DIGITALCOMMONS
@WAYNESTATE

Wayne State University

Wayne State University Theses

1-1-2017

Expression, Purification And Characterization Of Lysine Methyltransferase Smyd5

Wen Xue
Wayne State University,

Follow this and additional works at: https://digitalcommons.wayne.edu/oa_theses

 Part of the Biochemistry Commons, and the Molecular Biology Commons

Recommended Citation
Xue, Wen, "Expression, Purification And Characterization Of Lysine Methyltransferase Smyd5" (2017). *Wayne State University Theses.* 596.
https://digitalcommons.wayne.edu/oa_theses/596

This Open Access Thesis is brought to you for free and open access by DigitalCommons@WayneState. It has been accepted for inclusion in Wayne State University Theses by an authorized administrator of DigitalCommons@WayneState.

Expression, Purification and Characterization of Lysine Methyltransferase SMYD5

by

Wen Xue

THESIS

Submitted to the Graduate School

Of Wayne State University,

Detroit, Michigan

In partial fulfillment of the requirements

for the degree of

MASTERS OF SCIENCE

2017

MAJOR: BIOCHEMISTRY AND MOLECULAR BIOLOGY

Approved by:

_____

Advisor Date

© COPYRIGHT BY

WEN XUE

2017

All Rights Reserved

DEDICATION

To everyone who contributed to the better version of myself

ACKNOWLEDGEMENTS

I would like to thank all of my committee members, Dr. Zhe Yang, Dr. Ladislau Kovari, Dr. David Evans, for the academic support and advice they have provided to me throughout the last two years. I am very grateful to have met such an inspirational and kind group of individuals. All of you truly make the Wayne State University feel like a community for me. I would also like to say a very special thanks to my mentor Dr. Zhe Yang, who has taught me how to learn things and solve problems.

Secondly, I want to thank all of my previous colleagues, whom I now have the privilege to consider close friends. Thank you for making me feel welcome and appreciated in the lab. I also want to thank all of you for training me and passing on your knowledge during our years of working together. I want to thank Josh and Nick for always being helpful and patient whenever I need. I really appreciate this. Also thank Junmei for sharing her good experience in Western Blot with me, Yingxue for showing me how to grow mammalian cells, Maysaa for placing all the orders, Stephanie for organizing the notebooks.

TABLE OF CONTENTS

DEDICATION.....ii

ACKNOWLEDGEMENTS.....iii

LIST OF TABLES.....v

LIST OF FIGURES.....vi

CHAPTER 1: Background.....1

CHAPTER 2: SMYD5 Molecular Cloning and Expression.....21

CHAPTER 3: SMYD5 Purification.....30

CHAPTER 4: GST-SMYD5 Molecular Cloning, Expression and Purification.....38

CHAPTER 5: SMYD5 Protein Characterization.....43

CHAPTER 6: Discussion.....48

CHAPTER 7: Future Direction.....50

CHAPTER 8: References.....51

ABSTRACT.....57

LIST OF TABLES

Table 1 SMYD Protein (Homo) Sequence Alignment.....9

Table 2: SMYD Proteins Percent Identity Matrix.....11

Table 3: Human SMYD5 DNA Sequence.....23

Table 4 Human SMYD5 Sequencing Result (pCDF-SUMO vector) 26

Table 5: Mitochondria Targeting Sequence.....46

LIST OF FIGURES

Figure 1:The catalysis of the methyl transfer 2

Figure 2: Histone lysine modification.....4

Figure 3: Human KMTs and their histone and non-histone substrate.....6

Figure 4: Overall structure of SMYD) proteins.....7

Figure 5: Role of H4K20me3 in the regulation of TLR4-responsive genes.....12

Figure 6: Bright-field microscopy of ES cells.....13

Figure 7: Knockdown of smyd5 results in normal growth of zebrafish embryo 15

Figure 8: Expression of markers for primitive hematopoietic lineages in smyd5 morphants 17

Figure 9: Expression of markers for definitive hematopoietic lineages in smyd5 morphants.....17

Figure 10:ErbB2-TGL cells stably transduced with empty vector or the indicated genes 19

Figure 11: hSMYD5 poly-E region.....20

Figure 12: hSMYD5 mitochondrial targeting sequence.....20

Figure 13: pCDF-SUMO vector map.....22

Figure 14: Agarose gel 24

Figure 15: Molecular cloning mechanism 26

Figure 16 : 15% SDS-page of large scale protein expression test 30

Figure 17: Histidine and Imidazole binds to Ni.....31

Figure 18 Gel filtration Column Mechanism.....32

Figure 19: Overview of SUMO-SMYD5 purification strategy 32

Figure 20: SUMO-SMYD5 Purification 1st Ni Column.....34

Figure 21: 15%SDS-PAGE of SUMO-SMYD5 Purification 1st Ni Column 34

Figure 22: SUMO-SMYD5 Purification Desalting Column.. 35

Figure 23: 15%SDS-PAGE of SUMO-SMYD5 Purification Desalting Column 35

Figure 24: SMYD5 Purification 2nd Ni Column Elution. 36
Figure 25: Gel filtration standard 37
Figure 26: SMYD5 Purification Gel Filtration Column. 37
Figure 27 : 15% SDS-PAGE of SMYD5 Purification Gel Filtration Column. 38
Figure 28: pGEX-6P-2 vector map 39
Figure 29: 12% SDS-PAGE of GST-SMYD5 expression test 40
Figure 30: GST-SMYD5 Purification strategy 41
Figure 31: GST-SMYD5 purification GST-trap column 42
Figure 32: 12%/15% SDS-PAGE OF GST-SMYD5 purification. 42
Figure 33: GST-SMYD5 purification desalting column 43
Figure 34: Sitting drop & Hanging drop 44
Figure 35: Crystal screen analysis 45
Figure 36: poly-E region of SMYD5 sequence 46
Figure 37: Ponceau S staining and western blot result of SMYD5 activity test 48
Figure 38: Predicted structure of SMYD5 (19-391) 50

CHAPTER 1: Background

1.1 Introduction

1.1.1 Protein post-translational modification

Protein post-translational modification (PTM) raises the functional variety of the proteome. Depending on the specific way of protein modification, such as extra bonding of function groups or proteolytic cleavage of subunits, it can be classified into phosphorylation, glycosylation, nitrosylation, ubiquitination, lipidation, acetylation, methylation, and proteolysis. Well understanding and identifying PTMs are important in the study of cell biology and disease treatment as these modifications impact almost every sides of pathogenesis and microbiology.

Protein methylation is a process that is transferring a one-carbon methyl group to an amino acid side chain of the protein. This transfer usually occurs on lysine or arginine, conjugating one, two or three methyl groups to residues. When methyl group bound to carboxylic acids, methylation neutralizes a negative amino acid charge and improves the hydrophobicity of protein. S-adenosyl methionine (SAM) is the primary methyl group donor, which has been suggested to be the second common used substrate in enzymatic reaction.

Lysine methylation is also a well-known machinery of epigenetic regulation, since histone lysine methylation and demethylation effects the transcriptional accessibility of DNA.

1.1.1.1 Non-histone Protein Lysine Methylation

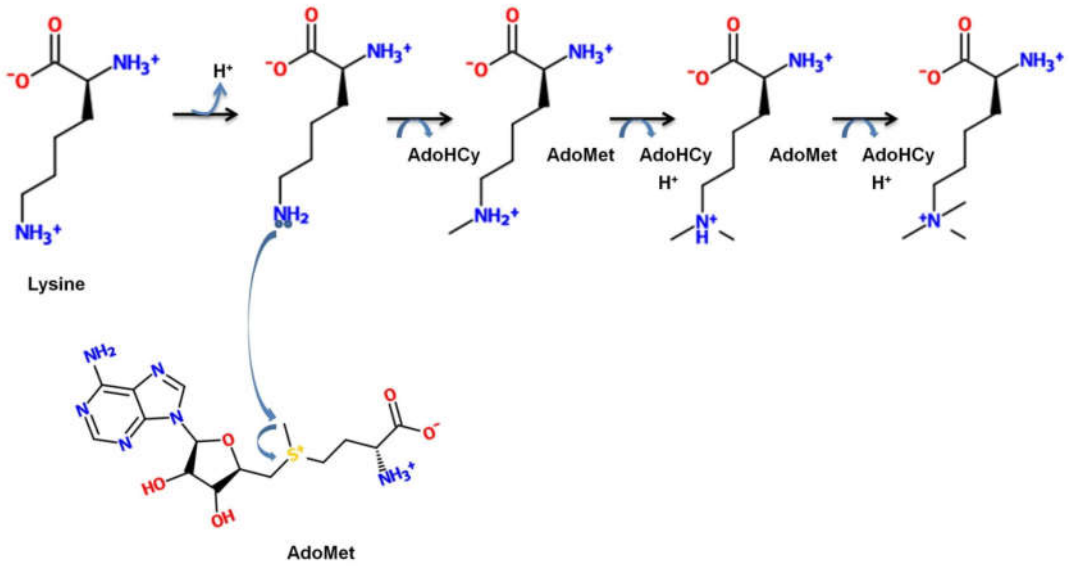


Figure 1:the catalysis of the methyl transfer from co-factor S-Adenosyl-Lmethionine (AdoMet) to the ε-amine of the lysine residue (1)

However, comparing to histone methylation, people are just getting to know non-histone methylation. The first reported event of non-histone protein methylation was mono-methylation of p53 protein by SET7 (2). Beside of protein p53, several other non-histone proteins have been studied as well. Most of those proteins are either histone-/DNA-modifying enzymes, like KMT1C and DNMT1, or transcriptional factors, like RB, E2F1, NFκB(3).

Methylation of lysine residues (Figure 1) on non-histone proteins has occurred as an important regulator of cellular signaling transduction that mediated by several signaling pathways, such as BMP, Hippo, JAK–STAT, MAPK, and WNT. Crosstalk between non-histone protein and histone methylation, and between methylation and other kinds of post-translational modifications, constantly emerges and affects biochemical cellular functions including chromatin remodeling, DNA repair, gene transcription, signal transduction and

protein synthesis. With recent progress in proteomic research and particular mass spectrometry, the stage is set to reveal the methylproteome and define its functions in health and disease.(4)

1.1.2 Epigenetic regulation

Epigenetic regulation consists of 3 highly interconnected epigenetic pathways: DNA methylation, histone translational modifications, and RNA modification, impacting on the chromatin structure and accessibility. Understanding the term of epigenetics requires a comprehension of chromatin structure. Chromatin, formed by repeating units called nucleosomes, which is a complex of protein and DNA/RNAs (5). A nucleosome are made up of double-stranded DNA wrapping around histone proteins, which is an octamer consists of two copies of each core histones H2A, H2B, H3, and H4. DNA can be chemically modified with epigenetic marks that affect the structure of chromatin by changing the affinity of interactions with proteins which bind to chromatin and by the chromatin electrostatic nature altering.

Comparing to DNA modification, histone modification is more dynamic. Histones are subjects to a few different covalent modifications, such as acetylation, methylation, sumoylation, phosphorylation and ubiquitination(Figure 2). Histone modifications could have diverse influences based on the location and type of modification on the histone. The best-identified histone modifications are acetylation and methylation. The enzymes that methylate histone are called histone methyltransferases. Histone methylation plays a vital role in transcriptional regulation, could be either suppressive or simulative, according to the location of the methylated residue. Methylation of the lysine at the fourth residue of histone H3

(H3K4me), for example, facilitates a transcriptionally active conformation, as opposite to H3K4me, where at the ninth lysine, H3K9me inhibit a transcriptionally conformation. Furthermore, H3K36me can be either suppressive or simulative, depending on proximity to a gene promoter region.

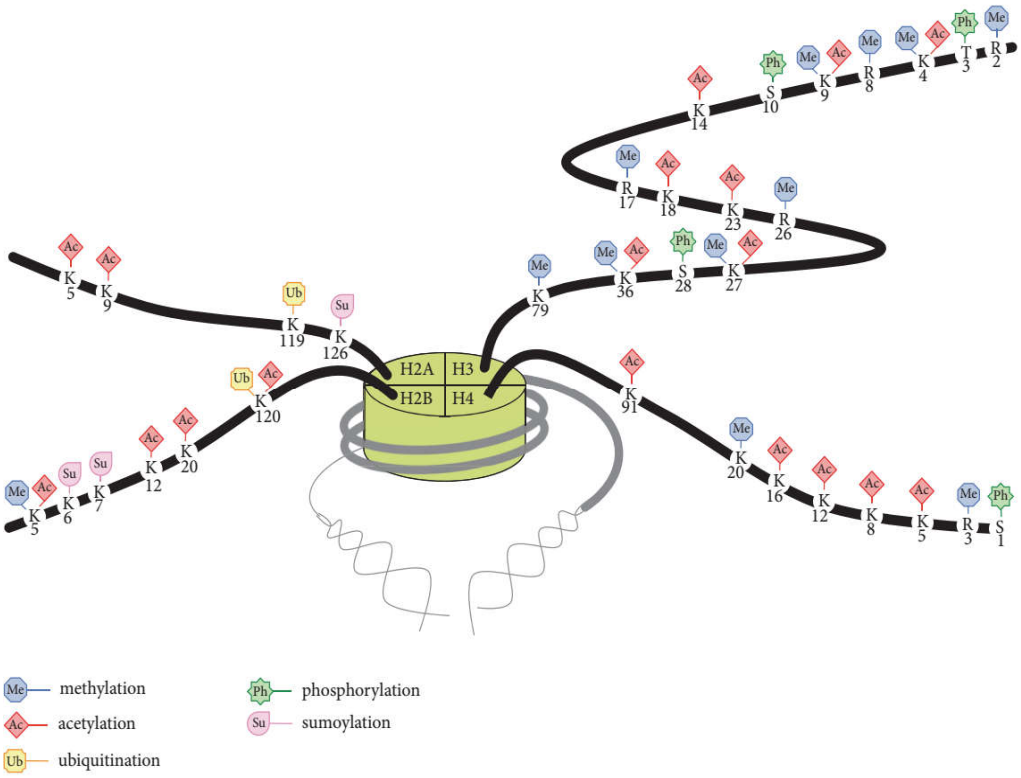


Figure 2: Histone lysine modification(6)

1.1.2.1 Histone Lysine Methylation

As this figure above, lysine is the majority of histone modification substrate. Lysine methylation is a covalent protein post-translational modification (PTM) that modulates a myriad of biological processes. The presence of this PTM on histone alters chromatin structure and function, which is involved in the regulation of DNA damage response,

transcription activity and cell cycle progression (5). Difficult to characterize in itself, methylation complexity also increases from the potential of mono-, di-, or tri- methylation of lysine side chain, producing multiple states for the single modification(7). The lysine methylation was identified to be localized on histone proteins in 1964 (8). While sites of methylation on histones were verified, the biological function of the histone methylation remained unknown. The biological consequence of the methylation of lysine residues on histones and the influences on genetic expression were first discovered through the formation of heterochromatin-mediated by the recruitment of HP1 to the H3K9 methyl mark (9), raising interest in lysine methylation on epigenetic regulation.

1.1.3 Lysine Methyltransferases

Lysine methyltransferases (KMTs) catalyze methylation by transferring one, two, or three methyl groups, from SAM to the ε-amino group of a lysine residue. Except for KMT4/DOT1L, all known KMTs are consist of a conserved SET (Su(var)3-9, Enhancer of Zeste, Trithorax) domain with the enzymatic activity (10). Besides the SET domain, most KMTs also have some other defined protein domain or homologous sequence to classify KMTs into distinct subfamilies (11). Naming enzymes were based on their relationship in sequence and domain structure, including the catalytic domain (Figure 3).The majority of KMT3 family focused on methylation of H3K36, whereas some SMYD proteins were observed to methylate H3K4 and other residues.

Family	Enzyme	Other name	Substrates		Reference[a]
			Histone	Non-histone	
KMT1	KMT1A	SUV39H1	H3K9me3		
	KMT1B	SUV39H2	H3K9me3		
	KMT1C	G9a/EHMT2	H3K9me2 H3K27me2 H1.4K26me2 H1.2K187me2	p53, G9a, C/EBPβ, Reptin, RARα, DNMT1, CDYL1, WIZ, ACINUS	[4–9]
	KMT1D	GLP/EHMT1	H3K9me2	p53	[4]
	KMT1E	ESET/SETDB1	H3K9me3	Tat	[10]
	KMT1F	SETDB2	H3K9me3		
KMT2	KMT2A	MLL	H3K4me3		
	KMT2B	MLL2	H3K4me3		
	KMT2C	MLL3	H3K4me3		
	KMT2D	MLL4	H3K4me3		
	KMT2E	MLL5	H3K4me3		
	KMT2F	hSET1A	H3K4me3	Dam1 (*Saccharomyces cerevisiae*)	[11]
	KMT2G	hSET1B	H3K4me3		
	KMT2H	ASH2	H3K4me3		
KMT3	KMT3A	SET2	H3K36me3		
	KMT3B	NSD1	H3K36me2 H4K20me2	NFκB	[12]
	KMT3C	SMYD2	H3K36me2 H3K4me	p53, RB	[13,14]
	KMT3D	SMYD1	H3K4me		
	KMT3E	SMYD3	H3K4me3	VEGFR	[15]
KMT4	KMT4	DOT1L	H3K79me2/3		
KMT5	KMT5A	SET8	H4K20me1	p53	[16]
	KMT5B	SUV420H1	H4K20me3		
	KMT5C	SUV420H2	H4K20me3		
KMT6	KMT6A	EZH2	H3K27me3		
	KMT6B	EZH1	H3K27me3		
KMT7	KMT7	SET7/9	H3K4me1	p53, TAF7, TAF10, ERα, AR, DNMT1, NFκB, PCAF, RB, E2F1, STAT3, Tat	[17–31]
KMT8	KMT8	PRDM2/RIZ1	H3K9me3		

Figure 3: Human KMTs and their histone and non-histone substrates(12)

1.2 SMYD protein family

The SMYD (SET and MYND domain) family proteins, which are a special group of protein lysine methyltransferases participated in methylation of histones and non-histone targets, plays pivotal roles in numerous cellular processes including gene expression regulation and DNA damage response*(13)*.

So far, SMYD family has five members, SMYD1–5(Figure 4)(14). Each member contains a conserved SET (Suppressor of variegation, Enhancer of Zeste, Trithorax) domain interrupted by an MYND (Myeloid-Nervy-DEAF1) domain(15). The SET domain, which is approximately 130 amino acid long, is a conserved catalytic unit responsible for lysine methylation and found in nearly all histone methyltransferase(16). The MYND domain is a zinc finger motif that basically functions as a protein–protein interaction module(17,18) and has been suggested that has a preference for binding with a proline-rich motif (PXLXP) in the

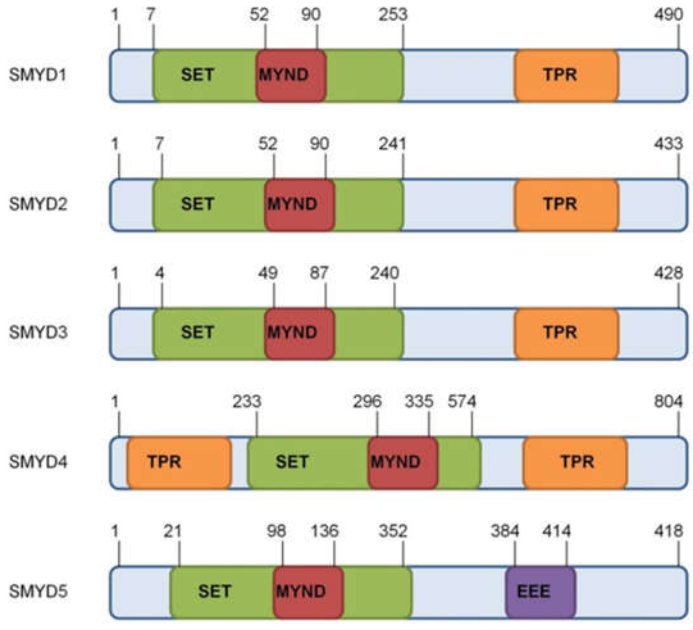


Figure 4: Overall structure of SMYD) proteins.(19)

interaction with other protein(20). The last feature is the TPR (Tetratrico Peptide Repeat) domain found in SMYD1–4 but not in SMYD5, which are important for the modulation of protein-protein interactions, and SMYD4 contains an additional TPR domain at its N-terminus.

1.2.1 SMYD protein sequence analysis

As shown in (Figure 4), the structures of SMYD proteins share a lot of commons. However, according to the alignment of the protein sequence of SMYD proteins, the percent identity of two SMYD protein (homo sapiens) members is approximately 20%-30% (Table 1), indicate the SMYD proteins are diverse in metazoans. So, to understand the history of evolutionary and functional conservation or diversification, a critical phylogenetic reconstruction of this complicated gene family is significant. The study of phylogenetic tree analysis of SMYD protein sequences has been started, and the results indicate that the existing metazoan SMYD genes turned into three main classes, SMYD4, SMYD5 and SMYD3 (including chordate-specific SMYD1, SMYD2 genes). SMYD1 and SMYD2, which are most close to each other in this five members, are most similar to SMYD3. The reason SMYD1 and SMYD2 are under SMYD3 class is they are respectively exclusive of chordates and vertebrates but SMYD3 exist in all species of metazoans. The evolutionary of the SMYD3 class and SMYD5 class is relatively simple. However, the SMYD4 class has passed through a few events such as gene duplication, gene loss, and lineage-specific expansions in the animal phyla. SMYD4 gene has dupicated into more sub-group but they are not redundant since four

SMYD4 genes in Drosophila melanogaster are expressed different patterns and deletion of individual genes behave diverse phenotypes (21).

Table 1: SMYD Proteins Percent Identity Matrix. According to protein sequence alignment of SMYD family, this table shows percent identity between two SMYD proteins.

"100" - 100% identity.

	SMYD5	SMYD4	SMYD1	SMYD3	SMYD2
1: SMYD1	25.10	20.86	100.00	31.12	30.61
2: SMYD2	27.86	21.81	30.61	32.37	100.00
3: SMYD3	26.09	23.35	31.12	100.00	32.37
4: SMYD4	22.95	100.00	20.86	23.35	21.81
5: SMYD5	100.00	22.95	25.10	26.09	27.86

1.2.2 SMYD family function in biological development

The SMYD family members SMYD1-4 are required in critical developmental processes, especially related to cardiac and muscle development.

SMYD1 is essential for heart development via cardiomyocyte differentiation and maturation(22) and its expression in the heart is directed by myocyte enhancer factor-2c (MEF2C)(23) and serum response factor (SRF)(24). In skeletal muscle cells, its transcription is regulated by myogenic differentiation antigen 1 (MYOD)(25). SMYD1 is expressed during the growth of zebrafish and is necessary for proper development of skeletal muscle, myofibril organization and muscle contraction(26), as well as sarcomerogenesis.

Additionally, SMYD2 is expressed in cardiac muscle as well during cardiogenesis (27) and both SMYD1 and SMYD2 enzymes are expressed during muscle development in

Xenopus laevis (28). SMYD2 is also reported to methylate HSP90 and altered chick muscle functions(29).

In zebrafish, SMYD3 is vital in heart and trunk muscle development. Knockdown of SMYD3 ends up in abnormal expression of heart chamber markers and myogenic regulatory factors(30).

The knockdown of SMYD4 in embryos of Drosophila cause 80% of knockdown flies remained trapped in the pupal case and died during the eclosion stage. Eclosion requires movement of the abdominal muscles to make flies be able to escape from the case. Suggesting that s4 is likely to play a role in controlling muscle development(31).

However, SMYD5 is not required for heart and skeletal muscle development. Recent research suggests that SMYD5 plays an important role in hematopoiesis(32).

1.2.3 SMYD family methylate histone and non-histone protein

SMYD proteins methylate a diverse group of histone and non-histone proteins that contribute to their critical roles in cell regulation such as chromatin remodeling, transcription, signal transduction, and cell cycle control. Methylation substrates of SMYD family members are summarized in the following table (Table 2).

Table 2: Summary of known substrates for the SMYD family of methyltransferases.(20,33)

Enzyme name	Substrate	
	Histone	Non-histone
SMYD1	H3K4	Myosin, skNAC HSP90
SMYD2	H3K4 H3K36	p53, RB1, PARP1, ERα HSP90
SMYD3	H3K4 H4K5	VEGFR1 MAP3K2 HSP90
SMYD4	?	?
SMYD5	H4K20	?

1.3 SMYD5

1.3.1 SMYD5 methylate H4K20

According to a recent report, histone H4 lysine 20 (H4K20) can be trimethylated by SMYD5(34). Associated with NcoR chromatin remodeling complexes, the methylation represses the inflammatory response through restricting of toll-like receptor 4 (TLR-4) mediated expression in mouse primary macrophages. Removal of H4k20me3 by PHF2 can activate the expression (Figure 5).

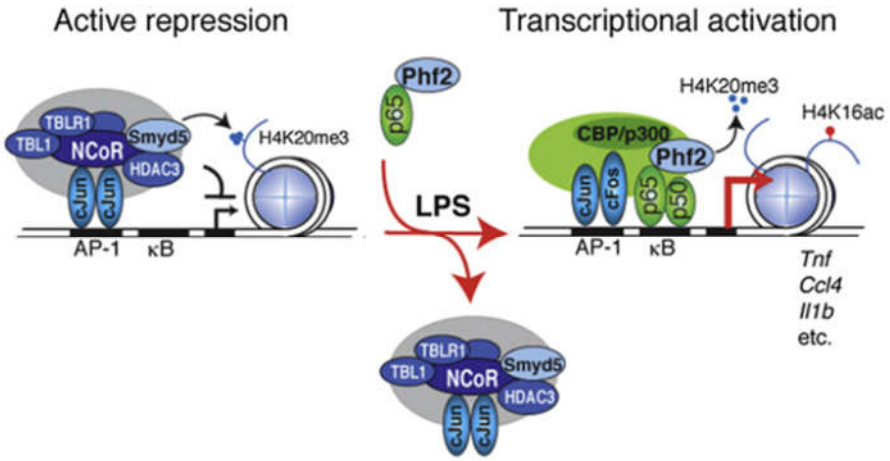


Figure 5: Role of H4K20me3 in the regulation of TLR4-responsive genes.(34)

However, the roles of this repressive histone modification in biological development and pluripotency are largely elusive.

Epigenetic regulations of chromatin states are assumed that involved in the self-renewal and differentiation of embryonic stem (ES) cells. The latest study showed that the histone lysine methyltransferase SMYD5 regulates ES cell self-renewal, alters cell differentiation (35).

Knockdown of SMYD5 leads to ES cell colony morphology abnormality, where shSmyd5 ES cell colonies became flat and scattered and lost contacting between cells at the colony periphery (Figure 6). Also, deletion of SMYD5 results in dysregulate expression of pluripotency regulators OCT4 targets, and disturbed ES cell differentiation.

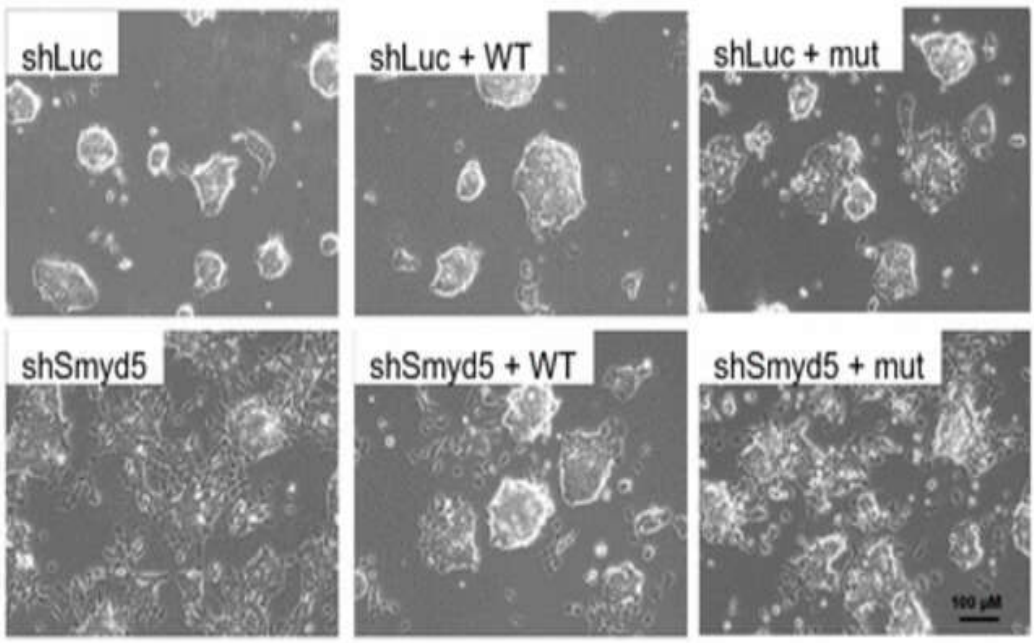


Figure 6: Bright-field microscopy of ES cells infected with shLuc or shSmyd5 lentiviral particles and wild-type (WT) SMYD5 or an enzymatically mutant (mut) version of SMYD5 (H315L and C317A) lentiviral particles and stably selected with puromycin and G418.(35)

This paper also reported that SMYD5 regulates H4K20me3 modification at heterochromatin in ES cells.

Heterochromatin is a tightly packed formation of DNA and localizes at the periphery of the nucleus. It was thought inaccessible of polymerases and not transcribed. However, since 2007, many studies have demonstrated that it is actually transcribed into noncoding RNA molecules containing telomeric and subtelomeric sequences(36). Heterochromatin plays an essential role in gene expression during development and

differentiation (37) and maintains genome integrity. It was identified that associate with H3K9 methylation, which is controlled by LSD1 and ESET/Setdb1(38), and H4K20 methylation.

Depletion of SMYD5 led to a downregulation in H4K20me3 levels, but not H4K20me2 or H4K20me1, demonstrating that SMYD5 endows H4K20me3 methyltransferase activity. Overexpressing a shRNA-resistant version of wild-type SMYD5 restored the level of H4K20me3. Knockdown of SMYD5 also reduced levels of H3K9me2/3 and HP1α. Because H4K20me3 is reported to co-localize with H3K9me3 at heterochromatic regions (5,39) and H3K9me3 is necessary for recruitment of HP1 and heterochromatin formation, it is possible that delocalization of HP1 and H3K9me3, depending upon the lacking of SMYD5 and H4K20me3, may tend to reduced heterochromatin.

Overall, a loss of SMYD5, which silence heterochromatin adjacent genic domains, results in increased expression of lineage-specific genes, contributing to the decreased self-renewal and disturbed differentiation of SMYD5-kockdown ES cells.(35)

1.3.2 SMYD5 regulate hematopoiesis in zebrafish embryos

Unlike SMYD1-4, Smyd5 is not required for heart and skeletal muscle development. Knockdown of SMYD5 in zebrafish embryo showed no abnormality (Figure 7)(32).

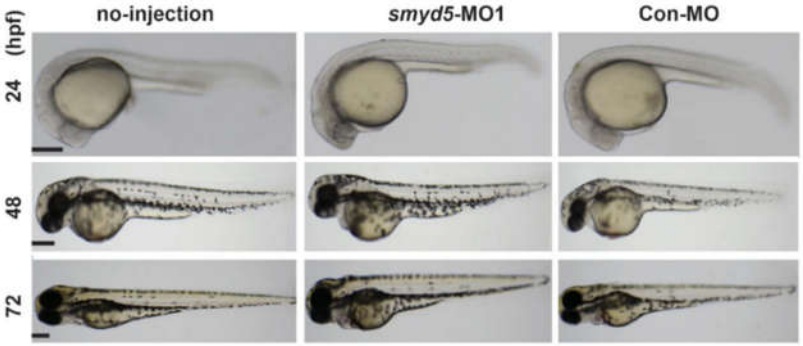


Figure 7: Knockdown of smyd5 results in normal growth of zebrafish embryo, including heart and skeletal muscle. (32)

Also, the gene expression patterns of cardiac and myogenic markers, the structure of sarcomere of heart, fast-and-slow skeletal muscle were indistinguishable between control embryo and knock down embryo(32). These results indicate that SMYD5 has physiological functions which are distinct from those played by the other members.

With the goal of defining the function of Smyd5 in zebrafish, scientists focused on the development of hematopoietic cells, derived from the mesoderm as heart and skeletal muscle. Hematopoietic cells generate blood cells of all lineages through the process of hematopoiesis. Similar to the other vertebrates, zebrafish hematopoiesis has two stages, primitive and definitive hematopoiesis(40).

To investigate the role of SMYD5, they tested the expression of genes related to hematopoiesis by performing a method of whole-mount in situ hybridization (WISH), to detect the target gene expression in the whole embryo. When knockdown of SMYD5, the expression of pu.1, mpx, l-plastin (Figure 8) increased during primary state and mpx, l-plastin, cmvb(Figure 9) increased during definitive state (32). The elevated expression of myeloid markers in SMYD5 loss-of-function zebrafish embryo, suggests that SMYD5 negatively regulates the expression of genes related to primitive and definitive myelopoiesis in zebrafish.

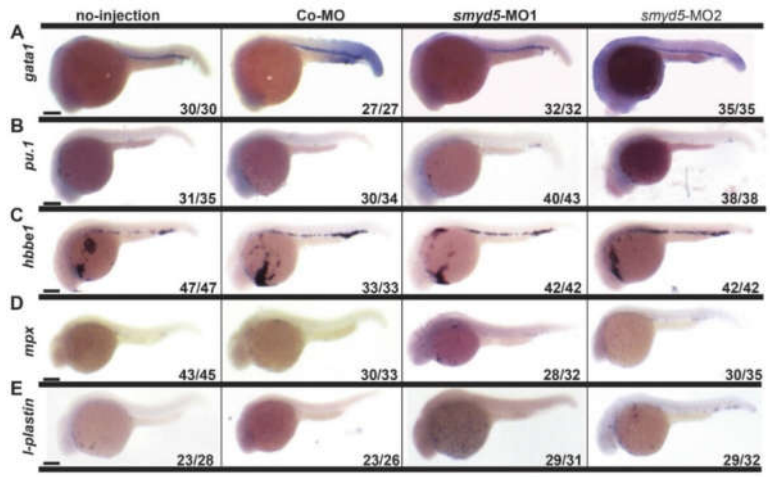


Figure 8: Expression of markers for primitive hematopoietic lineages in smyd5 morphants by WISH(32)

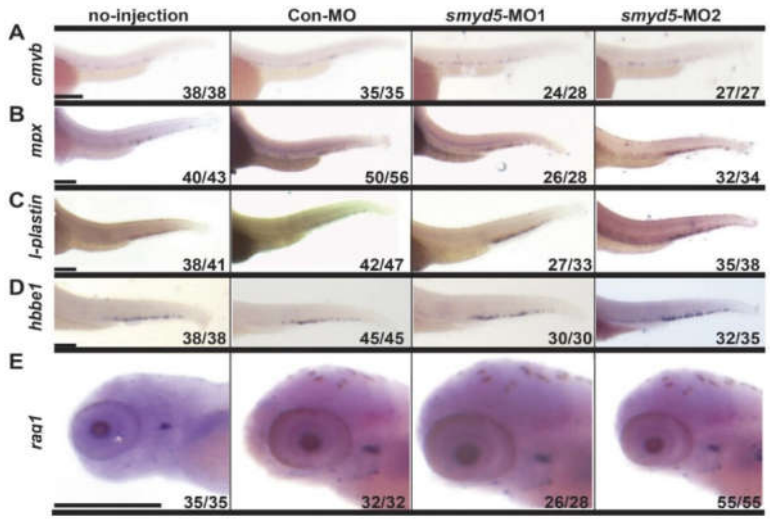


Figure 9: Expression of markers for definitive hematopoietic lineages in smyd5 morphants by WISH.(32)

1.3.3 SMYD5 in Cancer

SMYD5 was reported to involve in breast cancer, the most common among women and the second most frequent cancer worldwide.

In the present study, people explored the expression profile of SMYD5 in breast tumors and contiguous non-tumor samples, in breast cancer cells and some normal tissues. qPCR revealed expression level of SMYD5 was decreased in 7 breast cancer cell lines comparing to normal breast tissues. These findings led them to investigate the role of SMYD5 in cell proliferation. They observed that downregulation of SMYD5 expression through shRNA in HEK293 cells induced culture growth acceleration.(41)

Furthermore, the generation of MDA-MB-231 tumor cell sublines expressing different levels of SMYD5 revealed varying proliferative profiles. SMYD5 expression oppositely related with growth rate in highly proliferative cultures but slow proliferating cells do not seem to be altered by SMYD5 levels. This observation shows that specific subgroup of highly proliferative cells might depend on SMYD5 downregulation to display this phenotype. (41)

However, another study revealed that SMYD5 inhibited the capacity of metastatic breast cancer cells to colonize the lung.(42)

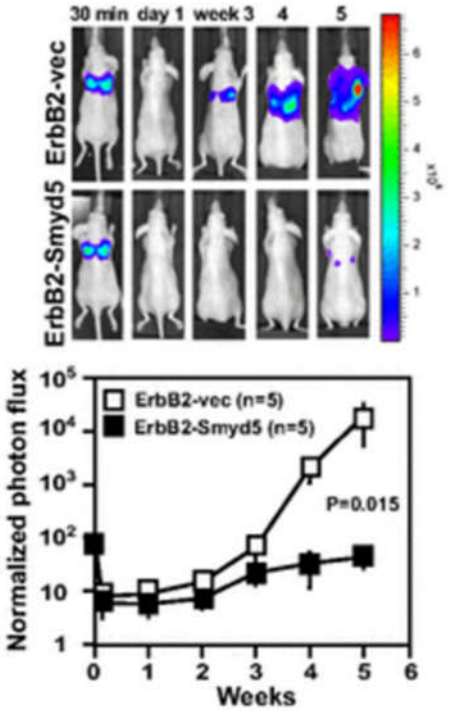


Figure 10: ErbB2-TGL cells stably transduced with empty vector or the indicated genes (Numb, Left; Smyd5, Right) were inoculated i.v. into syngeneic mice. Lung metastasis was measured by bioluminescent imaging. The panels show representative images (Top), and the graphs show the normalized photon flux at the indicated times (Bottom). (42)

1.4 Statement of Hypothesis

Biological roles of SMYD5 have started to emerge in the area of immune reaction, embryonic stem cell regulation, hematopoiesis regulation, and cancer inhibition, but the biochemical and structural features needed to facilitate SMYD5 mediated methylation of histone H4, as well as other potential substrates, has not been clearly explored.

The lacking of C-terminal TPR domain makes SMYD5 different from the other protein members while a poly-E region (Figure 11) at the end of C-terminal draw our attention, which tends to associate with metal ions, like calcium.



Figure 11: hSMYD5 poly-E region(19)



Figure 12: hSMYD5 mitochondrial targeting sequence

Moreover, we found that SMYD5 contains a mitochondrial targeting sequence (Figure 12) at the N-terminal indicating that SMYD5 could potentially be a mitochondrial protein. It is well-known that mitochondria is a calcium sensor which confirmed our assumption about the relation between SMYD5 poly-E tail and calcium ions. According to these, we assuming that SMYD5 might be a mitochondrial protein and its poly-E tail could associate with calcium thus affect the enzymatic activity of SMYD5.

Hypothesis:

- Crystal structure would provide better view of its specific biochemical function and structural features.
- SMYD5 has a potential to be a mitochondrial protein.
- The poly-E tail of SMYD5 might affect protein activity.

Aims:

- Overexpression of SMYD5 in E.coli.
- Obtain purified SMYD5 to perform biochemical and structural studies.
- Determine the crystal structure of SMYD5 by using X-ray crystallography.
- Characterize the enzymatic activity of methyltransferase SMYD5.

CHAPTER 2: SMYD5 Molecular cloning and expression

To perform biochemical and structural studies, the protein preparation must be pure and homogeneous, molecular cloning and expression are inevitable processes to get purified SMYD5 protein.

2.1 Molecular cloning

2.1.1 Material and Method

To get human SMYD5 gene, DH10B cells were purchased from the Open Biosystem company. Gene accession number is BC073806 (Table 3), and a full-length open frame sequence of SMYD5 is 1254 base pairs (bp). SMYD5 gene was in the POTB7 vector. To create a plasmid that SMYD5 gene will be inserted in pCDF-SUMO vector (Figure 13), primers were designed to curtail the interest gene from its original vector by PCR.

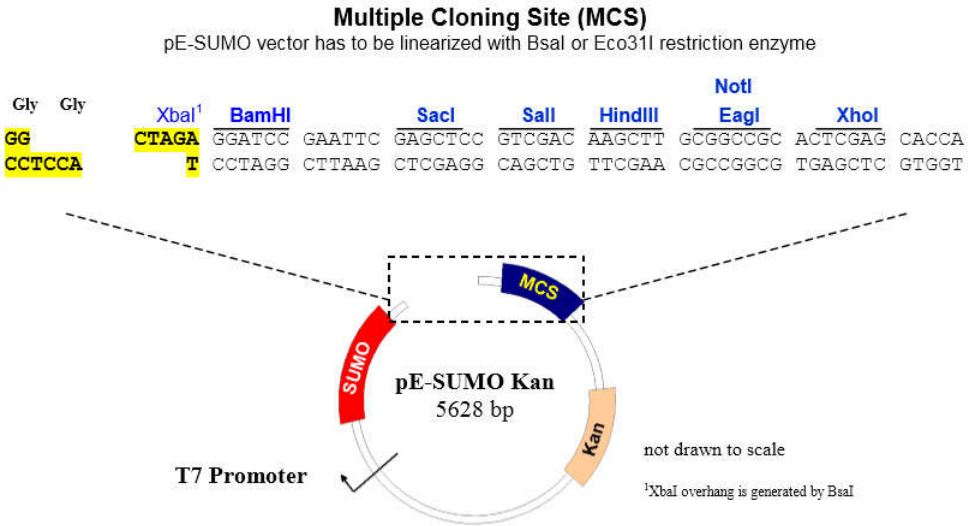


Figure 13: pCDF-SUMO vector map

5' primer length is 36 bp, GTATAAGAAGACATAGGTATGGCGGCCTCCATG TGC(Tm=64.5ºC), included a BbsI restriction site. 3' primer is 33 bp, CAC GCCTCGAGTTACACATCAGTCATCTCATC (Tm=63.6ºC), included a XhoI restriction site. Ordered primer from Invitrogen Company.

Table 3: human SMYD5 DNA sequence

CDS:

>gi|154689857:30-1286 Homo sapiens SMYD family member 5 (SMYD5), mRNA
ATGGCGGCCTCCATGTGCGACGTGTTCTCCTTCTGCGTGGGCGTGGCGGGCCGCGCGCGGGTCTCCGTGG
AAGTCCGTTTCGTGAGCAGCGCCAAGGGAAAGGGGCTGTTTGCCACACAGCTCATCCGGAAGGGGGAGAC
CATCTTCGTAGAACGGCCCCTGGTGGCTGCACAGTTTCTCTGGAATGCACTTTATCGCTACCGAGCCTGT
GACCACTGCCTTAGGGCACTAGAGAAGGCAGAGGAGAATGCCCAGAGGCTGACCGGGAAACCAGGCCAGG
TTCTGCCTCACCCAGAGCTGTGCACTGTGCGCAAAGACCTCCACCAGAACTGTCCCCATTGCCAAGTGAT
GTACTGCAGTGCAGAATGTCGGTTGGCAGCCACTGAGCAATACCACCAGGTCCTGTGCCCAGGCCCCTCC
CAGGATGACCCCTTGCATCCTCTCAATAAGCTTCAGGAGGCATGGAGGAGTATTCACTACCCACCTGAGA
CTGCAAGCATCATGTTGATGGCTAGGATGGTGGCCACAGTGAAGCAGGCGAAGGACAAGGACCGTTGGAT
CAGACTCTTTTCCCAGTTTTGTAACAAAACAGCCAATGAAGAGGAGGAAATTGTCCATAAACTTCTGGGA
GACAAATTCAAGGGCCAACTGGAACTTCTGCGGAGACTCTTCACAGAGGCCCTCTATGAGGAAGCAGTCA
GCCAGTGGTTCACTCCAGATGGATTCCGGTCTCTCTTTGCTCTTGTTGGGACCAATGGCCAAGGAATCGG
GACCAGCTCCCTAAGCCAGTGGGTCCATGCCTGTGACACTCTGGAGTTGAAGCCTCAGGACCGTGAGCAG
CTTGACGCCTTCATTGACCAGCTATACAAGGACATCGAGGCAGCAACTGGAGAGTTTCTTAACTGTGAAG
GATCTGGCCTCTTTGTGCTTCAGAGCTGCTGCAACCACAGTTGTGTGCCCAATGCAGAGACCTCCTTTCC
AGAAAACAACTTCCTTTTGCATGTCACTGCTCTGGAGGATATTAAGCCAGGAGAGGAAATTTGTATCAGC
TACTTGGACTGCTGTCAGCGGGAGCGCAGCCGCCACAGCCGCCACAAGATCCTCAGGGAGAACTATCTAT
TTGTCTGTTCCTGTCCCAAATGCCTGGCAGAGGCTGATGAACCCAATGTGACCTCAGAAGAGGAAGAGGA
AGAGGAGGAGGAGGAGGAAGGAGAGCCAGAAGATGCAGAGCTGGGGGATGAGATGACTGATGTGTGA

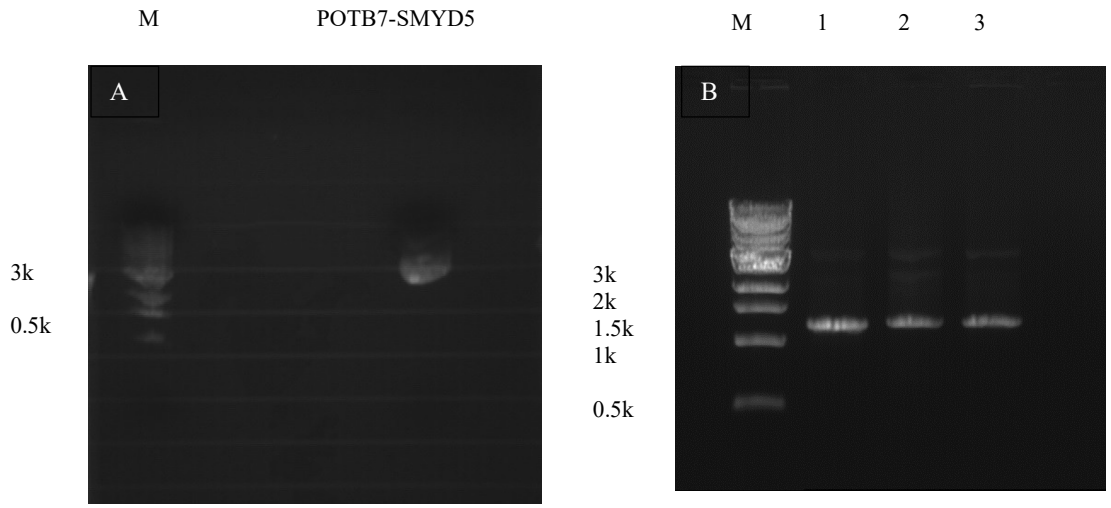


Figure 14: 0.8% Agarose gel (A) SMYD5 gene in POTB7 vector, bands located as expected (B) SMYD5 gene PCR product. 1/2/3 indicated 2Mm/3Mm//4Mm MgCl2

First, SMYD5 commercial plasmid is going to transferred into DH5α cell and saving as glycerol stock. DH10B cells grew in 5ml LB media contains 34 µg/mL chloramphenicol at 37 ºC overnight. The clones were purified by mini prep kit (Qiagen Company) and got purified POTB7-SMYD5 plasmid. Product was analyzed on a 0.8% agarose gel for 30mins at 100V to check the size, which should be around 3k as the size of SMYD5 is 1254 bp and POTB7 vector length is 1815 bp (Figure 14). 1.0 ng of vector was transformed into 20 µL of DH5α cells and heat shocked for 45 seconds at 42˚C. Then, a 250µL volume of warm SOC medium was added and transformed cells were shaken for 1hour. Next, 50µL volume of transformed cells were plated onto chloramphenicol resistant LB Agar plates. LB plates were grown overnight at 37˚C. One single colony was selected and grew in in 5mL LB media contains 34 µg/mL chloramphenicol at 37 ºC overnight. DH5α culture was aliquoted into 1mL with 300 µL glycerol in 2ml coring tubes, saved in -80 ºC for future use.

Secondly, SMYD5 gene is going to be inserted into pCDF-SUMO vector in DH5α cells. SMYD5 gene was amplified through PCR, then analyzed on a 0.8% agarose gel for 30mins at 100V to check the PCR product size (Figure 14). The bands were at the expected position. PCR product was purified using PCR clean up kit (Qiagen Company) and SMYD5 gene was obtained as an insert fragment. 2 restriction enzymes and their buffer were added into SMYD5 tube and vector tube respectively, 2 tubes were incubated for 1 hour at 37 ºC to let enzyme cut the sequences at restriction site to reveal the sticky ends (Figure 15).Then a 0.8% agarose "blue gel" contained crystal violet was made to use, which can stain DNA in the gel to make it visible to naked eyes, result in preventing DNA from UV damage of the imager. Both insert and vector were loaded on the gel, run for 30mins at 100V, and needed DNA segments were cut out from the gel. Gel pieces were placed in separate tubes and mass the segments, then genes were extracted from blue gel using gel extraction kit (Qiagen Company). Purified insert and purified vector were mixed with quick ligase and buffer, incubated at room temperature for 7mins then all ingredients were added into 20 µL DH5α cells. Transformation was performed using heat-shock method, transferring SMYD5-pCDF-SUMO plasmid into cells. Then, a 250µL volume of warm SOC medium was added and transformed cells were shaken for 1hour. Next, 100µL volume of transformed cells were plated onto streptomycin resistant LB Agar plates. LB plates were grown overnight at 37˚C. 3 colonies were selected and grew in in 5mL LB media contained 100 µg/mL streptomycin at 37 ºC overnight. Cell culture of each colony was aliquoted into 1mL with 300 µL glycerol in 2ml coring tubes, saved in -80 ºC for future use. The left clones were purified by mini prep kit (Qiagen Company). Plasmid samples were sent to Genewiz Company for sequence.

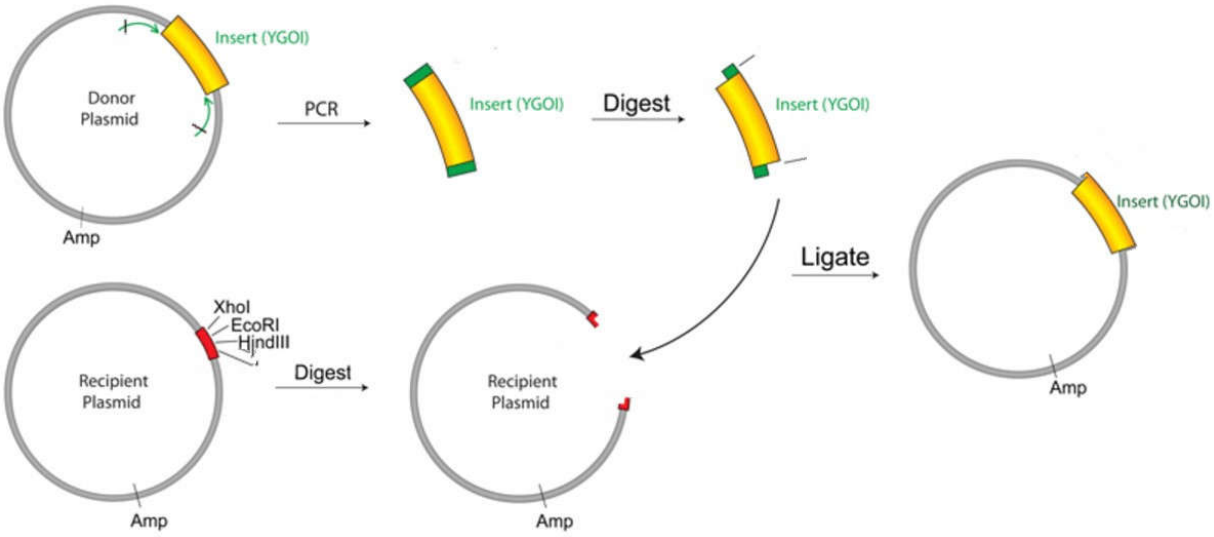


Figure 15: Molecular cloning mechanism

2.1.2 Molecular Cloning Result

SMYD5 gene was successfully inserted into pCDF-SUMO vector. According to the sequencing result (Table 4), this clone contains one silent mutant 132E→E (GAG→GAA), which raises the usage in E.coli. Conclusively, the SMYD5 gene was successfully cloned into the pCDF-SUMO vector and this plasmid can be used for the continuous experiment.

Table 4 Human SMYD5 sequencing result (pCDF-SUMO vector)

```
5'-3'
Query  76   ATGGCGGCCTCCATGTGCGACGTGTTCTCCTTCTGCGTGGGCGTGGCGGGCCGCGCGCGG  135
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1    ATGGCGGCCTCCATGTGCGACGTGTTCTCCTTCTGCGTGGGCGTGGCGGGCCGCGCGCGG  60
```

```
Query  136  GTCTCCGTGGAAGTCCGTTTCGTGAGCAGCGCCAAGGGAAAGGGGCTGTTTGCCACACAG  195
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  61   GTCTCCGTGGAAGTCCGTTTCGTGAGCAGCGCCAAGGGAAAGGGGCTGTTTGCCACACAG  120

Query  196  CTCATCCGGAAGGGGGAGACCATCTTCGTAGAACGGCCCCTGGTGGCTGCACAGTTTCTC  255
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  121  CTCATCCGGAAGGGGGAGACCATCTTCGTAGAACGGCCCCTGGTGGCTGCACAGTTTCTC  180

Query  256  TGGAATGCACTTTATCGCTACCGAGCCTGTGACCACTGCCTTAGGGCACTAGAGAAGGCA  315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  181  TGGAATGCACTTTATCGCTACCGAGCCTGTGACCACTGCCTTAGGGCACTAGAGAAGGCA  240

Query  316  GAGGAAAATGCCCAGAGGCTGACCGGGAAACCAGGCCAGGTTCTGCCTCACCCAGAGCTG  375
            ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  241  GAGGAGAATGCCCAGAGGCTGACCGGGAAACCAGGCCAGGTTCTGCCTCACCCAGAGCTG  300

Query  376  TGCACTGTGCGCAAAGACCTCCACCAGAACTGTCCCCATTGCCAAGTGATGTACTGCAGT  435
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  301  TGCACTGTGCGCAAAGACCTCCACCAGAACTGTCCCCATTGCCAAGTGATGTACTGCAGT  360

Query  436  GCAGAATGTCGGTTGGCAGCCACTGAGCAATACCACCAGGTCCTGTGCCCAGGCCCCTCC  495
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  361  GCAGAATGTCGGTTGGCAGCCACTGAGCAATACCACCAGGTCCTGTGCCCAGGCCCCTCC  420

Query  496  CAGGATGACCCCTTGCATCCTCTCAATAAGCTTCAGGAGGCATGGAGGAGTATTCACTAC  555
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  421  CAGGATGACCCCTTGCATCCTCTCAATAAGCTTCAGGAGGCATGGAGGAGTATTCACTAC  480

Query  556  CCACCTGAGACTGCAAGCATCATGTTGATGGCTAGGATGGTGGCCACAGTGAAGCAGGCG  615
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  481  CCACCTGAGACTGCAAGCATCATGTTGATGGCTAGGATGGTGGCCACAGTGAAGCAGGCG  540

3'-5'
Query  87    ACACATCAGTCATCTCATCCCCCAGCTCTGCATCTTCTGGctctccttcctcctcctcct  146
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1255  ACACATCAGTCATCTCATCCCCCAGCTCTGCATCTTCTGGCTCTCCTTCCTCCTCCTCCT  1196
```

```
Query  147   cctcttcctcttcctcttctGAGGTCACATTGGGTTCATCAGCCTCTGCCAGGCATTTGG  206
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1195  CCTCTTCCTCTTCCTCTTCTGAGGTCACATTGGGTTCATCAGCCTCTGCCAGGCATTTGG  1136

Query  207   GACAGGAACAGACAAATAGATAGTTCTCCCTGAGGATCTTGTGGCGGCTGTGGCGGCTGC  266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1135  GACAGGAACAGACAAATAGATAGTTCTCCCTGAGGATCTTGTGGCGGCTGTGGCGGCTGC  1076

Query  267   GCTCCCGCTGACAGCAGTCCAAGTAGCTGATACAAATTTCCTCTCCTGGCTTAATATCCT  326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1075  GCTCCCGCTGACAGCAGTCCAAGTAGCTGATACAAATTTCCTCTCCTGGCTTAATATCCT  1016

Query  327   CCAGAGCAGTGACATGCAAAAGGAAGTTGTTTTCTGGAAAGGAGGTCTCTGCATTGGGCA  386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1015  CCAGAGCAGTGACATGCAAAAGGAAGTTGTTTTCTGGAAAGGAGGTCTCTGCATTGGGCA  956

Query  387   CACAACTGTGGTTGCAGCAGCTCTGAAGCACAAAGAGGCCAGATCCTTCACAGTTAAGAA  446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  955   CACAACTGTGGTTGCAGCAGCTCTGAAGCACAAAGAGGCCAGATCCTTCACAGTTAAGAA  896

Query  447   ACTCTCCAGTTGCTGCCTCGATGTCCTTGTATAGCTGGTCAATGAAGGCGTCAAGCTGCT  506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  895   ACTCTCCAGTTGCTGCCTCGATGTCCTTGTATAGCTGGTCAATGAAGGCGTCAAGCTGCT  836

Query  507   CACGGTCCTGAGGCTTCAACTCCAGAGTGTCACAGGCATGGACCCACTGGCTTAGGGAGC  566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  835   CACGGTCCTGAGGCTTCAACTCCAGAGTGTCACAGGCATGGACCCACTGGCTTAGGGAGC  776

Query  567   TGGTCCCGATTCCTTGGCCATTGGTCCCAACAAGAGCAAAGAGAGACCGGAATCCATCTG  626
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  775   TGGTCCCGATTCCTTGGCCATTGGTCCCAACAAGAGCAAAGAGAGACCGGAATCCATCTG  716

Query  627   GAGTGAACCACTGGCTGACTGCTTCCTCATAGAGGGCCTCTGTGAAGAGTCTCCGCAGAA  686
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  715   GAGTGAACCACTGGCTGACTGCTTCCTCATAGAGGGCCTCTGTGAAGAGTCTCCGCAGAA  656

Query  687   GTTCCAGTTGGCCCTTGAATTTGTCTCCCAGAAGTTTATGGACAATTTCCTCCTCTTCAT  746
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  655   GTTCCAGTTGGCCCTTGAATTTGTCTCCCAGAAGTTTATGGACAATTTCCTCCTCTTCAT  596
```

```
Query  747   TGGCTGTTTTGTTACAAAACTGGGAAAAGAGTCTGATCCAACGGTCCTTGTCCTTCGCCT  806
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  595   TGGCTGTTTTGTTACAAAACTGGGAAAAGAGTCTGATCCAACGGTCCTTGTCCTTCGCCT  536
```

2.2 Protein Expression

2.2.1 Material and Method

First, SMYD5 plasmid is going to be transferred into host cell BL21 for protein expression. Once the co-construct vector was obtained, 1.0 ng of vector was transformed into 20μL of codon optimized BL21 cells and heat shocked for 45 seconds at 42˚C. Then, a 250μL volume of warm SOC medium was added and transformed cells were shaken for 1hour. Next, 100μL volume of transformed cells were plated onto streptomycin resistant LB Agar plates. LB plates were grown overnight at 37˚C. 1 colony was selected and grew in 5mL LB media contained 100 μg/mL streptomycin at 37 ºC overnight. BL21 cell culture was aliquoted into 1 mL with 300 μL glycerol in 2mL coring tubes, saved in -80 ºC for future use.

Second, a small scale and a large scale protein expression test is performed. 30μL BL21 overnight cell culture was added into 3mL fresh LB media, then growth continued at 37 ºC once the pre-induction OD600 reached 0.4-0.6. A 1.0 mL volume of culture was saved to test for non-induced expression and the remaining culture was slowly cool down to 15 ºC, sat for half hour, and induced with 0.1 mM IPTG. Induced culture was grown for 20-22 hours at 15˚C and collected for SDS-PAGE. Cells were sonicated and centrifuged to

make supernatant and pellet samples to test protein solubility. Samples were analyzed on 15% SDS-PAGE.

Large scale test is an 100 times amplification of small scale test. 5mL overnight BL21 culture was added into 500mL fresh LB media, then growth continued at 37 ºC once the pre-induction OD600 reached 0.4-0.6. A 1.0 mL volume of culture was saved to test for non-induced expression and the remaining culture was slowly cool down to 15 ºC, sat for half hour, and induced with 0.1 mM IPTG. Induced culture were grown for 20-22 hours at15˚C and collected for SDS-PAGE (Figure 16). Cells were French pressed to get cell lysate for purification.

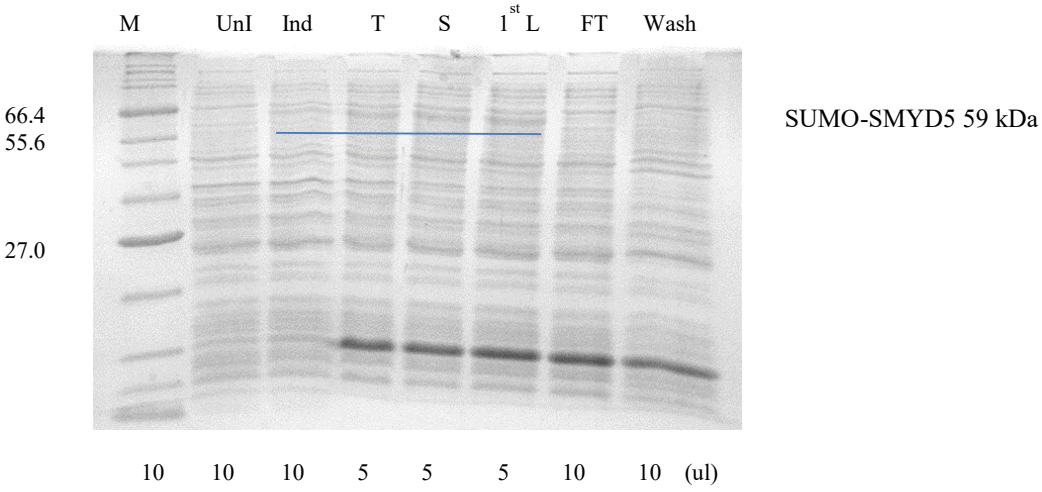


Figure 16 : 15% SDS-page of large scale protein expression test

"UnI"-un-induced cell, "Ind"-induced cell, "T"-total cell lysate, "S"-supernatant of cell lysate

CHAPTER 3: SMYD5 Purification

3.1 Material and Method

3.1.1 Purification Strategy

The SUMO protein has a His tag which can bind to the His-trap nickel column agarose matrix (GE Company). Loading cell lysate on the 1st Ni column and SUMO-SMYD5 bound to the column because of the SUMO His tag (Figure 17). However, imidazole is part of the structure histidine binding to Ni ions and it competes with His for Ni binding sites (Figure 17).

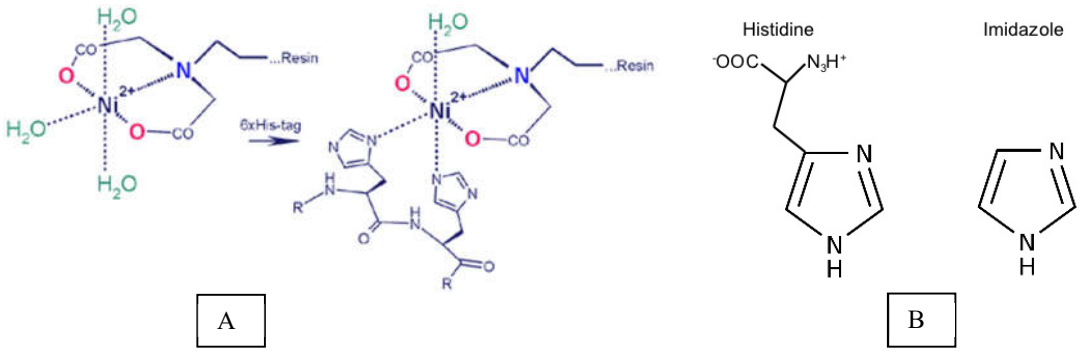


Figure 17: Histidine (A) and Imidazole (B) binds to Ni

Low concentration of imidazole prevents the non-specific and low affinity binding of background protein. After washing with low concentration imidazole buffer, SUMO-SMYD5 protein was eluted out by high concentration of imidazole buffer.

SUMO protease was added to cut protein at 4 ºC overnight. As protein is in elution buffer, next step is using desalting column to replace the elution buffer with binding buffer. When load protein samples, which have been cut, on the 2nd his trap, SMYD5 directly

passed through the Ni column whereas sumo protein bound to the column. Eventually, gel filtration column separates protein based on size (Figure 18). Larger proteins were eluted earlier while smaller proteins were eluted later and provide a better resolution.

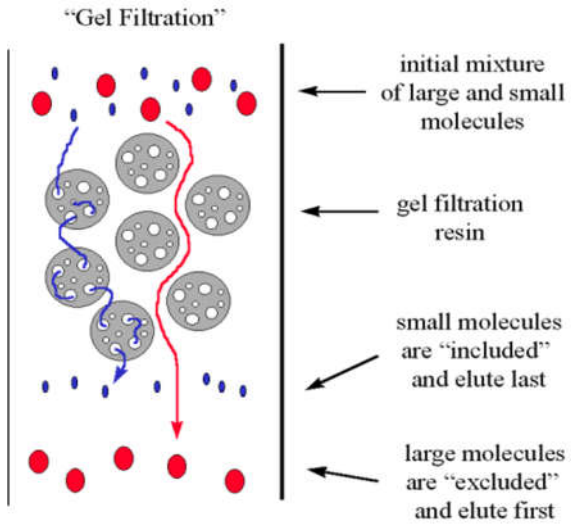


Figure 18 Gel filtration Column Mechanism

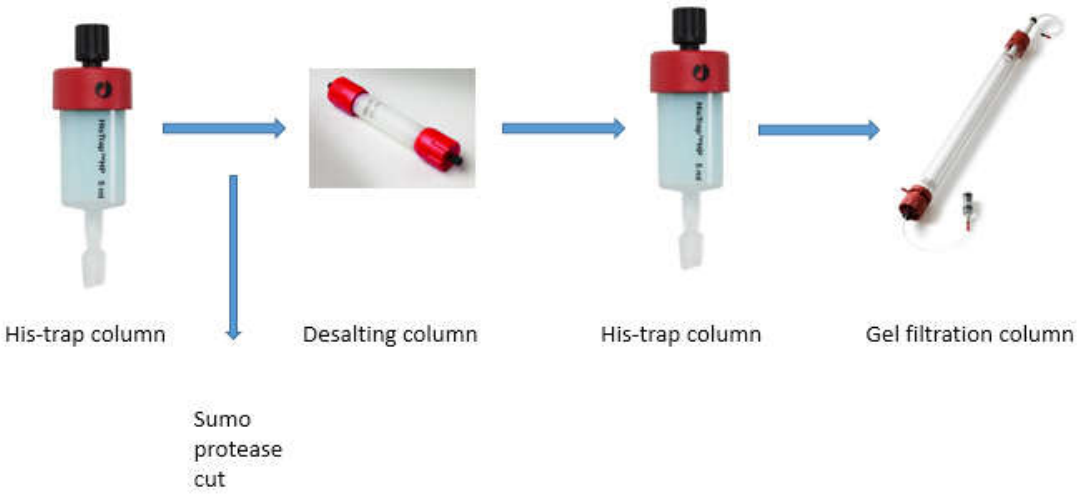


Figure 19: Overview of SUMO-SMYD5 purification strategy

3.1.2 Buffer preparation

Binding buffer - 20 mM Tris, pH7.4, 0.5 M NaCl, 20 mM imidazole, 5% glycerol, 5 mM BME; Elution buffer - 20 mM Tris, pH7.4, 0.5 M NaCl, 500 mM Imidazole, 5% glycerol, 5 mM BME; Gel Filtration Buffer - 20mM Tris pH8.0, 150 mM NaCl, 5% glycerol, 5mM BME. All the buffers were filtered and degassed.

3.2 Purification Result

3.2.1 1st Ni column

In the chromatogram (Figure 20), the green line represents continuous flow of increasing imidazole, (20mM/500mM) and the blue line represents the UV absorbance of eluted protein. The first peak indicated washed out heat-shock protein 70 with weak binding affinity to nickel column, and the second peak clearly showed SUMO-SMYD5 protein was eluted out, which is also shown on the SDS-PAGE gel (Figure 21), Fractions A9-B2 were pooled together. 15 μL SUMO protease was added into collection and incubated at 4 ºC overnight to cut SUMO protein off.

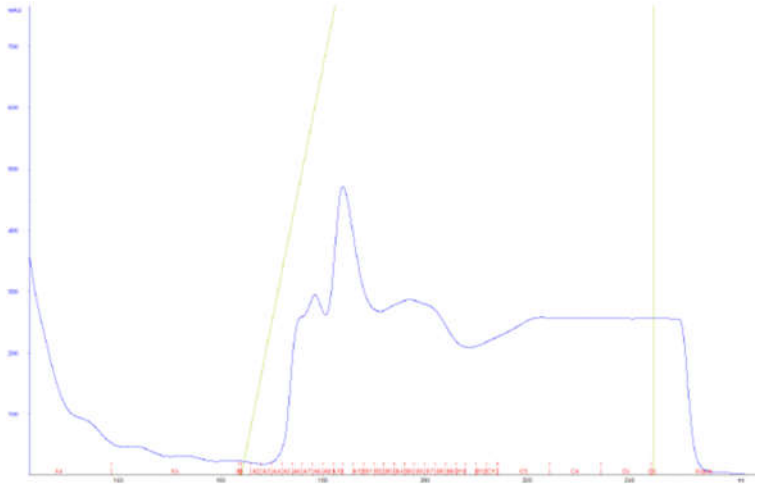


Figure 20: SUMO-SMYD5 Purification 1st Ni Column. 1st Peak-HSP 70, 300 mAU, eluted by 28% elution buffer. 2nd peak-SUMO-SMYD5, 471 mAU, eluted by 40% elution buffer.

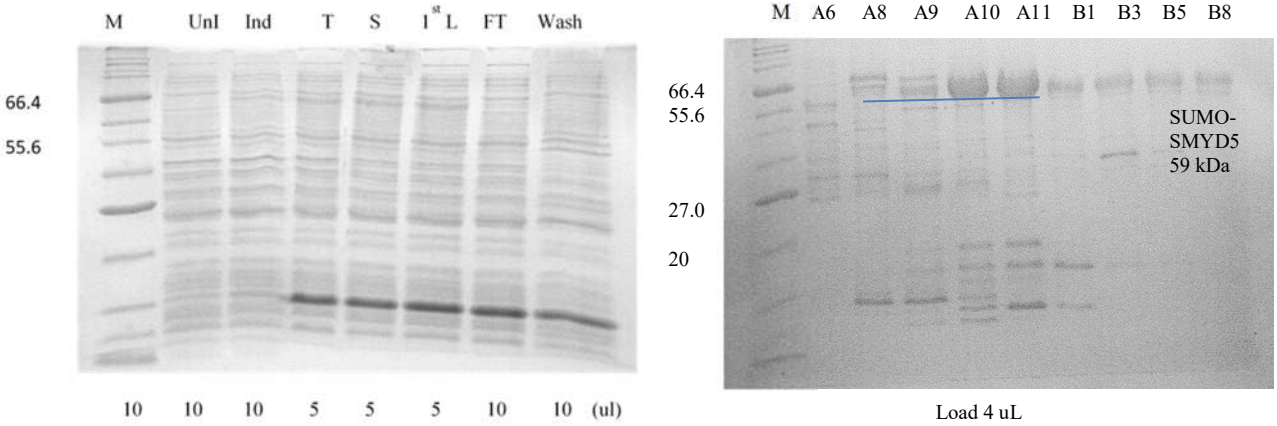


Figure 21: 15%SDS-PAGE of SUMO-SMYD5 Purification 1st Ni Column

"1st L"-1st Ni Column Loading, "FT"- Flow Through

3.2.2 Desalting column

The desalting column is used to separate soluble macromolecules from smaller molecules or replace the buffer system. Proteins flow pass the beads, salts and small molecules flow through the beads so proteins flow faster through the column. In the chromatogram (Figure 22), there was only one peak. According to this UV absorbance, fractions A5-B1 were collected, and analyzed on the 15% SDS gel (Figure 23)

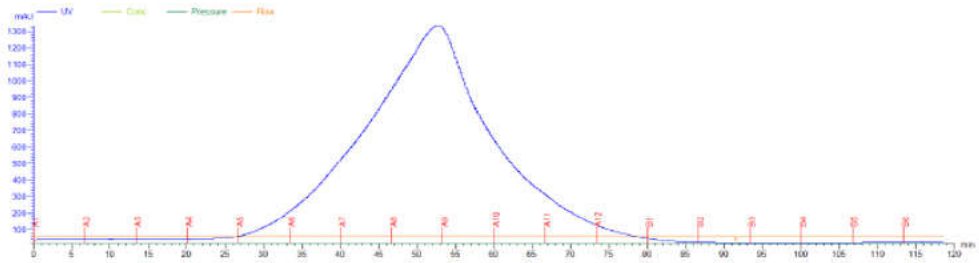


Figure 22: SUMO-SMYD5 Purification Desalting Column.Peak:1200mAU.

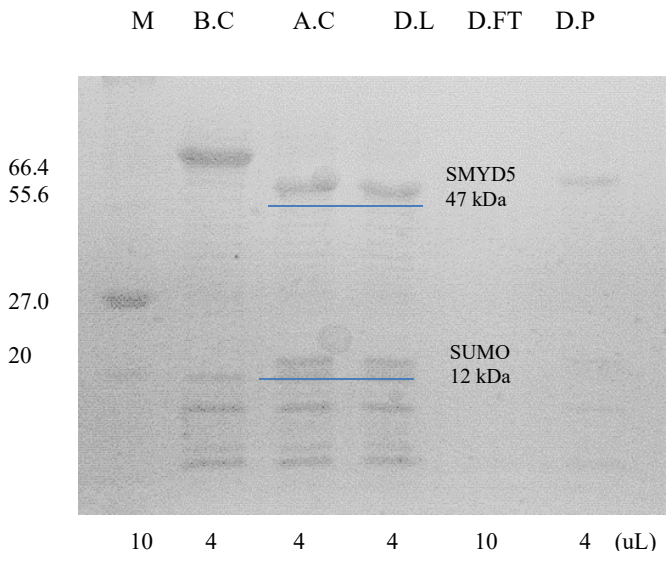


Figure 23: 15%SDS-PAGE of SUMO-SMYD5 Purification Desalting Column
"B.C"-before cut, "A.C"-after cut, "D.L"-desalting load, "D.FT"-desalting flow through, "D.P"-desalting pool

3.2.3 2nd Ni column

After desalting column and incubation with SUMO protease, SMYD5 and SUMO protein were separated in the binding buffer system. When loading the sample on the 2nd Ni column, SMYD5 directly passed column while SUMO protein bound to Ni column because of the His tag. SMYD5 was collected in the flow through fraction and SUMO protein was washed out by elution buffer (Figure 24).

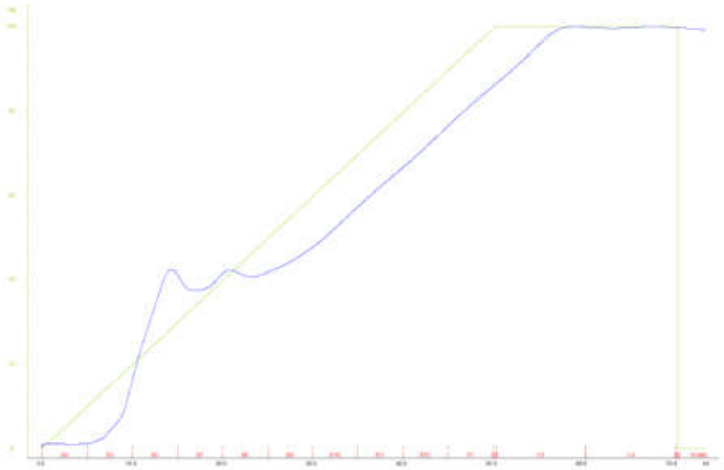


Figure 24: SMYD5 Purification 2nd Ni Column Elution.

1st peak-HSP 70, eluted by 28% elution buffer. 2nd peak-SUMO protein, eluted by 41% elution buffer.

3.2.4 Gel filtration column

SMYD5 was concentrated to 1mL, and loaded on gel filtration column. According to the gel filtration standard (Figure 25), the elution volume 79ml was consistent with SMYD5 molecular weight 47kDa (Figure 26). Also the bands on SDS-PAGE were on the expected position (Figure 27). Fractions B5-B10 were collected. SMYD5 protein was filtered and concentrated. 1.35mg purified SMYD5 protein was obtained.

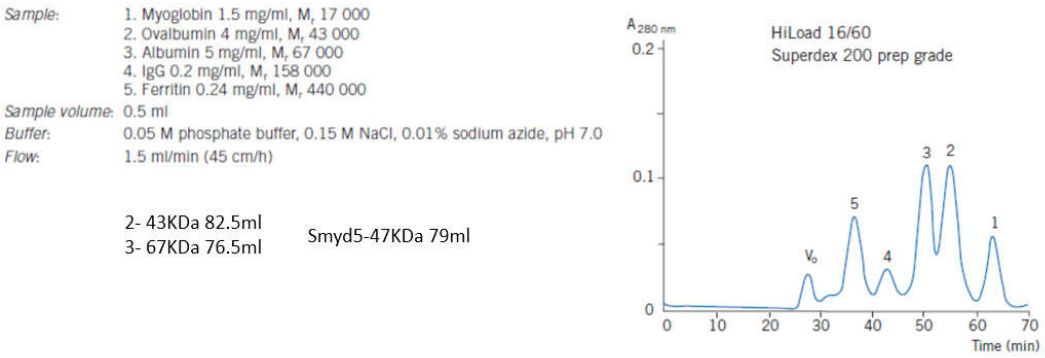


Figure 25: Gel filtration standard

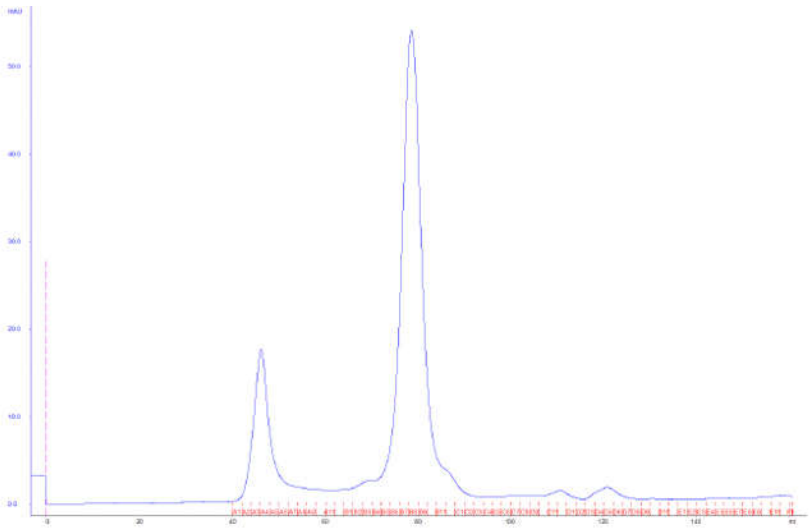


Figure 26: SMYD5 Purification Gel Filtration Column.

1st peak-void peak. 2nd peak-SMYD5, 54mAU, eluted at 79ml.

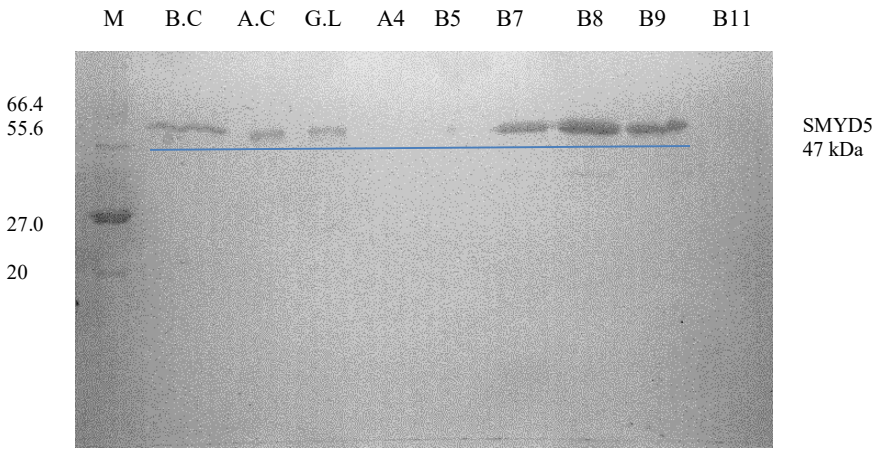


Figure 27 : 15% SDS-PAGE of SMYD5 Purification Gel Filtration Column."B.C"-before concentration, "A.C"-after concentration, "G.L"-gel filtration load.

CHAPTER 4: GST-SMYD5 Molecular Cloning, Expression and Purification

In order to get more variable references in SMYD5 activity assay, GST-SMYD5 purification was performed.

4.1 Molecular Cloning

Clone human SMYD5 gene into pGEX-6P-2 vector (Figure 28).

4.1.1 Material and Method

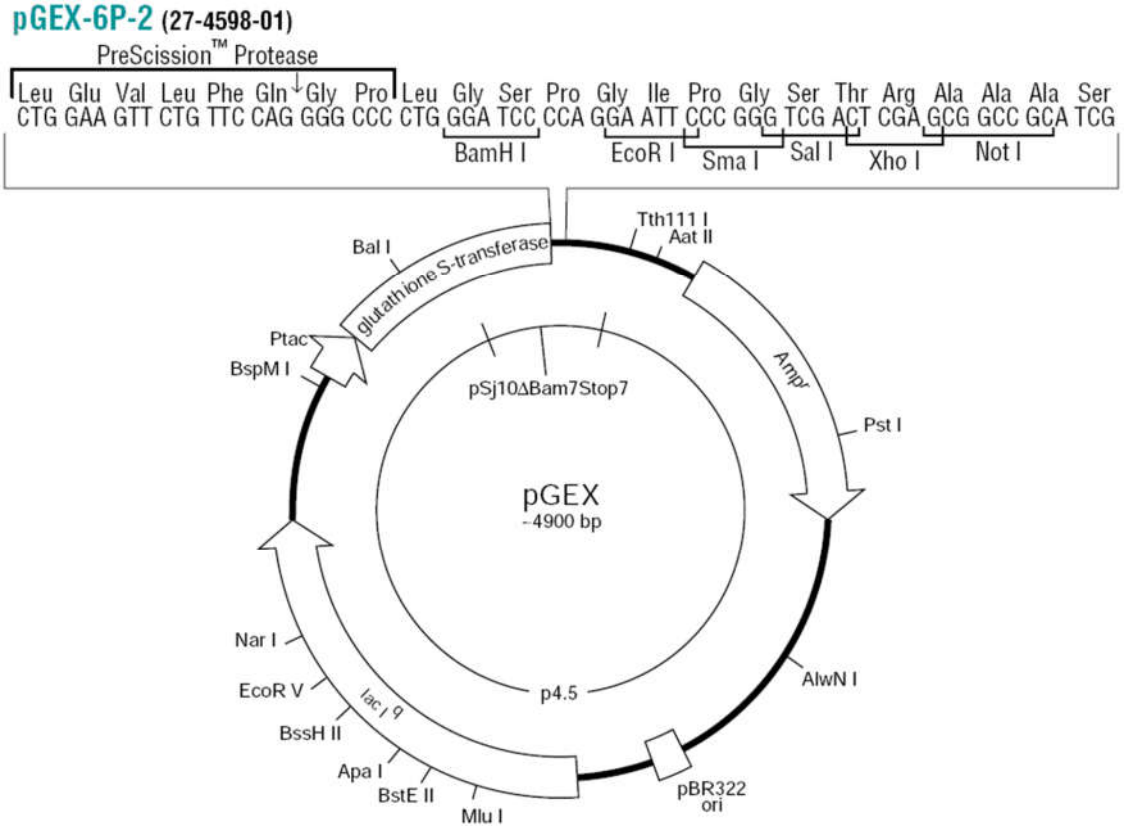


Figure 28: pGEX-6P-2 vector map

The whole procedure was basically as same as the approach of cloning SUMO-SMYD5.

Size of 5' primer is 33 bp, TACTTAGGATCCATGGCGGCCTCCATGTGCGAC (Tm=68.1ºC), included a BamHI restriction site. Size of 3' primer is 33 bp, CAC GCCTCGAGTTACACATCAGTCATCTCATC (Tm=63.6ºC), included an XhoI restriction site. Ordered primer from Invitrogen Company.

4.1.2 Molecular Cloning Result

SMYD5 gene was successfully inserted into pGEX vector, and no mutant occurred.

4.2 Expression

Since the GST-SMYD5 is 76kDa, bigger than 60kDa, lower percentage SDS-PAGE was selected to promote the 66-212 kDa range protein separation (Figure 29).

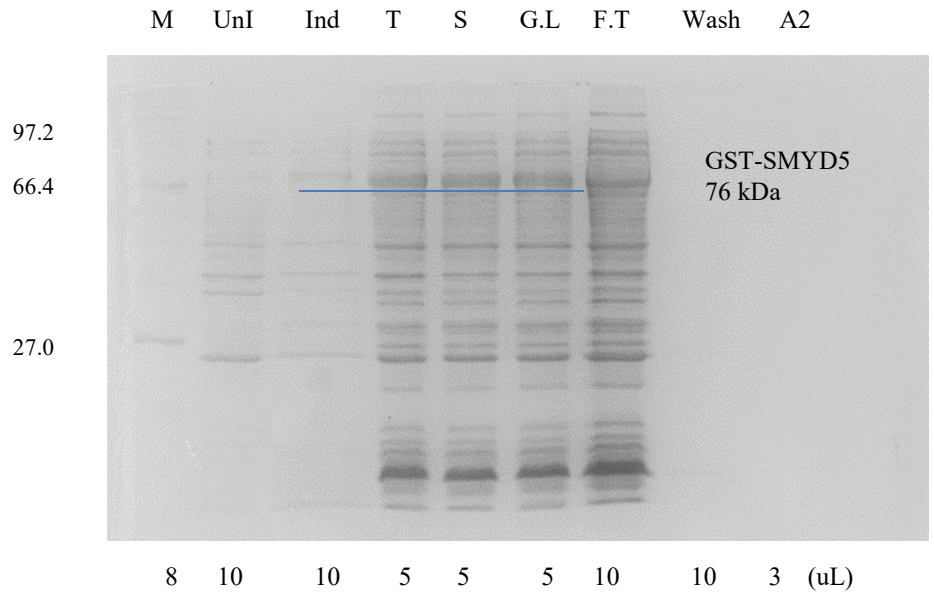


Figure 29: 12% SDS-PAGE of GST-SMYD5 expression test. *"UnI"-un-induced cell,"Ind"-induced cells, "T"-total cell lysate, "S"-supernatant of cell lysate*

4.3 Purification

4.3.1 Purification strategy

The purification strategy is based on the high affinity of GST for glutathione (Figure 30). To let GST tagged protein bind to matrix, we choose the condition that favor interaction of the GST with the glutathione column which is around PH7.5. Therefore, binding buffer should be between PH 7.2-7.5. When sample was applied to the affinity matrix, GST fusion proteins bound to the glutathione ligand, and impurities were removed by washing with binding buffer. When column was washed with a PH8.5 elution buffer contained glutathione, tagged proteins were eluted from the chromatography resin under mild, non-denaturing conditions that preserve both protein structure and function.

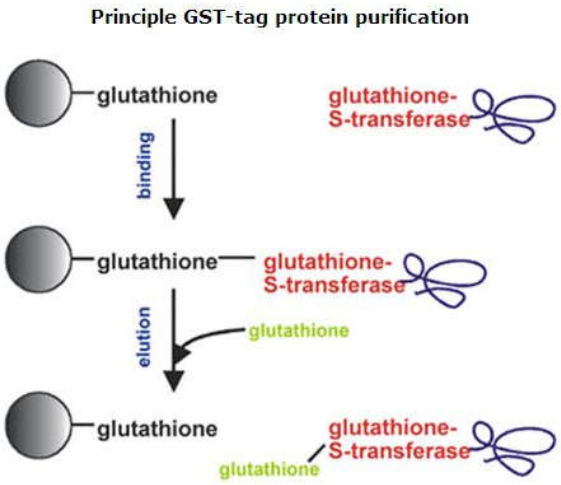


Figure 30: GST-SMYD5 Purification strategy

4.3.2 Buffer preparation

Binding buffer-50mM Tris pH 7.5, 200mM NaCl, 1mM EDTA, 1mM BME, 5% glycerol. Elution buffer-50mM Tris pH 8.5, 500mM NaCl, 1mM EDTA, 1mM BME, 5% glycerol, 10mM reduced glutathione. Gel Filtration Buffer-20mM Tris pH 8.0, 150 mM NaCl, 5% glycerol, 5mM BME. All the buffer was filtered and degassed.

4.3.3 Purification result

4.3.3.1 GST-trap column

GST-SMYD5 was eluted out quickly (Figure 31). GST-SMYD5 protein didn't show either in the flow through or wash according to the SDS-PAGE analysis.

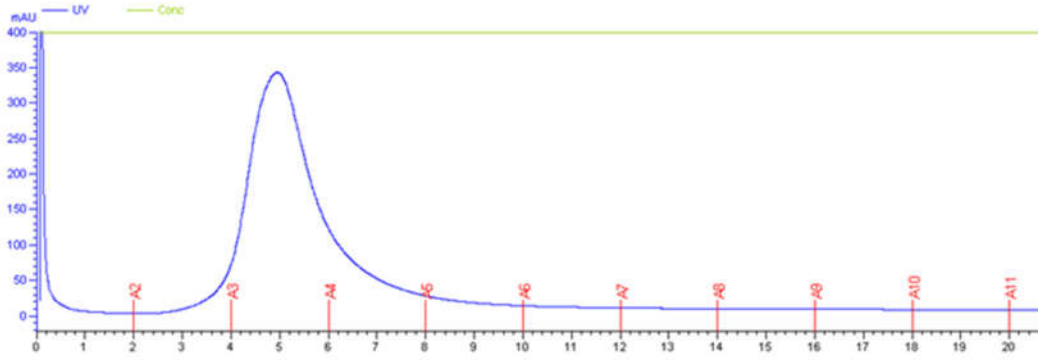


Figure 31: GST-SMYD5 purification GST-trap column. Peak-GST-SMYD5:350mAU.

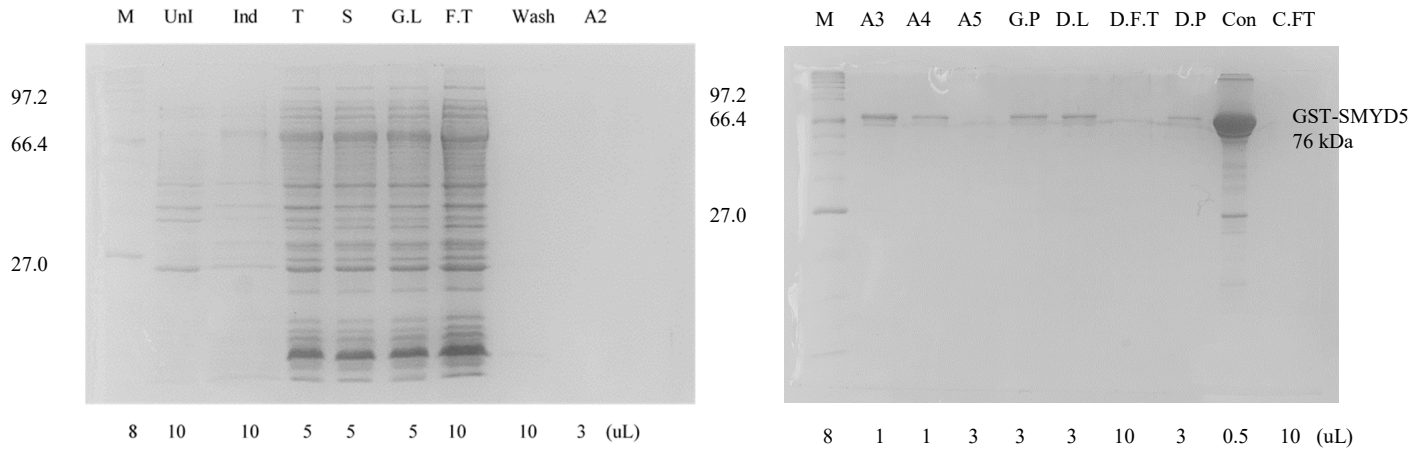


Figure 32: 12%/15% SDS-PAGE OF GST-SMYD5 purification.

"G.L"-GST-trap load, "FT"-flow through, "D.L"-desalting load, "D.F.T"-desalting flow through", "Con"-concentrated, "Con"-concentration flow through

4.3.3.2 Desalting column

Desalting column replaced GST elution buffer by gel filtration buffer (Figure 33).

1.6mg purified GST-SMYD5 protein was obtained.

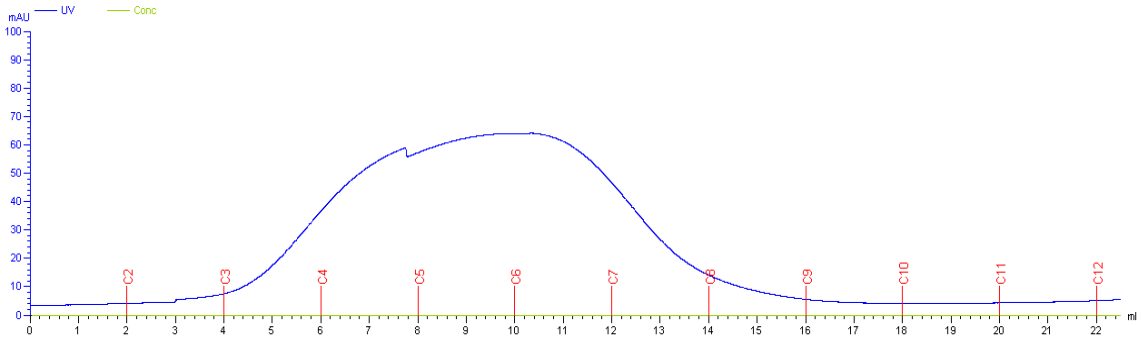


Figure 33: GST-SMYD5 purification desalting column

CHAPTER 5: SMYD5 Protein Analysis

5.1 Crystal Structure

Sequence alignment of SMYD5 with the other members of the SMYD family revealed several structural differences not found in the sequences of the other SMYD proteins. These observations indicate that SMYD5 recognizes its substrate using divergent structural determinants comparing to SMYD1-3. Obtaining a crystal structure of SMYD5 would reveal how these differences translate into specificity.

5.1.1 Material and Method

There are 2 common methods used in crystal screening, sitting drop and hanging drop (Figure 34), sharing the same principle which is the sample drop contains a lower reagent concentration than the reservoir. Water vapor leaves the drop and finally ends up in the reservoir to achieve equilibrium. Because water leaves the drop, the sample experiences an increase in relative supersaturation. Both the sample concentration and reagent concentration increase because water leaves the drop for the reservoir.

Sitting drop vapor diffusion: Time efficient. • Often easier when using detergents, organics and hydrophobic reagents. • Drops can be positioned in a stable sitting spot.

Hanging drop vapor diffusion: Easy access to crystals. • Could perform multiple drops with a single reservoir.

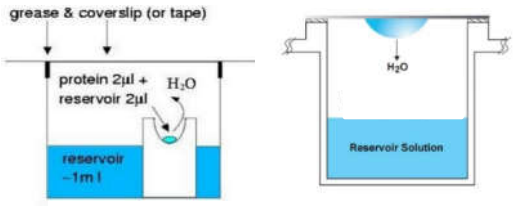
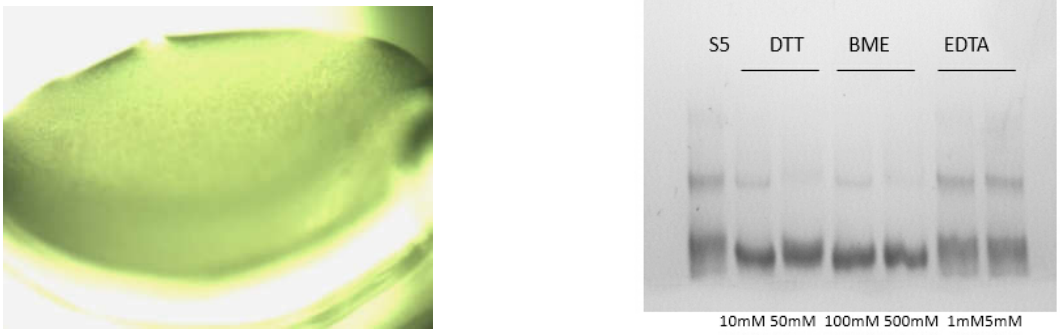


Figure 34: Sitting drop & Hanging drop

5.1.2 Result

During the approach to crystalize SMYD5, several promising conditions (Figure 35) were achieved in hundred contentions. To check the protein quality, SMYD5 protein samples were prepared to run on a 6% native gel. After staining by Coomassie blue, an additional upper band occurred (Figure 35). We assumed that the band might be a protein dimer formed by non-specific disulfide bonds. Therefore, some reducing agent were added to SMYD5 samples as reducing agent could break the bonds. An observation that the bands were removed by high concentration reducing agent confirmed the speculation that the band might be protein dimer formed by non-specific disulfide bonds.

More crystal screens were set up with conditions contained 50mM DTT. After introducing DTT, some previous precipitate conditions turned into clean or light precipitate. These phenomena suggested that high concentration DTT may increase protein solubility. Although crystallization of SMYD5 has not been successful, increasing DTT concentration may promote protein solubility and facilitate chance of crystallization.



A B

Figure 35: Crystal screen analysis.(A)promising crystal screen condition: face separation.(B)6% native gel of SMYD5 protein

5.2 SMYD5 Methyltransferase Activity Assay

The reason why SMYD5 is unique, is not only about lacking of C-terminus TPR domain, but also the additional poly-glutamate stretch tail (Figure 36), which brings more negative charge to SMYD5.



Figure 36: poly-E region of SMYD5 sequence

On other hand, SMYD5 has a potential to be a mitochondria protein, as it contains a mitochondria targeting sequence (Table 5) at the N-terminus, which consists of an alternating pattern of hydrophobic and positively charged amino acids to form an amphipathic helix, which direct a newly synthesized protein to the mitochondria.

Table 5: mitochondria targeting sequence

10	20	30	40	50
MAASMCDVFS FCVGVAGRAR VSVEVRFVSS AKGKGLFATQ LIRKGETIFV				

Moreover, as we all know, mitochondria functions as a calcium sensor and this poly-E region shows a potential of binding to Calcium as well. According to our predict structure, this region is close to SMYD5 C-terminal active site. Overall, we assumed that this poly-E region might effect S5 activity if it binds to calcium ion.

An antibody based activity assay (western blot) was performed to test SMYD5 enzymatic activity.

5.2.1 Material and Method

SMYD5 activity assay were carried out in methyltransferase reaction buffer, with histone H4 and SAM, incubated at 30 °C overnight. Reactions were stopped with the addition of SDS-PAGE loading buffer and boiled for 5 minutes, then were separated by SDS-PAGE. Proteins were transferred to nitrocellulose membrane using wet tank transfer method (Bio-rad), at 4°C 100V for1hour.

Membrane was incubated in 2%BSA or milk for 1hour, washed with 0.0.5% TBST buffer 2 *5 mins then gently shaken in primary antibody anti-tri-methylated lysine (1:1000, Abcam) at 4°C overnight. Then membrane was washed with 0.0.5% TBST for 3 *5 mins. Second incubation was shaking membrane in secondary antibody Goat-anti rabbit for 1-2hours. After this, membrane was washed again with 0.0.5% TBST buffer 3 *5 mins.

ECL reagent was added equably on the whole membrane to expose membrane in the dark room of UV imager.

Reaction buffer conditions were tested using 2mg/4mg/10mg SMYD5, 5mg GST-SMYD5 with 1mg/1.5mg/2mg histone H4, 50mM/100mM/200mM SAM in 50mM buffer (Tris pH 7.5, 8.0 and 9.0, Bicine pH 9.0) with 0.5mM/50mM/DTT, 0.1Mm/0.5mM/1mM/5Mm/10mM EDTA, 5mM/25mM $CaCl_2$, 5mM/25mM $MgCl_2$, 25mM/50mM/100mM NaCl, 5%/25% Glycerol.

5.2.2 Activity assay result

So far, no obvious signal has been observed from either SMYD5 or GST-SMYD5 in Tris buffer system. However, a weak signal showed in western blot result of Bicine buffer reaction (Figure 37).

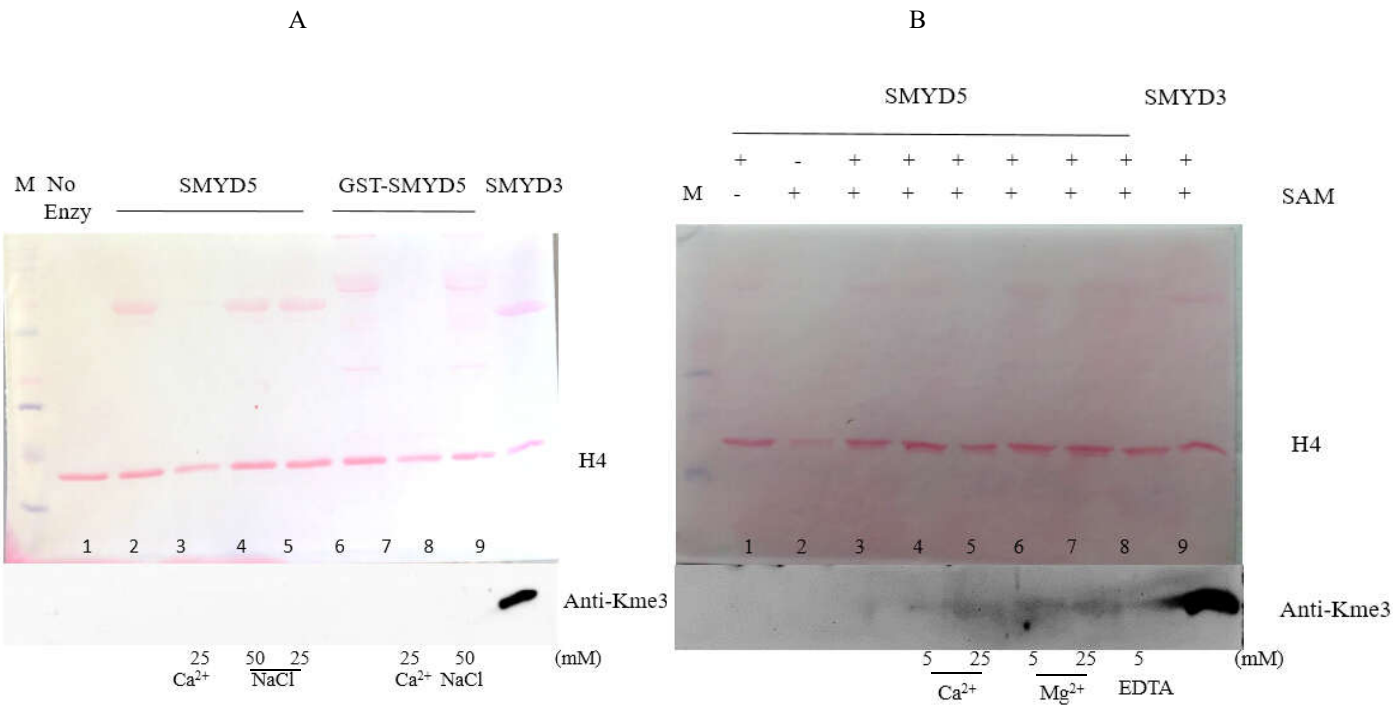


Figure 37: Ponceau S staining and western blot result (exposure time 5 mins)of SMYD5 activity test.(A)PH9.0 Tris buffer (B)PH 9.0 Bicine buffer

CHAPTER 6: Discussion

After successful expression and purification of SMYD5, to characterize SMYD5 biochemical features, I expected to see what could enhance or repress SMYD5 enzymatic activity. However, SMYD5 didn't show a strong enzymatic activity on either histone H4 or H4K20 peptide in comparison of SMYD3. I have raised several potentials.

First, the SMYD5 protein I purified lost activity during the purification. As we all known, the matrix of His-trap column is Ni-agarose. Ni ions could be toxic to some proteins, making them inactive. SMYD5 also could bind to some molecules or inhibitors during the purification process. Therefore, we produced another construct, GST-SMYD5, to avoid Ni-column during purification. Unfortunately, it didn't show a significant difference from previous purified SMYD5.

Second, SMYD5 need some co-factor to activate. SAM is the donor of methyl group in methylation reaction. The interact motif between SAM and SMYD5 still remain unknown. It is feasible that SAM is required a co-factor to achieve the binding site of SMYD5.

Third, SMYD5 is a context-substrate dependent enzyme. For example, respiratory nitrate reductase complex (NarGHI) could exhibit two catalytically distinct forms of the enzyme because of the influence of substrate, pH and inhibitors (43). Histone H4 was reported as the first substrate of SMYD5 recently, and more related mechanisms need to further explore.

However, according to the result I observed, the enzymatic activity of SMYD5 was improved by Ca and Mg ions in Bicine buffer and the higher concentration of metal ions provided a better result. This suggested that our hypothesis could be feasible and enabled

a deeper investigation on SMYD5 associated with Ca and mitochondria. Additionally, the predicted structure of SMYD5 (Figure 38) offered some clues. Although there is no predicted structure of SMYD5 poly-E tail so far, the last residue before it leans towards to the SET domain and remain close to the active site, raising a potential of altering the enzymatic activity of SMYD5.

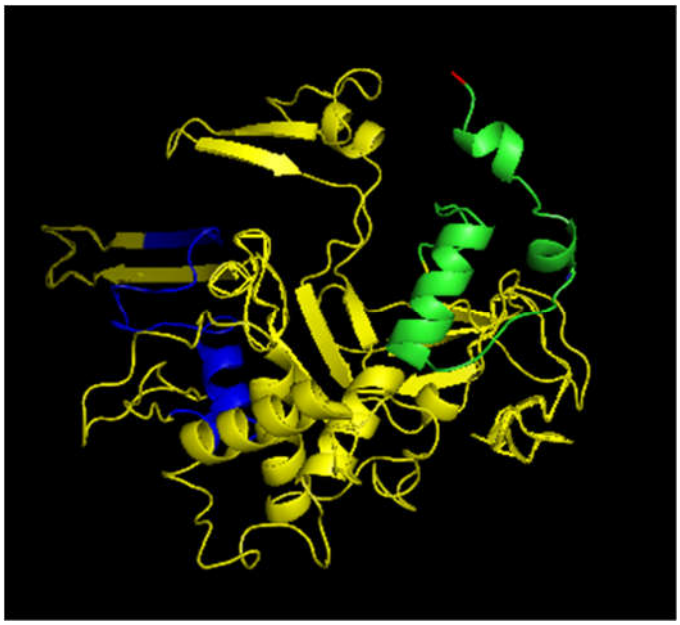


Figure 38: Predicted structure of SMYD5 (19-391). Yellow-SET domain, Blue-MYND domain, Green-C-terminal domain, Red-the last residue before poly-E tail.

CHAPTER 7: Future Direction

- Incubate peptides with SMYD5 to applied in crystal screens or develop fusion linker protein construct to facilitate crystallization of SMYD5
- Perform more crystal screens containing DTT and SAM or other cofactors
- Truncate poly-E tail of SMYD5 to test methyltransferase activity
- Explore more substrates of SMYD5 protein
- Investigate subcellular localization of SMYD5 to reveal its association with mitochondria

CHAPTER 8: References

1. Zhang, X., and Bruice, T. C. (2007) A quantum mechanics/molecular mechanics study of the catalytic mechanism and product specificity of viral histone lysine methyltransferase. *Biochemistry* **46**, 9743-9751
2. Chuikov, S., Kurash, J. K., Wilson, J. R., Xiao, B., Justin, N., Ivanov, G. S., McKinney, K., Tempst, P., Prives, C., and Gamblin, S. J. (2004) Regulation of p53 activity through lysine methylation. *Nature* **432**, 353-360
3. Pradhan, S., Chin, H. G., Estève, P.-O., and Jacobsen, S. E. (2009) SET7/9 mediated methylation of non-histone proteins in mammalian cells. *Epigenetics* **4**, 383-387
4. Biggar, K. K., and Li, S. S. C. (2015) Non-histone protein methylation as a regulator of cellular signalling and function. *Nat Rev Mol Cell Biol* **16**, 5-17
5. Allis, C., Jenuwein, T., Reinberg, D., and Caparros, M. (2007) Epigenetics Cold Spring Harbor Laboratory Press. Cold Spring Harbor, NY
6. Araki, Y., and Mimura, T. (2017) The Histone Modification Code in the Pathogenesis of Autoimmune Diseases. *Mediators of Inflammation* **2017**
7. Couture, J.-F., and Trievel, R. C. (2006) Histone-modifying enzymes: encrypting an enigmatic epigenetic code. *Current opinion in structural biology* **16**, 753-760
8. Allfrey, V., Faulkner, R., and Mirsky, A. (1964) Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proceedings of the National Academy of Sciences* **51**, 786-794
9. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., and Kouzarides, T. (2001) Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124
10. Qian, C., and Zhou, M.-M. (2006) SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cellular and molecular life sciences* **63**, 2755-2763
11. Allis, C. D., Berger, S. L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., and Shiekhattar, R. (2007) New nomenclature for chromatin-modifying enzymes. *Cell* **131**, 633-636
12. Zhang, X., Wen, H., and Shi, X. (2012) Lysine methylation: beyond histones. *Acta Biochim Biophys Sin (Shanghai)* **44**, 14-27
13. Spellmon, N., Holcomb, J., Trescott, L., Sirinupong, N., and Yang, Z. (2015) Structure and function of SET and MYND domain-containing proteins. *International journal of molecular sciences* **16**, 1406-1428
14. Laurell, C.-B. (1951) Analytical Review: What is the Function of Transferrin in Plasma? *Blood* **6**, 183-187
15. Worwood, M. (1990) Ferritin. *Blood Reviews* **4**, 259-269
16. Rouault, T. A. (2006) The role of iron regulatory proteins in mammalian iron homeostasis and disease. *Nature chemical biology* **2**, 406-414
17. Capozzi, F., Ciurli, S., and Luchinat, C. (1998) Coordination sphere versus protein environment as determinants of electronic and functional properties of

iron-sulfur proteins. in *Metal Sites in Proteins and Models Redox Centres*, Springer. pp 127-160

18. Johnson, D. C., Dean, D. R., Smith, A. D., and Johnson, M. K. (2005) Structure, function, and formation of biological iron-sulfur clusters. *Annu. Rev. Biochem.* **74**, 247-281
19. Mongeon, V. (2014) Elucidating the biochemical and structural features required for SMYD5 mediated methylation of histone H4 and other potential substrates *thesis*
20. Du, S. J., Tan, X., and Zhang, J. (2014) SMYD proteins: key regulators in skeletal and cardiac muscle development and function. *The Anatomical Record* **297**, 1650-1662
21. Calpena, E., Palau, F., Espinós, C., and Galindo, M. I. (2015) Evolutionary History of the Smyd Gene Family in Metazoans: A Framework to Identify the Orthologs of Human Smyd Genes in Drosophila and Other Animal Species. *PloS one* **10**, e0134106
22. Gottlieb, P. D., Pierce, S. A., Sims, R. J., Yamagishi, H., Weihe, E. K., Harriss, J. V., Maika, S. D., Kuziel, W. A., King, H. L., and Olson, E. N. (2002) Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nature genetics* **31**, 25-32
23. Phan, D., Rasmussen, T. L., Nakagawa, O., McAnally, J., Gottlieb, P. D., Tucker, P. W., Richardson, J. A., Bassel-Duby, R., and Olson, E. N. (2005) BOP, a regulator of right ventricular heart development, is a direct transcriptional target of MEF2C in the developing heart. *Development* **132**, 2669-2678
24. Li, D., Niu, Z., Yu, W., Qian, Y., Wang, Q., Li, Q., Yi, Z., Luo, J., Wu, X., and Wang, Y. (2009) SMYD1, the myogenic activator, is a direct target of serum response factor and myogenin. *Nucleic acids research*, gkp773
25. Tan, X., Rotllant, J., Li, H., DeDeyne, P., and Du, S. J. (2006) SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2713-2718
26. Rice, J. C., and Allis, C. D. (2001) Histone methylation versus histone acetylation: new insights into epigenetic regulation. *Current opinion in cell biology* **13**, 263-273
27. Diehl, F., Brown, M. A., Van Amerongen, M. J., Novoyatleva, T., Wietelmann, A., Harriss, J., Ferrazzi, F., Böttger, T., Harvey, R. P., and Tucker, P. W. (2010) Cardiac deletion of Smyd2 is dispensable for mouse heart development. *PloS one* **5**, e9748
28. Kawamura, S., Yoshigai, E., Kuhara, S., and Tashiro, K. (2008) smyd1 and smyd2 are expressed in muscle tissue in Xenopus laevis. *Cytotechnology* **57**, 161-168
29. Donlin, L. T., Andresen, C., Just, S., Rudensky, E., Pappas, C. T., Kruger, M., Jacobs, E. Y., Unger, A., Zieseniss, A., and Dobenecker, M.-W. (2012) Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization. *Genes & development* **26**, 114-119

30. Fujii, T., Tsunesumi, S.-i., Yamaguchi, K., Watanabe, S., and Furukawa, Y. (2011) Smyd3 is required for the development of cardiac and skeletal muscle in zebrafish. *PLoS One* **6**, e23491
31. Thompson, E. C., and Travers, A. A. (2008) A Drosophila Smyd4 homologue is a muscle-specific transcriptional modulator involved in development. *PLoS One* **3**, e3008
32. Fujii, T., Tsunesumi, S.-i., Sagara, H., Munakata, M., Hisaki, Y., Sekiya, T., Furukawa, Y., Sakamoto, K., and Watanabe, S. (2016) Smyd5 plays pivotal roles in both primitive and definitive hematopoiesis during zebrafish embryogenesis. *Scientific reports* **6**
33. Hamamoto, R., Saloura, V., and Nakamura, Y. (2015) Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nature Reviews Cancer* **15**, 110-124
34. Stender, J. D., Pascual, G., Liu, W., Kaikkonen, M. U., Do, K., Spann, N. J., Boutros, M., Perrimon, N., Rosenfeld, M. G., and Glass, C. K. (2012) Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Molecular cell* **48**, 28-38
35. Kidder, B. L., Hu, G., Cui, K., and Zhao, K. (2017) SMYD5 regulates H4K20me3-marked heterochromatin to safeguard ES cell self-renewal and prevent spurious differentiation. *Epigenetics & Chromatin* **10**, 8
36. Azzalin, C. M., Reichenbach, P., Khoriauli, L., Giulotto, E., and Lingner, J. (2007) Telomeric repeat–containing RNA and RNA surveillance factors at mammalian chromosome ends. *Science* **318**, 798-801
37. Avner, P., and Heard, E. (2001) X-chromosome inactivation: counting, choice and initiation. *Nature Reviews Genetics* **2**, 59-67
38. Matsui, T., Leung, D., Miyashita, H., Maksakova, I. A., Miyachi, H., Kimura, H., Tachibana, M., Lorincz, M. C., and Shinkai, Y. (2010) Proviral silencing in embryonic stem cells requires the histone methyltransferase ESET. *Nature* **464**, 927-931
39. Lachner, M., Sengupta, R., Schotta, G., and Jenuwein, T. (2004) Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. in *Cold Spring Harbor symposia on quantitative biology*, Cold Spring Harbor Laboratory Press
40. Jing, L., and Zon, L. I. (2011) Zebrafish as a model for normal and malignant hematopoiesis. *Disease models & mechanisms* **4**, 433-438
41. de Matos Neto, J. N., Araujo, M. d. A. F., do Amaral Rabello, D., Motoyama, A. B., de Oliveira, D. M., and Silva, F. P. (2015) Expression analysis of SMYD5 in breast cancer and normal tissues. AACR
42. Gao, H., Chakraborty, G., Lee-Lim, A. P., Mavrakis, K. J., Wendel, H.-G., and Giancotti, F. G. (2014) Forward genetic screens in mice uncover mediators and suppressors of metastatic reactivation. *Proceedings of the National Academy of Sciences* **111**, 16532-16537
43. Marangon, J., de Sousa, P. M. P., Moura, I., Brondino, C. D., Moura, J. J., and González, P. J. (2012) Substrate-dependent modulation of the enzymatic catalytic activity: reduction of nitrate, chlorate and perchlorate by respiratory nitrate

reductase from Marinobacter hydrocarbonoclasticus 617. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **1817**, 1072-1082

ABSTRACT

Expression, Purification and Characterization of Lysine Methyltransferase SMYD5

By

Wen Xue

May, 2017

Advisor: Dr. Zhe Yang

Major: Biochemistry & Molecular Biology

Degree: Master of Science

Methylation of histones and non-histone proteins play vital roles in numerous cellular processes including gene expression regulation and DNA damage response. The identifications of methyltransferase SMYD protein family are not well characterized. SMYD5 is a unique but critical member of SMYD family involved in immune response, stem cell renew, hematopoiesis regulation and cancer metastasis. Understanding its function and structure is monumental to human disease. With the achievement of SMYD5 expression and purification, the association between SMYD5 and its substrate histone H4 has been investigated. While possessing a poly-E tail instead of the TPR domain included in other members, SMYD5 is assumed that associated with Calcium and could potentially be a mitochondrial protein.

AUTOBIOGRAPHICAL STATEMENT

EDUCATION

M.S. Biochemistry and Molecular Biology (2015-2017) Department of Biochemistry and Molecular Biology, School of Medicine, Wayne State University

B.S. Bioengineering (2008-2012) Department of Chemical Engineering, Northwest University, Xi'an, China