends away from the dimer interface. The TPR C-terminal domains are thought to play major roles in mediating protein interactions within the APC.

The subunit APC3 plays major roles in regulating APC function. Within an APC3 dimer, each C-terminal domain recruits the Ile-Arg motifs of substrate coreceptors Cdh1 (or Cdc20) and APC10. Cdh1 and APC10 together recruit substrates for ubiquitination. Therefore, it is important to understand the structure of APC3, and how APC3 mediates interactions. To address this problem, I used a novel “hybrid TPR” technology, in which some TPRs from a distant relative of APC3 are fused upstream of the C-terminal domain from human APC3. This approach enabled determination of a 3Å resolution structure encompassing the sequence of the APC3 C-terminal domain. Interestingly, only a fraction of the structure resembles canonical TPR repeats. Interpretation of the crystal structure based on published structures of complexes between TPR proteins and their partners, and on published electron microscopy structures of APC-Cdh1-APC10, reveal that the region containing the Cdh1/APC10 binding site adopts 3 canonical TPR repeats. The remainder of the portion of the structure corresponding to human APC3 is folded into an alternative conformation, in which a helix from the atypical portion of APC3 buries the Cdh1/APC10-interacting groove within the crystal. Accordingly, unlike wild-type APC3, the hybrid TPR APC3 fails to bind Cdh1 and APC10. Nonetheless, the crystal structure of “hybrid TPR APC3 C-terminal domain” allows the prediction of potentially important residues for binding to Cdh1 and APC10. Taken together, the data reveal strengths and weaknesses of hybrid TPR technology for obtaining structural insights into TPR subunits of multiprotein assemblies such as APC.



## TABLE OF CONTENTS

<b>CHAPTER 1. INTRODUCTION .....</b>	<b>1</b>
APC Regulates Cell Cycle Progression .....	1
Cell cycle brief introduction .....	1
An overview of ubiquitin cascade.....	1
The APC directs cell division via substrate destruction .....	2
APC Architecture and Subunit Organization.....	5
APC subunit organization .....	5
APC recruits co-activators through TPR subunits .....	8
Co-activators interact with the APC through multiple interactions.....	9
Substrate recruitment to APC through the co-activators and APC10.....	11
The contribution of APC10 to APC processivity .....	13
TAME is an IR-tail mimic APC inhibitor.....	14
Crystallography study of APC subunits.....	14
APC Activity Regulation .....	18
Emi1 inhibition .....	18
MCC inhibition.....	18
Substrate specificity controlled through co-activators.....	19
APC activity regulation via UbcH10 auto-ubiquitination .....	20
Aims of the Study .....	21
<b>CHAPTER 2. CRYSTALLIZATION AND STRUCTURE DETERMINATION OF THE C-TERMINAL APC3: INSIGHTS INTO THE RECRUITMENT OF THE APC CO-ACTIVATORS AND APC10 BY APC3 .....</b>	<b>22</b>
Crystallization Strategy.....	22
Results and Discussion .....	24
The binding ability test of the chimeric APC3s.....	24
Improvement of bacterial expression of the chimeric APC3 .....	24
Purification improvement of the seleno-methionine labeled protein.....	30
Diffraction improvement and structure determination .....	33
Structure analysis of the chimeric APC3 atomic structure .....	33
Identify human APC3 C-terminal binding groove through model comparisons.....	41
Human APC3 binding groove has two conserved pockets at the inner surface, one negatively charged and the other hydrophobic .....	45
Validate human APC3 interactions with the IR-tails through mutations of the APC3 binding groove .....	45
A hypothesis of APC3 recruiting the IR-tail and implication for future study.....	50
TPR subunits models docking into APC EM map.....	50
Implication of the C-terminal APC3 alternative folding .....	55
Biological meaning of APC3 folding and implication to future study .....	60
Crystallization attempts of APC TPR subunits.....	62
Materials and Methods.....	64

Constructs for crystallization .....	64
Protein expression and purification from bacteria .....	65
Crystal screening and optimization.....	66
Crystallization of the chimeric APC3 modified by reductive methylation.....	67
Crystallization of the T4 Lysozyme, EGFP and MBP-(AAA) fused chimeric APC3.....	68
Data collection .....	68
Solution and phase determination.....	68
Structure refinement.....	69
Co-pulldown assays .....	69
The models of the APC subunits docking in the APC EM maps .....	71
<b>CHAPTER 3. INTERACTION STUDY OF THE CO-ACTIVATORS WITH APC3, APC8 AND THE APC PLATFORM.....</b>	<b>73</b>
Introduction.....	73
Results and Discussion .....	73
Cdh1 binds to APC3 mainly through its C-terminal IR-tail peptide .....	73
Study of the IR-tail APC3 interactions via an IR-tail mimicked compound .....	80
APC8 interacts with N-terminal Cdh1 .....	82
The interaction of N-terminal Cdh1 with the APC platform .....	85
Implication of Cdh1 N-terminal interactions .....	85
Materials and Methods.....	89
Constructs and insect cell infections.....	89
Protein expression and purification .....	91
Native gel shift assays.....	92
WaterLOGSY .....	92
<b>CHAPTER 4. IMPLICATIONS AND FUTURE DIRECTIONS .....</b>	<b>93</b>
Insights from the Hybrid Structure of the C-terminal Portion of Human APC3 .....	93
Strength and Weakness of Hybrid TPR Proteins/Technology.....	94
Brief introduction to the hybrid proteins strategy.....	94
Comparisons of LRR and TPR proteins .....	95
The strength and weakness of the TPR hybrid technology.....	98
<b>LIST OF REFERENCES.....</b>	<b>101</b>
<b>APPENDIX. SUPPLEMENTARY DATA FOR STRUCTURAL STUDIES OF THE TPR SUBUNITS OF THE HUMAN APC .....</b>	<b>108</b>
<b>VITA.....</b>	<b>115</b>

## LIST OF TABLES

Table 1-1.	The APC core and regulatory subunits .....	6
Table 2-1.	Summary of optimization approaches to improve the chimeric APC3 crystal diffraction .....	34
Table 2-2.	Structure data and refinement statistics of the chimeric APC3 .....	37
Table 2-3.	Correlation coefficients of EM-derived TPR structures to APC EM maps .....	53
Table 4-1.	Comparison of LRR (Leucine rich repeat) and TPR (Tetratricopeptide repeat) proteins.....	96
Table A-1.	A crystallization summary of the APC TPR subunits .....	111

## LIST OF FIGURES

Figure 1-1.	APC functions as an E3 ligase in ubiquitination cascade .....	3
Figure 1-2.	The order of APC substrate destruction.....	4
Figure 1-3.	The schematic of the APC subunit topology and the APC architecture .....	7
Figure 1-4.	Sequence alignment of C-terminal APC10, Cdh1 and Cdc20 .....	10
Figure 1-5.	Detail and schematic demonstration of Cdh1 and APC10 as APC substrate co-receptors.....	12
Figure 1-6.	Crystal structures of APC10, Cdh1, and Cdc20 .....	15
Figure 1-7.	APC6 adopts a superhelix structure.....	17
Figure 2-1.	The crystallization strategy of the human APC3 C-terminal domain.....	23
Figure 2-2.	Test the chimeric APC3 binding ability to Cdh1 by a co-pulldown.....	25
Figure 2-3.	Examine the chimeric APC3 binding ability with APC inhibitor via 1D WaterLOGSY .....	27
Figure 2-4.	Bacterial production of the chimeric APC3 is improved by a GST-MBP tag .....	31
Figure 2-5.	ATP incubation efficiently removed Chaperonin during the chimeric APC3 purification .....	32
Figure 2-6.	Original hits and improved crystals of the SeMet incorporated chimeric APC3 .....	35
Figure 2-7.	Representative electron density .....	36
Figure 2-8.	Analysis of the chimeric APC3 structure through sequence and structural comparisons .....	38
Figure 2-9.	Structural alignments of the chimeric APC3 with TPR proteins reveal a rigid binding groove and pseudo binding partner .....	42
Figure 2-10.	Pseudo binding partner $\alpha$ -11 interacts with two binding pockets in the APC3 binding groove .....	43
Figure 2-11.	Mutants of key residues inside the APC3 binding groove validated interactions between the APC3 and Cdh1 IR-tail .....	47

Figure 2-12.	A hypothetical model demonstrates IR-tail interacting with the APC3 binding groove .....	51
Figure 2-13.	TPR subunits form elongated V-shape homo-dimers in the human APC assembly.....	54
Figure 2-14.	Models of TPR subunits docked into the negative-stain EM map of the human APC <sup>Cdh1-Emi1</sup> complex .....	56
Figure 2-15.	The binding groove of human APC3 affects TPR11 adopts a canonical folding.....	59
Figure 2-16.	Human APC3 TPR7 folds in a TPR6-dependent manner.....	61
Figure 2-17.	The superhelix structure limits the accessibility of the APC3 binding groove .....	63
Figure 3-1.	Human Cdh1 binds to APC3 and APC8 <i>in vitro</i> .....	74
Figure 3-2.	Map human Cdh1 domains required to interact with APC3.....	76
Figure 3-3.	Identify Cdh1 binding domains to APC3.....	79
Figure 3-4.	Compound TAME mimics the IR tail interactions with APC3 .....	81
Figure 3-5.	Map human Cdh1 domains required to interact with APC8.....	83
Figure 3-6.	Identify Cdh1 binding domains to APC8.....	84
Figure 3-7.	N-terminal Cdh1 interacts with the APC platform .....	86
Figure 4-1.	Leucine rich repeat proteins have repeatable structural framework with rigid structural motifs.....	97
Figure 4-2.	TPR proteins have interaction networks involved in TPR motif packing.....	99
Figure A-1.	Intact masspectra of the SeMet incorporated and native chimeric APC3.....	108
Figure A-2.	Different crystal forms of the chimeric APC3 .....	109
Figure A-3.	The structural comparison of the N-terminal APC7 and APC3 .....	110
Figure A-4.	Crystallization of the human APC3 $\Delta$ L-APC16 <sup>74</sup> .....	112
Figure A-5.	The crystallization summary of APC3 $\Delta$ L-APC16 <sup>74</sup> -Cdh1-Hsl <sup>30</sup> .....	113
Figure A-6.	Human Cdh1 N-terminal secondary structure prediction from Psipred ..	114

## LIST OF ABBREVIATIONS

Å	Angstroem
AMP	Adenosine 5'- monophosphate
APC	Anaphase promoting complex
ATP	Adenosine 5'-triphosphate
BTP	Bis-tris propane
Cdc20	Cell division cycle 20
Cdh1	Cdc20 homolog 1
Cdk	Cyclin dependent kinase
ChA3s	Chimeric APC3s
CRL	Cullin RING ligase
C-terminal	Carboxy terminus
Da	Dalton
D- box	Destruction box (RxxLxxI/VxN)
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
E1	Ubiquitin- activating enzyme
E2	Ubiquitin- conjugating enzyme
E3	Ubiquitin ligase
EM	Electron microscopy
Emi1	Early mitotic inhibitor 1
Fl	Full- length
G1	Gap phase 1
G2	Gap phase 2
Hi-5 cells	<i>Trichoplusia ni</i> ovarian cells
His6	Hexahistidine
Hsp	Heat shock protein
IPTG	Isopropyl- $\beta$ - D- thiogalactopyranoside
IR- tail	Isoleucine- arginine C terminus
K11- linked	Ubiquitin chains amioesterified via lysine 11 of the ubiquitins
KEN- box	Destruction motif (KENxxxN/D)
LRR	Leucine rich repeat
MBP	Maltose binding protein
MCC	Mitotic checkpoint complex
MCM	Minichromosome maintenance
NEBD	Nuclear Envelope Breakdown
N-terminal	Amino terminus
PBS	Phosphate buffered sialine
PC repeat	Proteasome/cyclosome repeat
PEG	Polyethylene glycol
SAC	Spindle assembly checkpoint
SCF	Skp- culiln- F- box containing complex; E3 ubiquitin ligase
Sf9 cells	<i>Spodoptera frugiperda</i> ovarian cells
S phase	Synthesis phase

Strep tag	A synthetic peptide consisting of eight amino acids (WSHPQFEK)
TLR	Toll-like receptor
TPR	Tetratricopeptide repeat
Ub	Ubiquitin
v/v	Volume per volume
WaterLOGSY	Water-ligand observed via gradient spectroscopy
WD40 domain	WD40- repeat containing protein fold; also $\beta$ - propeller
ZBR	Zinc-binding region
3D	Three dimensional

## CHAPTER 1. INTRODUCTION

### APC Regulates Cell Cycle Progression

The anaphase-promoting complex (APC) is a multi-subunit Cullin-RING E3 ubiquitin ligase that regulates cell cycle progression through mitosis. One of the most important functions of APC is to decorate key regulator proteins of the cell cycle with polyubiquitin chains conjugated to initiate their proteolysis by the 26S proteasome.

#### Cell cycle brief introduction

The eukaryotic cell cycle is the process during which the cell replicates its genomic DNA, and divides the DNA equally between two newly replicated daughter cells through mitosis. The precise DNA duplication and cell division are critical to maintain normal function during cell reproduction. The cell cycle is divided into a long interphase and a relatively short mitotic phase. Interphase is the time that cells prepare themselves for the process of cell division. It is further divided into phases of G1 (1<sup>st</sup> gap phase), S (synthetic phase) and G2 (2<sup>nd</sup> gap phase). The cells grow in the G1 and G2 phases by producing proteins and cytoplasmic organelles. In S phase, cells replicate DNA and duplicate the genome. The duplicated genome is split during mitosis and separated into two identical sets. The division of the entire cell (cytokinesis) follows mitosis, and mitosis and cytokinesis together comprise the mitotic phase. Based on the order of mitotic progression, mitosis is further divided into prophase, prometaphase, metaphase, anaphase, and telophase prior to cytokinesis.

To ensure the appropriate complement of genetic material and equally distributed cellular components between daughter cells, cells use “checkpoint” control mechanisms to guarantee cell cycle events to occur in a precise order.

#### An overview of ubiquitin cascade

The ubiquitin cascade is an event prior to the major pathway for misfolded or unwanted protein degradation. Ubiquitin is a small protein (~8.6kDa) and it is normally covalently attached to proteins through its C-terminal glycine residues linked to the lysine residues of proteins via an isopeptide bond. This process is known as ubiquitination. Following the first ubiquitin, a second ubiquitin molecule is added to the first one by linking its last C-terminal glycine residue to a lysine of the first ubiquitin. The repeating of this step eventually yields a polyubiquitin chain. Ubiquitin has seven lysine residues that may serve as ubiquitination points, the K48-linked and K63-linked polyubiquitin chains being best characterized. The K48-linked polyubiquitin chain mediates proteasome-dependent degradation while the K63-linked chains are associated with cellular signaling (1-3).

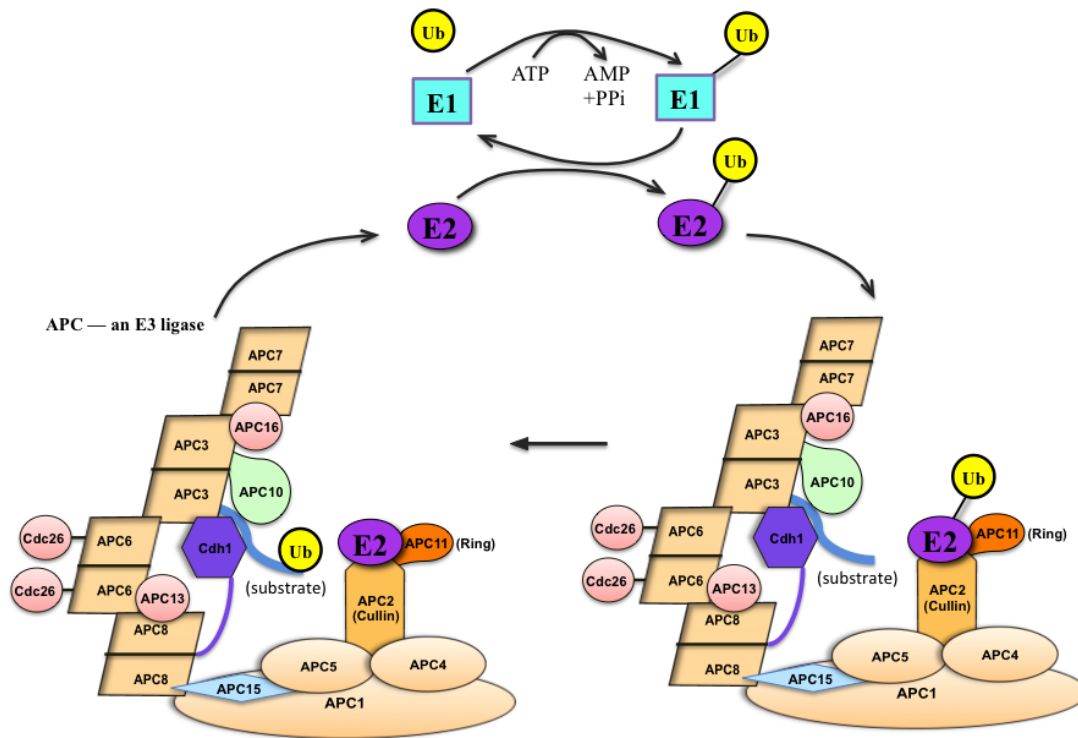


Three types of enzymes, E1s, E2s and E3s, mediate the ubiquitination cascade. They are also known as ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, and ubiquitin ligases, respectively. At the start of the ubiquitination cascade, ATP activates ubiquitin by forming a C-terminal acyl adenylated ubiquitin. The cascade continues with an E1 catalytic cysteine forming a thioester bond with ubiquitin C-terminal carboxyl group. Catalyzed by E2, the ubiquitin transfers from E1 to an active E2 cysteine through a trans-esterification reaction. At the final step, E3 specifically recognizes substrates and catalyzes ubiquitin transfer to N $\epsilon$ -amino group of substrate or ubiquitin lysines.

In the ubiquitination cascade pathway, E1 can bind with many E2s, which can bind with hundreds of E3s in a hierarchical way. Correspondingly, there are two E1s, tens of E2s and hundreds of E3s in the human proteome to mediate the ubiquitination targeting thousands of substrates. E3 ligases are generally divided into two categories: HECT (Homologous to E6-AP Carboxyl Terminus) E3s and RING (Really Interesting New Gene) E3s. HECT E3s form a covalent E3-ubiquitin intermediate (via a thioester bond with the ubiquitin C-terminus) to mediate ubiquitin transfer (4), whereas RING E3s directly transfer ubiquitin from E2 enzymes to substrate without a covalent intermediate. The RING domain does not directly participate in the reaction (5). RING E3s represent the largest structural class of ligases and APC (Anaphase-Promoting Complex) belongs to the RING E3 ligase family (**Figure 1-1**).

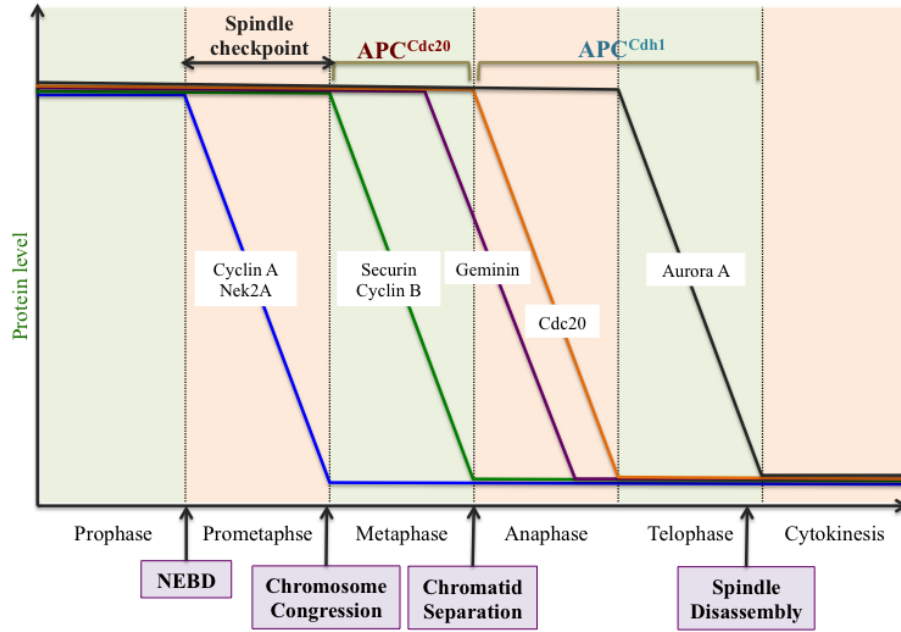
### **The APC directs cell division via substrate destruction**

The APC orchestrates mitosis and G1 by sequentially promoting degradation of key cell-cycle regulators. APC is found in its active form at the onset of mitosis (**Figure 1-2**). Some well-studied APC substrates, notably mitotic Cyclins, Securin, Geminin, AuroraA and Hsl1, could be divided as early substrates and late substrates based on their degradation timing. (i) Cyclin A and Nek2A are early substrates and are recognized by APC at prometaphase (6, 7). Cyclin A forms complexes with Cdk1 and Cdk2 that support S-phase and G2-phase progression. Nek2A is a mammalian kinase required to ensure the correct formation of mitotic spindles at mitosis onset. Decreasing the cellular concentrations of Cyclin A and Nek2A allows mitosis to enter into metaphase. (ii) To promote cell cycle transition from metaphase to anaphase (when sister chromatids separate), APC assembles polyubiquitin chains to degrade Securin. The degradation of Securin releases Separase, a protease, which advances sister chromatid separation by cleaving Cohesin (8, 9). (iii) Another substrate that is degraded rapidly during metaphase-anaphase transition is Geminin. Geminin inhibits DNA replication during S, G2, and M phases by preventing the incorporation of the MCM complex (mini-chromosome maintenance complex). Destruction of Geminin permits replication in the succeeding cell cycle (10). (iv) During late mitosis, APC also targets Cyclin B (and other mitotic Cyclins) for degradation, which then deactivates Cdk1. The decreased activity of Cdk1 plays a critical role in mitotic spindle disassembly and chromosome decondensation, which in turn promotes cell exiting from mitosis and entry into cytokinesis (11); (v) Aurora A is another late substrate which is quickly degraded after mitosis. Aurora A is essential for



**Figure 1-1. APC functions as an E3 ligase in ubiquitination cascade.**

A ubiquitin (yellow) is initially activated by an E1 (activating enzyme, cyan) and is driven by hydrolysis of an ATP molecule. The ubiquitin then forms a thioester linkage with the catalytic cysteine of the E1. Next, the E1 recruits an E2 (magenta) and transfers the ubiquitin to the E2 catalytic cysteine. Finally, with the aid of the E3, APC (light orange), the ubiquitin is transferred to a lysine on the target substrate.



**Figure 1-2. The order of APC substrate destruction.**

The activity of APC is first inhibited by a protein complex called the spindle assembly checkpoint (SAC), which interacts with Cdc20 and thereby blocks substrate recruitment by APC. Cyclin A and the kinase Nek2A are two of prometaphase targets and they are ubiquitinated by APC without being recognized by Cdc20. SAC is inactivated when all the sister-chromatid pairs are attached to the spindle, allowing an activated active APC to target metaphase substrates, like Securin and Cyclin B. Cyclin destruction leads to cyclin-dependent kinase (Cdk) inactivation, which results in the dephosphorylation and activation of the second APC activator, Cdh1. APC<sup>Cdh1</sup> triggers the destruction of various targets at late mitosis (from anaphase to telophase), like Geminin, Cdc20 and Aurora A. These substrates are degraded at different times, presumably providing mechanisms that order late mitotic events. Key mitotic events are labeled out in text boxes along with the cell-cycle time course.

proper centrosome separation once the mitotic spindle is formed (12). (vi) Hsl1 is a substrate of the budding yeast APC (13), which accumulates after G1 phase as cells begin to bud but disappears in late mitosis. Hsl1 promotes the degradation of yeast cytoskeletal proteins and allows cells to proceed to mitosis (13).

The APC adopts a similar two-step mechanism to other ubiquitin ligases to assemble a polyubiquitin chain. The E2 UbcH10 functions as a “priming” E2, whereby the APC specifically mono-ubiquitinates substrates (14). The second E2 Ube2S elongates the ubiquitin chain. The APC triggers substrate degradation by assembling K11-linked ubiquitin chains, and the specificity of this chain formation depends on a surface of ubiquitin (15). APC is the largest of the RING E3 family, which are characterized by direct ubiquitin transfer from E2 to substrates without forming a covalent intermediate (16, 17). APC positions both a substrate and a ubiquitin conjugated UbcH10 in close proximity to facilitate ubiquitin transfer to substrate lysines.

## **APC Architecture and Subunit Organization**

### **APC subunit organization**

The human APC consists of 14 subunits and has a molecular weight around 1.2MDa (**Table 1-1**). Based on biochemical and genetic data, the complex is considered to have two large domains, called the “arc lamp” and the “platform”, which together enclose a central cavity (**Figure 1-3**). The arc lamp is comprised of tetratricopeptide repeat (TPR) proteins APC7, APC3, APC6, and APC8 with small subunits APC16, Cdc26 and APC13. The arc lamp is therefore also called TPR arm. The platform consists of APC1, APC4, APC5, and APC15 (18). The catalytic core of APC is composed of subunits APC2 and APC11, which are analogous to the Cullin and Rbx1 subunits of Cullin-RING ligases in the SCF superfamily (19). The Cullin domain of APC2 interacts with the RING-finger APC11, which potentially mediates interactions with the ubiquitin conjugated E2s (UbcH10 and Ube2S). APC2 and APC11 are neighbored by APC1 inside the complex, whereas TPR subunits co-localize to a more distal region to APC1.

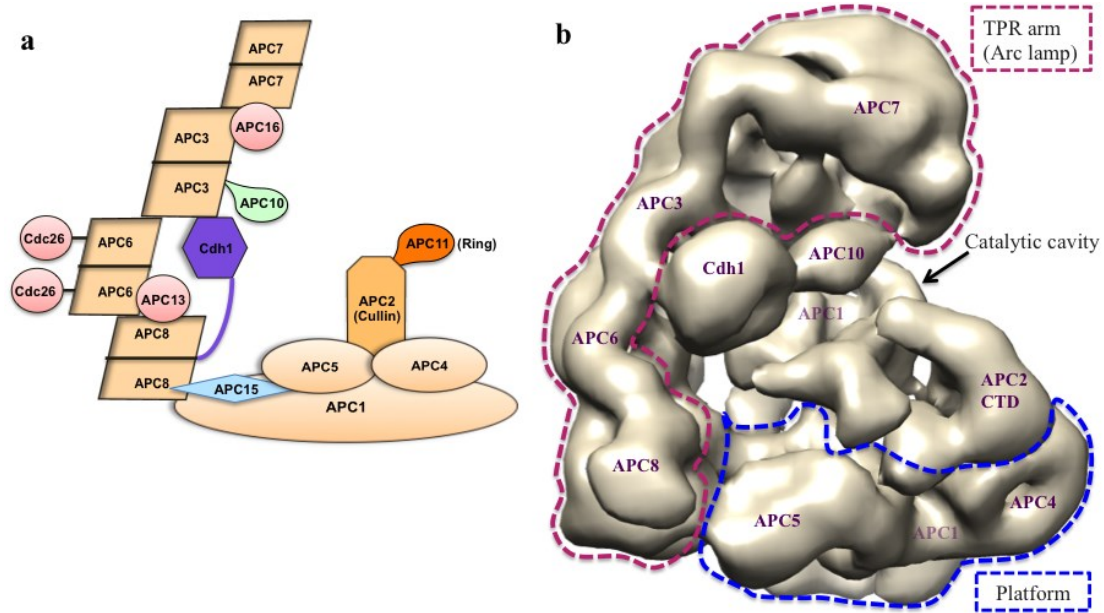
With respect to human APC, APC7 is the most peripheral TPR subunit of the “arc lamp”, and APC8 is the most internal subunit. APC6 stacks between APC3 and APC8, and APC3 stacks between APC7 and APC6. Yeast APC lacks the APC7 subunit, and therefore yeast APC3 (Cdc27) is the most peripheral TPR protein. Each TPR subunit forms a homo-dimer through their N-terminal domains, and four homo-dimers spirally stack together and constitute a TPR arm (20). The TPR arm has the flexibility to support APC switching overall conformation between open and closed status (21).

The small subunit APC16 was recently discovered and characterized (22) to be important for maintaining APC activity towards mitotic substrates. APC16 associates with both APC3 and APC7, and it may function to stabilize interactions between them. Another small subunit, APC13, interacts with APC8 and is reported to have a function of

**Table 1-1. The APC core and regulatory subunits.**

<i>H. sapiens</i>	<i>S. cerevesiae</i>	<i>S. pombe</i>	Molecular mass (kDa)	Stoichiometry	Structure motif	Function
APC1	APC1	Cut4	216.5	1	PC repeats	Scaffolding subunit
APC2	APC2	APC2	93.8	1	Cullin homology	Catalytic subunit
APC3	Cdc27	Nuc2	92.6	2	TPR	Scaffolding subunit, recruit substrate co-receptors
APC4	APC4	Lid1	92.04	1	$\beta$ -propeller, extended TPR	Scaffolding subunit
APC5	APC5	APC5	85.1	1	Extended TPR	Scaffolding subunit
APC6	Cdc16	Cut9	71.7	2	TPR	Scaffolding subunit
APC7	—	—	63.2	2	TPR	Scaffolding subunit
APC8	Cdc23	Cut23	68.3	2	TPR	Scaffolding subunit
APC10	Doc1	APC10	21.2	1	$\beta$ -barrel, IR tail peptide	Substrate co-receptor
APC11	APC11	APC11	9.84	1	RING-H2 finger	Catalytic subunit
APC13	Swm1	APC13	8.5	1	A few $\alpha$ -helix	Stabilize APC8
APC15	—	APC15	14.28	1	A few $\alpha$ -helix	Unknown
APC16	—	—	11.67	1	C-terminal long $\alpha$ -helix	Stabilize APC7 and APC3
Cdc26	Cdc26	Hcn1	9.78	2	Unstructured N-terminal domain, C-terminal $\alpha$ -helix	Stabilize APC6
Cdh1	Cdh1		55.18	1	C-box, WD40 repeat, IR tail peptide	Activator, substrate co-receptor
Cdc20	Cdc20		54.72	1	C-box, WD40 repeat, IR tail peptide	Activator, substrate co-receptor

Molecular mass corresponds to *H. sapiens* subunits.



**Figure 1-3. The schematic of the APC subunit topology and the APC architecture.**

(a) Schematic representation of human APC showing approximate subunit topology based on biochemical and genetic data cited in the text.

(b) 3D model of human APC. The ‘arc lamp’ and the ‘platform’ domains enclose a catalytic cavity. Each subunit is assigned to its corresponding density. The EM map, APC<sup>Cdh1-Emi1</sup> is reprinted from the open source EM databank (EMD 2354) (23, 24).

stabilizing association of APC3 and APC6 (25). Cdc26 was proven to be essential for proper folding of APC6 (26, 27). APC6 could not be expressed in vitro without Cdc26. In the budding yeast, the deletion of Cdc26 resulted in reduced levels of APC6 and APC3 incorporation into the APC (28).

The largest subunit in the platform of APC is APC1, which has a molecular weight more than 200kDa. APC1 C-terminal domain has nine pentatricopeptide repeats (PPR), which fold into the helix-turn-helix structure similar to TPR motifs. APC4 associates with APC5 (29) and subunits APC1, APC4, APC5 interdependently associate with each other to form the platform. APC8 connects the TPR arm with the platform.

### **APC recruits co-activators through TPR subunits**

Without the co-activators Cdh1 and Cdc20, APC is an inactive E3 ubiquitin ligase. Cdh1 and Cdc20 are not the constitutive subunits of APC and they are recruited to the complex alternatively during cell cycle phases. By switching between Cdh1 and Cdc20, APC specifically targets various cell-cycle regulators during cycle procession.

TPR subunits APC3 and APC8 contribute to the recruitment of Cdh1 and Cdc20. TPR structure motif is a protein-protein interaction module that consists of two antiparallel  $\alpha$ -helices. Most TPR proteins are characterized by continuous  $\alpha$ -helices as their secondary structures (30). Both APC3 and APC8 are predicted to have 14 TPR motifs and the N-terminal 6 TPR motifs mediates homo-dimerization.

APC3 is required for the Cdc20 and Cdh1 association with APC. Both yeast and human APC EM structures reveal that Cdh1 localizes adjacently to the C-terminal domain of a molecule within the APC3 homo-dimer (31, 32). Cdh1 could be specifically cross-linked to APC3 through multiple crosslinking compounds (33). In the APC3-deleted yeast APC, Cdh1 and Cdc20 binding to the complex are dramatically decreased compared to the wild type APC (18, 33). These APCs lost the ability to assemble long ubiquitin chains onto substrates (the ubiquitin number of most chains is less than three), which contributes to the mutated complex having a 100-fold lower activity than wild type APC (33, 34).

APC3 C-terminal TPR domains were predicted to recruit co-activators, Cdh1 and Cdc20 through their C-terminal Ile-Arg (IR) tail motif (35). The double mutation of two conserved residues (N548A and L579A) in yeast APC3 raised the Cdh1 dissociation rate from APC without affecting other subunit incorporation. The APC3 mutant (N548A/L579A) reduces the APC ubiquitination activity more than 5-fold once it is incorporated into APC (34). The conserved residues N548 and L579 sit on the  $\alpha$ -helices of TPR8 and 9 of yeast APC3, and therefore TPR8 and 9 are believed to form a binding groove responsible for co-activator recruitment.

## Co-activators interact with the APC through multiple interactions

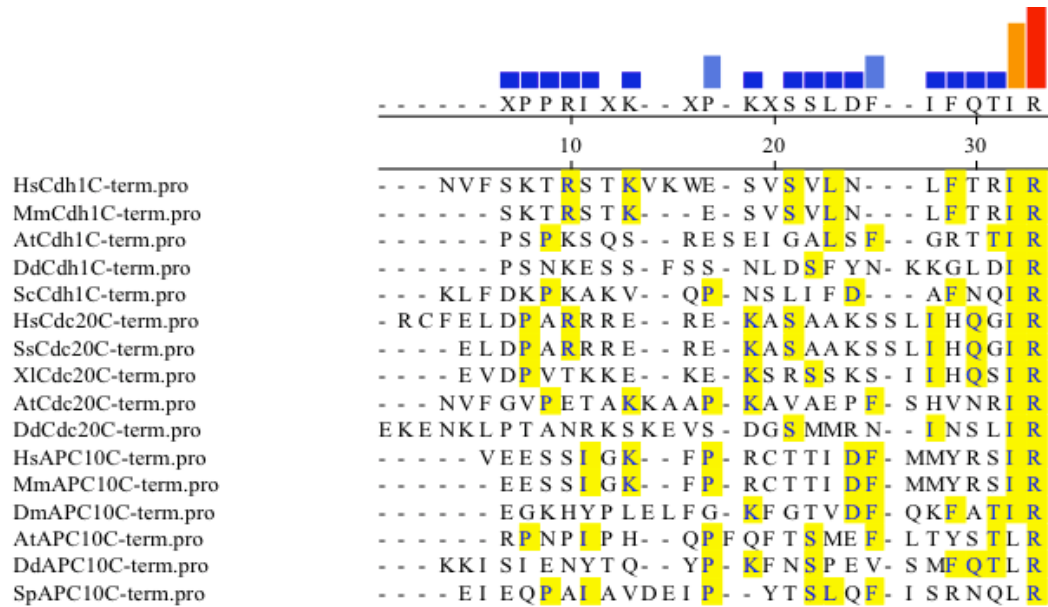
Cdh1 and Cdc20 are recruited to APC3 through its C-terminal Ile-Arg (IR) tail motif. The last two Isoleucine and Arginine residues of Cdh1, Cdc20 and APC10 are highly conserved across species (**Figure 1-4**). APC10 is also found to associate with APC3 through its C-terminal IR motif (34, 36). Multiple research groups have found that the IR-tail deleted/mutated Cdh1 or Cdc20 have APC-binding defects and contribute to APC-activation defects (18, 34, 37). The Arginine to Lysine mutation of the IR-tail disrupted the interaction between Cdh1 and APC (38). IR-tail peptides of Cdc20 and Cdh1 could bind to APC3 and the peptides inhibit *in vitro* APC ubiquitination by competing with co-activators recruitment to APC (37). The IR-tail peptide of Cdh1 is essential for budding yeast viability whereas C-terminal IR-tail peptide of Cdc20 is not (18, 33). This may reflect differences in affinities of multiple Cdh1 and Cdc20 binding sites on APC, although future studies will be required to know how much each binding site contributes.

Cdh1 and Cdc20 also have an eight-residue C-box motif near the N-terminus. The C-box enhances the ubiquitination activity of APC though the mechanism is not well understood. The C-box containing Cdc20 fragment (N-terminal fragment) enabled the ubiquitination of Nek2A to be mediated by the substrate interaction domain, WD40 domain (a domain comprised of tandem copies of WD40 repeat) (39). The C-box deleted Cdh1 caused a large reduction of APC ubiquitination activity. The co-activator C-box might promote a change in APC conformation that is independent of the activators' other domains.

A xenopus Cdc20 C-box containing protein fragment was discovered to interact with APC through subunit APC3, and the affinity is weaker than the IR motif-APC3 interaction (40). Multiple residues of Cdh1 C-box, WD40 domain and IR-tail could be cross-linked to APC3, and the residues of the IR-tail demonstrated much stronger cross-linking interaction than others. However, all the cross-linked interactions disappeared once the C-terminal IR peptide is deleted (33). The IR-tail deletion increases Cdh1 partial dissociation from APC, and double deletion of C-box and IR motif further increases Cdh1 dissociation (34, 37). Cdh1 and Cdc20 seem to have multiple sites mediating low-affinity interaction with APC and their IR motifs are responsible for anchoring the co-activators into the correct binding groove of APC3.

The mutation of yeast APC8 conserved residue (N405A) also affected Cdh1 recruitment. APC3 (N548A and L579A), APC8 (N405A) double-mutant APC caused cell cycle arrest in metaphase with high levels of Cyclins (34). This APC8 mutant also demonstrates a higher Cdh1 dissociation rate. The mutation of Cdh1 C-box greatly reduced the activity of APC8 mutant incorporated APC. Therefore, the residue Asn405 (N405) in APC8 is likely to interact with a third, unknown site on Cdh1.





**Figure 1-4. Sequence alignment of C-terminal APC10, Cdh1 and Cdc20.**

Sequence alignment of the C-terminal Cdh1, Cdc20 and APC10 from human (Hs), Mouse (Mm), *Drosophila melanogaster* (Dm), *Arabidopsis thaliana* (At), *Dictyostelium discoideum* (Dd), *Saccharomyces cerevisiae* (Sc), *Schizosaccharomyces pombe* (Sp). The last two residues, Isoleucine and Arginine, are highly conserved across different species.

## Substrate recruitment to APC through the co-activators and APC10

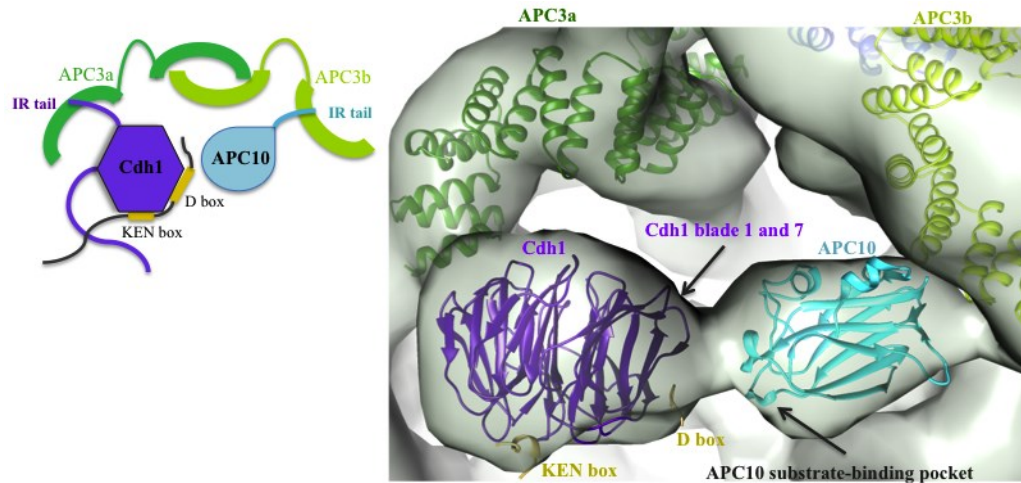
APC has weak interaction with substrates without its co-activators (38, 41). The role of the co-activators is to act as a substrate recognition subunit and recruit substrates, which is analogous to the F-box protein receptors of the SCF. Substrates of the APC contain conserved APC-targeting sequence elements: the KEN and destruction box (D-box). D-box and KEN box motifs are characterized with peptide sequences RxxLxxI/VxN and KENxxxN/D, respectively.

Cdh1 and Cdc20 each contain a WD40 domain which folds into a seven-blade  $\beta$ -propeller structure. This domain functions as a destruction box (D-box) receptor to bind substrates. The fact that WD40 domains of Cdc20 and Cdh1 could be cross-linked to substrates in a D-box-dependent manner confirms a role for co-activators in recruiting of substrates to the APC core (33, 42). Moreover, residue mutations of the evolutionarily conserved surface of WD40 domains abolish crosslinking interactions with substrates (33).

Human and yeast APC EM structures explicitly revealed the interaction between Cdh1 and substrates. The apo APC EM map demonstrated that Cdh1 and APC10 have separated density in close proximity. APC10 is an APC subunit localizing at the opposite side of the catalytic center from Cdh1. When the substrate Hsl1 is bound, the  $\beta$ -propeller domain (WD40 domain) of Cdh1 shifted approximately 7Å towards APC10. The repositioning of Cdh1 created a well-defined density to bridge the gap between them (**Figure 1-5**). The KEN box of Hsl1 alone also promoted repositioning of Cdh1 towards APC10, but there wasn't connecting density between KEN box and APC10. It seemed that only the D-box of substrates contributed to the physical connection between Cdh1 and APC10. Cdh1, Hsl1 and APC10 together form a three-layer structure (31, 32).

Arginine and Leucine residues appear as two invariant residues inside the D-box motif (RxxLxxI/VxN), and the C-terminal Asparagine residue is less conserved. Mutations of Arginine and/or Leucine eliminate the recruitment of this motif to the co-receptors. The D-box of the substrates directly interacts with the WD40 propeller structure of Cdh1, and this interaction is essential for ubiquitination processivity (33). The mutated D-box and KEN box also abolished substrate recognition and recruitment to the APC (34, 38, 41).

The use of two distinct motifs for substrate recognition introduces a greater degree of substrate specificity. APC<sup>Cdh1</sup> is also able to ubiquitinate substrates containing only a KEN box motif, such as Cdc20 and Sororin. Contrastingly, APC<sup>Cdc20</sup> recognition of substrates requires the D-box motif. The conformation of the KEN and D-box are different when they are bound to Cdc20. It has recently been reported that the KEN box folds into an underwound helix, whereas the D-box is potentially an extended structure (**Figure 1-5**) (43, 44). The optimal spatial arrangement of KEN and D-box is not well understood. But the Cdh1 inhibitor Acm1 and Mes1 have the spacing of 18 and 24 residues, respectively, between the KEN and D-box motifs, which optimize their affinity to the co-activators (45).



**Figure 1-5. Detail and schematic demonstration of Cdh1 and APC10 as APC substrate co-receptors.**

Left: Schematic diagram of the substrate-recognition module responsible for the D-box and KEN box binding, and Cdh1 and APC10 recruitment to the APC3 homo-dimer. D-box is represented as a box in the interface between Cdh1 and APC10. Right: Details of the Cdh1 and APC10 functioning as co-receptors to the D-box. Hsl1's D-box is recruited to Cdh1 between blade 1 and 7, while KEN box is recruited to the surface at the center of the bottom side of the WD40 domain. APC10 faces toward D-box with its substrate-binding pocket. APC<sup>Cdh1-Emi1</sup> EM map (EMD 2354) (23) is used for the structural demonstration.

It is also not well understood whether there is an optimal distance between a target Lysine (the ubiquitin acceptor residue) and the KEN or D-box motifs. Randall King's group reported that 15 residues between the D-box and the target Lysine is sufficient to support the degradation of the D-box fused proteinA (46). A C-terminal Lysine immediately following the D-box peptides (of CyclinB1 and Securin) functioned as a ubiquitin acceptor (33). Lysines preceding the D-box in Cyclin B from *S. pombe* also accepted ubiquitins and mutating them made a D-box containing peptide non-ubiquitinable (47).

### **The contribution of APC10 to APC processivity**

APC10 is a constituent subunit of the APC and binds to APC3 through its C-terminal IR motif. APC10 acts as another co-receptor for the APC substrates (31, 32, 48, 49). APC10 adopts a  $\beta$ -barrel (jellyroll) fold with a binding pocket structure, both structurally similar to galactose oxidase (36, 50). APC10 is necessary for optimal Cdh1/Cdc20-dependent substrate recognition, through which APC10 contributes to substrate processivity (38). The mutation of conserved residues in the APC10 binding pocket affected substrate processivity, and the deletion of the entire subunit impaired polyubiquitin assembly in a D-box-dependent manner (48, 49). The docking of the APC10 crystallographic model into APC EM maps displays that a conserved loop of APC10 faces toward the substrate D-box (**Figure 1-5**). However, the interaction affinity between APC10 and substrates is pretty low. Da Fonseca and Kong et al demonstrated the weak interaction of APC10 to the substrate D-box through HSQC-type NMR experiments, whereas the substrate KEN box didn't appear to interact (32).

There is a bipartite model to interpret the contribution to the substrate recruitment by APC10. In this model, the D-box of substrates may form a divalent connection to bridge a co-activator and APC10. APC10 potentially stabilizes the association of co-activator-substrate to APC through this bipartite interaction (divalent bridge), which probably results in a higher ubiquitination processivity. The mutant APC lacking the APC10 subunit displays a 50-fold reduction in the Cyclin recruitment and the Cyclin B peptides interacted more efficiently with Cdc20 when it is bound to the holoAPC (33, 49). Hsl1 enhanced the Cdh1 association to the APC in vivo and Cyclin B enhanced this association in vitro (34). Individually, both the co-activators and APC have low affinity and specificity for substrates (41). APC10 probably stabilizes Cdh1 binding to APC through the D-box of substrates.

APC10 promotes APC processivity in an E2 independent manner. APC $\Delta$ APC10 produces less ubiquitinated Cyclins due to the poorly bounded substrates to the complex (49). But the interactions between the APC and E2s are not compromised in the absence of APC10, and E2s show similar effects to promote the activity of both wild type APC and the APC $\Delta$ APC10.

## TAME is an IR-tail mimic APC inhibitor

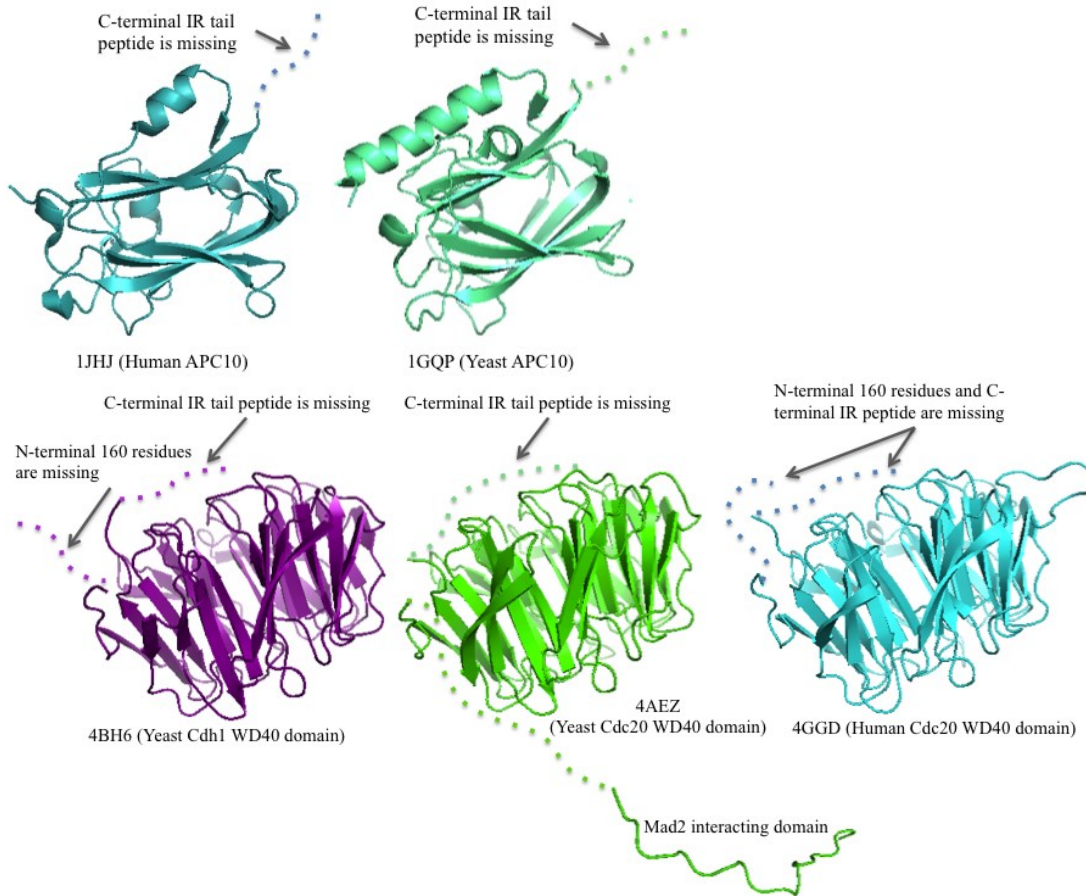
TAME (tosyl-L-arginine methyl ester), is a small molecule, which structurally resembles the IR motifs of the APC co-activators and APC10. The benzyl ring of the tosyl group mimics the hydrophobic interaction contributed from Isoleucine of the IR motifs. TAME is reported to inhibit the APC activation by preventing Cdc20 and Cdh1 binding. In an assay TAME almost completely disrupted Cdc20 association to the APC and prevented polyubiquitin chains from being assembled onto Cyclin B (51). TAME specifically antagonizes the IR motif-dependent interactions without inhibiting other interactions between the co-activators and APC. The compound disrupted the IR-tail peptides which were cross-linked to APC but didn't affect the recruitment of the IR motif deleted co-activators or the C-box only fragments of the co-activators (40). Similar to the IR-tail peptides reduction of the APC Kcat, TAME slows the Kcat of the APC<sup>Cdc20</sup> by a ~55%, while the amount of the Cdc20 associated to the APC is not reduced (51). Whether or not TAME affects APC10 is not well understood.

TAME induces Cdc20 dissociation from the APC by promoting Cdc20 ubiquitination. The ubiquitination of Cdc20 is upstream of the C-box, reducing its binding affinity for the APC. Cyclin B promotes the Cdc20 binding to the APC and suppressed Cdc20 ubiquitination (40, 51). A cell-permeable prodrug (proTAME) is capable of inhibiting APC-Cdh1 activation during S phase, and induces mitotic arrest in a SAC dependent manner. HeLa cells treated with 12  $\mu$ M proTAME were arrested in metaphase without mitotic spindle disruption. ProTAME also greatly increased mitotic duration in asynchronous hTERT-RPE1 cells (51).

## Crystallography study of APC subunits

It is still mysterious why the APC is such a large complex, comprised of a large number of subunits. Structural studies of the APC subunits have yielded crystallographic structures of numerous TPR subunits and substrate co-receptors: APC6-Cdc26 complex from human and yeast (PDB code: 3HYM, 2XPI), APC3 N-terminal dimer domain from *E. cuniculi* (PDB code: 3KAE), and APC8 N-terminal domain from *S. pombe*. The co-receptor APC10 structures were determined from both the yeast and human proteins, but the C-terminal peptides and IR motifs are missing in these atomic models (**Figure 1-6**). The WD40 domain of yeast Cdh1 was co-determined with Cdh1 inhibitor Acm1 (PDB code: 4BH6). Cdc20 of both human and yeast were determined together with the MCC complex subunits (PDB code: 4AEZ, 4GGD).

From these studies, APC3, APC6 and APC8 were found to form homo-dimers, with an interlocking dimer interface mediated by the self-association of their N-terminal TPR motifs. The overall structure of APC6 displays a TPR superhelix created by seven successive TPR motifs. The superhelix features each APC6 subunit/monomer in a rod-like structure of a continuous 14 TPR motifs (27). The N-terminal 13 residues of Cdc26 extend into the inner concave groove of the APC6 superhelix (from the C-terminal APC6) and stabilize the superhelix conformation through intimate interactions



**Figure 1-6. Crystal structures of APC10, Cdh1, and Cdc20.**

Cartoon views of the overall structures of all the available atomic models from APC10, Cdh1 and Cdc20. All the atomic models have C-terminal around 20 residues absent. Human and yeast APC10 both adopt  $\beta$ -barrel structure except yeast APC10 has a longer N-terminal  $\alpha$ -helix. The models of Cdh1 and Cdc20 miss the N-terminal  $\sim$ 160 residues, but Mad2 interaction domain of yeast Cdc20 is visible in structure.

(**Figure 1-7**). The rest of Cdc26 forms an  $\alpha$ -helix that packs against the APC6 C-terminal TPR  $\alpha$ -helix (26, 27).

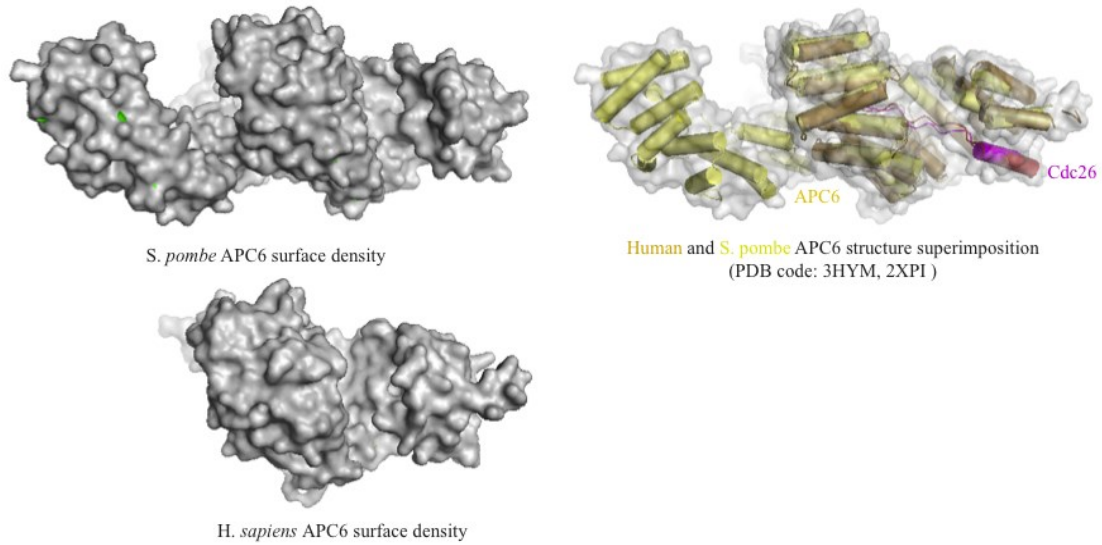
The C-terminal 9 TPR motifs of *E. cuniculi* APC3 are absent from its crystallized N-terminal dimer. The C-terminal APC3 was predicted to form a continuous TPR superhelix like APC6. Among all of the C-terminal TPRs, TPR8 and TPR9 are predicted to be the motifs recruiting the co-activators and APC10 through their C-terminal IR peptides. Matyskiela et al predicted the conserved Asparagine of TPR8 and Leucine of TPR9 both engaged in the interactions with the Cdh1 IR motif (34).

The structure of Cdh1-Acm1 demonstrates the D-box of the substrates and inhibitors binding to an APC co-activator. The D-box of Acm1 binds to an inter-blade groove on Cdh1 (between blade 1 and 7) and the KEN box binds to a conserved surface at the center of the topside of WD40 domain. The KEN box recognition sites on Cdh1 are likely to be universal through all species. Acm1 uses a motif A and an  $\alpha$ -helix, to specifically target Cdh1 but not Cdc20. The A motif aligned parallel with  $\beta$ -sheet blade 3 and interacts with another inter-blade channel (between blade 3 and 4), which is less well conserved than the D-box and KEN box interaction regions. Acm1-Cdh1 forms a heterodimer to inhibit the APC activation.

The structure of Cdc20 is determined within the MCC complex (mitotic checkpoint complex, PDB code: 4AEZ, 4GGD). Human Cdc20 is comprised of a WD40 domain forming a  $\beta$  propeller with a preformed KEN box binding site at its top surface. Both human and yeast Cdc20 use a deep hydrophobic pocket at the surface of blades 1 and 7, to interact with the D-box Leucine residue (RxxLxxI/VxN) (43, 44). The pocket is created from the conserved non-polar residues on the surface. The mutants of D-box interacting residues diminish the ubiquitination ability whereas the mutants of KEN box-binding site lack such deficiency. Blocking the Leucine-binding pocket with small molecules may be a potential mechanism to inhibit APC<sup>co-activator</sup> activity (43, 44).

The affinity of Cdc20 for the substrate D-boxes is weaker than Cdh1. Unlike Cdh1, Cdc20 residues could not crosslink to substrates the way Cdh1 could, although they localize at the equivalent sites of the WD40 domain (33, 43). KEN box contacting sites are also dispensable for APC<sup>Cdc20</sup> ubiquitination of Securin, where the same sites on Cdh1 are essential for APC<sup>Cdh1</sup> activity (43).

An alternative structural approach to study APC assembly is through Electron Microscopy. Yeast APC EM structures were determined using cryo-electron microscopy, whereas the human and Xenopus structures were determined by the approaches of both cryo- and cryo-negative stain electron microscopy. The effective resolution is of the order of 10-20Å.



**Figure 1-7. APC6 adopts a superhelix structure.**

Surface density of *S. pombe* and human APC6 structure (left) demonstrate that APC6 continuous TPR motifs form the superhelix tertiary structures. Two structures were superimposed onto each other with an r.m.s.d of 1.77Å (right), and N-terminal Cdc26 stabilizes APC6 tertiary structure through interactions with the superhelix inner groove.



## APC Activity Regulation

There are multiple complexes involved in cell-cycle control system to coordinate cell cycle progression. The Emi1 (early mitotic inhibitor 1) protein and mitotic checkpoint complex (MCC) inhibit the APC activity by mimicking substrate recognition motifs; the D-box and/or KEN box.

### Emi1 inhibition

Emi1 is an inhibitor of the APC and it was first discovered due to its role in preventing premature APC activation in early mitosis (52). The APC inhibition by Emi1 is accomplished through a combination of structure motifs: a D-box, a conserved zinc-binding region (ZBR), a linker linking D-box to ZBR and a highly conserved C-terminal LRRL tail (Leu-Arg-Arg-Leu) (53). The D-box of Emi1 allows the protein to mimic the substrates interaction with the APC and Emi1 is recruited between Cdh1 and APC10 (23). Comparing with an isolated D-box peptide from the yeast APC substrate Hsl1, the isolated Emi1 D-box is relatively weak, although Emi1 competes with the D-box-dependent substrates to be recruited to the APC (23, 54).

The ZBR of Emi1 provides an APC E3 ligase antagonizing activity (52, 54). Interestingly, the mutation of the Emi1 ZBR converts Emi1 to a D-box-dependent APC substrate (54). The linker between the D-box and ZBR contributes to inhibition with its specific side chains. The Emi1 C-terminal tail is a specific inhibitor to the Ube2S-dependent ubiquitin-chain formation, and ZBR and C-terminal tail synergize to block chain elongation. Much of the Emi1 is predicted to be intrinsically disordered, and this disorder has implications for an ability to span over a broad distance of the APC catalytic center, although this inhibitor has a relatively small size. The Emi1 D-box, linker, ZBR and tail synergistically block both the substrate binding to the APC and ubiquitin chain elongation by the APC and Ube2S (23).

### MCC inhibition

The MCC (Mitotic Checkpoint Complex) is a multi-subunit protein complex that is activated by the spindle assembly checkpoint (SAC). The MCC includes the co-activator Cdc20, which assembles with the other MCC proteins and loses its APC activation function. The blocked Cdc20 leads to an inhibited degradation of Securin and Cyclin B (substrates of APC<sup>Cdc20</sup>) until late mitosis. This regulatory network of the APC activity is contributed from three complexes to ensure that chromatids are attached to mitotic spindles and well aligned before onset of the anaphase. However, Nek2A and Cyclin A still escape MCC-mediated inhibition (55), being targeted instead at prometaphase for ubiquitin-mediated proteolysis.

A couple of EM studies have demonstrated that the MCC localizes to the front end of the platform domain (21), and Cdc20 is partially overlapped with MCC. MCC sits

in the lower region of the APC cavity. Mad2 contacts the TPR subunits APC8 and APC5, whereas Mad3 interacts with APC1 (44). Checkpoint inhibition of APC<sup>Cdc20</sup> requires the binding of a Mad3 (also called BubR1) KEN box to Cdc20. Due to this mechanism, Cdh1 is required for the exit of mitosis.

The MCC can be assembled *in vitro* from recombinant Mad2, Mad3 (BubR1), and Cdc20 proteins in the absence of kinetochores. Mad2 and Mad3 (BubR1) bind directly to Cdc20 and together they inhibit the APC<sup>Cdc20</sup> activity towards ubiquitinating Securin and Cyclin B. The conserved N-terminus of BubR1 that is essential for MCC assembly incorporates a KEN box, and the BubR1 N-terminus adopts a helix–loop–helix (HLH) structural motif. The HLH motif simultaneously binds Mad2 and Cdc20, orienting the KEN box towards its binding residues on the receptor Cdc20 (44). Binding of the N-terminal Mad3 to Cdc20 is KEN box-dependent (43), and the same KEN box also promotes Mad3-dependent APC-mediated Cdc20 degradation (56), which suggests that Mad3 plays a role in positioning Cdc20 for ubiquitination mediated by the APC catalytic subunits (44). The EM structures of the APC-MCC complex demonstrate that the D-box receptor residues are directed towards, but not in contact with, the substrate co-receptor APC10 (21, 44). Comparing with the Cdh1 position on APC<sup>Cdh1</sup>, Cdc20 is displaced downwards towards APC5, and it is lower than Cdh1. This lower position may facilitate Cdc20 ubiquitination. Furthermore, the lower position of Cdc20 prevents its D-box-binding site from generating a bipartite D-box co-receptor with APC10 (44).

### **Substrate specificity controlled through co-activators**

The APC-mediated coordination of cell-cycle progression is achieved through the temporal regulation of APC activity. Cdc20 activates the APC during early mitosis when Cdh1 is inhibited by its N-terminal phosphorylation. The N-terminus of Cdc20 alone, without the substrate-binding WD40 domain, is able to activate the APC catalytic activity. The APC<sup>Cdc20</sup> reduces mitotic Cdk activity through the degradation of Cyclins which are required for Cdk activity, which in turn stimulates the APC<sup>Cdh1</sup> activity. This then leads to the APC<sup>Cdh1</sup> ubiquitinating Cdc20 and deactivating the APC<sup>Cdc20</sup>. The APC<sup>Cdc20</sup> and APC<sup>Cdh1</sup> have different substrate specificities. Dynamic changes in the co-activator positions relative to core APC subunits may provide the combinations of substrate-recognition interfaces.

In addition to the D-box and KEN box binding sites, the APC<sup>co-activators</sup> utilize additional recognition sites to bind substrates and regulatory proteins. The C-terminal Met-Arg (MR) residues of Nek2A are structurally related to the IR-tails of the co-activators and APC10, and are directly recruited to APC in a MR-dependent manner (55). The Nek2A MR tail may engage the IR motif binding sites of APC3. Cyclin A is recruited to the APC through its binding partner Cks1(56) that recognizes the phosphorylated APC3 of the TPR arm (57).

The N-terminal Cdc20 facilitates Nek2A destruction in a C-box-dependent manner in Cdc20-depleted cells (39). Cdc20 is required for the degradation of Nek2A and

Cyclin A, but it does not mediate their binding to the APC at prometaphase. This renders their degradation insensitive to the mitotic checkpoint complex (39, 55, 57). The ubiquitination of Cdc20 is not required to release it from the checkpoint complex, but to degrade it to maintain mitotic arrest (56).

### **APC activity regulation via UbcH10 auto-ubiquitination**

Coupling of APC activity to E2s provides another mechanism of regulation. Once a substrate is bound to APC, several different ubiquitin-conjugating enzymes will catalyze the ubiquitin chain assembly. UbcH10 or UbcH5 (Ubc4 in yeast) catalyzes mono-ubiquitination onto the substrate lysines, followed by Ube2S (Ubc1 in yeast) assembling a Lys11-linked poly ubiquitin chain onto the pre-attached ubiquitins (58-60). The Ubc1 of *S. cerevisiae* assembles a Lys48-linked poly ubiquitin chain onto the yeast APC substrates (61).

The UbcH10 N-terminus sets a threshold for APC activation by limiting the APC activity for substrate selection and checkpoint control. Mutations of the highly conserved UbcH10 N-terminus increase substrate ubiquitination and the number of targeted lysines on substrates, allow ubiquitination of the APC substrates without a D-box, increase resistance to the APC inhibitors Emi1 and MCC *in vitro* and bypass the SAC (spindle assembly checkpoint) *in vivo* (14). These regulations guarantee that ubiquitin transfer by the E2 won't overcome the unmet criteria, in order to avoid the inappropriate ubiquitination and unwanted substrate destruction.

The UbcH10 is a target of APC<sup>Cdh1</sup>-mediated ubiquitination and degradation of UbcH10 inactivates the APC. Cyclin A inhibits the auto-ubiquitination of UbcH10, but not its E2 function. Therefore, the ubiquitination of UbcH10 will occur after high-affinity substrates, such as Cyclins, are degraded. Because of this mechanism, the APC activity is maintained as long as G1 substrates present (62). During G1 phase, the APC autonomously switches to a state permissive for Cyclin A accumulation. UbcH10 accumulates at a similar time to Cyclin A in late G1 phase and is degraded in mitosis (63). The APC is inactivated after mitosis exit.

Auto-ubiquitination of UbcH10 is regulated by a conserved N-terminal extension of UbcH10, which is unique among E2 enzymes. The deletion of the N-terminal UbcH10 impairs the formation of ubiquitin chains by APC<sup>Cdh1</sup> but simultaneously allows some UbcH10 ubiquitination by APC<sup>Cdc20</sup> (62). However, Pines' group questioned the auto-ubiquitination of UbcH10. They reported UbcH10 didn't show a rate-limiting role in mitosis towards APC substrate destruction. Instead they find that it is rate-limiting in late G1 phase where UbcH10 is required to destabilize Cyclin A and prevent premature DNA replication (63).

## **Aims of the Study**

The aim of this dissertation study was to gain a deeper understanding into how APC3 recruit substrate co-receptors: Cdh1 and APC10. Therefore, I studied their interactions through both a hybrid APC3 structure and biochemical assays. In Chapter 2, I determined the structure of the hybrid APC3 protein (a chimeric APC3) to gain further insights into how Cdh1 and APC10 were recruited by APC3 through their highly conserved C-terminus (IR-tails). The APC EM-derived structures of APC subunits (APC7, APC3, APC6, APC8) were also used to check the authenticity of an interaction model (of the IR-tails and APC3) from the structural studies. In Chapter 3, I performed biochemical assays to understand how Cdh1 mediates the interactions with other APC subunits. Chapter 4 serves as an overview and a discussion on the future implications of these studies.

## CHAPTER 2. CRYSTALLIZATION AND STRUCTURE DETERMINATION OF THE C-TERMINAL APC3: INSIGHTS INTO THE RECRUITMENT OF THE APC CO-ACTIVATORS AND APC10 BY APC3

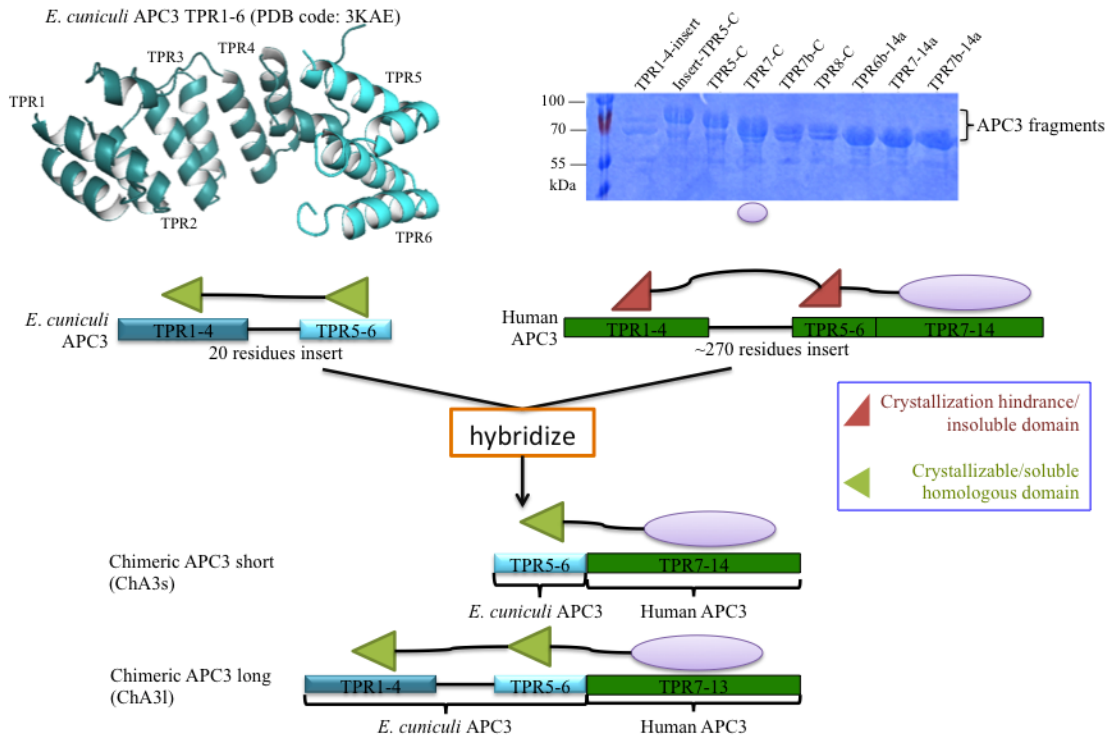
### Crystallization Strategy

Previous studies from multiple laboratories have shown that APC co-activators and APC10 are recruited to APC3 through their C-terminal highly conserved IR-tail peptides (33, 34, 36, 37). It is known that the APC3 C-terminal domain is responsible for recruiting APC co-activators and APC10, but there are no structural models to elaborate details for the interactions. This chapter mainly focuses on the crystallographic approaches to understand the mechanism of co-activator and APC10 recruitment by APC3.

The crystallization attempts of APC3 full-length protein and domains +/- co-activators/APC10 have been tried and none of them succeeded. When the atomic structure of N-terminal *E. cuniculi* APC3 was published (35), a strategy was developed by us to pursue the structure of human APC3 C-terminal domain by making hybrid *E. cuniculi*-human APC3 proteins. This hybrid protein idea was first reported to be a hybrid LRR (leucine rich repeat) technique and successfully applied to crystallize human TLR4 (Toll-like receptor), a member of the LRR family (64, 65). Recently the hybrid technique also successfully crystallized internalin B, another member of the LRR family (66). This hybrid strategy is to facilitate soluble expression and the crystallization of the protein of interest, by replacing an insoluble domain of the protein with a corresponding domain of a homologue. The substitute domain is more soluble and structurally similar to the domain that will be replaced. Similarly, the principle of our strategy is to facilitate human C-terminal APC3 crystallization through fusing it to another protein that has a high probability to crystallize.

Similar to *E. cuniculi* APC3, human APC3 is a TPR subunit with 14 predicted TPR motifs, and N-terminal TPR1-4 mediates homo-dimerization (35). Human APC3 has 270 residues predicted disordered between TPR4 and 5, whereas *E. cuniculi* APC3 only has 20 residues in the corresponding disordered region. The large disorderd region accounts for almost 1/3 of human APC3 protein sequence and it is a hindrance to the crystallization. In addition, human APC3 is a highly hydrophobic protein and the full-length protein easily aggregates during *in vitro* purification once the affinity/stabilizing tags are removed. The co-purification of APC3 with its co-activators or APC10 could not prevent the protein aggregation.

In order to increase protein crystallization potential and improve protein behaviors, two versions of the chimeric APC3 were constructed (**Figure 2-1**). The short version incorporated *E. cuniculi* APC3 TPR5-6 and the longer version includes TPR1-6, which includes all the crystallized domain in the published structure (35). Both versions fused *E. cuniculi* APC3 TPRs to human APC3 TPR7 to 14.



**Figure 2-1. The crystallization strategy of the human APC3 C-terminal domain.**

Illustration of the hybrid strategy to construct the chimeric APC3. The goal of the strategy is to replace the crystallization hindrance domains of the human APC3 with crystallizable/soluble, homologous domains of *E. cuculii* APC3. Two versions of the chimeric APC3 were generated from TPR5-6 or TPR1-6 of the *E. cuculii* APC3 N-terminal domain fused onto TPR7-14 of human APC3 C-terminal domain. The long one used N-terminal dimerization domain TPR1-4 and TPR5-6, whereas the short version only includes TPR5-6. The expression test of APC3 domains is shown on the top right, with each fragment fused onto a His6-MBP tag.

## Results and Discussion

### The binding ability test of the chimeric APC3s

To investigate whether the chimeric APC3a are able to recruit the APC co-activators, I used a co-pulldown experiment to test its binding ability with Cdh1. The co-pulldown results shows that both the long and short versions of the chimeric APC3 poorly interact with Cdh1 (**Figure 2-2**), which implies they lost the binding abilities to Cdh1 in solution. The co-expression of the chimeric APC3s with Cdh1 failed to restore them to the “binding” status. The IR-tail mimicking, APC inhibitor TAME was also used to test the chimeric APC3 binding ability via NMR WaterLOGSY. In the WaterLOGSY experiment, a free ligand gives multiple peaks pointing downward in the spectrum. Once the ligand binds to proteins, the peaks flipped upwards. No interaction signal with the chimeric APC3s was observed from the TAME 1D spectra, whereas TAME binds to the human wild type APC3 as a part of APC3-APC7-APC16 complex (**Figure 2-3**).

### Improvement of bacterial expression of the chimeric APC3

Although the chimeric APC3s didn't interact with the APC co-activator Cdh1, we were still interested in knowing what caused the change to the protein function. Both versions of the chimeric APC3 dramatically improved protein yield, solubility and reduced the amount of aggregation from their insect cell (Hi-5) expression. Both versions were purified through affinity chromatography, ion exchange and gel filtration, followed by robotic crystallographic screening. The short chimeric APC3 generated hits in conditions of 0.1M HEPES pH7.5, 1.6M AmSO<sub>4</sub>, 2% Dioxane at room temperature (around 25°C).

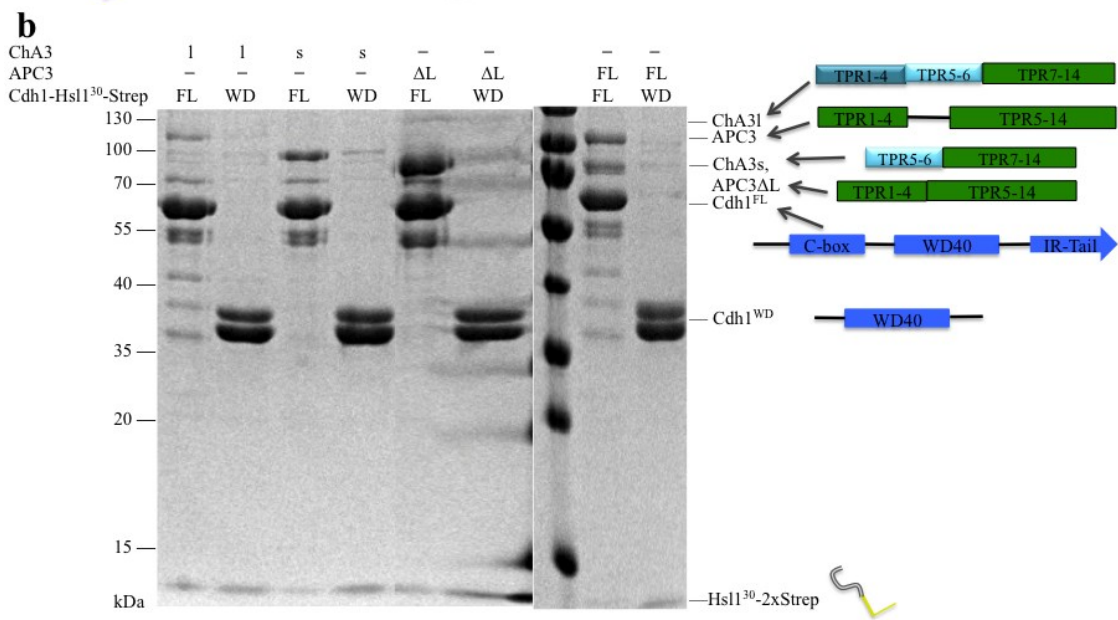
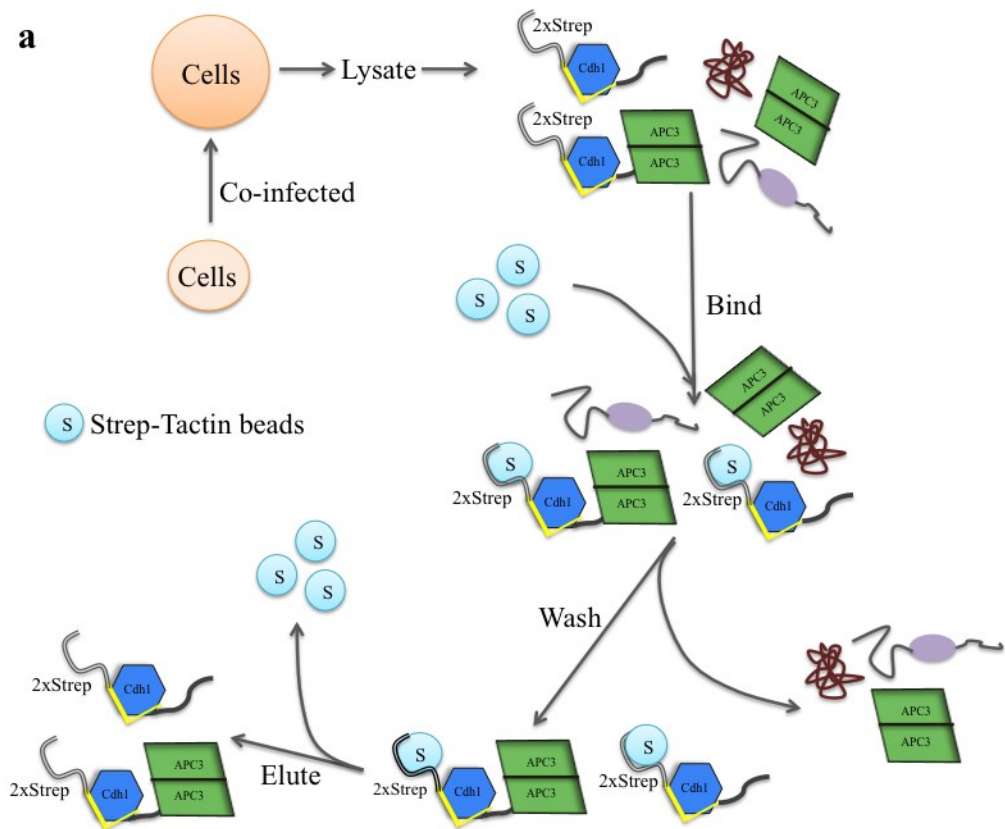
In order to determine the crystal structure of the chimeric APC3s, an experimental phase is required to calculate an electron density map. The electron density map presents a three-dimensional description of the electron density of the overall molecule structure and outlines the molecule surface. A good map will make atomic model building possible. Crystal diffraction data of high quality with phase information is key for the crystallography software to calculate high quality maps. The phase is an intrinsic defect in crystallography, because observed crystal diffraction is the summary of all the electrons in the molecule. Without correct phase information, the two-dimensional diffraction spots cannot be reconstituted back to three-dimensional coordinates to outline the molecular envelope. Generally there are two ways to get phases: one is through experiments (experimental phases) and the other is through homologous or predicted models. During the structure determination of chimeric APC3, I tried various models from published atomic models to synthetic models, and none of them was sufficiently accurate to generate a reasonable map. This fact indicates the chimeric APC3 structure is significantly different than the existing models tried.

**Figure 2-2. Test the chimeric APC3 binding ability to Cdh1 by a co-pulldown.**

(a) Flow chart to demonstrate the procedure of the co-pulldown experiment to test APC3 binding ability with the Cdh1. APC3 were co-expressed with the Cdh1 and Cdh1 substrate peptide Hsl1<sup>30</sup>-2xStrep in insect cell Hi-5 strain. The co-pulldown was performed on the Strep tag of Hsl1<sup>30</sup>. The amount of the APC3 co-eluted with Cdh1-Hsl1<sup>30</sup> was used to indicate their binding ability.

(b) SDS-PAGE to examine the co-pulldown results of (a). Four APC3 constructs were tested the binding ability with both the Cdh1 full-length and WD40 domain. APC3 $\Delta$ L: Disorder region between TPRs 4 and 5 deleted. Two versions of the chimeric APC3: human APC3 TPR 7-13 was fused with *E. cuniculi* APC3 TPR1-6 (long version) or TPR 5-6 (short version), respectively. FL: full-length protein. WD: WD40 propeller domain, the Hsl1<sup>30</sup>-2xStrep peptide-binding domain. Molecular weight standards are labeled on the left side of the gel and protein bands are labeled with their names and construct schematics.



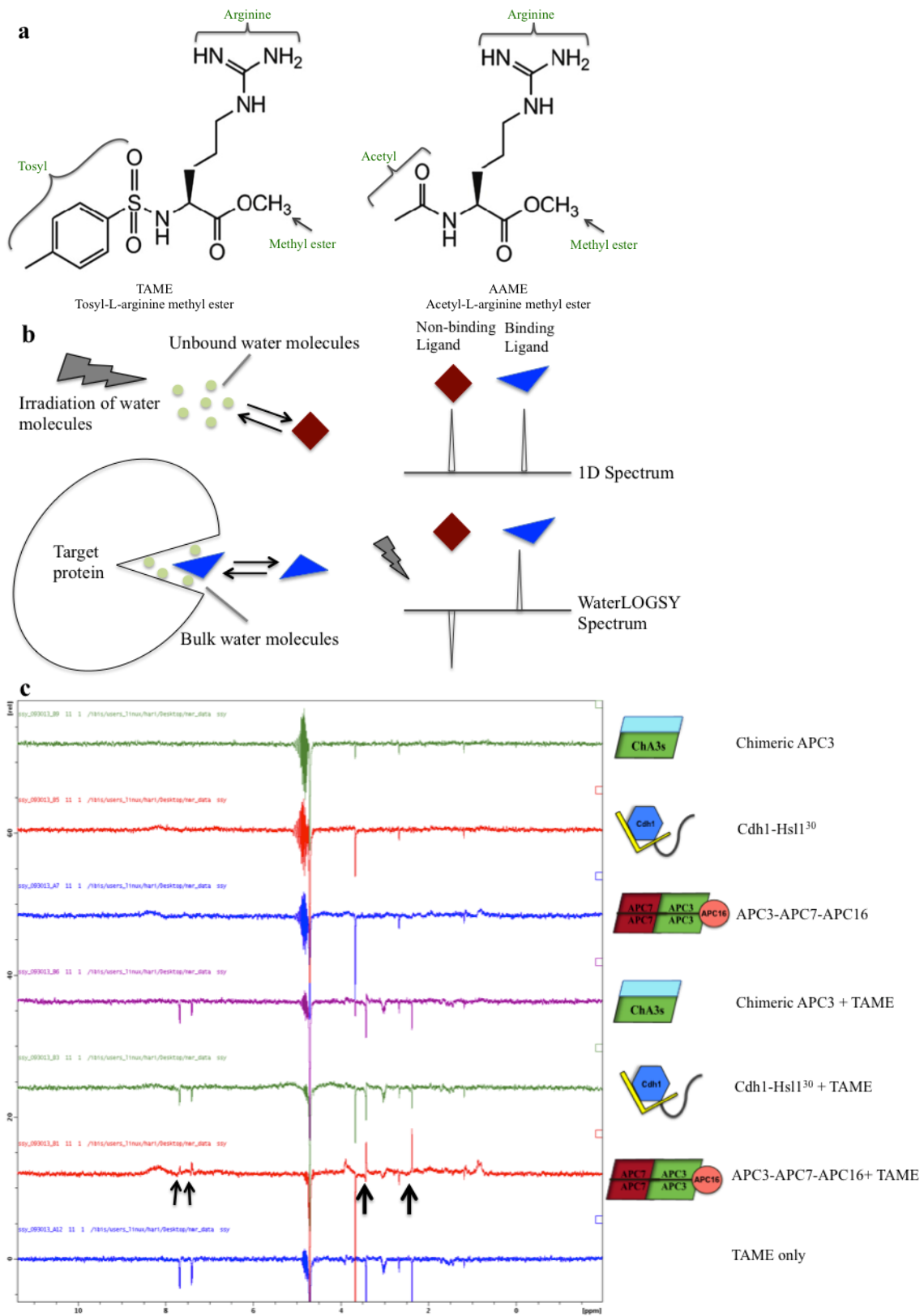


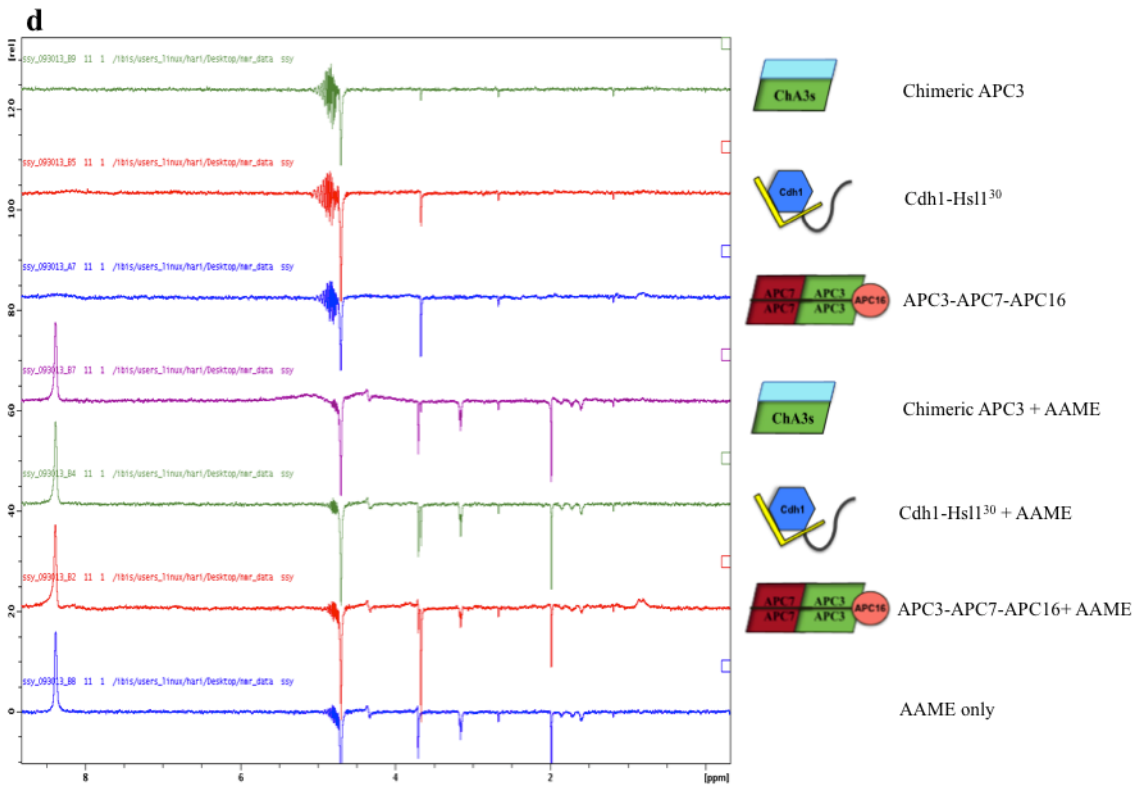
**Figure 2-3. Examine the chimeric APC3 binding ability with APC inhibitor via 1D WaterLOGSY.**

(a) Illustration of compounds TAME and AAME. Groups of each compound are illustrated in green color.

(b) Schematic diagram of WaterLOGSY basic principle. The red rectangles and blue triangles represent non-binding and binding ligands, respectively. Green spheres represent water molecules. Non-binding ligands only receive magnetism transferred from unbound water, whereas binding ligands will also receive magnetism transferred from target proteins via bulk water molecules. In the 1D spectrum, both non-binding and binding ligands give positive signals (upward peaks). In the WaterLOGSY spectrum, the signals from the binding ligands remain positive and signals of the non-binding ligands becomes negative (downward peaks).

(c,d) One-dimensional WaterLOGSY spectrum recorded of APC3 interaction with TAME (b) and AAME (c). The three spectra from top are the protein reference spectra of 2 $\mu$ M chimeric APC3, 2mM Cdh1-Hsl1<sup>30</sup>-Strep, and 2 $\mu$ M APC3-APC7-APC16, and one compound only reference of 0.2mM compound TAME (b) or AAME (c) at the bottom. The following three spectra from the top are the corresponding 2 $\mu$ M protein complexes in the presence of 0.2mM TAME (b) or AAME (c). The spectra were acquired with 10920 scans and protein signals were destroyed with the design of WaterLOGSY pulse sequence.





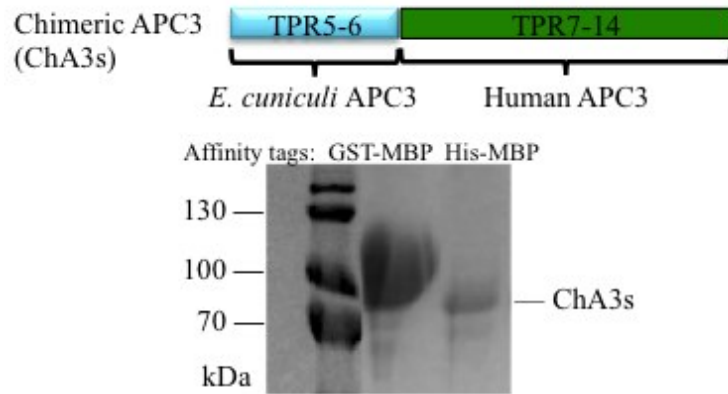
**Figure 2-3. (Continued).**

Another common approach to obtain experimental phases is through the diffraction that is derived from heavy atom labeled crystals. Theoretically, the covalently attached heavy atoms will slightly shift the crystal diffraction pattern without affecting the overall intensity. Then the crystallography software is able to calculate the map through the anomalous signals. Selenomethionine substitution is a common way to incorporate selenium into protein peptides and label protein crystals. The principle is to use selenium labeled methionine to substitute normal methionine in the bacteria culture for the cells expressing proteins to be studied. To make that method applicable, I reconstituted the expression construct of the chimeric APC3 for bacterial expression, followed by the expression test in various bacterial strains. In summary, the protein could be expressed in several strains but has a 10-fold lower yield, which could not meet the needs of the protein amount required for crystallization. In order to overcome this barrier, different affinity tags including GST, His-MBP, GST-MBP, GST-EGFP and GST-T4 Lysozyme, the chimeric APC3 mutants and the chimeric APC3 of different species were used to enhance protein expression. Of the entire search, the GST-MBP tag with the wild type chimeric APC3 increased the protein expression 5-fold in bacterial strain BL21 gold (**Figure 2-4**). The final yield of GST-MBP tagged chimeric APC3 (*in vitro* purified protein) was 2.6 mg per liter culture, whereas the His-MBP tagged fusion only produced 0.55mg per liter. The purified chimeric APC3 behaves as well as the one purified from insect cells and generated crystallography hits in a similar condition at room temperature.

### **Purification improvement of seleno-methionine labeled protein**

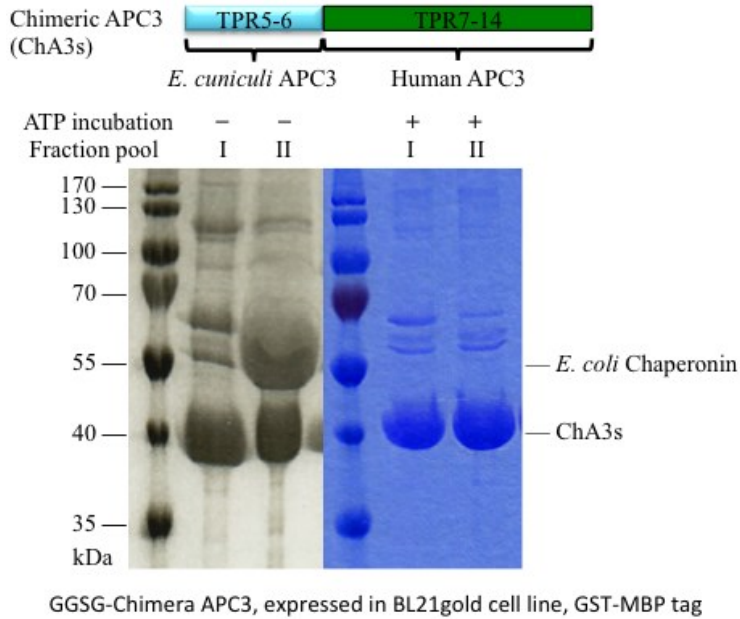
The selenomethionine (SeMet) substitution requires enriched media for bacteria growth, which benefits the protein expression. However, the *E. coli* 60kDa chaperonins were also enriched during the purification of SeMet incorporated protein. The chaperonins accounted for 50% of final product (**Figure 2-5**). These co-purified chaperonins were identified by mass spectrometry.

The chaperonins were reported to improve protein folding and they are driven by ATP hydrolysis. ADP stabilizes the chaperonins binding to its substrates, while ATP modifies its tertiary structure and promotes the chaperonin complex releasing the substrates (67). The co-purification of chaperonins with the chimeric APC3 suggested that they bind to the chimeric protein tightly. In order to separate APC3 from chaperonins, I included two-rounds of ATP incubation during the affinity chromatography and the incubation effectively removed the chaperonins (**Figure 2-5**). The selenomethionine incorporation was confirmed by intact mass spectrometry. Each selenomethionine has ~47Da more molecular weight than methionine. Through comparing the molecular weight shift between SeMet protein and native protein, 10 selenomethionines were calculated that were incorporated into chimeric APC3 (**Figure A-1**).



**Figure 2-4. Bacterial production of the chimeric APC3 is improved by a GST-MBP tag.**

The schematic map of the chimeric APC3 construct (top), and SDS-PAGE of the chimeric APC3 protein purified from GST-MBP and His-MBP affinity tags. Proteins in both lanes were loaded with the same amount of affinity chromatography elutions.



**Figure 2-5. ATP incubation efficiently removed Chaperonin during the chimeric APC3 purification.**

The schematic map of the chimeric APC3 construct (top), and SDS-PAGE of the purified chimeric APC3 protein purified with or without ATP incubation (bottom). The ATP incubation removed the *E. coli* Chaperonin (blue gel) that was presented in the previous purification (gray gel). *E. coli* Chaperonin were identified by mass spectrometry.

## Diffraction improvement and structure determination

The diffraction of the chimeric APC3 crystals started at 8Å and it was improved to 3Å after multiple rounds of optimization. The low resolution didn't have enough diffraction information to provide an accurate electron density map. A high mosaicity value was also present in the higher resolution dataset, which indicates the dataset has multiple overlapped crystals and crystallography software could not calculate the correct map. This problem also exists in the SeMet crystal dataset, which leads to SeMet crystal dataset failing to generate experimental phase. Many approaches were tried to improve the mosaicity, including screening at different temperatures, new crystals forms searching and optimization, heavy atom soaking and co-crystallization, seeding, TAME soaking, additive screening (**Table 2-1**). During the optimization process, the short version chimeric APC3 crystallized in multiple crystal forms (**Figure A-2**) and around 800 crystals in total from different crystal forms were harvested. All the crystals were shot at synchrotron for data collection. Eventually, the combination of seeding, additives, and lower temperature (18°C) helped crystals grow to more than 10 times of the original size (**Figure 2-6**). The dissection of crystals into small pieces separated the overlapped multiple crystals and a good-quality dataset was obtained for structure determination (**Figure 2-7**).

The atomic structure of the chimeric APC3 was determined by means of SAD (Single-wavelength anomalous diffraction) phasing from a SeMet dataset of 3Å. Data collection and refinement and statistics are summarized in **Table 2-2**.

## Structural analysis of the chimeric APC3 atomic structure

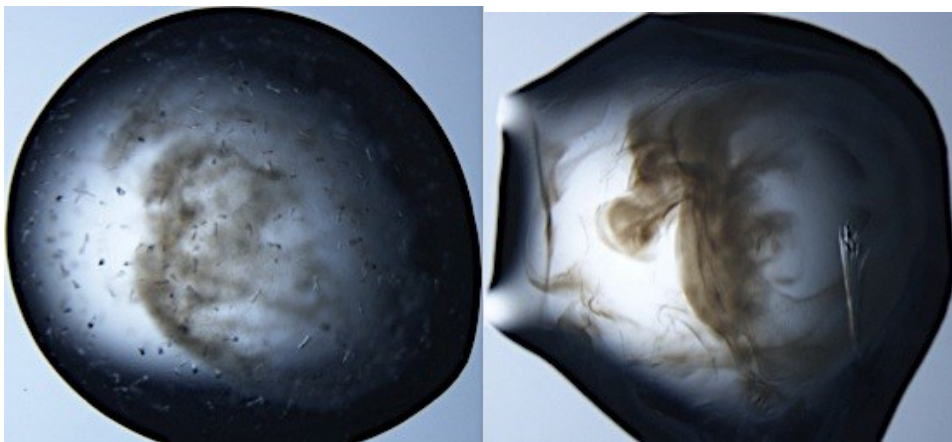
The chimeric APC3 consists of 16  $\alpha$ -helices, with 14 of them in helix-turn-helix (TPR) motif. Based on the structure superimposition and sequence alignment, 14 anti-parallel helices are TPR5-10 and TPR12, and they match the predicted TPR motif boundaries from APC3 secondary structure prediction from TPRpred and Psidepred (**Figure 2-8a**) (68, 69). Each TPR motif was then assigned to an *a* and a *b* helix, with the *a*-helices lining the inner concave surface and the *b*-helices forming the outer convex surface. There are two  $\alpha$ -helices:  $\alpha$ -11 and  $\alpha$ -13 are not in any predicted TPR motif. The helix  $\alpha$ -11 is flanked by two loops and it breaks the successive TPRs pattern by folding back to the concave surface around TPR8-10 (**Figure 2-8b**).  $\alpha$ -13 is a long  $\alpha$ -helix following TPR12 and projects towards N-terminal TPR5. The loop- $\alpha$ -11-loop and  $\alpha$ -13 are within the regions of predicted TPR11 and TPR13 domains, respectively, but neither folds into the canonical TPR (**Figure 2-8a**). The TPR proteins of APC complex were expected to form a superhelix with their consecutive TPRs as the tertiary structure (20). However, the overall shape of the chimeric APC3 molecule does not display this superhelix structure.

To interpret the folding mechanism, the chimeric APC3 structure was superimposed with *E. cuniculi* APC3 N-terminal structure and human APC3 predicted



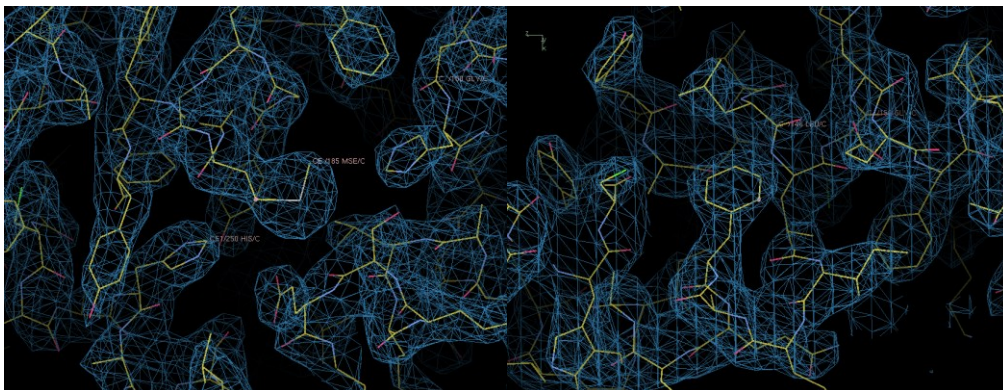
**Table 2-1. Summary of optimization approaches to improve the chimeric APC3 crystal diffraction.**

<b>Approaches</b>	<b>Description/Summary</b>	<b>Temp</b>	<b>Result</b>
Additive screening	110 additives screened	4°C & RT	Best diffraction at 3.5Å
Manually seeding	120 conditions manually seeded	RT	Diffraction around 4-5Å
Robotic seeding-screening	1,300 conditions robotically screened with crystal seeds	4°C & RT	Generate hits in new conditions for optimization
Heavy atoms co-crystallization	Co-crystallized with heavy atoms of Hg, Pt, (NH <sub>4</sub> ) <sub>2</sub> SeO <sub>4</sub> and lanthanides at 1mM and 10mM	RT	No crystals
Heavy atoms soaking	Crystals soaked with 1mM or 10mM Hg, Pt, Au, Pb, (NH <sub>4</sub> ) <sub>2</sub> SeO <sub>4</sub> and lanthanides within pH range 6.5-7.0	RT	Best diffraction at 4-5Å, but no anomalous signal
Low melting (LM) agarose growing	Crystallization in 36 optimized conditions with 1% and 2% LM agarose	RT	No crystals
TAME soaking	20 crystals soaked in 10mM TAME within pH range 5.5-8.5	RT	Diffraction around 4-5Å
Cryo buffer screening	30-40 cryo buffers tested during crystal harvest	RT	Best diffraction at 3Å
MBP fused chimeric APC3	Robotic screen around 1,300 conditions, followed by manual screening with 120 additive for optimization	4°C & RT	Diffraction around 8Å
T4 Lysozyme and EGFP fused chimeric APC3	Robotic screen around 3,000 conditions	4°C & RT	No crystals
Lys-methylated chimeric APC3	Robotic screen around 2,000 conditions	4°C & RT	Pom-pom shape crystals



**Figure 2-6. Original hits and improved crystals of the SeMet incorporated chimeric APC3.**

The crystals display the initial hits of the chimeric APC3 (left) and manual optimized hits (right). The initial hits grow in 0.1M MES pH6.5, 1.6M  $(\text{NH}_4)_2\text{SO}_4$ , 10% (v/v) Dioxane, RT. The optimization hits grew in 0.1M MES pH6.5, 1% Dioxane, 1.45M  $\text{AmSO}_4$ , 2% MPD, 0.2M  $\text{MgCl}_2$ , 0.1M  $\text{Li}_2\text{SO}_4$ , micro seeding 1:1000 at 18°C, seeded from native crystals.



**Figure 2-7. Representative electron density.**

Final 2Fo–Fc electron density map contoured at  $\sigma 1.4$  is shown over the chimeric APC3 structure. The blue mesh represents the calculated electron density map from the collected data and the yellow sticks indicate the peptide backbone, with oxygen atoms labeled in red and nitrogen atoms labeled in blue (70).

**Table 2-2. Structure data and refinement statistics of the chimeric APC3.**

Statistics	Chimeric APC3 (SeMet)
<b>Data collection</b>	
Beam line	APS 24-ID-C
Wavelength (Å)	0.9793
Space group	C121
Cell dimensions	
a, b, c (Å)	a=235.5, b=130.1, c=103.9
$\alpha$ , $\beta$ , $\gamma$ (°)	$\alpha=90$ , $\beta=110$ , $\gamma=90$
Resolution (Å)	80-3.0
Total reflections	303819
Unique reflections	111550
Rmerge (%)	11.4 (75.1)
Average I/ $\sigma$	12.2 (2.0)
Completeness (%)	97.9 (75.5)
Redundancy	4.2 (2.1)
Wilson B-factor	23.47
<b>Refinement</b>	
Resolution range (Å)	50-3
No. of reflections ( $\sigma \geq 0$ )	57195
R <sub>work</sub> (%)	22.24
R <sub>free</sub> (%)	25.49
Number of protein atoms	15930
Number of waters	43
Average B-factor (protein)	55.31
Average B-factor (water)	34.21
RMSD:	0.012
Bond lengths (Å)	1.216
<b>Ramachandran plot statistics</b>	
Residues in preferred regions (%)	98.03
Residues in additional allowed regions (%)	1.82
Residues in disallowed regions (%)	0.15

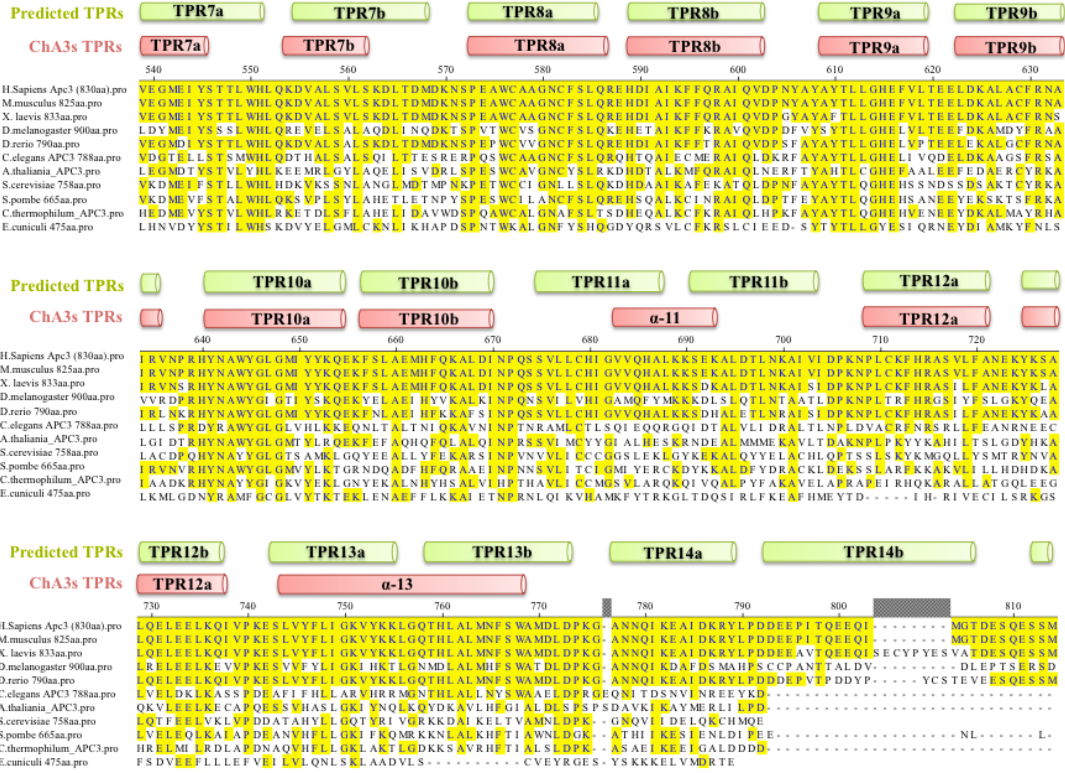
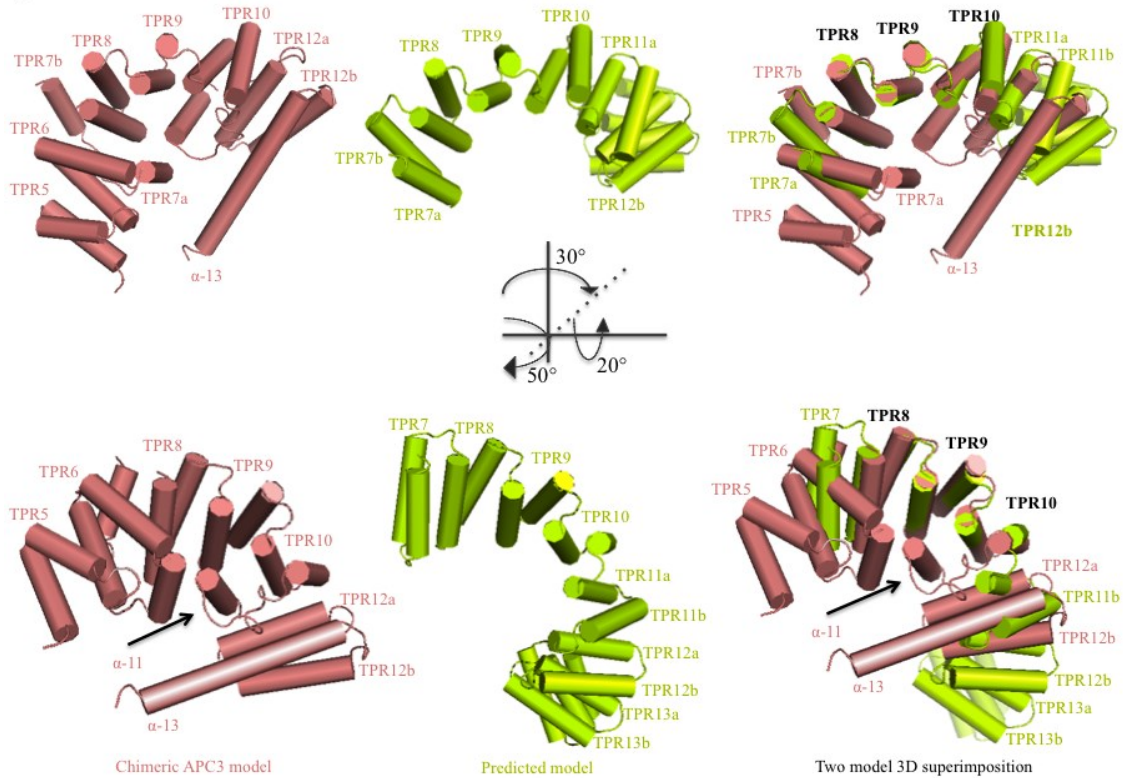
Highest resolution shell is shown in parenthesis.  $R_{work} = (|F_o - F_c| / F_o)$ .  $R_{free}$  is the crossvalidation of R-factor, with >5% of the total reflections omitted during model refinement.

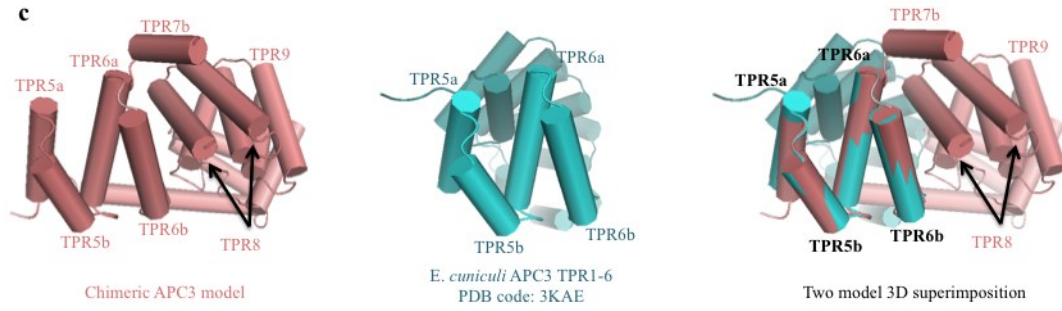
**Figure 2-8. Analysis of the chimeric APC3 structure through sequence and structural comparisons.**

(a) Multiple sequence alignment of human APC3 homologs. The residues that match human APC3 sequence are highlighted in yellow. Position of observed TPR motifs and  $\alpha$ -helix elements are labeled with salmon color, whereas predicted TPR motifs are in light green.

(b) Superimposition of the chimeric APC3 (salmon) onto a predicted APC3 C-terminal domain in light green. TPR8-10 superimpose closely onto the predicted model. Other TPR motifs aligned less well with the predicted motifs in light green.

(c) Superimposition of TPR 5-6 of the chimeric APC3 structure onto TPR 5-6 of *E. cuniculi* APC3 model. *E. cuniculi* APC3 residues 175-241, corresponding to TPR5-6 adopts identical folding in both structures.

**a****b**



**Figure 2-8. (Continued).**

model, with reference to APC3 sequence comparison across eukaryotic species. The TPR5 and 6 in the chimeric APC3 were well aligned onto their TPR5 and 6 in *E. cuniculi* APC3, with an r.m.s.d of 0.5Å (**Figure 2-8c**), which means the TPR5 and 6 of the chimeric APC3 kept the same folding as in *E. cuniculi* APC3. The following TPR7 forms two shorter  $\alpha$ -helices in the chimeric APC3 instead of the predicted canonical TPR7 helices. The loops following helices TPR7a and TPR7b, instead, are longer than the prediction. Two  $\alpha$ -helices of TPR7 spread across the concave and convex surfaces of TPR array instead of packing parallel to TPR8 (**Figure 2-8b**). The sequence boundaries of TPR8, 9 and 10 from the chimeric APC3 perfectly match their corresponding TPRs in prediction, and two structures of TPRs 8-10 are well aligned with an r.m.s.d of 1.46Å (**Figure 2-8a, b**).

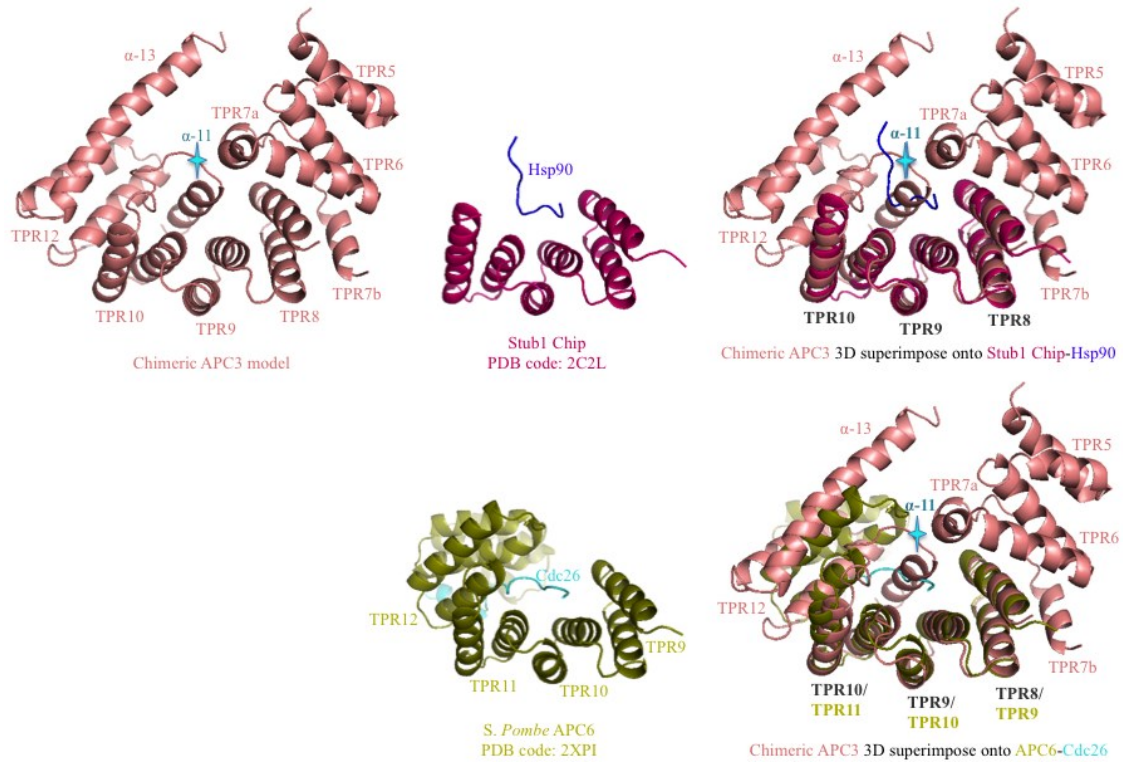
One unexpected mis-folding region is helix  $\alpha$ -11, which corresponds to the predicted TPR11 domain but folds into a short helix, and flanked by two longer loops. The predicted TPR11 parallel TPR10, and extends the  $\alpha$ -helices along the TPR array to form a superhelix. However, the helix  $\alpha$ -11 slides into a concave surface that is lined by the helices  $\alpha$  of TPR 8-10 (**Figure 2-8a, b**). The flanking loops triple the length of a predicted turn between two antiparallel helices. The orientation of the following TPR12 seems to be affected by the mis-folding of TPR11, although each TPR12 helix matches the expected boundary. TPR12 is packed in a perpendicular angle to TPR10 and breaks the pattern of the TPR arrangements (**Figure 2-8b**). Another affected region is the helix  $\alpha$ -13, which also forms within the expected TPR13 domain.  $\alpha$ -13 adopts a long  $\alpha$ -helix structure rather than two short anti-parallel  $\alpha$ -helices.  $\alpha$ -13 is parallel to TPR12 and is also perpendicular to TPR10. Meanwhile, the fact that  $\alpha$ -13 extends towards TPR5 make the overall structure a globular shape, which may explain that the chimeric APC3 have higher solubility than most other *in vitro* purified APC TPR proteins.

### **Identify human APC3 C-terminal binding groove through model comparisons**

In order to examine whether the structure of the chimeric APC3 is functional, The Dali server was used to search structurally similar TPR proteins through three-dimensional structure analysis. TPR7-12 domain of the chimeric APC3 was used as the search model, and three hits: Carboxyl terminus of Hsp70-interacting protein (Chip), and TPR9-11 of human APC6 and yeast APC6 all have highest structural similarity with human APC3 TPR8-10. All the models were superimposed onto APC3 TPR8-10 with an r.m.s.d smaller than 1.1Å and the helix  $\alpha$ -11 of the chimeric APC3 overlapped with all the peptides (binding partners) of three similar hits (**Figure 2-9**).

The structural superimposition suggests that the highly conserved TPR8-10 in the chimeric APC3 adopts canonical TPR protein folding (**Figure 2-8a, Figure 2-10a**) and structurally it is also highly similar to three peptide-interacting domains (from the hits). Beside, the mutagenesis study reported that the mutation (N581A/L612A of helix 8a and 9a) in this domain reduced Cdh1 association with APC with a decreased ubiquitination (34). All the above indicate TPR8-10 is the APC3 binding groove that is responsible for recruiting IR-tail peptides. The helix  $\alpha$ -11 is buried inside the binding groove and





**Figure 2-9. Structural alignments of the chimeric APC3 with TPR proteins reveal a rigid binding groove and pseudo binding partner.**

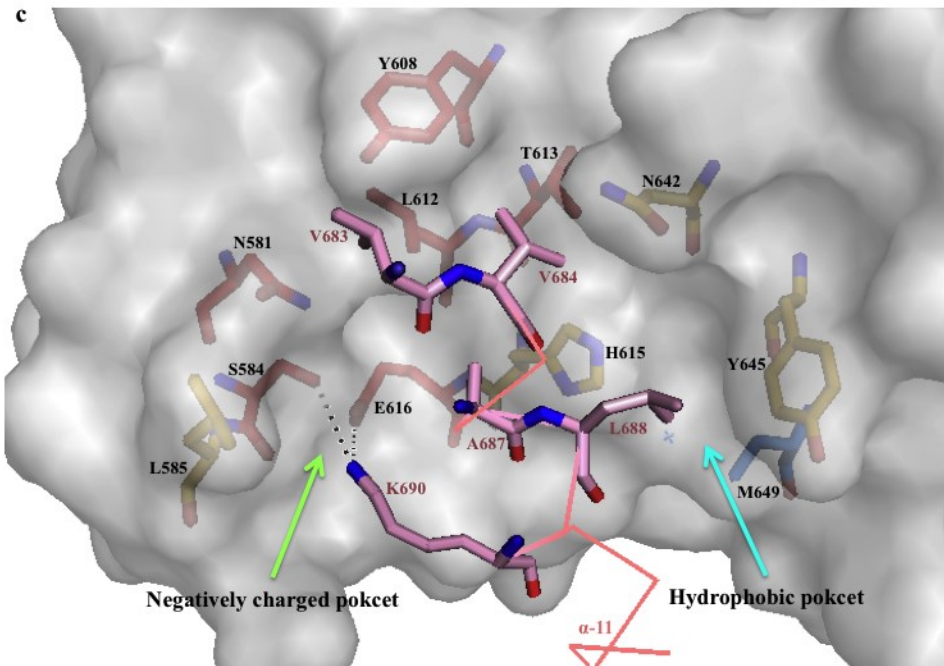
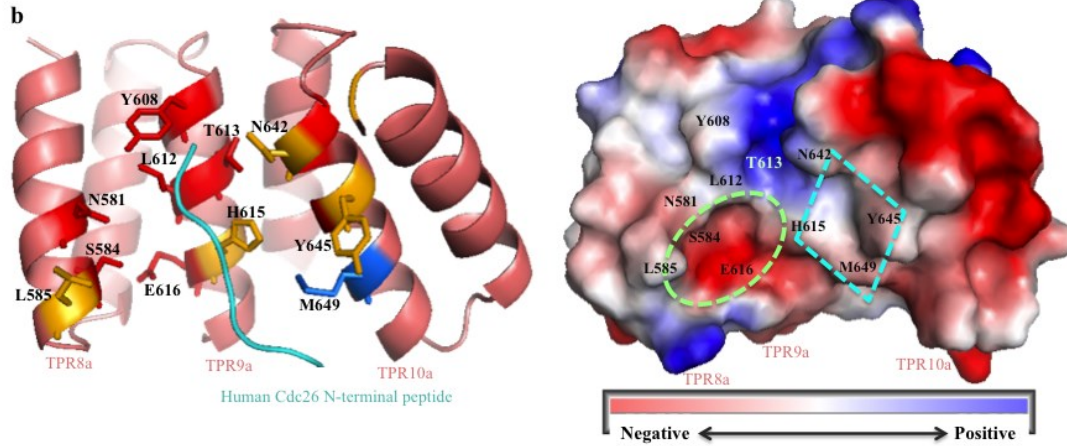
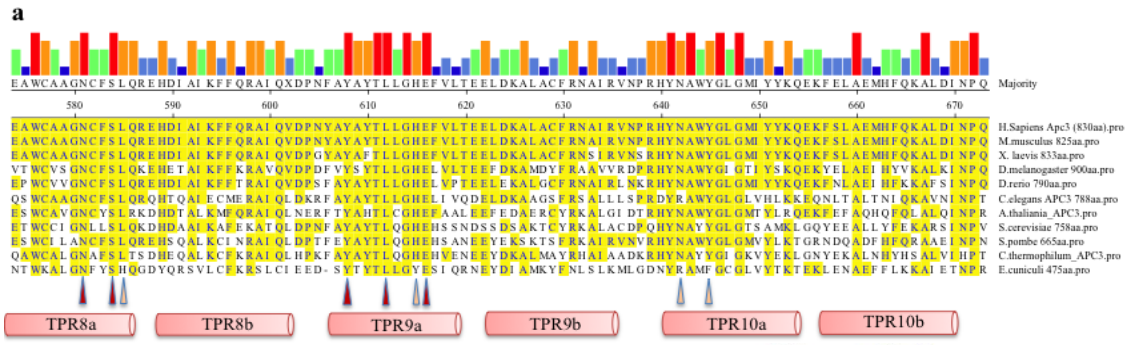
The chimeric APC3 (salmon) is superimposed onto the TPR domains of Stub1 Chip (warm pink) and the TPR9-11 of *S. pombe* APC6 (olive). The cyan asterisks labeled the interacting peptides: Hsp90 (blue) and Cdc26 (cyan), which overlapped with the helix  $\alpha$ -11.

**Figure 2-10. Pseudo binding partner  $\alpha$ -11 interacts with two binding pockets in the APC3 binding groove.**

(a) Multiple sequence alignment of TPR8-10 motifs of human APC3 homologs. The bars above consensus indicate the conservation strength: red, invariant residues; orange: conserved residues. Red or orange triangle legends indicate the residues labeled out in the model of (b). Positions of observed TPR 8-10 are labeled with salmon color, which matched the predicted TPR 8-10 in light green.

(b) Details of conserved residues at the inner face of the binding groove model (left) and their contribution to the surface charge (right). Left: Invariant residues are colored in red and conserved residues colored in orange. Human Cdc26 N-terminal peptide was superimposed onto the helix  $\alpha$ -11 position. Right: The bar at the bottom right indicates the surface electrostatic potential. Two binding pockets at the binding groove surface are outlined with green (negatively charged pocket) or cyan (hydrophobic pocket) dash lines.

(c) Close-up views of the key interactions between pseudo binding partner helix (light-pink ribbon and sticks) and two binding pockets (surface and sticks). The  $\alpha$ -11 residues interacting with the binding groove were displayed in pink sticks. Dotted lines indicate the electrostatic interaction and the hydrogen bond of Lys690 with Glu616 and Ser584, respectively.



overlapped with all three peptides from the hits, which suggests it folds into a pseudo binding partner for APC3 and blocked the IR-tail recruiting domain.

### **Human APC3 binding groove has two conserved pockets at the inner surface, one negatively charged and the other hydrophobic**

As every cloud has a silver lining, the fact that the chimeric APC3 is self-locked by its mis-folded TPR11 implies the functionality of the binding groove. Meanwhile, the pseudo binding partner  $\alpha$ -11 demonstrates the interactions that the binding groove potentially uses for recruiting IR tail peptides. Moreover, the chimeric APC3 structure also provides a clue for studying TPR protein folding mechanism.

A sequence comparison of the APC3 binding groove (TPR8-10) reveals a striking cluster of sequence conservation. There are 39 conserved residues among a total of 96 residues in the binding groove and 22 conserved residues locate on the inner surface formed from  $\alpha$ -helices 8a, 9a and 10a (**Figure 2-10a**). The highly conserved TPR8-10 probably is the reason that the binding groove folded correctly and independently in this structure without being affected by the mis-folded neighbor regions.

Highly conserved residues Asn581, Leu585, Ser584, L585, Tyr608, Leu612 and Glu616 were mapped onto the molecular surface of helices 8a and 9a, with their side chains facing toward superimposed peptides Cdc26 (**Figure 2-10b**). The conserved residues form a pocket-like surface with strong negative charges. The pocket surface around Ser584, Glu616 is most negative. Lys690 of helix  $\alpha$ -11 was recruited to this pocket through electrostatic interaction with Glu616 and hydrogen bonds with Ser584 (**Figure 2-10c**).

The second pocket-like region within the binding groove is formed through well-conserved residues His615, Asn 642, Tyr645 and maybe a non-conserved Met649, with their side chains facing towards the helix  $\alpha$ -11. His615 localizes on the helix 9a and the other three are mapped onto the helix 10a. The pseudo binding partner  $\alpha$ -11 establishes multiple hydrophobic interactions with the second binding pocket through Val684, Ala687 and Leu688 (**Figure 2-10c**). His615 packs its imidazole group toward Ala687 and L688 of helix  $\alpha$ -11 and the side chains of Val684 and Leu688 point to Asn642 and Tyr645, respectively. The second pocket creates a hydrophobic microenvironment along the concave inner face of the binding groove.

### **Validate human APC3 interactions with the IR-tails through mutations of the APC3 binding groove**

To validate the interaction mechanism observed from the chimeric APC3 structure, key residue mutations in the APC3 binding groove were constructed and expressed for the interaction assays. In order to have better behaviors of APC3 mutants and Cdh1, all the mutants were expressed and purified in the complex of human APC7-

APC3-APC16, and Cdh1 was co-expressed with its high-affinity substrate Hsl1<sup>30</sup>. Similar to the wild type APC3, all the APC3 mutants formed a stoichiometry complex with APC7 (**Figure 2-11a**, bottom panel), indicating that the mutated residues did not impair mutant interactions with APC7. Most mutants were defective in binding Cdh1 through co-pulldown experiment (**Figure 2-11a**, top panel), and only three mutants, Y608A/Y645A, H615A/Y645A and H615A/N642A/Y645A/M649A had no or opposite effects (**Figure 2-11a**, middle panel).

In order to identify the most informative mutants, the WaterLOGSY was also performed on the complexes APC7-APC3mutants-APC16. This approach is suitable for detecting binding affinity in the micromolar range (71), and thus it reveals more differences of the weak interactions. In the APC3 mutant WaterLOGSY experiment, the mutant of S584A/E616R directly affected the peaks 5 and 7, the interactions between the backbone of Arginine side chain and the binding groove. Peaks 5 and 7 disappeared in the mutant of S584A/E616R, and all other mutants that include these two mutations. This result matched the speculation that S584 and E616 mediate the salt bridge with the Arginine guanidinium group of the IR-tails (**Figure 2-11b**). Although salt bridge could not be directly detected by this approach due to their higher affinity, the disappeared interactions (of peaks 5 and 7) probably resulted from those missing interactions.

The interactions from tosyl group (Peaks 1 and 2) and from the backbone of Arginine side chain were almost completely wiped out by another mutant of H615A/N642A/M649A. It confirmed that the H615, N642 and M649 together form the hydrophobic core and potentially directly interacted with the Isoleucine of the IR-tails. Although the residues S584 and E616 were still available to mediate salt bridge, the Arginine side chain failed to stay bound. This result indicated that hydrophobic interactions from Isoleucine of the IR tails stabilized the salt bridge formed by the guanidinium group.

Other mutations of the binding groove showed more or less minor binding defectiveness towards IR-tails or TAME. The mutant of H615N/M649N lost the interactions with Cdh1 C-terminus in the co-pulldown but only had subtle effects on TAME (**Figure 2-11a, c**). However the mutant H615A/N642A/M649A became most defective to TAME when N642A mutation was incorporated. It seemed N642 was important for hydrophobic interactions of TAME, whereas the mutations of H615 and M649 were enough to disrupt the IR-tail interactions. TAME were reported as an APC inhibitor when Cdh1 was present (40, 51). Comparing to the Cdh1 IR-tail, it potentially mediated stronger interaction with APC3.

The Y645A mutation seemed to have opposite effect on Cdh1 interaction and no effect to TAME binding (**Figure 2-11a, c**). The mutants of H615A/Y645A and H615A/N642A/M649A/Y645A tended to increased the Cdh1 interaction in the co-pulldowns. The function of Y645 was not certain and the mutation of Y645A might create non-specific hydrophobic interactions.

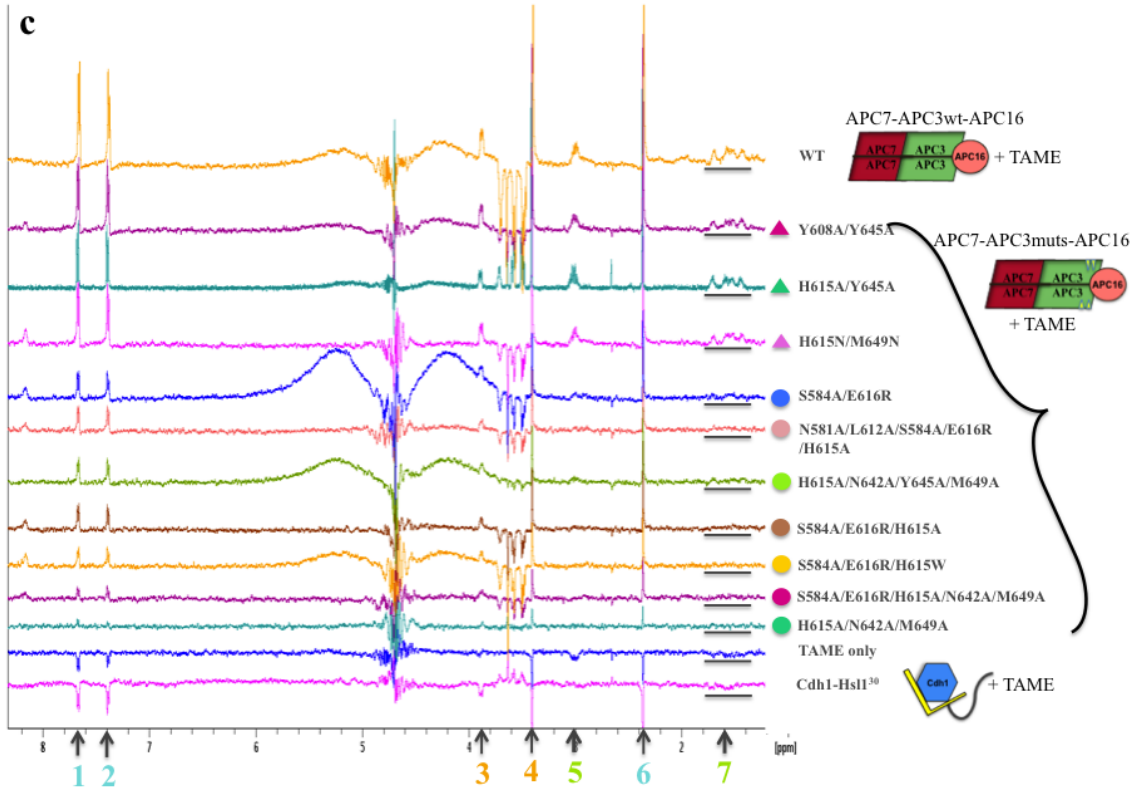
**Figure 2-11. Mutants of key residues inside the APC3 binding groove validated interactions between the APC3 and Cdh1 IR-tail.**

(a) Co-pulldown to examine APC3 mutants binding ability with Cdh1-Hsl1<sup>30</sup>-2xStrep in the complexes of GST-APC7-APC3 (wild type or mutants)-APC16. The co-pulldown was performed on the Strep tag of the substrate peptide Hsl1<sup>30</sup>-2xStrep (top two panels) and GST tag of APC7 (bottom panel), separately. The top panel shows the defective APC3 mutants and the middle panel groups three mutants of subtle or no effects. APC3 mutated residues in each co-pulldown are labeled on top of each lane, with APC3 and Cdh1ΔC as controls in the left four lanes of each gel. ΔC: IR-tail deleted Cdh1; fl: full length Cdh1; wt: wild type APC3. All the bands marked by asterisks were identified by mass spectrum. The red asterisks mark the bands electrophoresed from the expected proteins. The yellow asterisk labels containment protein bands of Tublin beta-1, which co-migrated with Cdh1fl protein through SDS-PAGE. The green asterisk indicates the breakdown of Cdh1ΔC. The spheres and triangles below the gels (top and middle panel) label the positions of the corresponding APC3 mutants in the WaterLOGSY spectrum (c), with triangles indicating the mutants of subtle or no effect towards TAME binding, and spheres marking the defective ones.

(b) Structure illustrations of similarity of the compound TAME to the IR-tail (left), and APC3 binding groove surface (right). The groups/residues in TAME/IR-tail are labeled in green color, with numbers marking the corresponding proton signals/peaks of the WaterLOGSY spectrum (c). Blue and red colors represent the electrostatic potential of the binding groove surface as positive and negative, respectively. Two binding centers: a negatively charged and a hydrophobic pocket at the surface are outlined with green and cyan dashed lines, respectively. The cyan colored numbers (1,2,6) indicate the corresponding protons potentially interact with the hydrophobic pocket (cyan dashed lines), while the green numbers (5,7) suggest the proton signals (of the Arginine side chain) are potentially affected by the negatively charged pocket (green dashed lines).

(c) One-dimensional WaterLOGSY spectrum recorded interactions of the APC3 mutants with TAME in the complexes of APC7-APC3mutants-APC16. All the spectra were scaled with the same magnitude. Three reference spectra include the top one of 2μM APC7-APC3wild type-APC16 with 200μM TAME and bottom two of 200μM TAME alone, and 2μM Cdh1-Hsl1<sup>30</sup> with 200μM TAME, respectively. The ten spectra in the middle are the 2μM APC7-APC3mutants-APC16 in the presence of 200μM TAME. The peaks labeled out by cyan, orange and green numbers are the signals of the protons in TAME, as labeled in (b). The cyan numbers (1,2,6) and green numbers (5,7) indicate the binding potentially affected by the mutants in the hydrophobic pocket (b, cyan dashed lines) and mutants in the negatively charged pocket (b, green dashed lines), respectively. The spectra were acquired with 10920 scans and protein signals were destroyed with the design of WaterLOGSY pulse sequence.





**Figure 2-11. (Continued).**



## **A hypothesis of APC3 recruiting the IR-tail and implication for future study**

The helix  $\alpha$ -11 binds to the groove through one lysine (Lys690) and one Leucine, and they bind to two pockets with multiple interactions. Besides, these two residues are structurally close to each other (**Figure 2-10c**). Lysine is structurally similar to Arginine and Leucine is the isoform of Isoleucine. In addition, the binding pockets consist of highly conserved residues across species, especially the negatively charged pocket (**Figure 2-10a**). The sequence conservation indicates that hydrogen bonds, electrostatic and hydrophobic interaction are evolutionarily important to bind the IR tail. Lys690 and Leu688 of helix  $\alpha$ -11 potentially mimic such interactions.

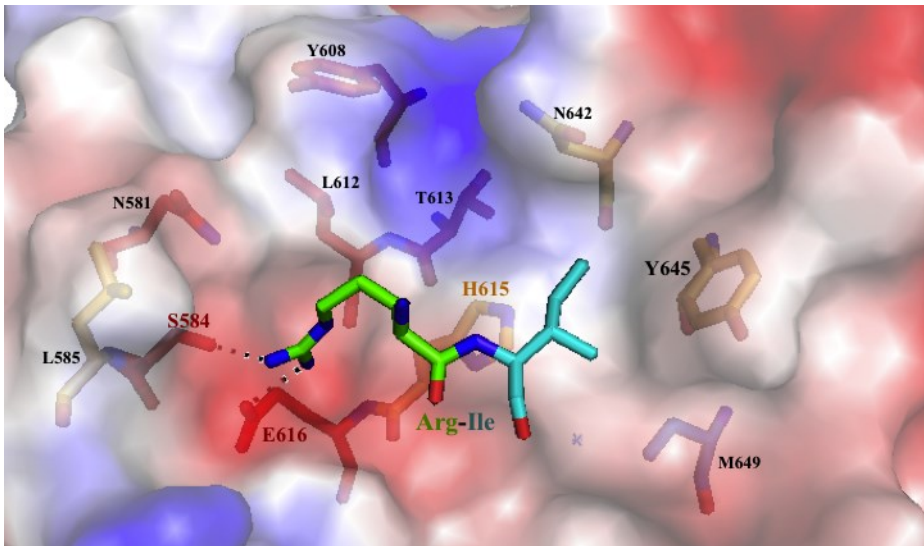
The first/negative charged pocket is surrounded with four conserved hydrophobic residues: Leu585, Tyr608, Leu612, His615 and one polar residue: Asn581, mapped onto helices 8a and 9a. Potentially they stabilize the helix  $\alpha$ -11 in the binding groove by contacting Leu683 and Ala687 of the helix  $\alpha$ -11 through hydrophobic interactions (**Figure 2-10c**). APC3 mutant N581A/L612A affected the hydrophobic interaction with TAME (Ile-Arg mimic compound) mediated by C $\beta$ , C $\gamma$ , C $\delta$  of TAME Arg side chain. IR-tail mutated to IK-tail leads to Cdh1 failing to associate with or activate APC (38). The protonated Arginine side chain has a pKa around 12.5 whereas Lysine's pKa is 10.5. The positive charge of guanidino group is extensively delocalized, and it is able to establish a stronger noncovalent interaction than Lysine.

Based on the above interaction analysis, a hypothetical model is proposed from this study (**Figure 2-12**). Arginine inserts into the first binding pocket, establishing hydrogen bonds with Ser584, and a salt bridge with Glu616. Because of the large mobility of the Arginine long side chain, the interaction contributed from surrounding residues Asn581 and Leu612, potentially restrict its side chain. Isoleucine is recruited to the second binding pocket mainly through His615 and Tyr645.

Biologically, the highly conserved residues of IR tails and APC3 binding groove indicate their interaction patterns are consistent for all species. IR tail deletion and mutation both reduced the co-activators affinity and their ability to activate APC. The IR tail interaction will anchor the co-activators and APC10 to the right location of APC3. Given the orientation change of the co-activators during the APC activity regulation and ubiquitination event (31, 32), its probably essential for the co-activators to be recruited to the correct location before being involved in substrate recruitment and poly-ubiquitin chain formation. The crystallographic model of APC3 with co-activators or APC10 will provide an accurate model to understand the interaction details. Alternatively, it could also be obtained through small compound TAME docking into APC3 binding groove, followed by the experimental validation.

## **TPR subunits models docking into APC EM map**

In the attempts to examine the authenticity of the APC3 binding groove model and study the APC assembly, APC negative-stain electron microscopy maps of the



**Figure 2-12. A hypothetical model demonstrates IR-tail interacting with the APC3 binding groove.**

A hypothetical model demonstrates the potential interactions between IR tail (green and cyan sticks) and APC3 binding groove (surface and sticks), with oxygen and nitrogen atoms of labeled residues in red and blue, respectively. The surface charge is identical as shown in **Figure 2-10b**. Dashed lines label the interactions of the salt bridge and hydrogen bond.

recombinant human APC and human APC<sup>Cdh1-Emil</sup> (EMD 2226 and 2354) (23, 72) were used for generating TPR subunit models. The Itasser predicted models of TPR subunits and APC1 were used to initiate model docking (73). N-terminal and C-terminal domains of each TPR subunit model were docked in separately. The N-terminal TPR motifs mediate dimerization and they were docked in as a dimer (instead of monomers). The TPR1-4 of both APC3 and APC7 form homodimers, whereas the N-terminal 6 TPR motifs (TPR1-6) mediate larger dimers of APC6 and APC8. The TPR4 and 5 of APC3 are separated by a disordered loop of about 270 residues, and the loop has no matched density on both EM maps. Moreover, the loops or short  $\alpha$ -helices of each model were deleted if there are no matched densities. The binding groove of the chimeric APC3 model agrees with the APC3 C-terminal predicted model with a superimposition r.m.s.d of 1.46Å (**Figure 2-6a, b**). Therefore, it is incorporated into the APC3 C-terminal predicted model for docking.

The correlation coefficients of models generated from EM map docking are listed in the **Table 2-3**. EMD 2354 (23) is the map of complex APC<sup>Cdh1-Emil</sup> and EMD 2226 (72) is from the APC sample without Cdh1. The coordinates of human APC10 and Yeast Cdh1 propeller domain were used as references and they both have higher correlation coefficients of 0.89 and 0.92, respectively. The high correlation coefficients between docked models and their corresponding densities of the EM maps indicate a good fit between the TPR models and the EM map.

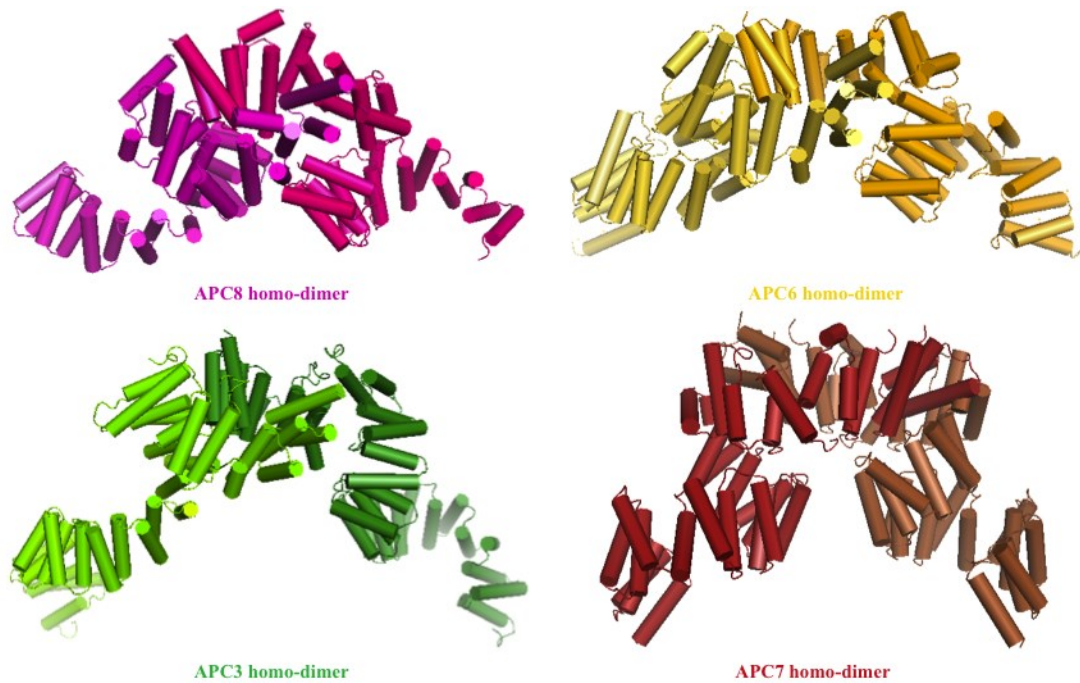
EM-derived dimers of APC7, APC3, APC6 and APC8 dimerize through their N-terminal domains and generate “V” shape molecules of all the APC TPR subunits (**Figure 2-13**). Each dimer has a unique overall shape with angles and dimensions different from others. APC3 dimer has a maximum dimension of 135 Å. Although there is a 270-residue loop separating TPR4 and 5, APC3 TPR motifs together with  $\alpha$ -helices between TPR4 and 5 adopt a linear superhelix measuring 110 Å in length and 40 Å in diameter.

All the prediction models are based on protein crystallographic or homologous modeling. N-terminal domains of human APC3 and APC6 are predicted based on the yeast APC6 atomic model (PDB code: 2XPI), whereas the human APC8 dimerization domain is based on the yeast APC8 structure (PDB code: 3ZNY). The EM-derived dimerization domains of APC3 and APC6 have 0.4Å and 0.8Å deviation with their homologous model (yeast APC6), respectively, and APC8 N-terminal domain deviates around 0.4Å from the yeast APC8 structure. The atomic structure of human APC7 TPR1-3 has been determined (PDB code: 3FFL), although the dimerization interface of the structure is biologically irrelevant. The model of APC1 PC repeats are predicted from 26S proteasome subunit Rpn2, and its overall structure looks like a closed toroid of two concentric  $\alpha$ -helical rings. The PC repeat consists of tandem copies of 35-40 amino acid repeat motifs, with each motif forming into  $\alpha$ -helix-turn- $\alpha$  helix.

TPR 1-4 of the EM-derived APC3 model were superimposed onto the corresponding TPRs of APC6, *E. cuniculi* APC3 and APC8 structures with r.m.s.d of

**Table 2-3. Correlation coefficients of EM-derived TPR structures to APC EM maps.**

<b>The models of EM docking</b>	<b>Correlation coefficient in APC<sup>Cdh1-Emi1</sup> EM map (EMD 2354)</b>	<b>Correlation coefficient in APC EM map (EMD 2226)</b>
APC7 N-terminal dimer (TPR1-4)	0.87	0.859
APC3 N-terminal dimer (TPR1-4)	0.8865	0.8868
APC6 N-terminal dimer (TPR1-6)	0.8983	0.9043
APC8 N-terminal dimer (TPR1-6)	0.8415	0.862
APC7 C-terminal domain (TPR5-C)	0.8337	0.8745
APC3 C-terminal domain (TPR5-C)	0.8706	0.858
APC6 C-terminal domain (TPR7-C)	0.8311	0.8926
APC8 C-terminal domain (TPR7-C)	0.8507	0.7949
APC1 PC repeats	0.8452	0.8131
APC10 (PDB code: 1JHJ)	0.8951	0.8905
Cdh1 (PDB code: 4BH6)	-	0.9182
Cdh1 with substrate KEN box and D box (PDB code: 4BH6)	-	0.9242



**Figure 2-13. TPR subunits form elongated V-shape homo-dimers in the human APC assembly.**

APC7, APC3, APC6, APC8 all form “V”-shaped homo-dimers and assemble in parallel in a hierarchical fashion to build the arc lamp (TPR arm) of the human APC. All the models are predicted by Itasser and adjusted based on the APC<sup>Cdh1-Emi1</sup> EM map (EMD 2354) (23). The more globular dimerization modules form the apex of the ‘V’, and the narrower C- terminal TPR superhelices project away from the dimer interfaces.

0.6Å, 2.2Å and 1.8Å, respectively. Additionally, the TPR1-4 in APC3 and APC7 structure (PDB code: 3FFL) are 2.8Å deviated from each other (**Figure A-3**). The first four TPR motifs in all EM-derived structures all adopt canonical folding and are similar to each other. The following secondary structures, TPR5-6, has different orientation and helix length among four TPR subunits and could not be aligned (**Figure 2-14a, b**). The C-terminal TPR motifs make the differences among V-shape dimers of four TPR subunits (**Figure 2-13**).

All the TPR subunit dimerization domains localize at the backside of APC complex, where all TPR proteins assemble into the TPR arm. The atomic fitting of the TPR subunits to the EM map accounts for the major density of the TPR arm (more than 90% of the density), and rationalized its repetitive layered architecture. Two copies of APC3 TPR5-14 spiral along two opposite directions to the front side of the APC, where Cdh1 and APC10 contact APC3 (**Figure 2-14c**). Each co-receptor (Cdh1 or APC10) is attached to one copy of the C-terminal APC3, respectively. In the EM-derived structures of the APC3 C-terminal TPR5-14, TPR8-10 is contributed from the chimeric APC3 atomic model. All the  $\alpha$ -helices present in the predicted models match the secondary structure prediction (**Figure 2-6**).

Although the C-terminal IR tail peptides of both Cdh1 and APC10 atomic structures are absent, EM maps demonstrate additional electronic densities which extend from the C-terminal end of APC10  $\beta$ -barrel folds. The densities connect APC10 to the APC3 binding groove TPRs (TPR8-10) and it ends at the docked IR-tail model inside the groove (**Figure 2-14d**). Potentially these densities are contributed from the C-terminal peptides of APC10. The distance between APC10 model C-terminus and Ile of the IR-tail model is around 30-40Å, and 20 residues of C-terminal APC10 will be more than enough to bridge this distance.

### **Implication of the C-terminal APC3 alternative folding**

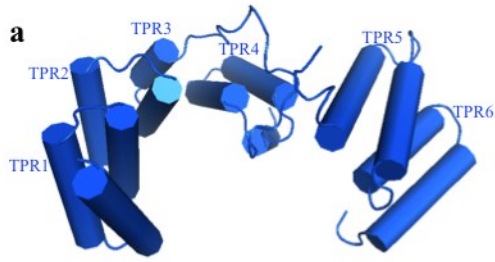
The chimeric APC3 self-lock is an interesting phenomenon during the protein folding. In order to explore the mechanism that accounts for this phenomenon, the APC3 sequence flanking the binding groove was analyzed through the comparison between EM-derived model and the chimeric APC3 model in the corresponding domains. For most canonical TPR proteins, the packing between  $\alpha$ -helices requires intra-helices interaction (20, 26, 27, 35, 37). In the EM-derived model, five residues —Val676, His680, Ile681, Val683 and Val684 — of the helix 11a (TPR11a) are involved in mediating such intra-helices interaction (**Figure 2-15a, left**) with the hydrophobic residues Tyr, Phe and Leu of TPR10. These hydrophobic interactions maintain the helix 11a packing in parallel to TPR10. However, this canonical folding pattern is disrupted in the chimeric APC3 model (**Figure 2-15a, right**), where residues Val683, Val684, Ala687, Leu688 and Lys690 interact with the binding groove and mediate the formation of the helix  $\alpha$ -11.

**Figure 2-14. Models of TPR subunits docked into the negative-stain EM map of the human APC<sup>Cdh1-Emi1</sup> complex.**

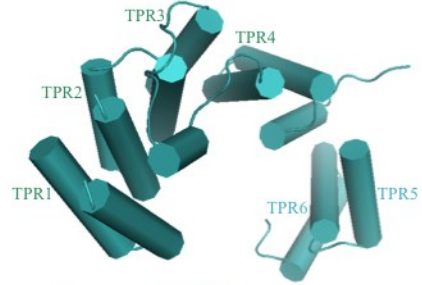
(a, b) The comparisons of EM-derived N-terminal models (TPR1-6) of human APC3, APC6 and APC8, with *E. cuniculi* APC3 N-terminal structure (a, PDB code: 3KAE)(a) and *S. pombe* APC8 N-terminal model (b, PDB code: 3ZNZ). The  $\alpha$ -helices are displayed with the shapes of cylinders in (b).

(c) Two stereo views showing the APC models closely fit in the molecular envelope of APC<sup>Cdh1-Emi1</sup> complex (EMD 2354) (23).

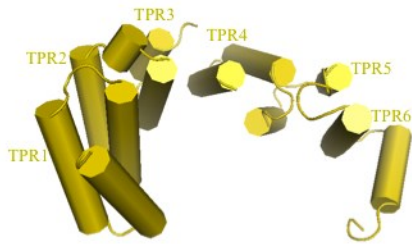
(d) Details of the interactions between APC3 and C-terminal APC10. The model of human APC10 is used for docking, with a red asterisk indicating its C-terminal end. The hypothetical IR tail interaction model (**Figure 2-12**) is incorporated into the EM-derived APC3 structure, and together they are docked into the APC<sup>Cdh1-Emi1</sup> EM map. The model of the IR tail is showed as cyan sticks, with the interacting residues labeled in red and orange sticks as showed in **Figure 2-11**. The black arrow points at the density contributed from APC10 C-terminal unstructured peptides, with the dot line representing its possible orientation. The red asterisk indicates C-terminal end of APC10  $\beta$ -barrels. The corresponding APC10 sequence to the labeled density is marked out with dotted lines at bottom. The unstructured C-terminal peptides of APC10, Cdh1 and Cdc20 all have similar number of residues.



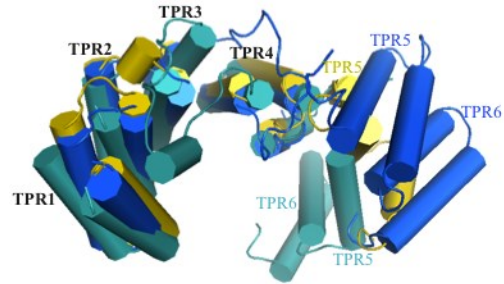
Human APC3 N-terminal TPR1-4 EM-derived model



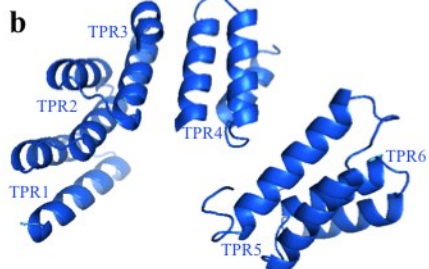
*E. cucurbit* APC3 TPR1-6 PDB code: 3KAE



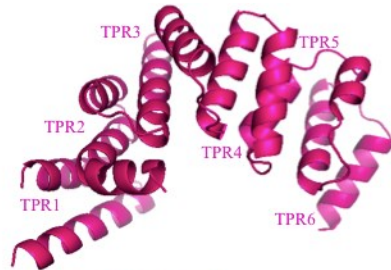
Human APC6 N-terminal TPR1-6 EM-derived model



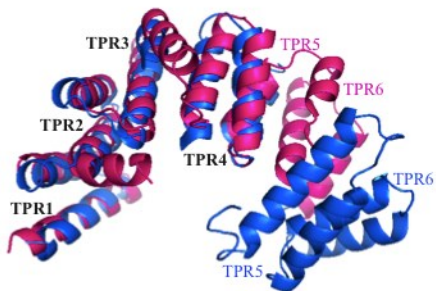
*E. cucurbit* APC3 and human APC6 superimpose onto human APC3 EM-derived model



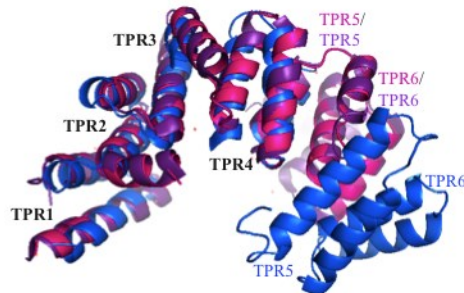
Human APC3 N-terminal TPR1-4 EM-derived model



Human APC8 N-terminal EM-derived model



Human APC8 superimpose onto human APC3 EM-derived model



Human APC8 and *S. pombe* APC8 superimpose onto human APC3 EM-derived model



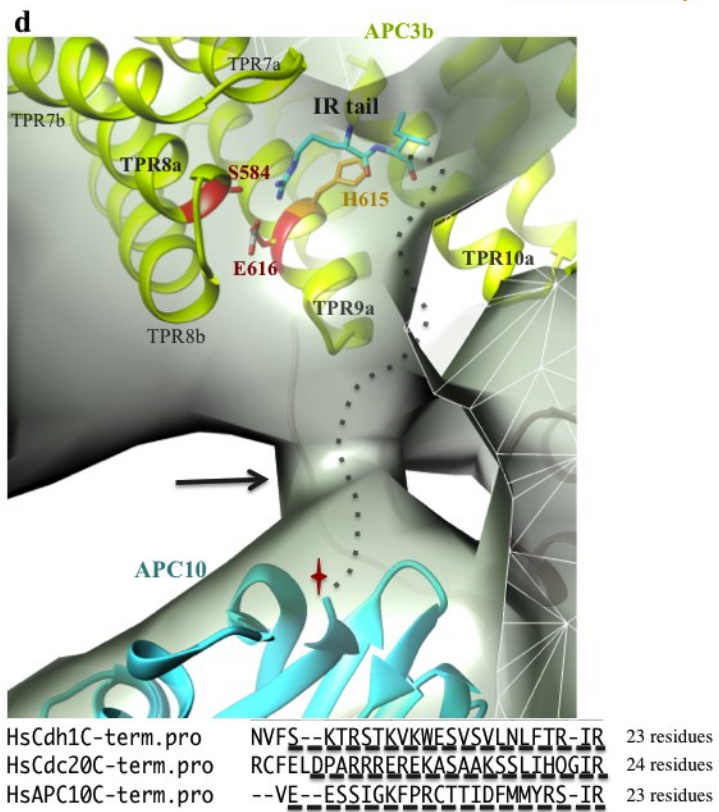
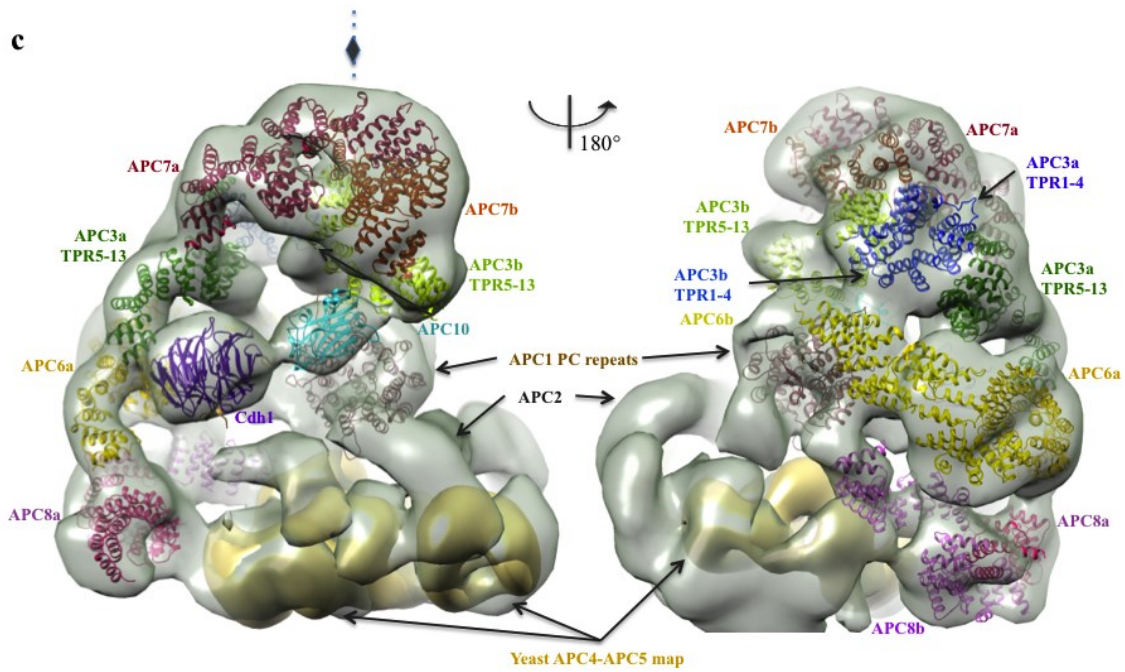
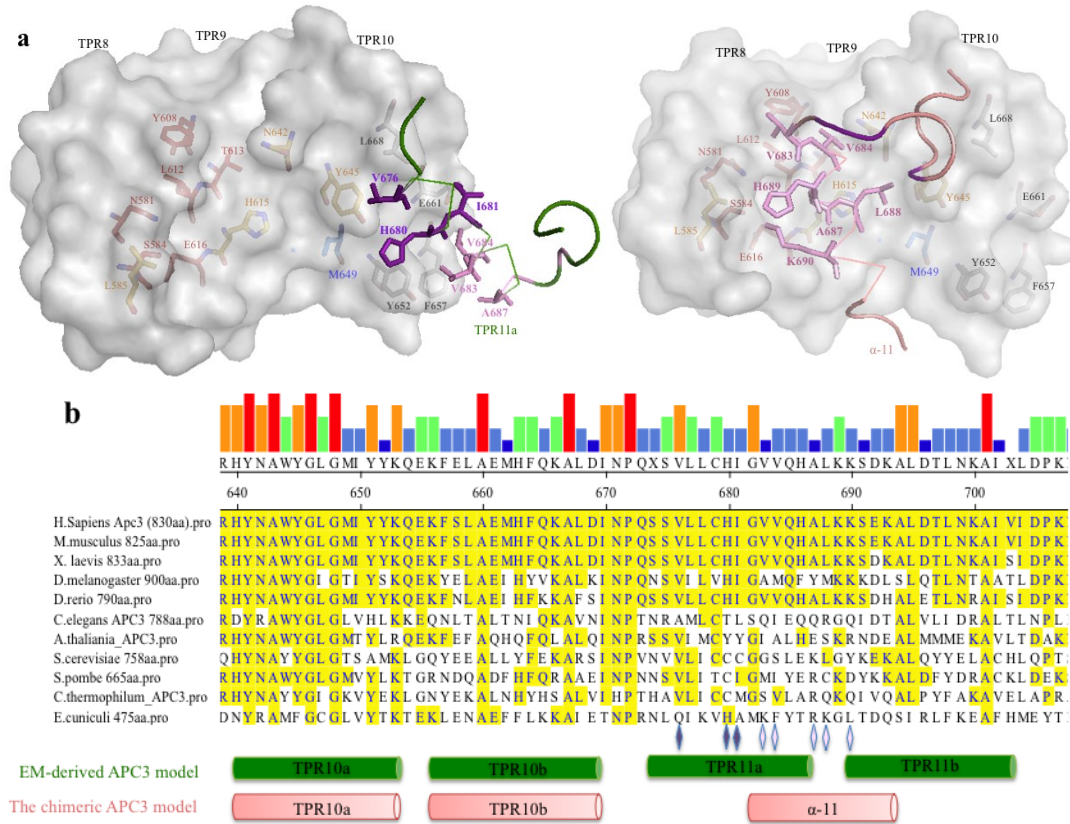


Figure 2-14. (Continued).



**Figure 2-15. The binding groove of human APC3 affects TPR11 adopts a canonical folding.**

(a) Views of TPR11 folding in EM-derived model (left panel) and the chimeric APC3 structure (right panel). The binding grooves (TPR8-10) are presented as surface density. Purple labeled residues mediating TPR11 packing in parallel with TPR10 (left) while they are in a disordered loop in chimeric APC3. Pink color denotes the residues that theatrically stabilizing the TPR11a (left) but mediating  $\alpha$ -11 formation in the chimeric APC3 (right).

(b) Multiple sequence alignment at TPR10-11 motifs of human APC3 orthologs. The blue bars above consensus indicate lower conservation strength than the red and orange bars. (Red, invariant residues; orange: conserved residues). The purple and pink diamonds denotes the residues that are colored or labeled out by sticks (with the same colors) in (a). Position of the chimeric APC3 TPR 10 and  $\alpha$ -11 are colored in salmon, and TPR10 and 11 in EM-derived APC3 model are labeled in green.

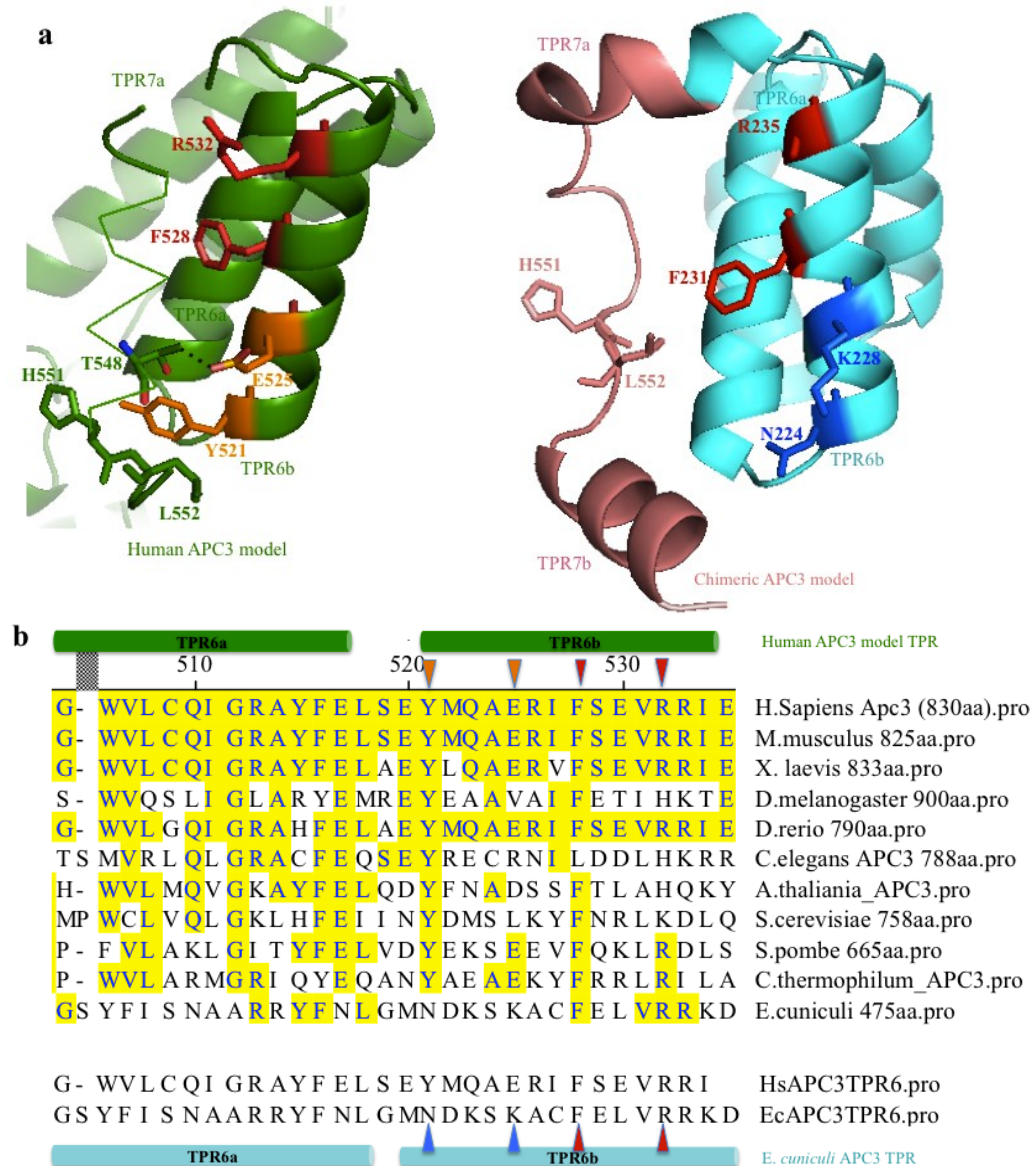
It is important to understand what triggers TPR11 folding into a pseudo binding partner instead of a canonical TPR motif, since no literature reports APC3 or TPR protein self-lock phenomenon. The sequence of APC3 TPR11 shows it is relatively less conserved compared to the sequence of other APC3 C-terminal TPR motifs (**Figure 2-15b, Figure 2-10a, Figure 2-8a**). This may suggest that this region is less stable in adopting the canonical TPR motifs. Meanwhile, the fact that TPR11 folds into another helix ( $\alpha$ -11) inside APC3 binding groove also indicates that the binding groove is capable of mediating strong hydrophobic interactions. Combining the above two aspects together, it seems human APC3 needs assistance from other domains or proteins to maintain canonical TPR protein folding.

In addition to non-canonical folding of TPR11, TPR7 in the chimeric APC3 structure doesn't keep canonical folding either. Instead it folds into a short helix-long loop-short helix secondary structure. The EM-derived APC3 model provides a clue to interpret the mechanism of the human APC3 canonical folding. The TPR7 of the EM-derived model is parallel with TPR6b. To study the factors that leads to this alternative folding, the interaction between TPR6 and TPR7, and protein sequence comparison are examined. The stereoviews of canonically folded TPR6 and 7 from the EM-derived model display the interactions between them (**Figure 2-16a**). The packing of helix 7a (TPR7a) to the helix 6b (TPR6b) is mediated through four conserved residues R532, F528, E525, Y521 of the helix 6b, with one residue assigned to one turn of the helix and lined up to face towards TPR7a (**Figure 2-16a**, left). Two of the four conserved residues: Arginine and Phenylalanine are the same as the *E. cuniculi* APC3 TPR6b at the equivalent positions (**Figure 2-16b**). The other two residues, Glutamic Acid and Tyrosine, potentially stabilize the helix 7a through hydrogen bonds with Threonine (T548 of TPR7a) and hydrophobic side-chain interactions with Histidine and Leucine (H551, L552) at the last turn of the helix 7a. However, Glutamic Acid and Tyrosine (E525, Y521) are not present in the TPR6 of *E. cuniculi* APC3. The disrupted interactions potentially cause the lost the canonical TPR folding of the helix 7a. Instead it becomes a loop in the structure, with Histidine and Leucine (H551, L552) exposed in the solvent (**Figure 2-16a**, right).

Unlike the TPR11, there is no other contact to TPR7 from other helices inside APC3 or other subunits of APC (demonstrated through EM map docking). It seems that TPR7 adopts correct folding in a TPR6-dependent manner, and TPR7 requires the whole set of conserved residues from its neighbor helix to stabilize its correct secondary structures. The sequence of both TPR7 and TPR11 are less conserved than those of the binding groove. Without the correct stabilizing interactions around TPR7, it forms into two shorter helices within the TPR motif boundary instead of a canonical  $\alpha$ -helix-turn- $\alpha$ -helix motif (**Figure 2-16a**).

### **Biological meaning of APC3 folding and implication to future study**

The protein superhelical conformation was first discovered in Phosphatase 5 and it is believed to coordinate multi-subunit assembly (74). The canonically folded APC3



**Figure 2-16. Human APC3 TPR7 folds in a TPR6-dependent manner.**

(a) Two stereoviews to demonstrate the interaction between the helices TPR6b and TPR7a in the EM-derived APC3 model (left), and the corresponding but disrupted interactions in the chimeric APC3 model (right).

(b) The cross-species comparison of sequence corresponding to the TPR6 motif of APC3. Red residues are conserved in both the human and *E. cuniculi* APC3; orange and blue residues vary. The yellow highlights mark residues identical among different species.

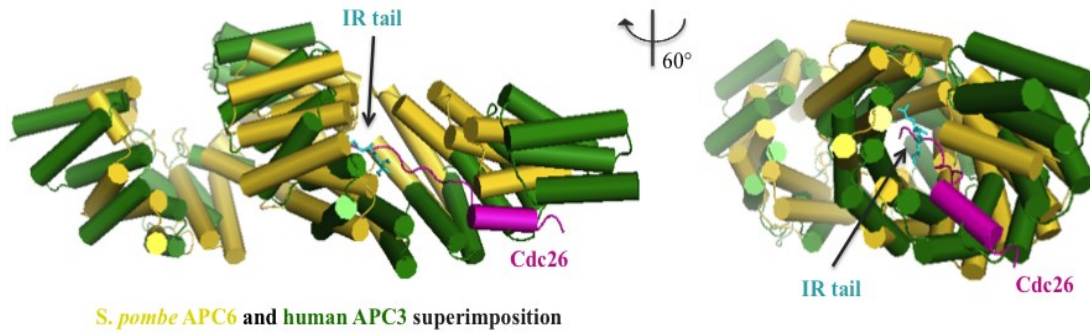
adopts the right-handed superhelical structure. The superhelix makes it possible to establish new contacts among sequence distant helices.

It is also interesting that the independently folded APC3 binding groove is flanked by two TPR motifs (TPR 7 and 11) whose secondary structures are sensitive to the preceding TPR motifs and/or the tertiary structure. The human APC3 seems to require the whole C-terminal TPR domain (TPR5-13) to adopt and maintain the canonical folding. APC3 itself is located in a highly hydrophobic microenvironment (the TPR arm) that is created from all other TPR subunits. The inner surface of the binding groove is encircled by the superhelix of TPR5-11, which potentially repel unexpected helices interacting with the binding groove but keep the groove accessible for the unstructured IR tail peptides, or small molecules like TAME. Given that APC3 is sandwiched in the TPR arm by another two subunits: APC7 (top) and APC6 (bottom), it is probably critical to prevent the nonspecific interactions (from unexpected  $\alpha$ -helices) locking the binding groove. Similar to APC3, APC6 also localizes in the hydrophobic microenvironment, and C-terminal superhelices of APC3 and APC6 have similar dimensions (**Figure 2-17**). The C-terminal APC6 accommodates the N-terminal peptide of Cdc26, which is the key to stabilize APC6 superhelical structure (26, 27). The APC3 superhelix potentially limits the space around the binding groove surface and this limited space is more accessible to peptides or small compounds than  $\alpha$ -helices.

Although the crystallographic model of the chimeric APC3 adopts a self-lock structure, it reveals reliable surface details of human APC3 binding groove, for further recruitment mechanism study. The self-lock structure also partially displays the binding groove mediated interaction. Being incorporated by the chimeric APC3 structure, the EM-derived model becomes a guide for further assay design to validate interactions of APC3 with the co-activators and APC10. This information could also benefit APC inhibitor, like the compound TAME, optimization to improve their affinity to the binding groove. Additionally, this model provides clues that allow speculation the role of APC3 tertiary structure in maintaining its function.

### **Crystallization attempts of APC TPR subunits**

Crystallizations of other APC TPR subunits were tried (approached), as listed in **Table A-1**. It includes the complex of the full-length proteins APC8-APC6-Cdc26 from different species, and a truncated APC3 in the complex of APC3 $\Delta$ L-APC16 with or without Cdh1. The full-length complex APC8-APC6-Cdc26 didn't crystallize, but the APC3 $\Delta$ L-APC16 crystallized with or without Cdh1 (**Figure A-4, Figure A-5**). Both crystals showed the optimization potential by diffracting to 6Å and 7Å, respectively. Until the time this dissertation is written, they are still being pursued.



**Figure 2-17. The superhelix structure limits the accessibility of the APC3 binding groove.**

Structural superimposition of *S. pombe* APC6 (yellow, PDB code: 2XPI) and EM-derived human APC3 (green) demonstrates that two structures have similar dimensions of superhelices. The docked IR tail model is shown as a cyan stick. Both IR tail and N-terminal Cdc26 are buried inside the groove of the superhelices.

## Materials and Methods

### Constructs for crystallization

All the protein expression constructs were made by standard PCR/ligation procedures, and sequences were verified by automated sequencing procedures. The APC3 deletion and mutation were designed with reference to the human APC3 830 a.a sequence isoform (uniprot code: P30260). The *E. cuciniculi* APC3 sequence used for designing the APC3 chimeras refers to the reported one in (35) (uniprot code: Q8SQV4). The TPR7-14 of the human APC3 refers to residues 539-830 of the full-length protein. TPR1-6 of *E. cuciniculi* APC3 refers to residues 1-241 and TPR 5-6 refers to 175-241 of the parasite APC3 sequence. The chimeric APC3s were cloned to an N-terminal His6-MBP- tagged insect or bacterial expression vector, and a GST or GST-MBP tagged bacterial vector.

The crystallization tags, including an enhanced green fluorescent protein (EGFP), a T4 lysozyme and a MBP-(AAA), used for crystallizing the chimeric APC3s were inserted between the N-terminal GST tag and the chimeric APC3. The EGFP residues 1-230 were used for crystallization (uniprot code: C5MKY7), and the T4 lysozyme (uniprot code: P00720) was linked to the N-terminus of the chimeric APC3s with three Alanines residues, similar to the MBP-(AAA) tag. The MBP-(AAA) is the crystallization version of a maltose binding protein (PDB code: 3RUM\_A). Alternative APC3 chimera constructs were the TPR7-14 of two yeasts and a worm APC3 fused to the *E. cuciniculi* APC3 TPR1-6 and TPR5-6, respectively. The TPR7-14 of different species refers to the residues: *S. pombe* 397-665 (NCBI code: NP\_594604.2), *S. cerevisiae* 506-758 (uniprot code: P38042), *C. elegans* 527-788 (NCBI code: NP\_001021714). The chimeric APC3 mutant refers to the mutations of V181K and I185D of the *E. cuciniculi* APC3.

The Cdh1 truncation and deletion constructs have been described in the Material and Methods of Chapter three. The human Cdh1, Cdc20 and APC10 C-terminal peptide (also named the IR-tail peptides) constructs were made by the above cloning procedure. The Cdh1 IR-tail peptide refers to the residues 479-496. Protein sequences of the human Cdc20 and APC10 are referred from Uniprot (uniprot code: Q12834, Q9UM13). The APC10 IR peptides have two versions: residues 161-185 and 171-185. The Cdc20 IR-tail peptides include residues 474-499. All of the peptides were fused to an N-terminal GST, His6-MBP-, and GST-MBP-(AAA) within the bacterial expression vectors.

The bacterial expression vectors for the N-terminal GST-, His6-MBP-, GST-MBP, or His6- fusions were from either a regular or modified version of pGEX4T1 (GE lifesciences), pRSF1b (Novagen) vectors. The insect cell expression vectors for the above fusions were constructed from the vectors pFastBac1, pFastBac-GST, and modified pFastBac-His6MBP (Invitrogen).

## Protein expression and purification from bacteria

The *E. coli* expression strains, BL21(DE3) Gold (Novagen), codon enhanced RIL (Stratagene) and Rosetta (Novagen) were used for expression tests of the protein constructs. Most of the large-scale protein production from the bacteria was done in BL21(DE3) Gold, with the cells cultured in LB broth Miller (EMD) supplemented with the appropriate antibiotics.

The chimeric APC3 proteins were expressed from the bacterial vector pGEX-4T1 (GE). The starter media was incubated at 37°C overnight in a rotary shaker at 200 revolutions per minute (rpm). The overnight culture 10 ml was then used to inoculate into 12L LB media that contains a final concentration of 200 µg/ml ampicillin. The BL21(DE3) Gold strain harboring the expression constructs were cultured at 37°C until reaching an OD600 of ~0.8. After cooling the media to 18°C for 1hr, the expression was induced by the addition of IPTG (isopropyl β-D-1-thiogalactopyranoside) to a final concentration of 0.6 mM, followed by overnight culture at 18°C. The next day, the cells were resuspended in a buffer of 50 mM Tris-HCl pH7.6, 0.3 M NaCl, 5 mM DTT, and supplemented with 2.5 mM PMSF, and lysed by sonication on ice. The GST-MBP tagged chimeric APC3s were purified by glutathione-affinity chromatography, and eluted protein fractions were treated with the homemade TEV protease at 4°C overnight at 1:50 ratio. The cleavage by TEV protease left 2 extra residues (Gly-Ser) on the N-terminus of proteins prior to the protein linker Glu-Phe-Ser-Gly.

For co-crystallization, the chimeric APC3s were purified by Ion Exchange chromatography (IEC) to separate the cleaved affinity tags from the purified proteins, based on the isoelectric point difference. IEC is done with a 5ml Hi-Trap column (GE) in a buffer of 50mM Tris-Cl pH 8.0, and the protein fractions were concentrated to 20-40mg/ml (Bio-Rad Protein Assay), followed by a gel filtration chromatography. The gel filtration was performed to further remove contaminants based on the molecular shapes and sizes, using a Superose 6 column (GE) in a buffer of 50 mM Tris-HCl pH 7.6, 0.3 M NaCl and 5 mM DTT. After gel filtration, the remaining GST-MBP and uncleaved chimeric APC3 protein were removed with glutathione-affinity resins before the purified protein was concentrated to 20-25 mg/ml (A280 measured), aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C until latter use.

The Selenomethionine labeled chimeric APC3 was expressed in the BL21(DE3) Gold cells cultured in the autoinduction media. The cells grew in a 24L-culture at 37°C till they passed the exponential phase and then the temperature was reduced to 18°C overnight for protein production.. The SeMet incorporated protein was purified through GST chromatography. Before the 10mM glutathione elution, a wash buffer containing 25mM ATP was incubated twice with protein-bound glutathione sepharose/resins for 15 mins/per wash, followed by the additional resin wash to remove the extra ATP. The elutions of the SeMet chimeric APC3 were applied the same purification procedure as the native protein. Following size exclusion chromatography, the protein was concentrated and stored in a buffer containing 50 mM Tris-Cl pH 7.6, 300 mM NaCl and 5mM DTT.



The protein concentration used for crystallography varied within the range of 12.5-25 mg/ml (A280 measured).

### Crystal screening and optimization

Crystallography screenings were performed using a Mosquito crystallization robot (TTP Labtech) with the commercial 96-well screens. Over 2,000 commercially available and homemade conditions were screened at 4°C and room temperature (RT), and one initial hit was found two days later in needle shaped clusters. This hit grew in the condition of 0.1M MES pH6.5, 1.6M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, and 10% (v/v) Dioxane at RT. All the robotic screening used the hanging drop vapor diffusion method against 100µl of well solution with a drop ratio 1:1 (200 nl protein: 200 nl well solution), with a protein concentration of 15-20mg/ml.

To improve crystal size, manual screenings were performed using the ratios of 1:1, 1:2 and 2:1 (1µl:1µl or 2µl:2µl, 1µl:2µl, and 2µl:1µl protein:well solution). Crystal drops were equilibrated by vapor diffusion against 500 µl-1ml of precipitant solutions at RT. Meanwhile, I optimized the screening around the initial conditions with different precipitant concentrations, different pH buffers, and different commercial and homemade additives under different temperatures, combining with the streak and micro seedings. The optimization screenings were summarized in the **Table 2-3**. The optimized condition of 0.1M MES pH6.5, 1.8M (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 2% (v/v) Dioxane and 2% MPD tripled the crystal size. Multiple rounds of cryo buffer screenings were performed to obtain qualified diffractions. More than 50 cryo buffers of 23 cryoprotectants were screened, and the best cryo-buffer is a buffer of 0.1M MES 6.5, 10% Dioxane, 1M AmSO<sub>4</sub>, and 6.5M Am-formate.

Another attempt to obtain high quality diffractions was to search new crystal forms. Around 1,000 commercial conditions were used as additives (10% commercial conditions +90% the optimized hit condition) and also as new precipitants (90% commercial +10% the optimized hit condition). MPD was screened out as a new precipitant that supports crystals rapidly growing larger (the crystals grew across a drop), in either long needle shaped clusters or thin plates. The optimized conditions of 0.1M MES 6.5/ 0.1M HEPES pH 7.0/ 0.1M Bicine pH 8.0, 0.2M AmSO<sub>4</sub>, and 48-51% MPD generated SeMet and native crystals with a length of ~0.4mm and in a shape of long-needles or thin-quadrilaterals. Interestingly, the larger crystals were dehydrated and quickly dissolved within two minutes when exposed to the air. The flash-frozen crystals in the drop solution (50%MPD is a cryoprotectant) diffracted to 4Å but in an anisotropic diffraction pattern (a diffraction pattern of stripes instead of lunes).

The further attempts focused on improving the diffraction quality of the SeMet crystals. The solution failed to be calculated at the beginning because of the complexity that is contributed from both 70 Selenium sites inside one asymmetric unit (ASU) and the overlapped crystal lattices. The streak seeding, macro-seeding and lower temperature (to slow down crystals growing) didn't help the crystals grow single. Eventually the single-

lattice dataset was collected from a small crystal (~0.2mm) that was gently dissected from a 0.4mm long/large crystal, growing in the condition of 0.1M MES pH6.5, 1% Dioxane, 1.45M AmSO<sub>4</sub>, 2% MPD, 0.2M MgCl<sub>2</sub> and 0.1M Li<sub>2</sub>SO<sub>4</sub>, micro seeded from native crystals in a ratio of 1:1000 at 18°C.

During the phase information pursuit, alternative approaches including heavy atom co-crystallization and soaking were also applied. The co-crystallization is performed manually by setting the screenings with the mixture of the protein and heavy atoms. The soaking is to soak native crystals in different heavy atoms, combined with the optimization of crystallization or cryo buffers, temperatures and wash times to remove the background heavy atoms. The soaked crystals were harvested for data collection at APS (Advanced Photon Source) and ALS (Advanced Light Source), and collected diffraction data were used for the heavy atom incorporation analysis.

### **Crystallization of the chimeric APC3 modified by reductive methylation**

Reductive methylation was applied to improve the solubility of the chimeric APC3 through dimethylation of all the surface-exposed lysines (75, 76). In order to get the complete methylation, the purified chimeric APC3 was thawed and dialyzed in a buffer of 50mM HEPES pH7.5, 0.3M NaCl and 5mM DTT to replace the protein storage (Tris) buffer. Freshly prepared 1M dimethyl amine borane complex (DMAB) and 1M formaldehyde were gently added into the protein solution, with the amount of 20µl DMAB per ml protein and 40µl formaldehyde per ml protein. This mix was gently repeated once and the protein was incubated at 4°C for 2 hours to be methylated. Then a final aliquot of 10µl DMAB per ml protein was added into the mixture to methylate the protein overnight. The methylation mixture was gel filtrated the next day with a Superose6 column in a buffer of 50mM Tris-Cl pH7.6, 0.3M NaCl, 5mM DTT to quench the reaction and remove the excess DMAB and formaldehyde. The elution was examined by SDS-PAGE and then was concentrated to 12.5-25 mg/ml (A280 measured), aliquots, flash-frozen in liquid nitrogen, and stored at -80°C. In order to confirm the protein methylation, one aliquot of the frozen protein was analyzed with Intact Massspectrometry to detect the protein mass changes contributed from the methyl groups. The chimeric APC3 had 6-7 Lysine methylated based on the Intact Massspectrometry results.

The purified Lys-methylated protein was thawed on ice and robotic screening of the protein was performed with the commercial 96-well screens. Around 1,000 commercial conditions were screened at 4°C and RT, respectively. One initial hit was generated in the condition of 0.1M CAPS pH10.5 and 40% MPD at 4°C. The following optimization was performed through robotic and manual screening, but the crystallization couldn't be optimized.

## **Crystallization of the T4 Lysozyme, EGFP and MBP-(AAA) fused chimeric APC3**

Robotic screenings were performed on the T4Lysozyme, EGFP and MBP-(AAA) fused chimeric APC3. The proteins were concentrated to 15-20mg/ml (A280 measured) after the purification from the affinity chromatography, ion exchange and size exclusion chromatography. For each protein, 2,000 of total commercially available conditions were screened at 4°C and RT. Only the MBP-(AAA) fusion was found two hits in the conditions of 0.1M BTP pH7.5, 20% PEG3350, and 0.2M NaI or KSCN. The MBP-(AAA) fusion mixed with maltose of a final concentration of 1mM before screening. The crystals grew in a shape of small rods (shorter than 0.1mm) or small plates. The PEG3350 of a serial of concentrations in the pH 6.5-8.5 was manually screened with the 0.2M additives of NaI or KSCN, using 96-well homemade blocks. The best crystallization buffer turned out to be Tris-Cl pH7.5. More additives screenings were performed using the commercial additive screens, and 20 additives out of all positive hits were further screened manual (with the original hit condition) to optimize the crystal size. The crystals grew into the shape of longer rods (0.1-0.2mm) or rhomboids, and were harvested for data collection at Argonne national laboratory (APS), but no crystals diffracted.

### **Data collection**

The diffraction datasets of high quality were collected at APS, 24-ID-C. The dataset was indexed into a space group of C2, and integrated and scaled with HKL2000. A fluorescence scan was performed prior to data collection (to another crystal) to determine the peak energy/wavelength for Selenium. The dataset was collected at single wavelength of  $\lambda = 0.9798\text{\AA}$  with an exposure time of 0.5 seconds per frame, and a 0.5 degree oscillation angle for 360 degree collection. The single-wavelength anomalous dispersion (SAD) benefits the low anomalous signals to be detected and scaled.

For heavy atoms screening, both the SAD and multiple-wavelength anomalous dispersion (MAD) were applied. A few fluorescence scans prior to data collection were performed to detect the heavy atom incorporation and optimize the collection strategy. The profile of the scans would provide the readouts of the absorption vector  $f'$ , dispersion vector  $f''$  and the collection wavelenghtes for MAD.

### **Solution and phase determination**

The crystallographic structure program phenix.autosol was used to generate the solution of the collected dataset and phase the chimeric APC3 structure, through a process of automatically determining the Selenium positions, calculating and refining the phases, and finally generating an electron density map. The scale file of the input for Autosol was from HKL2000 and included the information of a space group (C2), unit cell parameters, redundancy (2.2), wavelength (0.9798Å) and the scattering factors ( $f'$  -8.0 and  $f''$  4.5). In order to obtain enough information for the solution calculation, the

collected dataset was initially scaled at 2.8Å with a lower Isigma/error of 1. Autosol found 70 out of 72 expected Selenium sites and 58 sites have higher than 0.5 occupancy scores. In spite of a marginal value of 0.34 of the figure of merit (FOM, the marginal range is 0.25 - 0.45), the abundant heavy atoms sites provided a relatively strong phase, and led to a continuous experimental map with the clear solvent boundaries. The initial map has a deviation value (skew) of 0.16 at electron density distribution.

## Structure refinement

The initial structural model of the chimeric APC3 was built *de novo* in the program Autosol, with poly-Alanines modeled in the initial experimental map, followed by the protein sequence docking into the peptide chains. A refinement procedure was automatically initiated after docking and generated a refined map with a built-in model. The initial model has an Rfree of 0.30. This was followed by repeated cycles of manual model building in Coot (70). Then the side chains were placed into the autosol refined experimental map, and large errors from the initial models were manually fixed. The iterative refinement was performed with phenix.refine (77). The refinement aimed at two cycles of xyz coordinates, real-space, global non-crystallographic symmetry (NCS) restraints, and automatic Asn/Gln/His errors correction. Those refinement parameters are all suitable for diffraction of moderate resolution (lower than 2.5Å). R-free dropped from 0.3 to 0.28, which is calculated from randomly selected five percent of all the independently measured reflections to prevent the artifact errors introduced by models. The refined 2Fo-Fc and Fo-Fc Fourier maps, combining with the ramachandran constraints were used for correcting the model errors from the side chains and backbones.

The optimization of x-ray term/stereochemistry and x-ray ADP were incorporated into the following 3 cycles of refinement, and 70 water molecules were added/updated to the structural model (with the minimal resolution of 3.2Å). The refined model containing the newly added water molecules was manually checked after the 3 cycles of refinement. Residues of the model and the water molecules that have less than 30% occupancies were removed or side-chains truncated before the final refinement cycle. The final cycle of refinement of the chimeric APC3 was carried out with water molecules by minimization refinement (occupancies) and individual B-factor refinement, resulting in an Rwork and Rfree of 22.24% and 25.49%, respectively. All the statistic details of the refined structure are available in **Table 2-2**.

The Ramachandran plot was generated to check the main-chain torsion angles of structure model by program Coot (78). The Ramachandran plot has 98% of residues in the most favored region and 0.15% in the disallowed region.

## Co-pulldown assays

The binding capacity tests of the chimeric APC3 to Cdh1 were performed through the insect cells Hi-5 *in vitro* expression. The chimeric APC3s (both long and short

version) were fused to a N-terminal His6-MBP tag and cloned to the insect cell expression vector pFastBac. Both the full-length APC3 and APC3 $\Delta$ L constructs were tagged with the GST tags. The constructs of Cdh1 and Cdh1<sup>WDAC</sup> refer to the full-length protein and Cdh1 residues 162-462, respectively. Both these constructs were tagged with a N-terminal 3Myc-His6 and the Myc tag refers to a peptide sequence of N-EQKLISEEDL-C. Hsl1<sup>30</sup> refers to yeast Hsl1 residues 770-790 fused to residues 818-842, and together this sequence reconstitutes the substrate KEN and D-box motifs, which are recognized by the APC co-receptors. Hsl1<sup>30</sup> was tagged with a N-terminal twin strep tag (two copies), and each strep tag consists of residues N-WSHPQFEK-C. All the sequences of these constructs were verified by DNA sequencing.

The pFastBac constructs were then transformed into the competent cells of *E. coli* EMBacY strain (79), which generated the recombinant Bacmids (a baculovirus shuttle vector) to carry the genes for co-pulldown experiments. The transformed competent cells were cultured in 1ml Super Optimal Broth (SOC) media at 37°C for 6hrs before they were plated onto LB-agar plates containing 100 $\mu$ g/ml ampicillin, 50 $\mu$ g/ml kanamycin, 10 $\mu$ g/ml tetracycline and 7 $\mu$ g/ml gentamycin. All the LB-agar plates were supplemented with 100 $\mu$ g/ml Xgal and 40 $\mu$ g/ml IPTG in advance to enable the lacZ blue-white selection to identify the white colonies being incorporated with the gene of interest. The white colonies were re-streaked onto fresh plates with the same supplements and incubated at 37°C for another day. The colony-PCR valued positive colonies were amplified overnight for preparing Bacmid DNA. The over-night culture medium kept the same antibiotics as the LB-agar plates.

After cells were harvested from the over-night cultures, the cell pellets were resuspended in 300 $\mu$ l of Qiagen buffer P1, lysed by addition of 300 $\mu$ l of Qiagen buffer P2 at room temperature for five minutes and recovered on ice for ten minutes by incubating with another additional 300 $\mu$ l of Qiagen buffer P3. The precipitation was pelleted at room temperature by centrifugation at 13,000rpm for 10 minutes. The cleared lysate 700 $\mu$ l was added to 800 $\mu$ l of ice cold isopropanol and incubated for 5 minutes on dry ice. The precipitated Bacmid DNA was pelleted at room temperature by centrifugation at 13,000rpm for 10 minutes and the pelleted DNA was washed with 70% (v/v) Ethanol. The clean Bacmid DNA was resuspended in 50 $\mu$ l of sterile elution buffer (Qiagen) inside a laminar flow hood (sterile conditions).

Insect cell Sf9 was used to produce and amplify the baculovirus of infection. Bacmid DNA 1-2 $\mu$ g were diluted into 100 $\mu$ l serum-free media and then mixed with 10 $\mu$ l of FuGene HD transfection reagent. The transfection mix was then incubated at 27°C for 30 minutes before being diluted by 1ml serum-free medium. The diluted mix was then added into a 6-well plate that is freshly coated with  $2 \times 10^6$  healthy Sf9 cells and was incubated with Sf9 cells at 27°C for 6 hours. After transfection, another 2ml of serum-free medium was added back into the Sf9 cells for continuous incubation for 4-5 days at 27°C. The baculovirus were harvested when infected Sf9 cells began to produce green fluorescence signal (under UV excitation). The harvested virus was the 1<sup>st</sup> generation and it needed to be amplified for another two rounds before infection. The virus was amplified in 25mm cell culture dishes with each dish freshly coated by  $\sim 20 \times 10^6$  healthy

Sf9 cells. Each amplification mix was made by mixing 500µl of 1<sup>st</sup> generation virus with 5ml of serum-free medium and then the mix was added into 25mm dishes before incubation at 27°C for 2 hours. After 2 hours, another 15ml of serum-free medium was added into each dish for 3-day incubation at 27°C before 2<sup>nd</sup> generation of baculovirus was harvested. The 3<sup>rd</sup> generation of baculovirus was prepared with the same procedure.

Co-infections were performed on Hi-5 insect cells at a density of 10-20×10<sup>6</sup> cells/ml by adding the 3<sup>rd</sup> generation of baculovirus. The infected cells were cultured at 27°C for 2 hours at a speed of 120 revolutions per min (rpm) and then followed by incubation at 20°C for 2 days. The co-infection was set up in 200ml for each co-pulldown, and the baculovirus used for infection was kept at the ratio of 1:100 (the volume of viruses: infection). The infected Hi-5 cells were harvested and cell pellets were resuspended in a buffer of 50mM Tris-HCl pH7.6, 0.2M NaCl and 5mM DTT (wash buffer), supplemented with 10ug/ml Aprotinin, 5ug/ml Leupeptin, one tablet of the protease inhibitor per 50ml of the wash buffer (Sigma) and 2.5 mM PMSF, and lysed by sonication on ice. The lysates were spun at 4°C twice, at 15,000rpm for 20 minutes and then the supernatant were incubated with strep-tactin sepharose (IBA) on a rotary shaker at 4°C for 1 hour. Half-milliliter slurry of the strep-tactin beads (0.25ml) was used for each co-pulldown. The protein-bound strep beads were then washed twice with two bead-volume of the wash buffer before loading on the column, followed by a column wash with 20 bead-volumes of the wash buffer. The proteins of each co-pulldown were eventually removed from the beads with an elution buffer of 50mM Tris-HCl pH7.6, 0.2M NaCl, 5mM DTT and 2.5mM dethiobiotin, and each elution was examined by SDS-PAGE.

### **The models of the APC subunits docking in the APC EM maps**

Human APC10 model (PDB code: 1JHJ) was used for docking in APC10 density with the N-terminal 9 residues removed from the atomic coordinates. The N-terminal loop (9 residues) has no matching density in both EM maps, EM2226 and EM 2354. The density of the Cdh1 propeller domain was docked with *S. cerevisiae* Cdh1 molecular model (PDB code: 4BH6), with the coordinates of the Cdh1 inhibitor Acm1 removed. The coordinates of the inhibitor KEN box and D-box were kept with the Cdh1 docking model. The WD40 propeller domains of the human and yeast Cdh1 had a high sequence identity of 49%. All the models for the EM map docking are originally predicted by I-tasser server (73). The TPR7-14 of the human APC3 was modeled based on the atomic model of human Acetyl-Glucosaminyl Transferase (PDB code: 1w3b) with an overall sequence identity of 17%. The TPR1-6 of APC3 was modeled based on *S. pombe* APC6 molecular model (PDB code: 2xpi) with a 15% sequence identity. The other TPR subunits, APC7, APC6 and APC8 are all modeled from *S. pombe* APC6 structure, with the sequence identities of 15%, 36% and 18%, respectively. The crystallographic models of human C-terminal APC6 and N-terminal APC7 were incorporated into the corresponding docking models, respectively (PDB code: 3HYM, 3FFL).

All the model docking were performed through the software Chimera (80). The model editing and superimposition were done through COOT (81) and Pymol (82). Each model was manually docked into the corresponding density of the APC EM maps with a potential orientation, followed by the optimized local fitting through Chimera “fit in map” function, which adjusted the orientations of the filled models based on the calculated correlation coefficients. The docking was repeated through multiple rounds of iteration process including model re-editing and repeated docking till the most reasonable and/or the best fitting was achieved. The N-terminal and C-terminal coordinates of all the TPR subunits were docked in separately. After docking, each EM-derived structure was calculated for the correlation related to the EM maps.

The editing of the docking models was also guided from the model superimpositions. The superimposition of each model was adjusted with reference to the sequence comparisons of the APC subunit orthologs, and the 2<sup>nd</sup> structures corresponding to the same TPR motifs were manually aligned together in Coot.

## CHAPTER 3. INTERACTION STUDY OF THE CO-ACTIVATORS WITH APC3, APC8 AND THE APC PLATFORM

### Introduction

During the cell cycle, APC co-activators Cdh1 and Cdc20 are recruited to the complex through their C-terminal highly conserved IR-tails (33, 34, 36, 37). Whereas on the APC scaffold, the homo-dimeric TPR subunit APC3 is responsible for the co-activator recruitment. Cdh1 and APC10 also function as substrate receptors. APC substrates are recruited between Cdh1 and APC10 in a D-box (destruction motif) dependent manner. D-box is a common sequence within substrates that allows the APC to identify the substrates. An APC EM study revealed that Cdh1 and APC10 are connected through the substrate D-box once the substrate is bound (31, 32), although APC10 itself has weak affinity to substrates. Moreover, the ratio of co-activators and APC10 is maintained at 1:1 in an active APC complex.

Besides IR-tail mediated interaction, Cdh1 and Cdc20 were also found to have multiple interactions with the APC, and together these interactions generate high-affinity binding of the activator to the APC core. Cdh1 interacts with another TPR subunit, APC8 at C-terminal unknown sites (34) and also with APC2, since the removal of APC2 leads to reduced Cdh1 association (18). Cdh1 substrate-recruitment domain, the WD40 propeller, also shows some weak interactions with APC3 (33). It is likely that the high-affinity interactions keep activators bound during multiple substrate-binding of ubiquitination events (34).

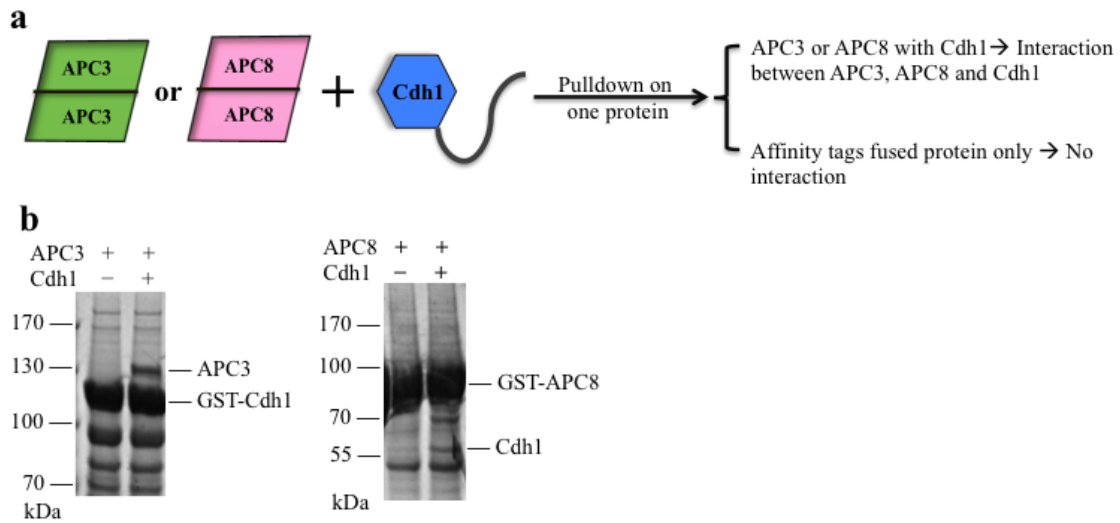
The TPR subunits APC3 and APC8 both have 14 TPR motifs, with N-terminal TPR1-4 mediating TPR subunit dimerization. The study of APC recruiting co-activators will help understand the APC ubiquitination mechanism and APC activity regulation during the cell cycles.

### Results and Discussion

#### Cdh1 binds to APC3 mainly through its C-terminal IR-tail peptide

In order to understand the recruitment mechanism of activators to APC, I reconstituted the interaction *in vitro* between the activators and APC3 or APC8. It was shown by previous lab members that most APC subunits could not be expressed in *E. coli* strains, and thus an insect cell expression system was chosen to perform experiments in this chapter. One approach used to investigate their interaction is co-pulldown assays. Principally, I co-expressed the various proteins fused with GST, Strep, His-MBP affinity tags in insect cells and detected their interaction through affinity purification, followed by examination of the purified products through SDS-PAGE (**Figure 3-1a**).





**Figure 3-1. Human Cdh1 binds to APC3 and APC8 *in vitro*.**

(a) Schematic representation of co-purification/co-pulldown procedure.

(b) Left: SDS-PAGE co-pulldown result of the GST-Cdh1 and APC3 (left), Cdh1 and GST-APC8 (right) from insect cell co-infections. All the bands labeled were identified by mass spectrometry. Molecular weight standards are labeled on the left side of the gel.

*In vitro* purified Cdc20 and APC10 behave poorly; therefore, Cdh1 was used for all the interaction studies in this chapter. Previously, the interactions of Cdh1 with APC3 and APC8 were tested by co-pulldown from Hi-5 insect cells. APC3 and APC8 could be co-purified with Cdh1 (**Figure 3-1**). GST (Glutathione S-Transferases) affinity tag was used for this experiment. Although the stoichiometry between APC3, APC8 and Cdh1 are not 1:1 in this initial experiment, the results demonstrated that the association between APC3 or APC8 and Cdh1 were reconstituted *in vitro*. Cdh1 expression was previously shown to be very impure, characterized by a lot of contaminants associated during purification (**Figure 3-1b**, left). Compared to most other subunits of the APC, APC8 expressed relatively well by itself. The unequal stoichiometry is probably because the GST-tagged proteins expressed better than the un-tagged binding partners (**Figure 3-1b**, right).

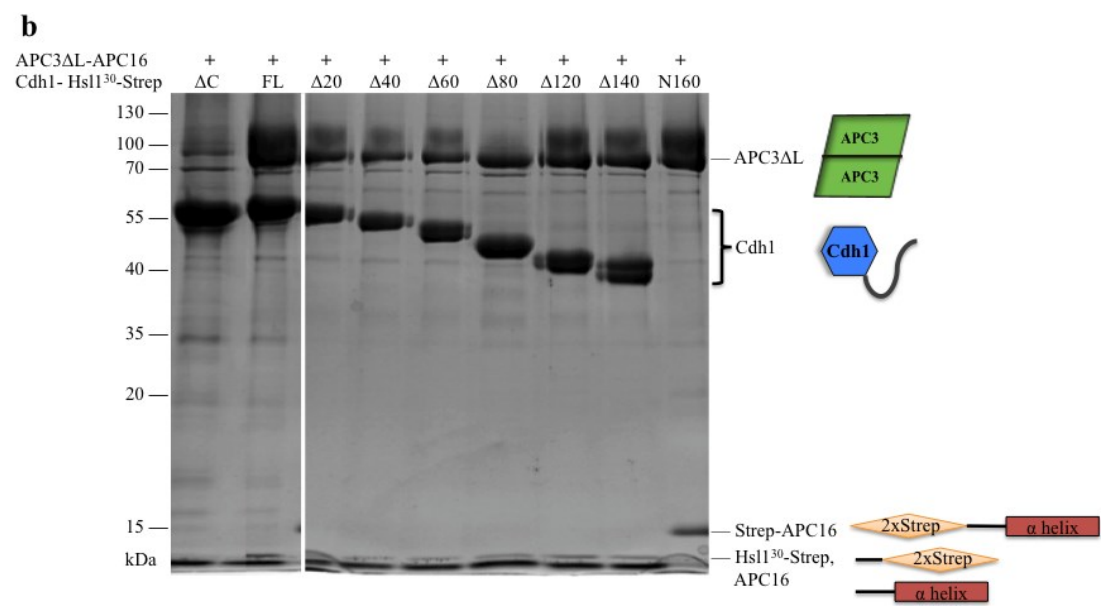
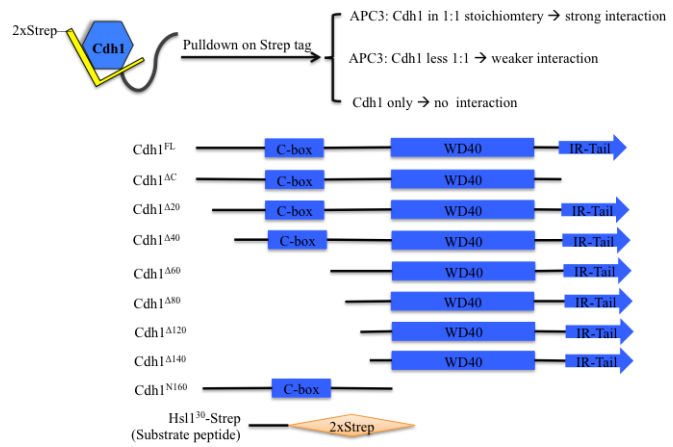
The co-pulldown experiments of Cdh1 with APC3 and APC8 were optimized through two strategies. One is to improve Cdh1 behavior by co-purification with yeast APC substrate peptide Hsl1<sup>30</sup> and the other strategy is to perform a co-pulldown on the twin-Strep tag that was fused to the Hsl1<sup>30</sup> C-terminus. Hsl1 is a yeast APC substrate and it associates with Cdh1 through its D-box (destruction box) and KEN box motif (83, 84). The D-box and KEN-box are the sequence elements inside most APC substrates that bind to the Cdh1 WD40 propeller domain. The Cdh1 WD40 domain functions as a D-box and KEN box receptor (33). Hsl1<sup>30</sup> is a shorter version of Hsl1 that only contains its D-box and KEN-box. The substrate peptide stabilizes Cdh1, making it behave better. The co-purified Cdh1-Hsl1<sup>30</sup> has less associated contaminants and adopts more homogenous conformations based on Cdh1 purification experiments from previous lab members.

In order to explore the ability of each Cdh1 domain to interact with APC3, the Cdh1 protein sequence was analyzed and serial deletions of Cdh1 were constructed for co-pulldown experiments (**Figure 3-2a**). The activator Cdh1 has similar structure domains/motifs: C-box, WD40 propeller and C-terminal IR tail peptides as Cdc20. N-terminal Cdh1 was serially truncated towards the WD40 domain in units of 20 amino acids, and the C-terminal IR tail peptide was also deleted. The N-terminal 160 residues of Cdh1 without the WD40 domain are not able to bind substrate Hsl1<sup>30</sup> by themselves (**Figure 3-2b**). Except the Cdh1<sup>N160</sup> pulldown, which was performed on the twin-strep tag of APC16, all other co-pulldown experiments were performed on the twin-strep tags of the Hsl1<sup>30</sup>.

To improve the behavior of human APC3, APC3ΔL substituted for the full-length protein for the experiment. APC3 is a highly hydrophobic protein due to its high  $\alpha$ -helix content. The full-length protein aggregates in the absence of stabilizing partners. APC3 has a large disordered loop between TPR 4 and 5 and deletion of that region — APC3ΔL effectively limits the aggregation. The disordered loop was discovered as a highly phosphorylated region (85) and it mediates the APC3 interaction with APC7. APC16 was recently identified as a small subunit of the APC to stabilize APC3 association with APC7 (86). APC16 binds to both APC3 and APC3ΔL, and the disordered loop of APC3 has no contribution towards binding with APC16 (**Figure 3-2b**).

**Figure 3-2. Map human Cdh1 domains required to interact with APC3.**

- (a) Illustration of the domain mapping strategy for APC3 $\Delta$ L, Cdh1 co-pulldown experiments. APC3 is represented in a green parallelogram, APC16 in a red sphere and Cdh1 in a blue octagon, with their construct schematics listed below. A GST affinity tag was fused to the APC3 $\Delta$ L N-terminus and a Strep tag was fused to the Hsl1<sup>30</sup> C-terminus. Each co-expression had the protein components of APC3 $\Delta$ L, APC16, Cdh1<sup>(full length or truncated)</sup>, Hsl1<sup>30</sup>. Strep tagged APC16 was used for the APC3 $\Delta$ L co-expression with Cdh1<sup>N160</sup> and the non-tag APC16 was used in the other co-expressions. Human Cdh1 constructs included C-terminal IR peptide deletion ( $\Delta$ C), a serial of N-terminal truncation ( $\Delta$ 20-140, residues number) and N-terminal fragment 1-160. WD40 domain is the Cdh1 substrate-binding region. Insect cell Hi-5 strain was used for co-expression and the experiment was performed on a Strep tag.
- (b) SDS-PAGE of purified product from APC3 $\Delta$ L, Cdh1 co-pulldown experiments. Protein loading for each lane was normalized to Cdh1. Molecular weight standards are labeled on the left side of the gel, and the protein bands on the gel were labeled with their names and their schematic illustrations.



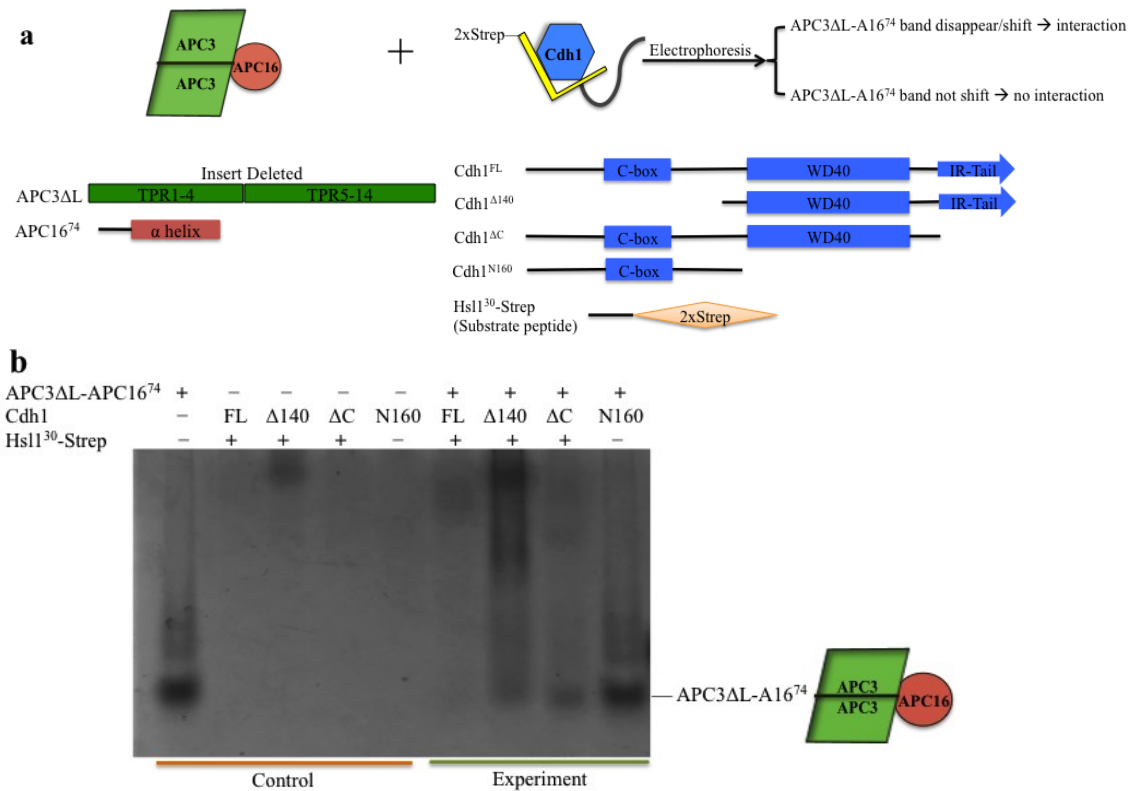
Cdh1 deletions were well expressed in insect cells (Hi-5 strain) and all the deletions were pulled down by Hsl1<sup>30</sup>-Strep except Cdh1<sup>N160</sup>. Compared to the deletion of Cdh1 IR tail peptide, the N-terminal deletions mildly affect Cdh1 interaction with APC3ΔL (**Figure 3-2b**). The amount of APC3ΔL co-purified with N-terminal deletions dropped down to half the amount of APC3ΔL co-pull down with full-length Cdh1. Without the C-terminal IR tail, APC3ΔL could barely interact with Cdh1. This result matches literature reports that the Cdh1 IR tail is critical for mediating APC3 interaction (37). There was no significant difference among the co-pulldowns from Cdh1 N-terminal deletions, which could be because co-pulldown assays are not sensitive enough to differentiate weak interactions.

To compensate for the limitation of co-pulldown assays and also to measure the kinetics of APC3 and Cdh1 interaction, other approaches including ITC (Isothermal Titration Calibration), SPR (Surface Plasmon Resonance) and native gel shift assay were applied. APC3, APC3ΔL and Cdh1 only behave well under the experimental conditions of native gel shift assays (4°C), but not at room temperature. The interaction between APC3 and Cdh1 did not produce a significant thermal change to be measured through ITC. The native gel shift assay is able to detect nM binding affinities of protein-protein interactions under native conditions through a polyacrylamide gel. It turned out to be the approach that produced detectable results in this study. The interacting protein pairs have a different mobility on PAGE compared to the non-binding proteins in the control.

Except Cdh1<sup>N160</sup>, all other Cdh1 deletions used for native gel shift assay were co-purified with Hsl1<sup>30</sup> from double affinity tag chromatography, followed by gel filtration. APC3ΔL-APC16 is purified and stored as a 1:1 stoichiometric protein complex and this complex forms a band on the native gel at pH 8.2, 4°C (**Figure 3-3b**). All the Cdh1-Hsl1<sup>30</sup> complexes and Cdh1<sup>N160</sup> were unable to electrophorese into native gels because of their basic pI value. Therefore, the gel shift experiment was designed to detect the band shift of the APC3ΔL-APC16 complex.

No APC3ΔL-APC16 band was detected in the presence of Cdh1<sup>FL</sup>-Hsl1<sup>30</sup>, indicating that all the loaded APC3ΔL-APC16 (*in vitro* purified complex) was bound to Cdh1. When Cdh1<sup>Δ140</sup>-Hsl1<sup>30</sup> interacts with APC3, APC3ΔL-APC16 runs as a smear on the native gel instead of a band or disappearance. It might be because the conformation of APC3ΔL-Cdh1 complex is different when the Cdh1 N-terminal 160 residues are missing. Cdh1's N-terminal 140 residues by themselves don't shift the APC3ΔL-APC16 band on the gel. It seems that N-terminal Cdh1 is not necessary to mediate the interaction with APC3 although the co-pulldown experiment showed less APC3ΔL bound to the N-terminal deletion.

Cdh1<sup>ΔC</sup>-Hsl1<sup>30</sup> also demonstrates some weak interactions with APC3 in this approach, which could be contributed from the WD40 propeller domains. The WD40 propeller domains of yeast Cdh1 were observed to have multi-site weak cross-linking interactions with APC3, and C box also has a one site weak-interaction with APC3 (33).



**Figure 3-3. Identify Cdh1 binding domains to APC3.**

(a) Illustration of experimental procedure for the native gel shift assay. APC3 is represented in a green parallelogram, APC16 in a red sphere and Cdh1 in a blue octagon, with their construct schematics listed below. Purified proteins or protein complexes were used for the experiment. The protein complexes APC3ΔL-APC16 and Cdh1<sup>(full length or truncated)</sup>-Hsl1<sup>30</sup>-Strep were purified separately from insect cell Hi-5 expression. Cdh1<sup>N160</sup> was purified by itself without Hsl1<sup>30</sup>-Strep. Beside a full-length construct used for this assay, human Cdh1 constructs also included C-terminal IR peptide deletion (ΔC), a N-terminal 140 residue truncation (Δ140) and N-terminal fragment 1-160. WD40 domain is the Cdh1 substrate-binding region.

(b) Native gel shift to examine the interaction between APC3ΔL and Cdh1. The control lanes are marked with an orange line at the bottom of the gel and the experimental lanes are marked with a green line. APC3ΔL-APC16 was mixed with Cdh1-Hsl1<sup>30</sup>-Strep (full length or truncated) at a 1:10 molar ratio and incubated on ice for 20 minutes before loading. 21 μM APC3ΔL-APC16 was used as a control and for each mixture. The protein complex band on the gel was labeled with its name and a schematic illustration.

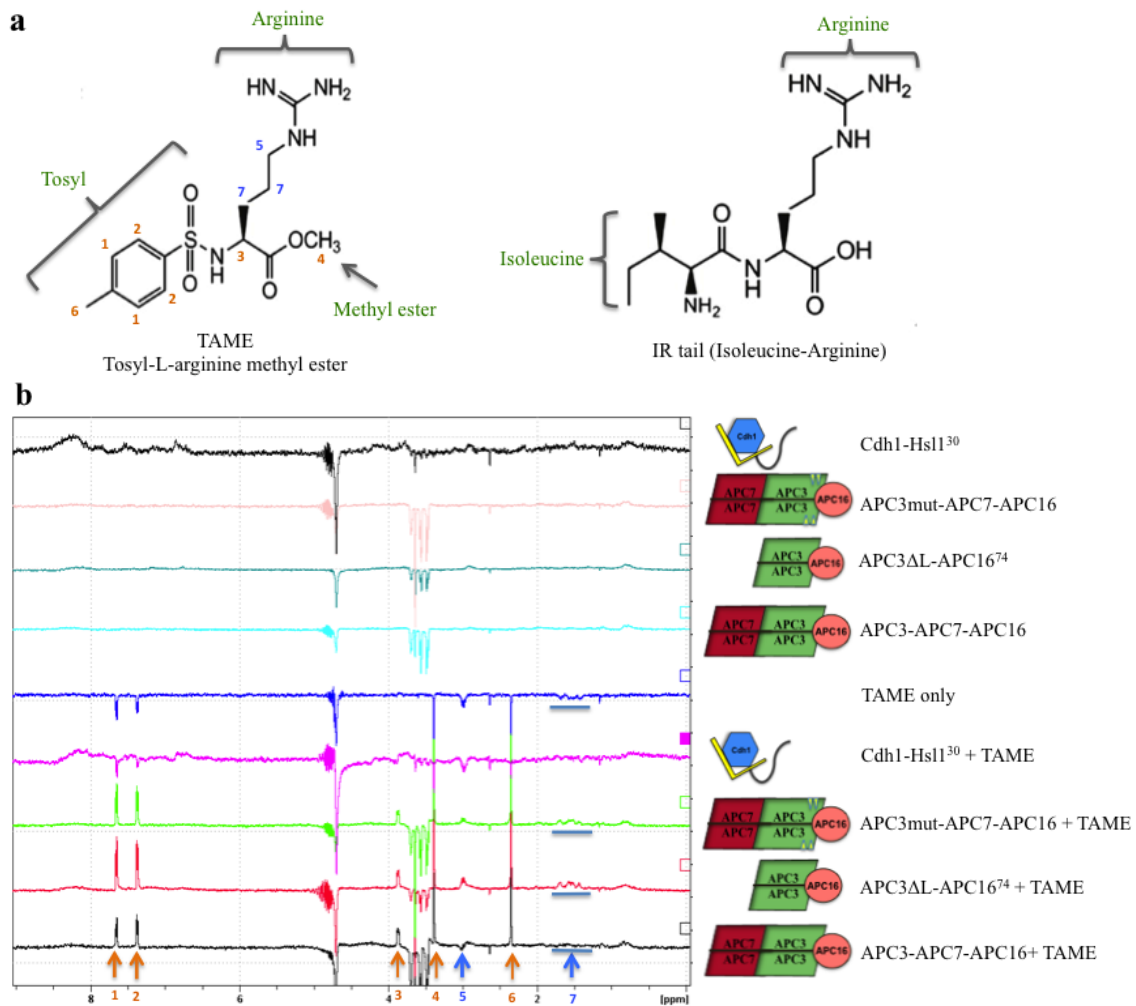
## Study of the IR-tail APC3 interactions via an IR-tail mimicked compound

The compound TAME is designed to inhibit APC activity by mimicking the IR tail structure of activators Cdh1, Cdc20 and co-receptor APC10, which all interact with APC3 through their C-terminal highly conserved IR tail. TAME disrupts the activator association with APC and inhibits the APC activity (40, 51). TAME uses a tosyl group to mimic Ile, and it also has a methyl ester group on the Arg carboxyl.

The interaction between TAME and APC3 protein complexes were detected by WaterLOGSY (water-ligand observed via gradient spectroscopy). The experiments were performed at room temperature after proteins were mixed with TAME in pH 7.0 phosphate-buffered saline (PBS) with 10% D<sub>2</sub>O. In WaterLOGSY experiments, the interacting ligands receive magnetization transferred from bulk water via the water molecules buried in binding pockets and labile protein protons, through the mechanism of NOEs (Nuclear Overhauser Effect) and chemical exchange. The experiments are designed to achieve efficient selective water excitation, which tends to detect the signal from interacting ligands by inverting the water signal (71, 87). The intensity of non-binding and tight binding ligands are characterized as negative amplitudes and those of weak interacting ligands are positive. For the tight binding ligands, the negative intensity results from either less water molecules inside unfilled binding pockets or a slower proton-exchange rate.

During any WaterLOGSY screening, the solvent signals (~5ppm) were suppressed to limit the artifacts in the spectrum. The peaks in this experiment were distributed from 1.5ppm to 7.5ppm. The downward peaks from the compound reference spectrum are signals from free ligands — non-binding TAME. The corresponding peaks of waterLOGSY spectrum inverted (upwards) in the presence of APC3 protein complex imply the interactions between APC3 and TAME. TAME specifically binds to APC3 complexes but not to Cdh1-Hsl1<sup>30</sup>, and each peak was assigned to a proton of TAME according to their ppm value (**Figure 3-4a**).

WaterLOGSY is a sensitive approach to detect relatively weak interaction between proteins and ligands. Signals indicating weak interaction have positive amplitude in the spectrum. Peaks one to four and peak six of the experimental spectrums revealed weak interaction between TAME and APC3, which were not affected by APC3 mutant and APC3ΔL (disorder loop deletion) (**Figure 3-4b**). Four of the interactions are contributed by tosyl protons and one is from methyl protons (**Figure 3-4a**). Meanwhile, peak five and seven pointed downwards in the spectrum of wild type APC3 with TAME, and they inverted upwards in the spectrum of APC3ΔL and APC3mut. Peak five and seven were the proton signal of the Arg group, and therefore the inversion indicated the Arg interaction with APC3mut or APC3ΔL is weaker than with wild type APC3. Peak three was detected in the spectrum of APC3 complexes with TAME but not in the compound reference spectrum. It is probably because the signal from proton three is very subtle if TAME is the free ligand in solution. Combining all the peaks in the three experimental spectrums, I can see that TAME interacts with APC3 with its tosyl group, Arg group and methyl ester, with the Arg group binding relatively tighter than the other two.



**Figure 3-4. Compound TAME mimics the IR tail interactions with APC3.**

(a) Structure illustration of compound TAME and IR tail. Groups are illustrated in green color.

(b) One-dimensional WaterLOGSY spectra record of APC3 interaction with TAME. The five spectra from the top are the reference spectrum of three APC3 complexes (2 $\mu$ M each), 2 $\mu$ M Cdh1-Hsl1<sup>30</sup>-Strep complex and 0.2mM compound TAME. Another four spectra from the bottom are the corresponding 2 $\mu$ M protein complexes in the presence of 0.2mM TAME. The peaks labeled out by orange and blue arrows are the signals of the protons in TAME, with the same set of numbers as in (a). The peaks labeled out with blue arrows indicate stronger interactions than those with orange arrows. The spectra were acquired with 10920 scans and protein signals were destroyed with the design of WaterLOGSY pulse sequence.



The mutated residues of APC3 are the residues that are conserved in the APC3 binding groove (**Figure 2-10**), which have been reported to reduce the association of Cdh1 with APC (34). The NMR WaterLOGSY results showed N581A/L612A decrease the affinity of the Arg (of the IR tail) to APC3 but do not completely disrupt the interaction. Given ten conserved residues of the binding groove surface potentially mediate IR tail recruitment, the mutation of two residues is probably not sufficient to eliminate the interaction. APC3 $\Delta$ L has no mutation in the binding groove but still has a similar defect as the APC3 mutants. The wild type APC3 of the APC3-APC7-APC16 complex appears to have better overall structure to stabilize the interactions with the IR tails.

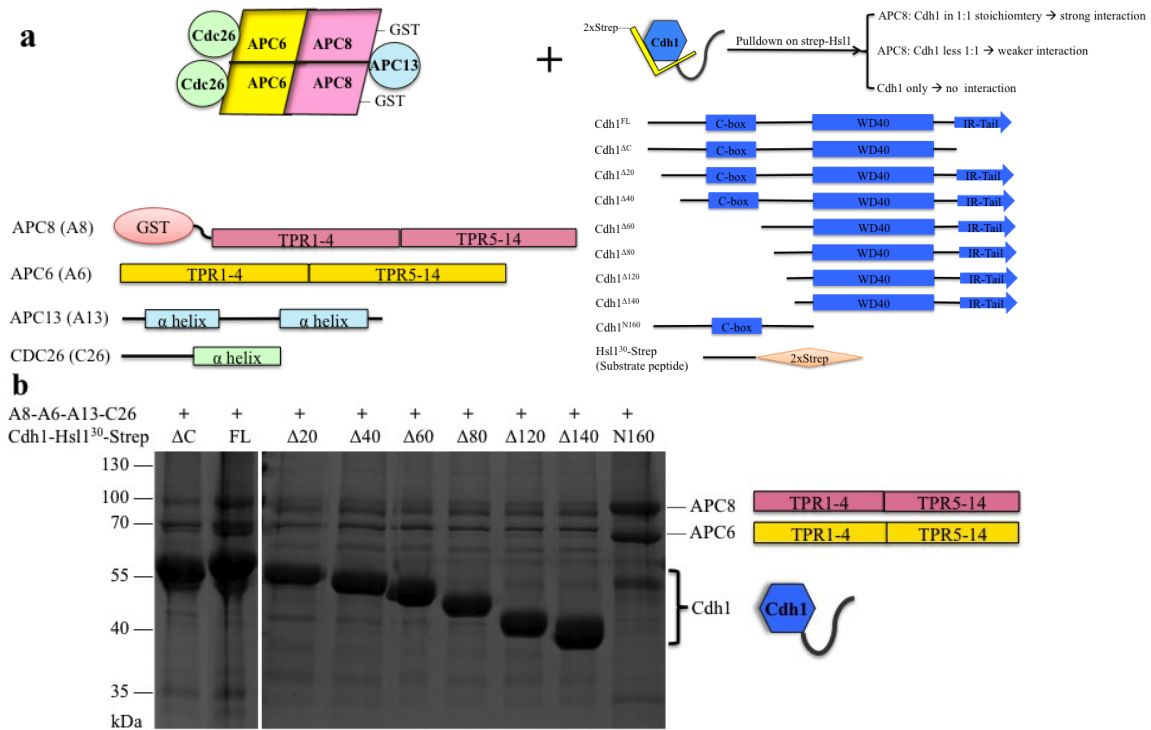
### **APC8 interacts with N-terminal Cdh1**

APC8 has similar aggregation behavior as APC3 if it is purified by itself. The strategy to solve the aggregation problem is to co-purify it with APC13 and APC6-Cdc26 complex. APC6 is an APC TPR subunit and it requires Cdc26 to fold into the correct conformation (26, 27). APC6 associates with APC8 tightly *in vitro*. APC13 is a small subunit that promotes the stable association of APC6 and APC3 to the APC complex (25), originally identified in yeast. Human APC13 also binds to APC8 *in vitro* but could not rescue APC8 alone from aggregation. Therefore, APC6-Cdc26 and APC13 are both used to stabilize APC8.

Compared to the interaction between APC3 and Cdh1 in their co-pulldown experiment, APC8 turned out to be a much weaker interacting partner for full-length Cdh1 (**Figure 3-5**). The Cdh1 N-terminal deletions appeared defective in pulling down APC8, and the difference among all the deletions also became indistinguishable. Nonetheless, the N-terminal Cdh1 showed a greater propensity towards binding APC8 over APC3.

To identify the Cdh1 binding domains to APC8 and compare the results to the interactions between Cdh1 and APC3, the same set of Cdh1 deletions were used for native gel shift assays. It appears that Cdh1 is able to bind to the APC8-APC6 complex only when the N-terminal 140 residues are present, and the interaction disappears when the N-terminal 140 residues were deleted (**Figure 3-6**). However, those N-terminal residues alone could not interact with APC8 if they are expressed/purified by themselves (without the WD40 domain).

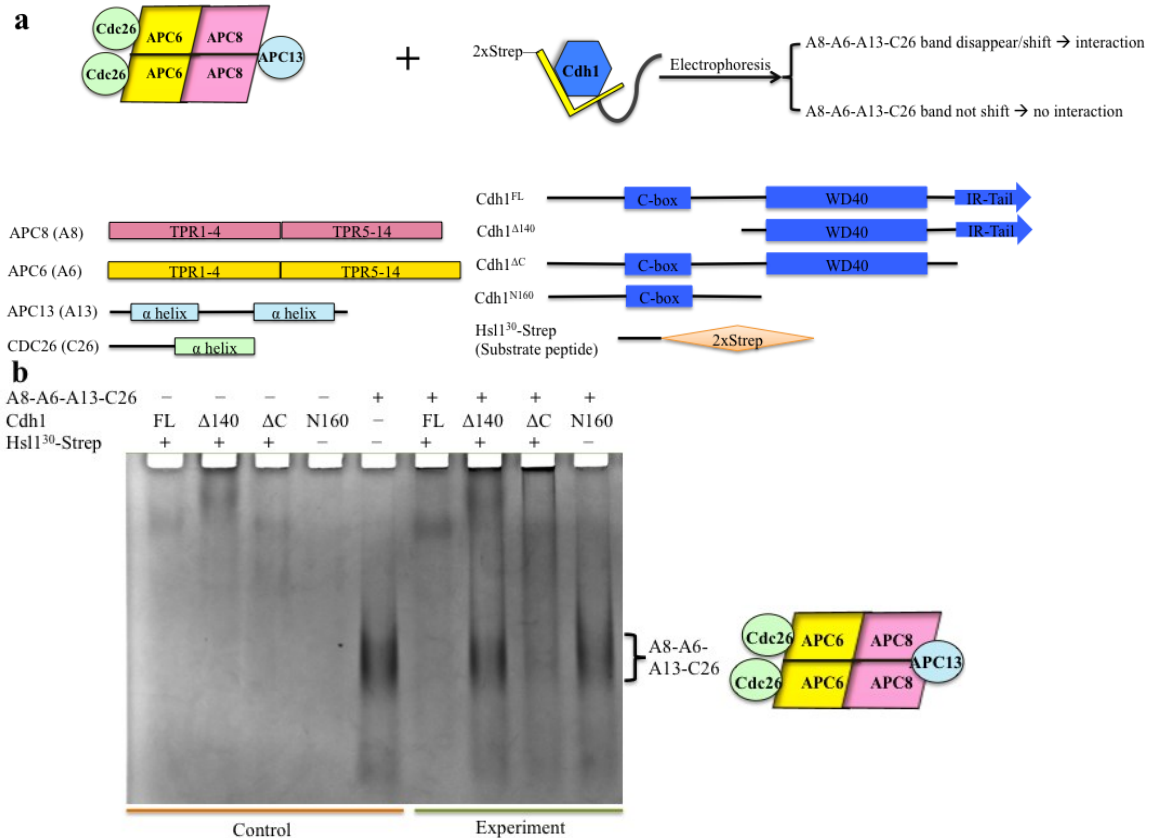
The APC8-APC6-Cdc26-APC13 complex for native gel shift assay is prepared through affinity chromatography on the GST affinity tag fused to APC8, followed by ion exchange chromatography and gel filtration. These gel shift results are reproducible.



**Figure 3-5. Map human Cdh1 domains required to interact with APC8.**

(a) Schematic illustration of the domain mapping strategy for APC8, Cdh1 co-pulldown experiments. APC8 and APC6 are represented in a pink and a yellow parallelogram, respectively, with APC13 and Cdc26 separately shown in a cyan and a light-blue sphere. Cdh1 is demonstrated with a blue octagon, and the schematics of all the constructs used for the co-pulldown are listed below. Human Cdh1 constructs included C-terminal IR peptide deletion ( $\Delta C$ ), a serial of N-terminal truncation ( $\Delta 20-140$ , residues number) and N-terminal fragment 1-160. WD40 domain is the Cdh1 substrate-binding region. A GST affinity tag was fused to the APC8 N-terminus and a Strep tag was fused to the Hsl1<sup>30</sup> C-terminus. APC8, APC6, APC13, Cdc26, Cdh1 (full length or truncated), and Hsl1<sup>30</sup>-Strep were all co-expressed in insect cell Hi-5 strain. The experiments were performed on a Strep tag except the one with Cdh1<sup>N160</sup>, which was performed on a GST tag.

(b) SDS-PAGE of purified product from APC8, Cdh1 co-pulldown experiments. Protein loading for each pulldown was normalized to Cdh1. Molecular weight standards are labeled on the left side of the gel, and the protein bands on the gel were labeled with their names and their schematic illustrations.



**Figure 3-6. Identify Cdh1 binding domains to APC8.**

(a) Schematic illustration of constructs used for APC8, Cdh1 native gel shift experiments. Purified proteins were used for the experiment. APC8 and APC6 are represented in a pink and a yellow parallelogram, respectively, with APC13 and Cdc26 separately shown in a cyan and a light-blue sphere. Cdh1 is demonstrated with a blue octagon, and the schematics of the constructs used for the assay are listed below. Human Cdh1 constructs included C-terminal IR peptide deletion ( $\Delta C$ ), a N-terminal truncation of 140 residues ( $\Delta 140$ ) and N-terminal fragment 1-160. WD40 domain is the Cdh1 substrate-binding region. The protein complexes APC8-APC6-APC13-Cdc26 and Cdh1<sup>(full length or truncated)</sup>-Hsl1<sup>30</sup>-Strep were purified separately by co-expression in insect cell Hi-5 strain. Cdh1<sup>N160</sup> was purified by itself without Hsl1<sup>30</sup>-Strep.

(b) Native gel shift to examine the interaction between Cdh1 and APC8 with APC8 as part of the APC8-APC6-APC13-Cdc26 complex. The control lanes are marked with an orange line at the bottom of the gel and the experimental lanes are marked with a green line. The APC8 complex and Cdh1<sup>(full length or truncated)</sup>-Hsl1<sup>30</sup>-Strep were mixed on ice for 20 minutes before loading. 18 $\mu$ M APC8-APC6-APC13-Cdc26 was used in both the control and in each mixture. The molar ratio of Cdh1<sup>(full length or truncated)</sup> to APC8 complex is 10:1. The protein complex band on the gel was labeled with its name and a schematic illustration.

## The interaction of N-terminal Cdh1 with the APC platform

N-terminal Cdh1 is predicted to fold into  $\alpha$ -helices (by Psipred, **Figure A-6**) (68) and the corresponding densities are identified towards the inside of the APC. But the molecular envelopes of all the published EM maps are not clear/detailed enough for the N-terminal secondary structure assignment. The APC quaternary structure displays both the APC8 and APC1 repeats are connected with N-terminal Cdh1 by strong densities (**Figure 3-7a**). The green spheres mark the corresponding mutation in APC8 (N339A) that was reported to reduce yeast Cdh1 association to APC *in vivo* (34). The APC topological structure displays the APC8 mutation is in close proximity to N-terminal Cdh1 density, suggesting it mediates interactions between these two subunits. The disruption to the interaction of APC8 with Cdh1 begins with Cdh1 N-terminal 20 residues deletion, and N-terminal 20 residues are predicted to form two  $\alpha$ -helices (**Figure 3-5b, Figure A-6**). It implies that the N-terminal 20 residues of Cdh1 contact APC8.

To test the interaction between Cdh1 and other APC subunits (mainly platform subunits), more native gel assays were performed (**Figure 3-7b**). The recombinant APC platform, which includes all the non-TPR subunits, appeared as a clearly visible band on a native gel with or without APC8-APC13. The clear bands from the native gel indicate that the majority of the complexes have a rigid and homogenous conformation. The same set of Cdh1 deletions from APC8 and APC3 binding test were used for the assay.

The full-length and C-terminal peptide deleted Cdh1 both bind to APC8/APC13 +/- platform (**Figure 3-7b**). The interaction disappears without N-terminal Cdh1 (before WD40 domain). The gel shift patterns are similar to the ones of the APC8 TPR complex interacting with Cdh1 (**Figure 3-6b**). Cdh1<sup>(fl or  $\Delta$ 140)</sup> attached to the APC8-APC13-platform forms a shifted band on the gel, whereas the majority of the Cdh1<sup>(fl or  $\Delta$ 140)</sup>-platform disappears or becomes a smear. This may mean the Cdh1<sup>(fl or  $\Delta$ 140)</sup>-platform has a more flexible conformation without APC8-APC13. The interaction between APC8 and Cdh1 might stabilize the association of Cdh1 with other platform subunits.

Same as the native gel shift results of Cdh1 with APC8 and APC3, the recombinant N-terminal fragment of Cdh1 did not show any interaction with either platform complex. There could be many explanations for this. The Cdh1 N-terminal fragment seems to have flexible folding, and half of the fragment is predicted to be a disordered loop (**Figure A-6**). The EM-derived N-terminal helices only cover 60% of the fragment sequence and other residues of the disordered loops could not be identified through the current EM maps. One possibility is that the Cdh1 N-terminal fragment will fold accordingly once it comes in contact with the APC platform subunits.

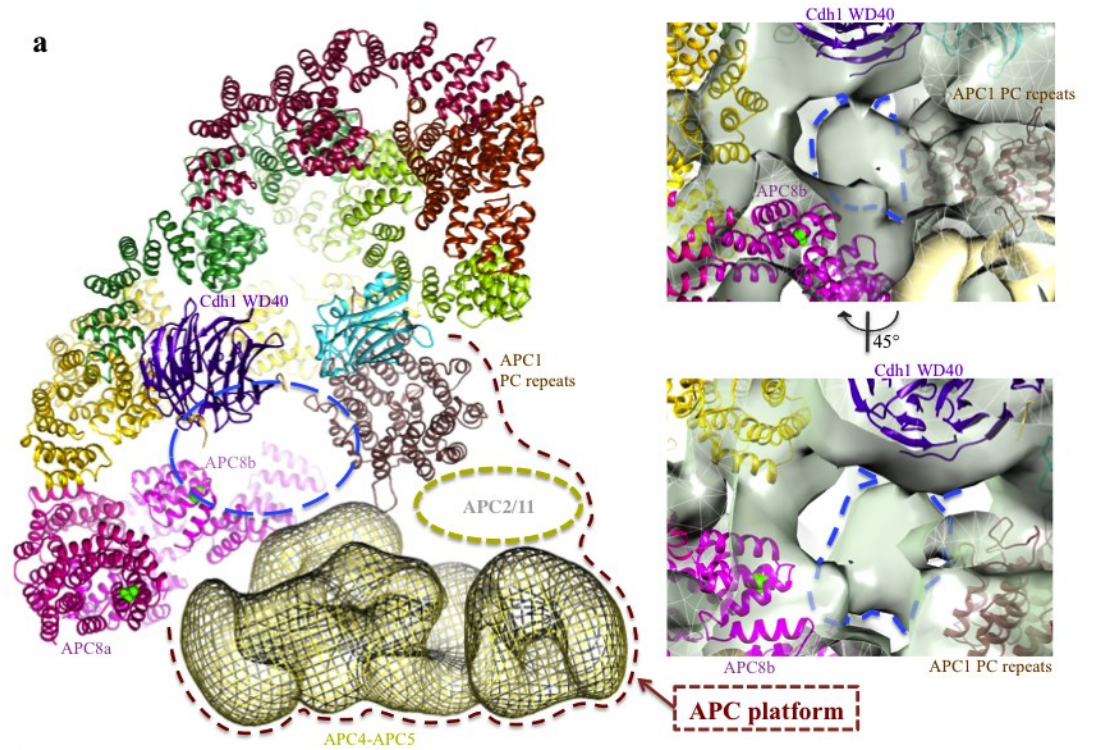
## Implication of Cdh1 N-terminal interactions

The amounts of co-activators bound to APC are rate-limiting for ubiquitination reactions (88), and therefore it is important and interesting to understand how co-activator association supports APC ubiquitination activity. Multiple previous studies

**Figure 3-7. N-terminal Cdh1 interacts with the APC platform.**

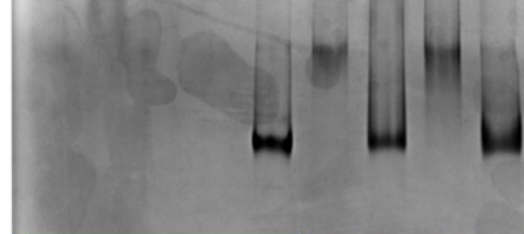
(a) Human APC architecture demonstrated by EM-derived structures and APC4-APC5 EM map (yellow mesh, EMD 1843) (89) docked into the corresponding densities on the APC EM map (EMD 2354) (23). The locations of APC2 and APC11 are annotated as a yellow circle and the APC platform is outlined with a red dash line. The model of the N-terminal Cdh1 includes a bunch of  $\alpha$ -helices extending towards the inside of the APC (left). The green spheres indicate the corresponding APC8 residue that decreased Cdh1 association with the yeast APC (32). The blue circle mark the localization of the density contributed by the N-terminal Cdh1, with two close views demonstrating the interactions of N-terminal Cdh1 with APC8 and APC11 repeats.

(b) Native gel shift assays to identify the Cdh1 domains required for APC8 +/- platform interaction (top, middle panels), with GST-ubiquitin as an interaction control (bottom panel). The control lanes in each gel are marked with an orange line at the bottom of the gel and the experiment lanes are marked with a green line. APC8 +/- platform or GST-ubiquitin were mixed with Cdh1<sup>(full length or truncated)</sup>-Hsl1<sup>30</sup>-Strep, respectively, on ice for 20 minutes before loading. 5 $\mu$ M APC8-platform, 3.8 $\mu$ M platform and 37 $\mu$ M GST-ubiquitin were used as controls and for each mix. The molar ratio of Cdh1 (full length or truncated) to APC8 +/- platform or GST-ubiquitin is 10:1. The protein complex band on the gel was labeled with its name and a schematic illustration.



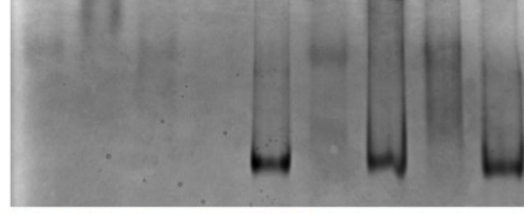
**b**

A8-Platform	-	-	-	-	+	+	+	+	+
Cdh1	FL	Δ140	ΔC	N160	-	FL	Δ140	ΔC	N160
Hs11 <sup>30</sup> -Strep	+	+	+	-	-	+	+	+	-

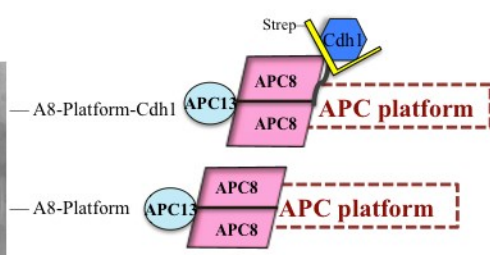


Control Experiment

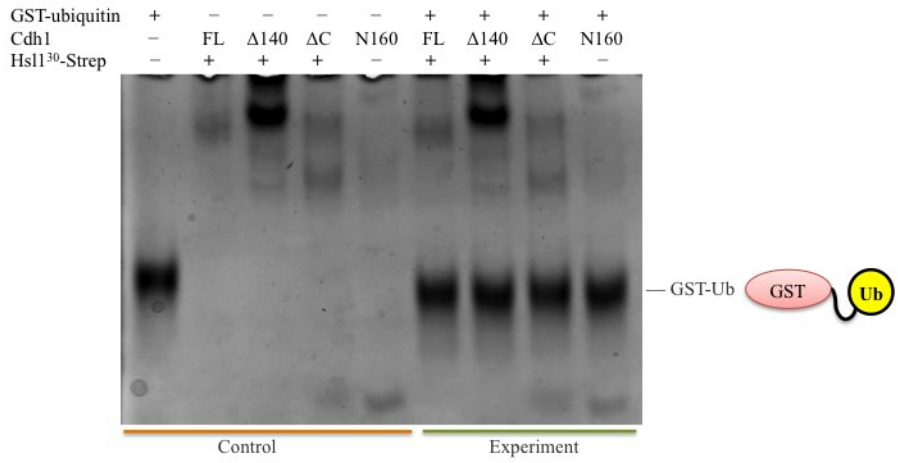
Platform	-	-	-	-	+	+	+	+	+
Cdh1	FL	Δ140	ΔC	N160	-	FL	Δ140	ΔC	N160
Hs11 <sup>30</sup> -Strep	+	+	+	-	-	+	+	+	-



Control Experiment



Platform APC platform



**Figure 3-7. (Continued).**

showed that the attachment of Cdh1 to APC is stabilized by multiple interactions, which likely supports that Cdh1 binds when ubiquitinated substrates are released from the APC (18, 34).

The interactions between APC3 and the IR-tail peptides of the co-activators or APC10 have been well characterized by biochemical and structural approaches, although their atomic details remain unknown. In contrast, the Cdh1 association to other APC subunits is still not well understood. In addition to the reported interactions of Cdh1 with APC being confirmed in this chapter (18, 33, 34, 37), the EM-derived structures are also incorporated to shed light on revealing the mechanism. Potentially, the N-terminal Cdh1 helices interact with APC8, APC1 PC repeats and APC4-APC5

Given that the disordered regions of N-terminal Cdh1 couldn't be identified through the EM map, biochemical assays will be required to determine the binding details between the Cdh1 N-terminal regions and the platform subunits. To characterize the role of the N-terminus of Cdh1 in the multi-interaction, more thorough Cdh1 N-terminal deletions/mutations would be required to test binding with the APC platform, recombinant APC1-APC4-APC5, and *in vitro* purified single platform subunits. Since APC2-APC11 couldn't form a visible band in a native gel, this approach is not applicable for the catalytic subunits. Meanwhile, the crystallization of the APC platform with Cdh1 would be an alternative approach to reveal the details of their interaction.

Human Cdh1 is subject to phosphorylation *in vivo* by Cdk (Cyclin-dependent kinases), and the Cdk-dependent phosphorylation sites are responsible for blocking the APC interaction with Cdh1. APC<sup>Cdc20</sup> inhibits the recruitment of Cdh1 to the APC by phosphorylating Cdh1 until the latter stages of mitosis (88, 90). The majority of the phosphorylation sites are at the N-terminal fragment (before the conserved WD40 domain) (85, 88), and therefore, phosphorylation interrupts Cdh1 binding to APC. *In vitro*, phosphorylated Cdh1 binds to or activates the APC less efficiently when compared to non-phosphorylated Cdh1, whereas non-phosphorylatable Cdh1 mutants constitutively activate APC *in vitro* and *in vivo* (90). Combining the effects from Cdh1 phosphorylation with the data on Cdh1's multi interactions with the APC may further elucidate the mechanism of how Cdh1 association regulate APC activity.

## Materials and Methods

### Constructs and insect cell infection

Protein expression constructs were made by standard PCR/ligation procedures, and sequences were verified by automated sequencing procedures. The APC3 deletion and mutation were designed in reference to the human APC3 830 a.a sequence isoform (uniprot code: P30260). The disordered region of residues 182-453 were deleted for APC3 $\Delta$ L, and APC3mut refers to APC3 with Asn581 and L612 both mutated into Ala. The APC16 construct includes the full-length protein sequence of APC16 (uniprot code:



Q96DE5), and APC16<sup>74</sup> refers to residues 74-110. The constructs of APC8, APC6, Cdc26 and APC13 also used the full-length protein sequence (uniprot code: Q9UJX2, Q13042, Q8NHZ8 and Q9BS18). Human APC8 refers to the isoform with 597a.a. For the chimeric APC3, there is a long version and a short one. The long version fused the N-terminal residues 1-241 of *E. cuniculi* APC3 to Human APC3 residues 539-830. The short one used the *E. cuniculi* residues 175-241 to fuse to Human APC3 residues 539-830. Hsl1<sup>30</sup> was made by cloning yeast Hsl1 residues 770-790, which includes a KEN box, and residues 818-842, which includes a high affinity D-box, into the substrate peptide. Human Cdh1 has 496a.a. (uniprot code: Q9UM11). Cdh1 $\Delta$ C refers to the last 17 residues deleted from the C-terminus, and Cdh1 $\Delta^{20-140}$  means the corresponding number of residues deleted from N-terminal Cdh1. The Cdh1 WD40 propeller domain includes residues 162-479, and the C box refers to residues 45-52.

All the constructs described in this chapter were expressed using the Bac-to-Bac baculovirus expression system (Invitrogen). The sequences-verified pFastBac constructs were transformed into *E. coli* DH10B or *E. coli* EMBacY competent cells, which contain a baculovirus shuttle vector — a bacmid to generate the recombinant bacmids carrying the genes of interest (79). Transformed cells were incubated at 37°C for 4hrs to allow the pFastBac expression cassette to be incorporated into bacmids by the Tn7 transposon. The culture was plated onto LB-Agar afterwards that contained 100 $\mu$ g/ml ampicillin, 50 $\mu$ g/ml kanamycin, 10 $\mu$ g/ml tetracycline, 7 $\mu$ g/ml gentamycin, with 100 $\mu$ g/ml Xgal and 40 $\mu$ g/ml IPTG to enable lacZ blue-white selection of the clones containing the gene of interest (white colonies). The plates were incubated at 37°C for 1.5-2 days and then the white colonies were re-streaked and re-evaluated by colony PCR. Positive clones were amplified in an over-night LB medium containing 100 $\mu$ g/ml ampicillin, 50 $\mu$ g/ml kanamycin, 7 $\mu$ g/ml gentamycin and 10 $\mu$ g/ml tetracycline at 37°C.

The recombinant pFastbac plasmids were purified via the Qiagen miniprep protocol, while a modified version of this protocol was used to purify the recombinant bacmid DNA. Cell pellets were resuspended in 300 $\mu$ l of Qiagen buffer P1, lysed by addition of 300 $\mu$ l of Qiagen buffer P2 at room temperature for five minutes and then incubated with 300 $\mu$ l of Qiagen buffer P3 on ice for ten minutes. The lysate was cleared at room temperature by centrifugation at 13,000rpm for 10 minutes. The cleared lysate 700 $\mu$ l was added to 800 $\mu$ l ice cold isopropanol and incubated for 5 minutes on dry ice. Precipitated bacmid DNA was pelleted by centrifugation at 13,000rpm for 10 minutes at room temperature and the DNA pellet was washed with 70% (v/v) Ethanol. Eventually the DNA was resuspended in 50 $\mu$ l of sterile elution buffer (Qiagen) inside a laminar flow hood (sterile conditions).

The baculovirus of the genes of interest were generated for protein production, and the Sf9 insect cell strain was used to amplify the baculovirus. Transfection mix was made by adding 10 $\mu$ l of FuGene HD transfection reagent into 100 $\mu$ l of serum-free media that contained 1-2 $\mu$ g of Bacmid DNA. The transfection mix was incubated at 27°C for 30 minutes before being diluted with 1ml serum-free medium and incubated with  $2 \times 10^6$  healthy Sf9 cells in a 6-well plate. After transfection, another 2ml of serum-free medium was added back into the Sf9 cells for continuous incubation at 27°C for 4-5 days before the baculovirus was harvested. The 1<sup>st</sup> generation of virus needs to be amplified twice

before being used for infection. The amplification was performed in 25mm cell culture dishes with  $\sim 20 \times 10^6$  healthy Sf9 cells per dish. Each amplification mixture was made by mixing 500 $\mu$ l of 1<sup>st</sup> generation virus with 5ml of serum-free medium. Then the amplification mix was added into 25mm cell culture dishes before incubation at 27°C for 2 hours. After 2 hours, another 15ml of serum-free medium was added into each dish for a continuous 3-day incubation at 27°C before the 2<sup>nd</sup> generation of baculovirus was harvested. The 3<sup>rd</sup> generation of baculovirus was prepared with the same procedure.

Hi-5 insect cells were infected or co-infected by the 3<sup>rd</sup> generation of baculovirus at a density of 10-20 $\times 10^6$  cells/ml, and the infection was shaken at a speed of 120 revolutions per min (rpm) and incubated at 27°C for 2 hours in a shaker incubator. After incubation, serum-free medium was added back to the infection and the cells were diluted to a density of 1-2 $\times 10^6$  cells/ml. The diluted infections were shaken at a speed of 155rpm/min and incubated at 27°C for one day and then at 20°C for another two days before harvesting. 1ml of virus was used for each 100ml infection culture and the viruses used for co-infection were at a 1:1 ratio.

### **Protein expression and purification**

Most of the protein complexes in this chapter were purified through GST and Strep affinity tags. For these purification, Hi-5 insect cell pellets were resuspended in a suspension buffer of 50mM Tris-HCl pH7.6, 0.3M NaCl, 5mM Dithiothreitol (DTT) supplemented with 10 $\mu$ g/ml Aprotinin, 5 $\mu$ g/ml Leupeptin, protease inhibitor 1 tablet/50ml buffer (Sigma) and 2.5 mM phenylmethylsulphonyl fluoride (PMSF), and lysed by sonication on ice. Lysates were cleared at 15,000rpm for 20 minutes at 4°C twice, and the cleared lysates were incubated with glutathione-affinity resins (Qiagen) or strep-tactin sepharose (IBA). Lysates were gently shaken on a rotary shaker at 4°C for 1 hour and then the beads were washed twice with two bead-volumes before loading onto the column. Beads were washed in the column with at least 20 bead volumes before elution. Proteins were eluted out of affinity beads by 10mM GSH (Glutathione) or 2.5mM dethiobiotin supplemented wash buffer (elution buffer), and the results were visualized through SDS-PAGE. For the protein complexes that required removal of the affinity tag, home made TEV protease, thrombin, or prescission was added to the elutions at a mass ratio of 1:50, 1:100, 1:100 (protease: eluted proteins), respectively. The eluted protein was concentrated to above 3mg/ml (measured by Bio-Rad Protein Assay) for the tag cleavage.

For the native gel shift assay, protein complexes were purified by gel filtration chromatography using a Superose6 column (GE) in a buffer of 50mM Tris-HCl pH7.6, 0.3M NaCl, 5mM DTT. The fractions of gel filtration were examined by SDS-PAGE to select the protein fractions of interest. The fractions of interest were pooled, concentrated with Amicon concentrational filter units, aliquotted, flash-frozen in liquid nitrogen, and stored at -80°C until later use.

## Native gel shift assays

The electrophoresis for the native gel shift assays was performed with 1×TB buffer (90mM Tris-borate) pH 8.2. Native gels are a 4.5% poly-acrylamide gel with 2% glycerol, made of 40% acrylamide/bis solution (37.5:1, Biorad), 100% glycerol and 5×TB buffer. Cdh1<sup>(full length or truncated)</sup>-Hsl1<sup>30</sup> was mixed with APC3ΔL-APC16 or APC8-APC6-Cdc26-APC13 at a molar ratio of 10:1. APC3ΔL-APC16 or APC8-APC6-Cdc26-APC13 were loaded with 21μM or 18μM for each mix, respectively. The mixes were diluted with the gel filtration buffer, 50mM Tris-Cl pH7.6, 0.3M NaCl, 5mM DTT, to a final volume of 8μl with 6.25% glycerol. For the control group, the same concentration of the protein complexes were used as those in the experimental group and were directly diluted by gel filtration buffer to 8μl with 6.25% glycerol. The diluted samples and mix were incubated on ice for 20 minutes before electrophoresis was performed in the cold room for 160mins (under 130V). The protein bands were visualized by commassie staining.

## WaterLOGSY

The purified protein complexes for WaterLOGSY screening were buffer exchanged into 100mM phosphate-buffered saline (PBS) solution pH7.4 by NAP-5 columns (GE). Protein complexes were then mixed with TAME and diluted with 100mM Na<sub>3</sub>PO<sub>4</sub> pH7.0 buffer to a final volume of 500μl of 2mM protein complexes, 0.2mM TAME and 10% D<sub>2</sub>O. The samples were transferred to NMR tubes at room temperature and screened with a WaterLOGSY pulse sequence in Jet 600MHz (Bruker). All the preparation before the sample transfer was done on ice or in the cold room. For each sample, a reference spectrum and a 1D WaterLOGSY spectrum were recorded. Each spectrum was scanned for 30 rounds, and each round includes 8 dummy scan and 32 scans.

## CHAPTER 4. IMPLICATIONS AND FUTURE DIRECTIONS

### Insights from the Hybrid Structure of the C-terminal Portion of Human APC3

The structural study of the human C-terminal APC3 leads to its structure determination. In the hybrid construct (the chimeric APC3), the C-terminal human APC3 folded into  $\alpha$ -helices with the predicted TPR boundaries, whereas two TPR regions folded into two  $\alpha$ -helices instead of the canonical the  $\alpha$ -helix-turn- $\alpha$ -helix (a TPR motif). However, the binding groove of APC3 (TPR8-10) independently adopted a canonical TPR folding without being affected by its neighbor mis-folded TPR7 and TPR11. The details demonstrated by the atomic model of the APC3 binding groove provided an accurate view to interpret current assays and help guide the design for later mutagenesis study to validate the structure.

Although a mis-folded structure was generally considered to be uninformative, the atomic model of the chimeric APC3 turned out to be useful for structure analysis. The key information lies with a mis-folded helix  $\alpha$ -11. The short alpha helix  $\alpha$ -11 is folded from the predicted TPR11. Instead of packing parallel to TPR10, it was recruited to the binding groove. This phenomenon indicated the APC3 binding groove is functional even it is surrounded with five mis-folded or mis-orientated  $\alpha$ -helices. The residues mediating the interactions between  $\alpha$ -11 and the binding groove are different from those stabilizing the canonical TPR11 packing (**Figure 2-15**). Although it couldn't be concluded that the formation of  $\alpha$ -11 resulted from its recruitment to the binding groove, the fact of this mis-folded helix indicates the binding groove is capable to mediate strong interactions and affect the canonical folding pattern.

As to the binding groove TPRs 8-10, three dimensional structure alignments showed that the whole binding groove well matched other canonical TPR motifs and the  $\alpha$ -11 overlapped with multiple binding partners of the matched hits (**Figure 2-9**). Furthermore, the fact that the  $\alpha$ -11 was aligned onto the binding partners implied that  $\alpha$ -11 potentially played a role of a pseudo binding partner in the chimeric APC3 (**Figure 2-9**).

Based on the fact that  $\alpha$ -11 locked the chimeric APC3 binding groove as a potential pseudo binding partner, the interactions between them partially represented the recruitment mechanism of the APC co-activators. Similarly, two residues: a Lysine and a Leucine, mimicked and represented the IR-tail (Isoleucine-Arginine) interactions, although the Lys and Leu were along a turn of  $\alpha$ -11 instead of a peptide. The carboxyl group of the Lysine also formed a peptide bond. But these two residues (the Lysine and Leucine) were recruited to two featured binding pockets respectively and demonstrated the same types of the interactions as the IR-tails (**Figure 2-10**). The IR-tails have a free carboxyl group, which most likely contributed to the recruitment by interacting with a basic residue surrounding the negatively charged pocket. A highly conserved His551 of TPR7a (**Figure 2-16a**), a well conserved His589 and Arg587 of TPR8a (**Figure 2-8a**, **Figure 2-10**) are all close to the carboxyl group of IR-tail and could potentially mediate

another salt bridge. The IR-tail probably established stronger interactions with the binding groove than what were demonstrated by  $\alpha$ -11 in the chimeric APC3.

The chimeric APC3 structure provided clues to speculate the biological meaning of their interactions and a hypothetical model. Both the IR-tail residues and the APC3 binding groove residues are conservative across species. This may indicate that it is critical to anchor the IR-tails (of the APC co-receptors) to the correct domain/region on APC3. The IR-tail interactions would potentially be much stronger comparing to the interactions of Cdh1 with APC8 and the APC platform (**Figure 3-5, Figure 3-6, Figure 3-7**).

The strong IR-tail interactions potentially keep the co-activators (Cdh1 and Cdc20) and APC10 bound to the APC during APC active stages. Considering that Cdh1 repositioned itself on the APC during substrate recruitment and ubiquitination (31, 32), the strong IR-tail interactions with APC3 would keep Cdh1 associating with the APC during its orientation change. As for APC10, a constituent subunit of the APC, it has a role for optimal Cdh1/Cdc20-dependent substrate recognition and substrate ubiquitination processivity (38). However, it is not well understood how APC10 contributes to the formation of poly-ubiquitin chains and whether it also has a movement on the APC. But the IR-tail is the only domain of APC10 that connects to APC3, which potentially requires the interactions strong enough to support the APC10 function.

The mis-folded TPRs inside the chimeric APC3 also provided clues for speculation of TPR protein folding mechanism to understand the biological meaning of their superhelix tertiary structure. The superhelix ensures the binding groove more accessible to unstructured the peptides (etc. IR tails) rather than numerous, random  $\alpha$ -helices surrounding APC3 (**Figure 2-17**). The correctly folded TPRs also served as an accurate model for APCEM docking.

## **Strength and Weakness of Hybrid TPR Proteins/Technology**

### **Brief introduction to the hybrid proteins strategy**

The first successful hybrid protein was reported on the human Toll-like receptor 4. Toll-like receptors are a class of proteins that play a key role in the innate immune system, and they belong to the Leucine rich repeats (LRR) protein family. LRR is a protein structural motif composed of repeating 20–30 amino acids that are unusually rich in the hydrophobic amino acid Leucine. Commonly a LRR motif forms an  $\alpha$  helix-turn- $\beta$  strand, and these repeats pack parallel to each other and form a horseshoe-like solenoid.

The basic concept of hybrid proteins is to replace a crystallization hindrance region of one protein with a homologous but crystallizable domain from its homologs (**Figure 2-1**). The homologous domains were supposed to benefit crystallization without changing or affecting the protein folding. The structure of human Toll-like receptor 4 was

determined through a series of hybrids (65), because the receptor itself (full-length or truncated) could not generate crystals. In order to help crystallization, the modules of hagfish variable lymphocyte receptors (VLR) were fused to substitute the insoluble part of the human toll-like receptor 4. These hybrid receptors crystallized and led to the structure determination.

In pursuit of human APC3 structure, both the full-length or truncated APC3 could not crystallize. The N-terminal *E. cuniculi* APC3 atomic structure was published around that time, which includes the N-terminal TPRs 1-6 (PDB code: 3KAE). TPRs 1-4 mediate the *E. cuniculi* APC3 dimerization, followed by a 20-residue insert and TPR5-6 (**Figure 2-1**). However, the N-terminal dimerization domain of human APC3 expressed poorly *in vitro* and the following disordered insert has 270 residues (**Figure 2-1**). These two regions were both crystallization hindrance. Since the corresponding region in *E. cuniculi* APC3 was crystallizable, it was used to substitute the insoluble domains of human APC3.

### Comparisons of LRR and TPR proteins

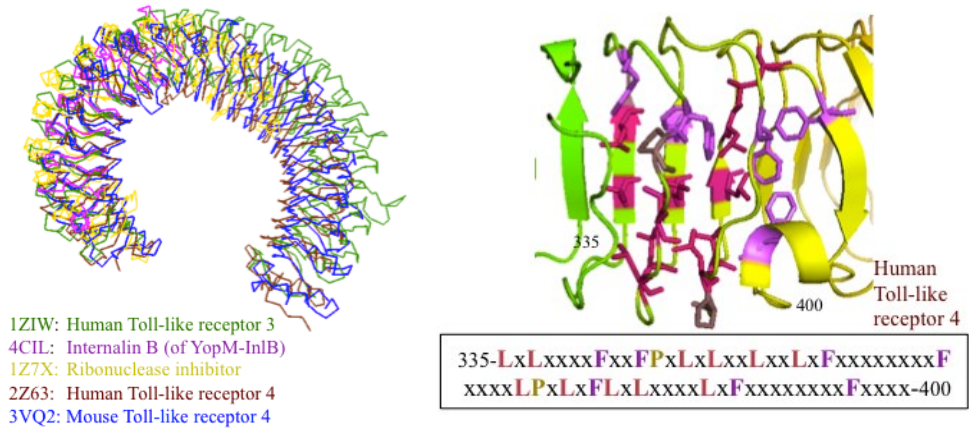
To this day, the hybrid technology has a couple successes in crystallizing LRR (Leucine rich repeats) proteins. Another example is a hybrid InlB variant YopM-InlB. InternalinB is an agonist of *L. monocytogenes* to the human receptor tyrosine kinase MET (91). Whereas YopM is a surface effector of *Y. pestis* that binds caspase-1 to inhibit its activity and sequesters it to block formation of a mature inflammasome (92). YopM and InlB are two functionally non-related LRR proteins. The cap domain of YopM replaced the cap of InlB protein, which originally folded into two  $\alpha$ -helices but became shorter in the hybrid protein. The residues at the hybrid interface also mildly shifted from original positions, but the interface didn't cause substantially global changes of the InlB structure (66). The hybrid YopM-InlB addressed the biological function of the N-terminal cap of internalins, which could not be studied with simple domain deletion constructs due to its indispensability to correct protein folding.

Although both LRR and TPR proteins are featured with repeat structure motifs, they are quite different. The difference of LRR proteins and TPR proteins are listed in the **Table 4-1**. Repeat proteins are involved in many important protein-protein interactions in most organisms, and LRR proteins have more rigid structural motifs to make their tertiary structures more repeatable among each other. Most LRR proteins form solenoid structures of a horseshoe with the motifs of an  $\alpha$ -helix/ $\beta$ -strand. Structures of four functionally non-related LRR proteins are well superimposed onto each other and the overlapped horseshoes have elongated and curved shapes with similar radii, although the length of each protein vary (**Figure 4-1**, left).

The conserved structural packing of the LRR family is probably contributed from their consensus sequence of the structural motifs, which is enriched with Leucine and a few other hydrophobic residues. For example, the consensus sequence in the LRR motif of human TLR4 (Toll-like receptor 4) consists of Leucine and Phenylalanine. These two

**Table 4-1. Comparison of LRR (Leucine rich repeat) and TPR (Tetratricopeptide repeat) proteins.**

<b>Comparisons</b>	<b>LRR proteins (Leucine rich repeats)</b>	<b>TPR proteins (Tetratricopeptide repeats)</b>
Motifs	An $\alpha$ -helix-turn- $\beta$ -strand	Two antiparallel helices
Packing	Stacking to form solenoid structures of a horseshoe	Paralleling to solenoid structures of a super helix
Folding	Independently folding, long or short hybrids folds correctly)	Less conserved TPRs, tend to depend on the neighbor TPRs or the protein tertiary structures
Consensus sequence	More conserved consensus sequence of structural motifs (from a large number of functionally unrelated proteins)	Consensus sequence of structural motifs are not identified
Repeat unit	$3_{10}$ -helices in a strict length (of one motif)	Repeat units of an average 34 residues (one TPR motif)
Function	Protein scaffold, mediate protein-protein interactions	Protein scaffold, mediate protein-protein interactions



**Figure 4-1. Leucine rich repeat proteins have repeatable structural framework with rigid structural motifs.**

Left: A structural superimposition of five different LRR proteins, with the name and PDB code of each protein labeled in the same color as the backbone in the structures.  
 Right: Illustration of the interactions that stabilize the packing of structure motifs in the human toll-like receptor 4. The warm pink sticks indicate Leucine and purple sticks indicate Phenylalanine.



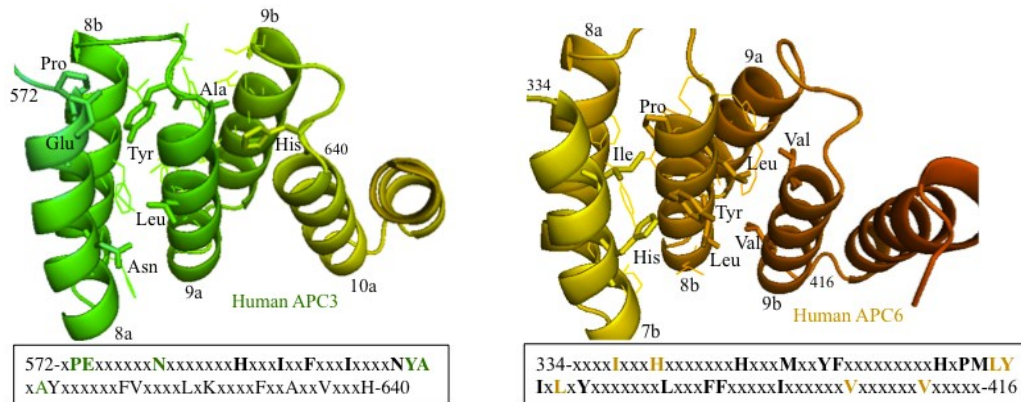
residues mainly mediate the inter-motif and intra-motif packing (**Figure 4-1**, right) and the simple interaction pattern is potentially responsible for rigid repeat units, which results in human TLR4 has a repeatable scaffold (tertiary structure) as other LRR proteins

Comparing to the LRR proteins interaction pattern, TPR proteins demonstrate complicate interaction “networks”. TPR protein is featured by TPR motifs packed parallel to each other. Instead of stacking to form an elongated horseshoe, most TPR proteins fold into a superhelix with their TPRs. The “networks” are possible because one  $\alpha$ -helix of a TPR motif is usually surrounded by four other  $\alpha$ -helices, rather than being simply flanked by an upstream and a downstream  $\alpha$ -helices. This local architecture allows one helix close to four  $\alpha$ -helices and the center  $\alpha$ -helix could establish interactions with all four of them. Both human APC3 and APC6 demonstrate this structure (**Figure 4-2**).

Sharing the complexity inside the TPR interaction networks, TPR protein residues that mediate the interactions show high diversity comparing to the LRR proteins. The rigid LRR motifs have consensus sequence, which make the LRR protein folding more predictable. Contrastingly, no consensus sequence could be identified for TPR motifs. Instead, TPR motifs use various hydrophobic residues to organize/support/build up the interaction networks. For instance, human APC3 TPR9a contacts TPR 8a and 10a besides establishing packing interactions with two neighbor  $\alpha$ -helices TPR 8b and 9b (**Figure 4-2**, left). A Tyr and a Leu of human APC3 TPR9a interact with a Glu and an Asn of the TPR8a, respectively, and an Ala from TPR9a interacts with a His of TPR10a. In another example of human APC6, APC6 TPR8b contacts an Ile and a His of TPR7b with a Pro and a Tyr, whereas two Leu on the other side of the TPR8b contact with two Val of TPR9b (**Figure 4-2**, right). The residues to mediate intra-TPR interactions could include all hydrophobic residues instead of a few. Furthermore, all these residues show no repeatable or predictable positions in the protein sequence. There are more complexities in TPR protein folding with respect to interaction networks and protein sequence.

### **The strength and weakness of the TPR hybrid technology**

There are advantages of hybrid protein technology. Generally, the successful hybrid proteins have better behaviors and showed the higher potential for crystallization, which offer the opportunities for structure studies and determination. Through the chimeric APC3 structures, it seems like the highly conserved TPR domains have higher chance to follow canonical folding in the hybrid proteins. The structures of these conserved domains provide the guidance for further studies and experimental designs. On the other hand, it is interesting to know that TPR proteins have the potential to adopt alternative folding to be soluble. Biologically, this could lower down the cell stress when the gene mutations happen or protein-folding machineries make mistakes. Furthermore, the domain swapping strategy also creates an approach to test the function of an essential domain which could neither be separated from the original protein nor be functional on its own (66).



**Figure 4-2. TPR proteins have interaction networks involved in TPR motif packing.**

Illustration of TPR protein interaction networks from human APC3 (left) and APC6 (right), with the residues mediating the interactions labeled out with lines and sticks. The lines indicate that residues mediate the center  $\alpha$ -helix packing to its neighbor  $\alpha$ -helices, and those residues are shown in the bottom sequence in black color. Similarly, sticks indicate the residues involved in the center  $\alpha$ -helix contacting distant  $\alpha$ -helices and they are shown at the bottom in green (human APC3, left) and orange (human APC6, right).

However, the risks of TPR hybrids are higher than the LRR hybrids. The interface of a hybrid protein potentially affects the downstream protein peptide folding since both LRR and TPR proteins pack based on the interaction among their consecutive structural units. Because TPR proteins are featured with more complicate interactions (cross-TPRs) mediated by a variety of residues, they are more sensitive to the domain swap. The hybrid YopM-InlB has the motif immediately following the YopM cap affected, while the rest of the hybrid protein kept the same scaffold (66). In the chimeric APC3, TPR7 failed to be stabilized by TPR6 and leads to the orientation of the whole protein changed.

Combining the consensus sequence and available atomic structures, the hydrophobic cores of LRR proteins are more stable and predictable than TPR proteins. Synthetic libraries of designed LRR proteins have been designed and used for generating artificial binders that replace antibodies (93). So far there is no such libraries for TPR proteins. However, the available atomic models from TPR proteins are increasing, which help to improvement from the protein structure prediction and analysis. As for TPR protein hybrids, the combination of multiple approaches will be required to design successful fusion structures, i.e. the homologous atomic models, structure prediction, EM-derived models. Meanwhile, the following strict and careful analyses are also essential to examine the authenticities of atomic structures.

## LIST OF REFERENCES

1. C.-J. Wu, D. B. Conze, T. Li, S. M. Srinivasula, J. D. Ashwell, Sensing of Lys 63-linked polyubiquitination by NEMO is a key event in NF- $\kappa$ B activation. *Nat Cell Biol* **8**, 398 (2006).
2. T. Tenno *et al.*, Structural basis for distinct roles of Lys63- and Lys48-linked polyubiquitin chains. *Genes to Cells* **9**, 865 (2004).
3. J. S. Thrower, L. Hoffman, M. Rechsteiner, C. M. Pickart, Recognition of the polyubiquitin proteolytic signal. *The EMBO Journal* **19**, 94 (2000).
4. D. Rotin, S. Kumar, Physiological functions of the HECT family of ubiquitin ligases. *Nat Rev Mol Cell Biol* **10**, 398 (2009).
5. R. J. Deshaies, C. A. P. Joazeiro, RING Domain E3 Ubiquitin Ligases. *Annual Review of Biochemistry* **78**, 399 (2009).
6. R. S. Hames, S. L. Wattam, H. Yamano, R. Bacchieri, A. M. Fry, APC/C-mediated destruction of the centrosomal kinase Nek2A occurs in early mitosis and depends upon a cyclin A-type D-box. *The EMBO Journal* **20**, 7117 (2001).
7. N. den Elzen, J. Pines, Cyclin a Is Destroyed in Prometaphase and Can Delay Chromosome Alignment and Anaphase. *The Journal of Cell Biology* **153**, 121 (2001).
8. O. Cohen-Fix, J. M. Peters, M. W. Kirschner, D. Koshland, Anaphase initiation in *Saccharomyces cerevisiae* is controlled by the APC-dependent degradation of the anaphase inhibitor Pds1p. *Genes & Development* **10**, 3081 (1996).
9. R. Ciosk *et al.*, An ESP1/PDS1 Complex Regulates Loss of Sister Chromatid Cohesion at the Metaphase to Anaphase Transition in Yeast. *Cell* **93**, 1067 (1998).
10. T. J. McGarry, M. W. Kirschner, Geminin, an Inhibitor of DNA Replication, Is Degraded during Mitosis. *Cell* **93**, 1043 (1998).
11. B. A. Buschhorn, J.-M. Peters, How APC/C orders destruction. *Nat Cell Biol* **8**, 209 (2006).
12. H. M. Kei Honda, Yuza Kato, Akio Yamaguchi, Hirofumi Tanaka, Hideyo Yasuda, Koichi Furukawa, Takeshi Urano, Degradation of human Aurora2 protein kinase by the anaphase-promoting complex-ubiquitin-proteasome pathway. *Oncogene* **19**, 8 (2000).
13. J. L. Burton, M. J. Solomon, Hsl1p, a Swe1p Inhibitor, Is Degraded via the Anaphase-Promoting Complex. *Molecular and Cellular Biology* **20**, 4614 (2000).

14. M. K. Summers, B. Pan, K. Mukhyala, P. K. Jackson, The Unique N Terminus of the UbcH10 E2 Enzyme Controls the Threshold for APC Activation and Enhances Checkpoint Regulation of the APC. *Molecular Cell* **31**, 544 (2008).
15. L. Jin, A. Williamson, S. Banerjee, I. Philipp, M. Rape, Mechanism of Ubiquitin-Chain Formation by the Human Anaphase-Promoting Complex. *Cell* **133**, 653 (2008).
16. C. A. P. Joazeiro *et al.*, The Tyrosine Kinase Negative Regulator c-Cbl as a RING-Type, E2-Dependent Ubiquitin-Protein Ligase. *Science* **286**, 309 (1999).
17. K. L. Lorick *et al.*, RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proceedings of the National Academy of Sciences* **96**, 11364 (1999).
18. B. R. Thornton *et al.*, An architectural map of the anaphase-promoting complex. *Genes & Development* **20**, 449 (2006).
19. N. Zheng *et al.*, Structure of the Cul1-Rbx1-Skp1-F boxSkp2 SCF ubiquitin ligase complex. *Nature* **416**, 703 (2002).
20. Z. Zhang *et al.*, The Four Canonical TPR Subunits of Human APC/C Form Related Homo-Dimeric Structures and Stack in Parallel to Form a TPR Suprahelix. *Journal of Molecular Biology* **425**, 4236 (2013).
21. F. Herzog *et al.*, Structure of the Anaphase-Promoting Complex/Cyclosome Interacting with a Mitotic Checkpoint Complex. *Science* **323**, 1477 (2009).
22. G. J. P. L. Kops *et al.*, APC16 is a conserved subunit of the anaphase-promoting complex/cyclosome. *Journal of Cell Science* **123**, 1623 (2010).
23. J. J. Frye *et al.*, Electron microscopy structure of human APC/CCDH1: EMI1 reveals multimodal mechanism of E3 ligase shutdown. *Nat Struct Mol Biol* **20**, 827 (2013).
24. C. L. Lawson *et al.*, EMDatabank.org: unified data resource for CryoEM. *Nucleic Acids Research* **39**, D456 (2011).
25. M. Schwickart *et al.*, Swm1/Apc13 Is an Evolutionarily Conserved Subunit of the Anaphase-Promoting Complex Stabilizing the Association of Cdc16 and Cdc27. *Molecular and Cellular Biology* **24**, 3562 (2004).
26. J. Wang, B. T. Dye, K. R. Rajashankar, I. Kurinov, B. A. Schulman, Insights into anaphase promoting complex TPR subdomain assembly from a CDC26-APC6 structure. *Nat Struct Mol Biol* **16**, 987 (2009).

27. Z. Zhang, K. Kulkarni, S. J. Hanrahan, A. J. Thompson, D. Barford, The APC/C subunit Cdc16/Cut9 is a contiguous tetratricopeptide repeat superhelix with a homo-dimer interface similar to Cdc27. *The EMBO Journal* **29**, 3733 (2010).
28. W. Zachariae *et al.*, Mass Spectrometric Analysis of the Anaphase-Promoting Complex from Yeast: Identification of a Subunit Related to Cullins. *Science* **279**, 1216 (1998).
29. H.-J. Yoon *et al.*, Role of Hcn1 and Its Phosphorylation in Fission Yeast Anaphase-promoting Complex/Cyclosome Function. *Journal of Biological Chemistry* **281**, 32284 (2006).
30. G. L. Blatch, M. Lässle, The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. *BioEssays* **21**, 932 (1999).
31. B. A. Buschhorn *et al.*, Substrate binding on the APC/C occurs between the coactivator Cdh1 and the processivity factor Doc1. *Nat Struct Mol Biol* **18**, 6 (2011).
32. P. C. A. da Fonseca *et al.*, Structures of APC/CCdh1 with substrates identify Cdh1 and Apc10 as the D-box co-receptor. *Nature* **470**, 274 (2011).
33. C. Kraft, H. C. Vodermaier, S. Maurer-Stroh, F. Eisenhaber, J.-M. Peters, The WD40 Propeller Domain of Cdh1 Functions as a Destruction Box Receptor for APC/C Substrates. *Molecular Cell* **18**, 543 (2005).
34. M. E. Matyskiela, D. O. Morgan, Analysis of Activator-Binding Sites on the APC/C Supports a Cooperative Substrate-Binding Mechanism. *Molecular Cell* **34**, 68 (2009).
35. Z. Zhang *et al.*, Molecular Structure of the N-terminal Domain of the APC/C Subunit Cdc27 Reveals a Homo-dimeric Tetratricopeptide Repeat Architecture. *Journal of Molecular Biology* **397**, 1316 (2010).
36. K. S. Wendt *et al.*, Crystal structure of the APC10/DOC1 subunit of the human anaphase-promoting complex. *Nat Struct Mol Biol* **8**, 784 (2001).
37. H. C. Vodermaier, C. Gieffers, S. Maurer-Stroh, F. Eisenhaber, J.-M. Peters, TPR Subunits of the Anaphase-Promoting Complex Mediate Binding to the Activator Protein CDH1. *Current Biology* **13**, 1459 (2003).
38. L. A. Passmore *et al.*, Doc1 mediates the activity of the anaphase-promoting complex by contributing to substrate recognition. *The EMBO Journal* **22**, 786 (2003).
39. Y. Kimata, J. E. Baxter, A. M. Fry, H. Yamano, A Role for the Fizzy/Cdc20 Family of Proteins in Activation of the APC/C Distinct from Substrate Recruitment. *Molecular Cell* **32**, 576 (2008).

40. X. Zeng *et al.*, Pharmacologic Inhibition of the Anaphase-Promoting Complex Induces A Spindle Checkpoint-Dependent Mitotic Arrest in the Absence of Spindle Damage. *Cancer Cell* **18**, 382 (2010).
41. E. Eytan, Y. Moshe, I. Braunstein, A. Hershko, Roles of the anaphase-promoting complex/cyclosome and of its activator Cdc20 in functional substrate binding. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2081 (2006).
42. Z. Hilioti, Y.-S. Chung, Y. Mochizuki, C. F. J. Hardy, O. Cohen-Fix, The anaphase inhibitor Pds1 binds to the APC/C-associated protein Cdc20 in a destruction box-dependent manner. *Current Biology* **11**, 1347 (2001).
43. W. Tian *et al.*, Structural analysis of human Cdc20 supports multisite degron recognition by APC/C. *Proceedings of the National Academy of Sciences* **109**, 18419 (2012).
44. W. C. H. Chao, K. Kulkarni, Z. Zhang, E. H. Kong, D. Barford, Structure of the mitotic checkpoint complex. *Nature* **484**, 208 (2012).
45. J. He *et al.*, Insights into Degron Recognition by APC/C Coactivators from the Structure of an Acm1-Cdh1 Complex. *Molecular Cell* **50**, 649 (2013).
46. R. W. King, M. Glotzer, M. W. Kirschner, Mutagenic analysis of the destruction signal of mitotic cyclins and structural characterization of ubiquitinated intermediates. *Molecular Biology of the Cell* **7**, 1343 (1996).
47. H. Yamano, C. Tsurumi, J. Gannon, T. Hunt, The role of the destruction box and its neighbouring lysine residues in cyclin B for anaphase ubiquitin-dependent proteolysis in fission yeast: defining the D-box receptor. *The EMBO Journal* **17**, 5670 (1998).
48. C. W. Carroll, M. Enquist-Newman, D. O. Morgan, The APC Subunit Doc1 Promotes Recognition of the Substrate Destruction Box. *Current Biology* **15**, 11 (2005).
49. C. W. Carroll, D. O. Morgan, The Doc1 subunit is a processivity factor for the anaphase-promoting complex. *Nat Cell Biol* **4**, 880 (2002).
50. S. W. N. Au, X. Leng, J. W. Harper, D. Barford, Implications for the Ubiquitination Reaction of the Anaphase-promoting Complex from the Crystal Structure of the Doc1/Apc10 Subunit. *Journal of Molecular Biology* **316**, 955 (2002).
51. X. Zeng, R. W. King, An APC/C inhibitor stabilizes cyclin B1 by prematurely terminating ubiquitination. *Nat Chem Biol* **8**, 383 (2012).

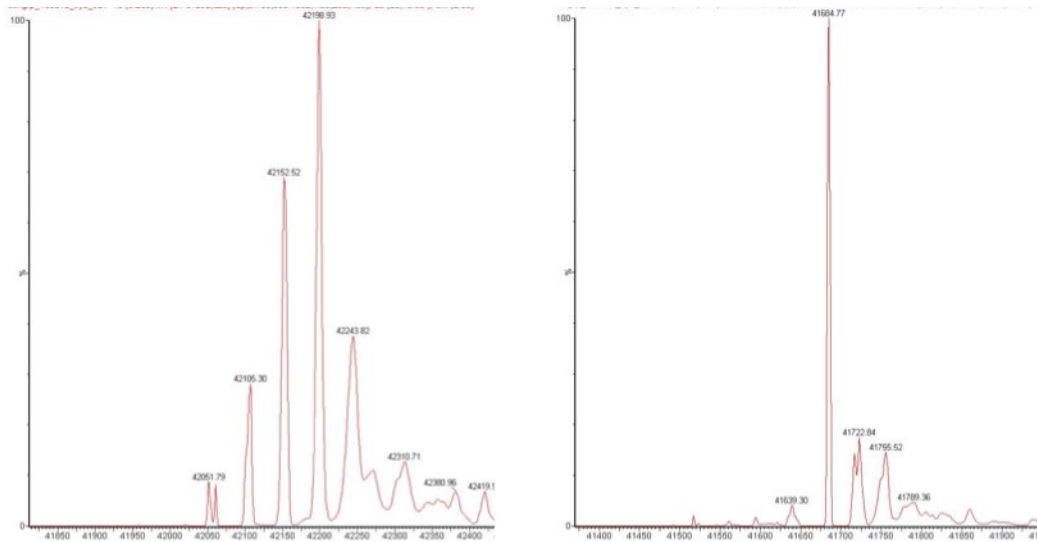
52. J. D. R. Reimann, B. E. Gardner, F. Margottin-Goguet, P. K. Jackson, Emi1 regulates the anaphase-promoting complex by a different mechanism than Mad2 proteins. *Genes & Development* **15**, 3278 (2001).
53. M. Ohe *et al.*, Emi2 Inhibition of the Anaphase-promoting Complex/Cyclosome Absolutely Requires Emi2 Binding via the C-Terminal RL Tail. *Molecular Biology of the Cell* **21**, 905 (2010).
54. J. J. Miller *et al.*, Emi1 stably binds and inhibits the anaphase-promoting complex/cyclosome as a pseudosubstrate inhibitor. *Genes & Development* **20**, 2410 (2006).
55. M. J. Hayes *et al.*, Early mitotic degradation of Nek2A depends on Cdc20-independent interaction with the APC/C. *Nat Cell Biol* **8**, 607 (2006).
56. J. Nilsson, M. Yekezare, J. Minshull, J. Pines, The APC/C maintains the spindle assembly checkpoint by targeting Cdc20 for destruction. *Nat Cell Biol* **10**, 1411 (2008).
57. D. Izawa, J. Pines, How APC/C-Cdc20 changes its substrate specificity in mitosis. *Nat Cell Biol* **13**, 223 (2011).
58. T. Wu *et al.*, UBE2S drives elongation of K11-linked ubiquitin chains by the Anaphase-Promoting Complex. *Proceedings of the National Academy of Sciences* **107**, 1355 (2010).
59. M. J. Garnett *et al.*, UBE2S elongates ubiquitin chains on APC/C substrates to promote mitotic exit. *Nat Cell Biol* **11**, 1363 (2009).
60. A. Williamson *et al.*, Identification of a physiological E2 module for the human anaphase-promoting complex. *Proceedings of the National Academy of Sciences* **106**, 18213 (2009).
61. M. C. Rodrigo-Brenni, D. O. Morgan, Sequential E2s Drive Polyubiquitin Chain Assembly on APC Targets. *Cell* **130**, 127 (2007).
62. M. Rape, M. W. Kirschner, Autonomous regulation of the anaphase-promoting complex couples mitosis to S-phase entry. *Nature* **432**, 588 (2004).
63. A. Walker, C. Acquaviva, T. Matsusaka, L. Koop, J. Pines, UbcH10 has a rate-limiting role in G1 phase but might not act in the spindle checkpoint or as part of an autonomous oscillator. *Journal of Cell Science* **121**, 2319 (2008).
64. M. L. Jin, JO, Application of hybrid LRR technique to protein crystallization. *BMB Reports* **41**, 5 (2008).
65. H. M. Kim *et al.*, Crystal Structure of the TLR4-MD-2 Complex with Bound Endotoxin Antagonist Eritoran. *Cell* **130**, 906 (2007).



66. D. Breitsprecher, E. Gherardi, W. Bleymuller, H. Niemann, Crystal structure of an engineered YopM-InlB hybrid protein. *BMC Structural Biology* **14**, 12 (2014).
67. Z. Xu, A. L. Horwich, P. B. Sigler, The crystal structure of the asymmetric GroEL-GroES-(ADP)<sub>7</sub> chaperonin complex. *Nature* **388**, 741 (1997).
68. D. T. Jones, Protein secondary structure prediction based on position-specific scoring matrices. (1999).
69. M. Karpenahalli, A. Lupas, J. Soding, TPRpred: a tool for prediction of TPR-, PPR- and SEL1-like repeats from protein sequences. *BMC Bioinformatics* **8**, 2 (2007).
70. T. Jones, Interactive electron-density map interpretation: from INTER to O. *Acta Crystallographica Section D* **60**, 2115 (2004).
71. C. Dalvit *et al.*, NMR-Based Screening with Competition Water-Ligand Observed via Gradient Spectroscopy Experiments: Detection of High-Affinity Ligands. *Journal of Medicinal Chemistry* **45**, 2610 (2002).
72. Z. Zhang *et al.*, Recombinant expression, reconstitution and structure of human anaphase-promoting complex (APC/C). *Biochemical Journal* **449**, 365 (Jan 15, 2013, 2013).
73. A. Roy, A. Kucukural, Y. Zhang, I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protocols* **5**, 725 (2010).
74. A. K. Das, P. T. W. Cohen, D. Barford, The structure of the tetratricopeptide repeats of protein phosphatase 5: implications for TPR-mediated protein-protein interactions. *The EMBO Journal* **17**, 1192 (1998).
75. T. S. Walter *et al.*, Lysine Methylation as a Routine Rescue Strategy for Protein Crystallization. *Structure* **14**, 1617 (2006).
76. I. Rayment, in *Methods in Enzymology*, Charles W. Carter, Jr., Ed. (Academic Press, 1997), vol. Volume 276, pp. 171-179.
77. P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallographica Section D* **66**, 213 (2010).
78. A. L. Morris, E. G. MacArthur Mw Fau - Hutchinson, J. M. Hutchinson Eg Fau - Thornton, J. M. Thornton, Stereochemical quality of protein structure coordinates. (1992).
79. S. Trowitzsch, C. Bieniossek, Y. Nie, F. Garzoni, I. Berger, New baculovirus expression tools for recombinant protein complex production. *Journal of Structural Biology* **172**, 45 (2010).

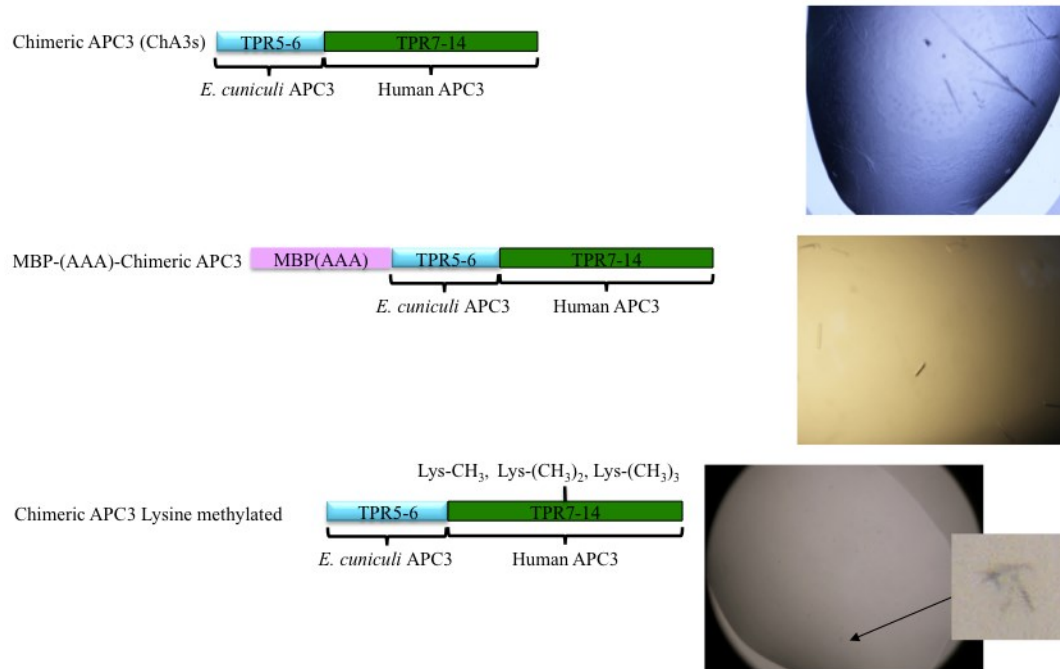
80. E. F. Pettersen *et al.*, UCSF Chimera—A visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605 (2004).
81. P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D* **60**, 2126 (2004).
82. Schrodinger, LLC. (2010).
83. M. Glotzer, A. W. Murray, M. W. Kirschner, Cyclin is degraded by the ubiquitin pathway. *Nature* **349**, 132 (1991).
84. C. M. Pflieger, M. W. Kirschner, The KEN box: an APC recognition signal distinct from the D box targeted by Cdh1. *Genes & Development* **14**, 655 (2000).
85. B. Hegemann *et al.*, Systematic Phosphorylation Analysis of Human Mitotic Protein Complexes. *Sci. Signal.* **4**, rs12 (2011).
86. J. R. A. Hutchins *et al.*, Systematic Analysis of Human Protein Complexes Identifies Chromosome Segregation Proteins. *Science* **328**, 593 (2010).
87. C. Dalvit, G. Fogliatto, A. Stewart, M. Veronesi, B. Stockman, WaterLOGSY as a method for primary NMR screening: Practical aspects and range of applicability. *J Biomol NMR* **21**, 349 (2001).
88. E. R. Kramer, C. Gieffers, G. Hölzl, M. Hengstschläger, J.-M. Peters, Activation of the human anaphase-promoting complex by proteins of the CDC20/Fizzy family. *Current Biology* **8**, 1207 (1998).
89. A. Schreiber *et al.*, Structural basis for the subunit assembly of the anaphase-promoting complex. *Nature* **470**, 227 (2011).
90. E. R. Kramer, N. Scheuringer, A. V. Podtelejnikov, M. Mann, J.-M. Peters, Mitotic Regulation of the APC Activator Proteins CDC20 and CDH1. *Molecular Biology of the Cell* **11**, 1555 (2000).
91. J. L. Gaillard, C. Berche P Fau - Frehel, E. Frehel C Fau - Gouin, P. Gouin E Fau - Cossart, P. Cossart, Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from gram-positive cocci.
92. C. N. LaRock, B. T. Cookson, The *Yersinia* virulence effector YopM binds caspase-1 to arrest inflammasome assembly and processing. (2012).
93. M. Stumpp, P. Forrer, H. Binz, A. Pluckthun, Designing repeat proteins: modular leucine-rich repeat protein libraries based on the mammalian ribonuclease inhibitor family. *J Mol Biol* **332**, 471 (2003).

**APPENDIX. SUPPLEMENTARY DATA FOR STRUCTURAL STUDIES OF THE  
TPR SUBUNITS OF THE HUMAN APC**



**Figure A-1. Intact masspectra of the SeMet incorporated and native chimeric APC3.**

The molecular weight of the SeMet incorporated and native chimeric APC3 are 42198.93Da and 41684.77 Da, detected by the Intact masspectrometry. A SelenoMethione has a molecular weight of 196.11Da and a Methionine weight 149.21Da. The incorporated Selenium sites are about 10 for each purified protein, which is calculated from  $(42198.93-41684.77)/(196.11-149.21)=10.35$ .



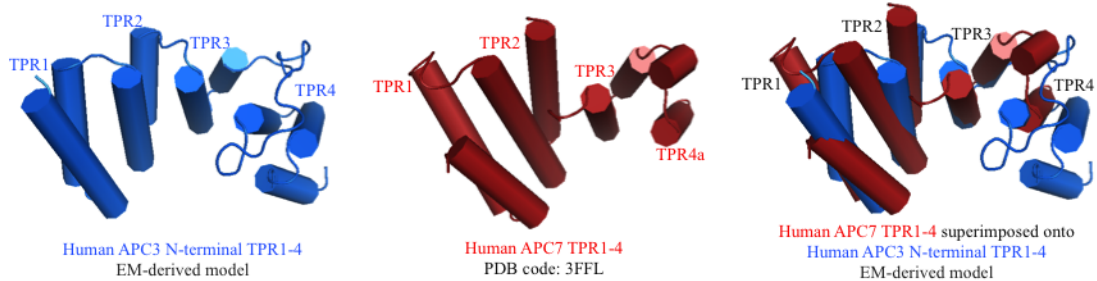
**Figure A-2. Different crystal forms of the chimeric APC3.**

Different crystal forms of the SeMet labeled (top), MBP-(AAA)- fused (middle), and Lysine methylated chimeric APC3, with the schematics of the corresponding constructs on the left. The crystallization conditions are:

Top: 0.1M HEPES pH 7.0, 0.2M AmSO<sub>4</sub>, and 48-51% MPD, at room temperature.

Middle: 0.1M BTP pH7.5, 20% PEG3350, and 0.2M NaI, at room temperature.

Bottom: Lysine methylated 0.1M CAPS pH10.5 and 40% MPD, at 4°C.

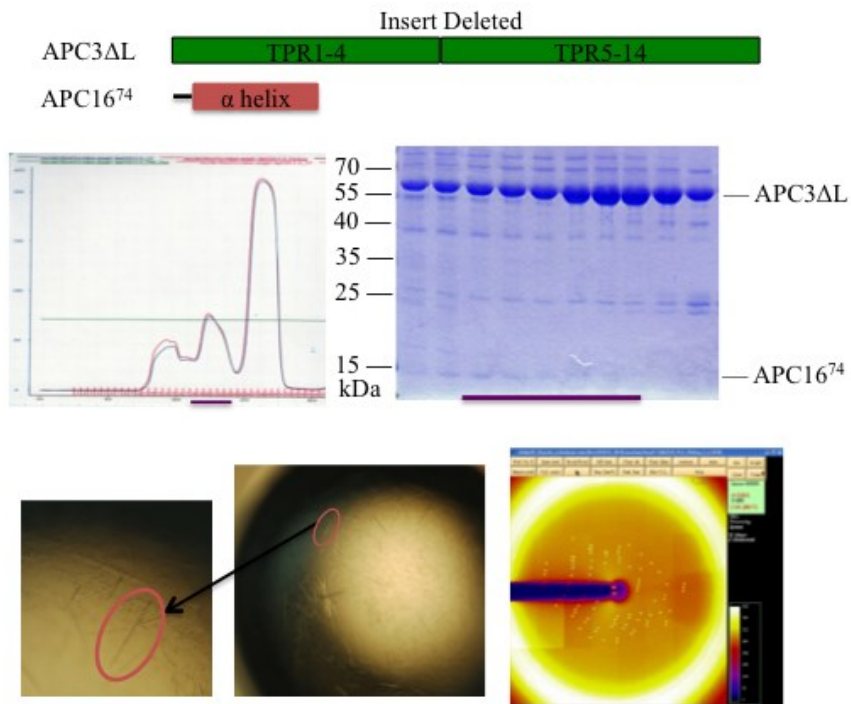


**Figure A-3. The structural comparison of the N-terminal APC7 and APC3.**

The 3D superimposition of the N-terminal human APC7 TPR1-4 (PDB code: 3FFL) onto the EM-derived human APC3 (TPR1-4). The  $\alpha$ -helices are displayed in the shapes of cylinders with the TPR numbers labeled.

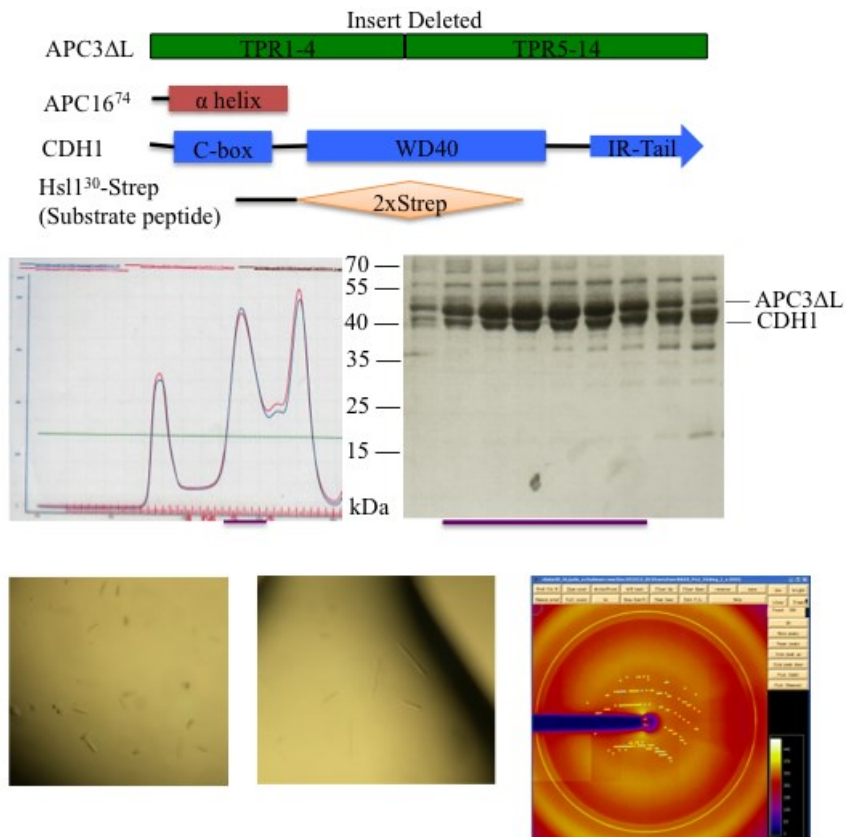
**Table A-1. A crystallization summary of the APC TPR subunits.**

<b>Complex</b>	<b>Details</b>	<b>Crystallization screening description</b>	<b>Hits</b>	<b>Crystal optimization summary</b>	<b>Results</b>
APC3ΔL-APC16 <sup>74</sup>	Human APC3Δ(183-452), with APC16 residues 74-110, co-purified, stoichiometry 2:1	Robotic screen around 2,000 conditions at 4°C & RT	Yes	Robotic screening of 250 conditions and manual screening/optimization in 500 conditions	Diffraction around 5Å
APC3ΔL-APC16 <sup>74</sup> -Cdh1fl-Hsl1 <sup>30</sup>	Human APC3Δ(183-452), APC16 residues 74-110, Cdh1 full-length and yeast Hsl1 includes KEN and D box, co-purified, stoichiometry 2:1:2:2	Robotic screen around 2,000 conditions at 4°C & RT	Yes	Robotic screening of 96 additives and manual screening/optimization in 290 conditions	Diffraction around 7Å
APC8fl-APC6s-Cdc26s	Human APC8 full-length with APC6residues 212-539, Cdc26 residues 1-29, co-purified, stoichiometry 1:1:1	Robotic screen around 2,600 conditions at 4°C & RT for each complex 4°C & RT	Yes	Robotically and manually screen 360 conditions	Diffraction around 3Å, but APC8 doesn't pack in the crystals
APC8fl-APC6l-Cdc26s	Human APC8 full-length with APC6residues 1-539, Cdc26 residues 1-29, co-purified, stoichiometry 1:1:1	Robotic screen around 2,800 conditions at 4°C & RT for each complex 4°C & RT	No	—	—
APC8fl-APC6fl-Cdc26fl	Human and zebrafish full-length proteins of APC8, APC6, Cdc26, co-purified, stoichiometry 1:1:1	Robotic screen around 2,400 conditions at 4°C & RT for each complex	No	—	—



**Figure A-4. Crystallization of the human APC3 $\Delta$ L-APC16<sup>74</sup>.**

A brief summary of the purification and the crystallization of the human APC3 $\Delta$ L-APC16<sup>74</sup>. Each component of the protein complex is schematically illustrated (top panel) The middle panel is a gel filtration profile from the purification of the protein complex (left), with the complex fractions examined by SDS-PAGE (right). The pooled fractions are labeled out with a purple line at the bottom of the gel. The purified APC3 $\Delta$ L-APC16<sup>74</sup> generated hits in the condition of 0.1M BTP pH 7.5, 10% PEG3350 at room temperature (bottom left) and the crystals diffracted to  $\sim 6\text{\AA}$ .



**Figure A-5. The crystallization summary of APC3 $\Delta$ L-APC16<sup>74</sup>-Cdh1-Hsl<sup>30</sup>.**

A brief summary of the purification and the crystallization of the human APC3 $\Delta$ L-APC16<sup>74</sup>-Cdh1-Hsl<sup>30</sup>. Each component in the protein complex is schematically illustrated (top panel). The middle panel is a gel filtration profile from the purification of the protein complex (left), with the complex fractions examined by SDS-PAGE (right). The pooled fractions are labeled out with a purple line at the bottom of the gel. The purified APC3 $\Delta$ L-APC16<sup>74</sup>-Cdh1-Hsl<sup>30</sup> generated hits at 4°C in the conditions of 0.1M BTP pH 7.0, 0.4M Mg-Formate (bottom left) and 0.1M MES pH 6.5, 0.5M AmSO4 (bottom middle). The crystals diffracted to  $\sim 7\text{\AA}$ .





## VITA

Shanshan Yu was born in Wuhan, Hubei Province, China in 1981. She graduated from Wuhan No.6 High School in Wuhan, Hubei Province in June 1999. The following August, she enrolled in the department for bioengineering at China Agricultural University. After achieving her Bachelor of Engineering degree in 2003, she directly joined in the graduate school at the same university and graduated with a Master of Biological Science in 2006. In 2008, she joined the Interdisciplinary Program at The University of Tennessee Health Science Center for further graduate studies. She is expected to obtain her Ph.D. degree in December 2014.