STRUCTURE-FUNCTION STUDIES ON NOVEL SIGNALING REGULATORS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Yeyun Zhou

May 2012

STRUCTURE-FUNCTION STUDIES ON NOVEL SIGNALING REGULATORS

Yeyun Zhou, Ph. D.

Cornell University 2012

This dissertation studies the structures and functions of novel signaling regulators, with Chapter 2-4 focusing on sirtuins and Chapter 5 on the DOCK180 family of Rho GEFs. Sirtuins are NAD-dependent deacetylases that regulate important biological processes. Mammals have seven sirtuins, SIRT1-7. SIRT4-7 have undetectable or weak deacetylase activity. In Chapter 2 we identified SIRT5 as an efficient protein lysine desuccinylase and demalonylase. The preference for succinyl and malonyl groups is accomplished by their interactions with residue Tyr102 and Arg105 of SIRT5. Lysine malonylation and succinylation were identified in mammalian proteins. Additionally, SIRT5 can reverse succinylation *in vivo*. In Chapter 3 I delineated the desuccinylation reaction of SIRT5 in crystals, including the complex structure of SIRT5 with a bicyclic intermediate. The SIRT5 complex structures will provide insights to the design of SIRT5-specific inhibitors to investigate its biological functions. In Chapter 4, we investigated a sirtuin homologue from the malaria parasite *Plasmodium falciparum*, PfSir2A, which regulates the expression of surface antigens to evade the detection by host immune surveillance. We present enzymology and structural evidence supporting that PfSir2A preferentially hydrolyzes medium and long chain fatty acyl groups from lysine residues. This would facilitate the

development of PfSir2A inhibitors as potential drugs in malaria treatment.

As the Guanine nucleotide Exchange Factors (GEFs) of Rho GTPases, the DOCK180 family proteins are key regulators of cell motility, phagocytosis, and adhesion. Mammals have 11 members, DOCK1-11, which are classified into four subfamilies, A through D. The DOCK-C subfamily that includes DOCK6-8 has been proposed to activate both Cdc42 and Rac1. In Chapter 5 I show that DOCK7 promotes very weak activation of non-prenylated Cdc42 or Rac1 in solution, but robust activation of prenylated Cdc42 and Rac1 on model liposomes, demonstrating that the prenylation and membrane localization of GTPases are essential for the activation by DOCK7. Additionally, DOCK7 harbors residues that impart GTPase specificity and these can be mutated to shift a given DOCK7 activity profile. Finally, the DOCK7 possesses a distal site that binds preferentially to the active forms of Cdc42 and Rac1 and thereby forms a possible positive feedback loop in activating GTPases.

BIOGRAPHICAL SKETCH

Yeyun Zhou was born as the fourth child of Zhongmou Zhou and Jinzhu Zeng in the countryside of Xinhua County of Hunan Province, P. R. China, on December 15, 1982. She entered University of Science and Technology of China (USTC) in August 2001 and graduated with a Bachelor degree in Biological Sciences in June, 2005. She then worked as a research assistant in the laboratory of Dr. Congzhao Zhou at USTC for one year. In August 2006, she was enrolled into the graduate field of Biophysics in the Department of Molecular Biology and Genetics at Cornell University. Ms. Zhou was initially under the supervision of Dr. Quan Hao and then was advised by Dr. Richard A. Cerione after Dr. Hao's career relocation to University of Hong Kong. She defended her Ph.D dissertation on January 30, 2012. Afterwards she will work as a Postdoc Associate in the laboratory of Dr. Thomas Cech at University of Colorado at Boulder.

To my parents and husband, sisters, and brother for their support and love

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my committee chair, Dr. Richard A. Cerione, for his contributions of time and ideas, his guidance and support, and his help in writing this dissertation. His enthusiasm and dedication for his research are motivational and contagious to me.

I am grateful to my committee member Dr. Quan Hao for his excellent mentoring in structural biology. During my study he has encouraged me to be creative and independent which will greatly benefit me in future career development.

I also want to thank my collaborator and committee member Dr. Hening Lin. Without his knowledge and insights, our collaboration could not have been so productive and successful.

I would like to express my gratitude to my committee member Dr. Sol M. Gruner for his help and kindness, especially for introducing his expertise in high pressure freezing to attack the problem I encountered in my research.

This dissertation work could not have been completed without the help of others at Cornell University, in particular, the former Hao group, the Lin group, the Cerione group, and MacCHESS.

In the Hao group, Qingqiu Huang selflessly shared his expertise and experience with me in protein expression, purification, and crystallization. Qun Liu taught me how to handle crystals, collect data, and to solve crystal structures. I am also thankful to Jun Fan and Qiuyue Yang for their sharing of knowledge and humor.

In the Lin Group, I would like to acknowledge my collaborators, Jintang Du,

Xiaoyang Su, Anita Zhu, and Bin He, for their inspiring discussions and indispensable contributions to the work in Chapter 2-4 of this dissertation.

I am indebted to the Cerione group. Without the help and fruitful discussions with Jon Erickson, Jared Johnson, Miao-chong Joy Lin, and Sekar Ramachandran, the research in chapter 5 would not have been possible. I am particularly grateful to the helpful suggestions in making important decisions in my graduate study from my friend, Bo Li, a former graduate student in the group. Additionally, the student meeting crew also offered me many insightful suggestions. I still remember the good times hanging out with some group members, such as Mark Antonyak, Jingwen Zhang, Lindsey Boroughs, Chenyue Wang, Garima Singh, Yunxin Li.

I want to acknowledge MacCHESS not only for the financial support, but also for the assistance the MacCHESS people provided. Specifically, Irina Kriksunov shared her expertise in crystallization and data collection. David Schuller taught me many useful computational skills.

Finally, I would like to thank my parents, sisters, and brother for their unconditional support in every step of my life. My husband, Rong Long, has given me nothing but love and support, which fills my graduate life with happiness. My little daughter, Carolyn Long, shines up my days of dissertation writing with her sweetest smiles.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ADP . . . . . Adenosine 5'-Diphosphate

BSA . . . . . Bovine Serum Albumin

CHES . . . . . *N*-cyclohexyl-2-aminoethanesulfonic acid

DH . . . . . Dbl Homology

DHR . . . . . DOCK-Homology Region

DNA . . . . . Deoxyribonucleic acid

DOCK . . . . . Dedicator of Cytokinesis

GAP . . . . . GTPase-Activating Protein

GDI . . . . . Guanine nucleotide Dissociation Inhibitor

GDP . . . . . Guanosine 5'-Diphosphate

GTP . . . . . Guanosine 5'-Triphosphate

GEF . . . . . Guanine nucleotide Exchanger Factor

HEPES . . . . . 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid

Mant-GDP . . . . . 2'-(or-3')-*O*-(*N*-methylanthraniloyl) Guanosine 5'-Diphosphate

NAD . . . . . Nicotinamide Adenine Dinucleotide

PH . . . . . Pleckstrin Homology

Chapter 1 Introduction

1.1 Sirtuins

The History and Characteristics of Sirtuins

Sirtuins are named after its founding member Sir2, short for *s*ilent *i*nformation *r*egulator *2*, that was discovered in yeast about two decades ago (1, 2). Sir2 was originally found to suppress the recombination of ribosomal RNA as well as to silence the genes near the telomere (1). In 1995, Brachman *et al*. discovered four additional homologues of Sir2 in yeast, Hst (homologues of Sir2) 1-4, which were all associated with gene silencing, cell cycle progression, and genomic stability. The discovery of Sir2 homologues in yeast led to the finding that sirtuins were conserved in every domain of life, from bacteria to mammals (3). Mammals have seven sirtuins: SIRT1-7, which have attracted increasing attention because of their profound roles in human health and diseases (2).

One significant characteristic of the sirtuin proteins is that they all have a highly conserved catalytic domain of about 275 amino acids and variant N- and/or C-termini of different lengths. According to the phylogenetic analysis of the catalytic domains from more than 60 sirtuins, ranging from prokaryotes to eukaryotes, sirtuins were classified into five classes: class I-IV and Class U. All the yeast sirtuins fall into Class I which also includes mammalian SIRT1-3 as well as at least one Sir2 homologue from most eukaryotes. Class II sirtuins come from various organisms, ranging from bacteria, to insects, nematodes, fungus and protozoans. Mammalian SIRT4 is a member of Class II sirtuins. Sirtuins from prokaryotes form the majority of Class III, with SIRT5 as a mammalian member. All Class IV members are from

eukaryotes, including mammalian SIRT6 and SIRT7. Class U contains those sirtuins which do not belong to Classes I-IV (4).

The unique variable N- and/or C- termini of sirtuins determine their sub-cellular localization. Yeast Sir2 was discovered to localize to the nucleolus where DNA-breaks reside (5). The seven mammalian sirtuins differ in their localizations (Table 1.1). SIRT1, SIRT6, and SIRT7 reside in the nucleus. Specifically, most SIRT1 is associated with euchromatin while SIRT6 is associated with heterochromatin, and SIRT7 localizes in the nucleolus. SIRT2 predominantly distributes in the cytoplasm. SIRT3-5 have been found in the mitochondria (6). Their different sub-cellular localizations are the major determinants of their cellular functions.

Sirtuins Are NAD-dependent ADP-ribosyltransferases, Deacetylases, Depropoinylases, and Debutyrylases.

Sirtuin was initially found as an NAD-dependent ADP-ribosyl transferase which transfers a single ADP-ribose from NAD to an arginine or lysine residue of a substrate protein. In 1999, Roy Frye discovered that human SIRT2 was able to transfer a mono-ADP ribose to BSA (7). The yeast Sir2 was then reported to ADP-ribosylate itself and histones, while a histidine to tyrosine mutation abolished its ADP-ribosylation activity (8). This led to the thought that the ADP-ribosylation activity of Sir2 was related to the output of gene silencing. In 2005, the mouse SIRT6 protein was shown to be self-ADP-ribosylated (9). In 2006, Haigis *et al*. demonstrated that mouse SIRT4 could downregulate the activity of glutamate dehydrogenase (GDH) by catalyzing its ADP-ribosylation, thereby suppressing insulin secretion (10). However,

Table 1.1. Summary of the seven mammalian sirtuins.

		Deacetylation activity	Deacetylation substrates
SIRT1	nucleus	robust	FOXO1[34], FOXO3a[22], FOXO4[22], H1K26[31], H3K9[13], H3K14[13], H4K16[13], Ku70[36], p300/CBP[21], p53[20], PCAF, PGC-1α[24], TAFI68[23], AceCS1/2[48], HMGCS1/2[48]
SIRT2	cytoplasm	robust	α-tubulin[16], FOXO1[33], FOXO3a[33], H3K14[13], H4K16[13], p53[39], p300[32]
SIRT3	mitochondrion	robust	AceCS2[17], H4K16, p53[40], ICDH2[18], GDH[18], LCAD[42], HMGCS2[48]
SIRT4	mitochondrion	none	none
SIRT5	mitochondrion	weak	Cytochrome C[18], CPS1[83]
SIRT6	nucleus	weak	H3K9[19], H3K56[36], CtIP[38]
SIRT7	nucleolus	none	none

recent studies suggested that the ADP-ribosylation activity of sirtuins was so low that it might be physiologically irrelevant (11, 12). Soon after the discovery of ADP-ribosylation, sirtuins were identified as NAD-dependent deacetylases, although distinct from the family of histone deacetylases (HDACs). Yeast Sir2 could deacetylate histone 3 at lysine residues 9 and 14, as well as histone 4 at lysine residues 5, 8 and 16, causing a decrease in the global acetylation levels of histones. This led to the conclusion that deacetylation rather than ADP-ribosylation catalyzed by sirtuins was essential for gene silencing and rDNA recombination (13, 14). The *in vitro* NAD-dependent deacetylation activity was described for a number of other sirtuins, including yeast Hst2, bacterial CobB from *Salmonella typhinurium*, archaebacterial Sir2Af from *Archaeoglobus fulgidus*, and human SIRT1-3, SIRT5 and SIRT6 (15-19). Extensive studies have shown that the sirtuin substrates are not limited to histones. For example, human SIRT1 could deacetylate diverse substrates, including the tumor suppressor p53, the histone acetyltransferase p300, the transcriptional factors FOXO proteins, NF-κB, and PGC-1α (20-24). To date, most biological functions of sirtuins have been assigned to their deacetylation activity.

Recently, a few sirtuins have been demonstrated to show depropionylation and/or debutyrylation activities (25-29). Human SIRT2 and yeast Hst2 showed moderate depropionylation and little debutyrylation on histone peptides, compared to their robust deacetylation activity (25). The findings of the depropionylation and debutyrylation of sirtuins were coupled by the identification of two novel posttranslational modifications: lysine propionylation and butyrylation on some proteins, including histones, propionyl-CoA synthetase, and p53 (28, 30). The discovery of those new activities of sirtuins not only extends their physiological roles,

but also builds a broader perspective on the substrates and biochemical functions of sirtuins. Sirtuins with little or no deacetylation activity might function to remove other acyl groups from substrate lysine residues.

Cellular Functions of Sirtuins:

- *Transcriptional Regulation*

Histone acetylation is associated with active gene transcription by relaxing the condensed chromatin DNA. sirtuins were found to negatively regulate gene expression by deacetylating histones at some specific lysine positions, including positions 9 and 26 of H1, 9 and 14 of H3, and 16 of H4 (13, 31). Human SIRT1 can deacetylate the histone acetyltransferase p300 at Lys1020 and Lys1024 within the cell cycle regulatory domain 1 (CRD1 domain), therein repressing p300-mediated transcriptional activity (21). SIRT2 can also deacetylate p300 within the catalytic histone acetyltransferase (HAT) domain (32).

The FOXO proteins, the class O of forkhead box members, are a family of transcriptional factors that play pivotal roles in the regulation of genes involved in apoptosis, metabolism, and longevity. SIRT1 was demonstrated to deacetylate FOXO3 under oxidative stress, therein enhancing the expression of target genes in cell cycle arrest and stress resistance, while reducing the expression of those genes involved in apoptosis (22). In another study, SIRT2 was reported to deacetylate FOXO1 in adipocytes and promote its binding to PPARγ, decreasing PPARγ's transcriptional activity (33). FOXO1 can also be deacetylated by SIRT1, which inhibits FOXO1-mediated gene transcription (34).

- *DNA Repair and Genomic Stability*

Tsukamoto *et al*. found that yeast Sir2, forming a complex with Sir3 and Sir4, interacted with yeast Ku proteins and promoted the double-strand break (DSB) repair pathway via non-homologous end joining (NHEJ) (5). The yeast Sir2 homologues, Hst1-4, were also involved in DNA repair and genomic stability. Hst3/Hst4 double mutants displayed low viability and increased chromosome loss (3). Further study elucidated that Hst3 and Hst4 controlled DNA repair and genome integrity by deacetylating histone 3 at lysine 56 during the G2/M phase of the cell cycle (35).

Human SIRT1 was also found to deacetylate Ku70 protein under the stress of ionizing radiation (IR), thereby activating the DSB repair pathway (36). Another nuclear sirtuin SIRT6 in mammals also played a crucial role in preserving genomic stability and promoting DNA repair. SIRT6-deficient cells became notably more sensitive to DNA-damage agents, and showed accumulation of fragmented chromosomes, detached centromeres, and gaps, which was caused by the inactivation of base excision repair (BER) mediated by SIRT6 (37). Additionally, SIRT6 was identified as a Histone 3 lysine 9 deacetylase and to be involved in the regulation of telomere maintenance. Further study demonstrated that SIRT6 could deacetylate histone 3 at lysine 56, which linked SIRT6 to telomeric DNA repair (19). Recently, SIRT6 was reported to interact and deacetylate CtIP [C-terminal Binding Protein (CtBP) Interacting Protein], a DSB resection protein, indicating that SIRT6 is pivotal in DNA damage repair and in the preservation of genomic integrity (38).

- *Cell Survival and Apoptosis*

Sirtuins have been implicated in various cancers, thanks to their suppression of apoptosis and promotion of cell survival. Among all the seven mammalian sirtuins,

SIRT1-3 have been especially well studied with regard to their roles in cancer development. SIRT1-3 can deacetylate and downregulate the transactivation activity of the tumor suppressor p53, causing cells to bypass p53-dependent apoptosis under stress (20, 39, 40). SIRT1-deficient mice showed developmental defects and usually died shortly after birth. The SIRT1-deficient MEFs exhibited p53 hyperacetylation when exposed to DNA damage agents (41). SIRT2 was shown to interact with 14-3-3 β/γ which enhanced the deacetylation of p53 by SIRT2 in an AKT-dependent way (39). The p53 protein was recently identified as a SIRT3 substrate, and SIRT3 overexpression could rescue the p53-dependent cell cycle arrest (40). As mentioned above, SIRT1 can deacetylate FOXO3, which represses FOXO3's capability to induce cell death, thereby promoting cell survival (22).

- *Metabolism*

Mitochondria are the energy machinery of cells. In mammals, three sirtuins, SIRT3-5, predominantly reside in the mitochondria, which play important roles in regulating metabolism in response to nutrient changes. Acetyl-CoA synthetase 2 (AceCS2), the first SIRT3 substrate to be identified, is deacetyated and activated by SIRT3, thereby producing more acetyl-CoA for energy production in the citric acid cycle (17). SIRT3 has been found *in vitro* to deacetylate and activate isocitrate dehydrogenase 2 (ICDH2) which generates a key molecule, α-ketoglutarate, in the citric acid cycle. In the same study, SIRT3 was demonstrated to deacetylate and activate glutamate dehydrogenase (GDH) which converts glutamate to α-ketoglutarate for ATP production and plays a crucial role in the citric acid and urea cycles (18). Recently, long-chain acyl CoA dehydrogenase (LCAD) has been identified as a SIRT3 deacetylation target, indicating that SIRT3 is pivotal for regulating fatty-acid oxidation

(42). SIRT3 can deacetylate 3-hydroxyl-3-methylglutary-CoA synthetase 2 (HMGCS2) on the three lysines within its active site, and enhances HMGCS2 enzymatic activity under fasting conditions, thereby increasing the production of ketone bodies for direct use as an energy source in some tissues (43). Several research groups have shown that SIRT3 interacts with and deacetylates the mitochondrial respiration complex I-V in the electron transport chain, suggesting that SIRT3 helps regulate ATP homeostasis (44-47). However, opposite to SIRT3, SIRT4 has been reported to ADP ribosylate GDH and downregulate its activity, which inhibits ATP production and insulin secretion in pancreatic β cells (10).

Similar to SIRT3, SIRT1 has been shown to deacetylate and activate AceCS1, 2 and HMGCS1, 2, with AceCS1, HMGCS1 as the cytoplasmic homologues of mitochondrial AceCS2 and HMGCS2, respectively, directly linking SIRT1 to the citric acid cycle and fatty-acid oxidation for ATP generation (48). Additionally, SIRT1 can interact with and deacetylate the transcriptional coactivator PGC-1α, through which SIRT1 controls glucose homeostasis via the glycolytic pathway in response to fasting (24).

- *Neurodegeneration and Neuroprotection*

Sirtuins play important roles in neurodeneration and neuroprotection, as some sirtuins are highly expressed in the nervous system, including SIRT1 and SIRT2 (6). SIRT1 was highly expressed in neuronal bodies while SIRT2 was enriched in oligodendrocytes and in the Schwann cells of axons (49). Qin *et al*. found that SIRT1 prevented amyloid β peptide formation in Alzheimer disease in response to calorie restriction by suppressing the Rho-associated kinase ROCK1-mediated inhibition of α-secretase which produces the non-toxic α-amyloid instead of the pathogenic β-amyloid

peptide (50). SIRT2 deacetylates α-tubulin which is the key component of microtubules, suggesting that SIRT2 is pivotal in the modulation of neuronal differentiation and migration (16). The mechanism of sirtuin's neurodegenerative and neuroprotective is still unclear.

Sirtuins have been appreciated for their role in lifespan extension. Overexpression of Sir2 can extend the lifespan of yeast and worms under conditions of calorie restriction (51, 52). This beneficiary effect of lifespan extension under calorie restriction was also found in other organisms, including fruit flies, spiders, and rodents (53-55). Sirtuins are also involved in some age-associated diseases, such as metabolic disease (56, 57), Alzheimer's disease (50, 58, 59), Parkinson's disease (60, 61), and cancer (62-65).

The Deacetylation Mechanism of Sirtuins and Substrate Specificity

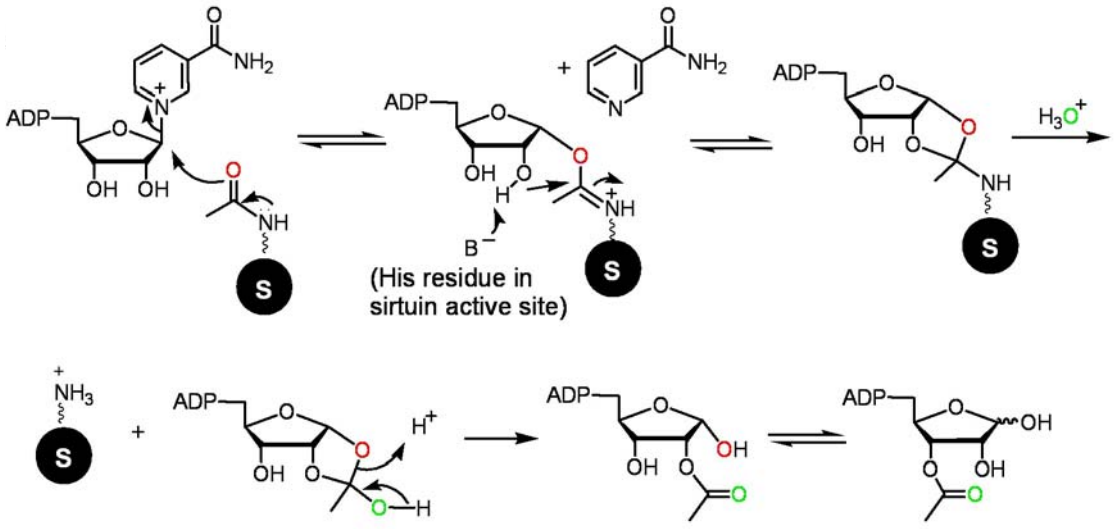
In the presence of NAD, sirtuins transfer the acetyl group from an acetylated substrate protein/peptide to the ADP-ribose portion of NAD, and produce nicotinamide, a deacetylated protein/peptide, and a novel molecule, O-acetyl-ADP-ribose (*O*AADPRr) (66-68). The two substrates, acetylated protein and NAD, bind to the active site of sirtuins in an ordered manner: the acetylated protein binds first, followed by the binding of NAD. Nicotinamide is then released. The deacetylated protein and *O*AADPr are released randomly when the reaction is completed (69). With the expanding understanding of the important cellular roles of sirtuins, the importance of understanding the enzymatic mechanism of deacetylation becomes more obvious, to facilitate the design of inhibitors and activators for pharmaceutical purposes. Extensive biochemical and structural studies have been conducted to elucidate the

enzymatic mechanism. Three alternatives have been proposed: the nucleophilic attack mechanism, the enzyme nucleophile mechanism, and the ADPR-peptidyl-imidate mechanism, with the latter being the most widely accepted (66).

In the ADPR-peptidyl-imidate mechanism (Fig.1.1), upon the binding of both acetylated protein and NAD, the nicotinamide group from NAD is first cleaved, forming an oxocarbonium-ion transition state; then, the carbonyl oxygen from the acetylated substrate protein attacks the positively charged 1'-carbon position of the nicotinamide ribose, forming the ADPR-peptidyl-imidate intermediate. The ribose 2'-OH, when deprotonated by the enzyme, attacks the amidate at the acetyl carbon, generating the 1'-2'-cyclic intermediate, followed by the hydrolysis of a water molecule, producing 2'-O-acetyl-ADP-ribose (2'- *O*AADPr) which can be non-enzymatically isomerized into 3'-*O*AADPr. The histidine residue that is highly conserved among the sirtuin family serves as a general base to deprotonate the 2'-OH for attacking the 1'-*O*-alkyl-amidate. Mutating the conserved histidine to tyrosine dramatically decreased the enzyme's deacetylation activity.

A series of structures of sirtuin complexes with either their substrate peptide or NAD, or both, as well as with deacetylated peptide product, have been solved (reviewed in 67). These structures have defined the binding sites for the two substrates and provided insights into the catalytic mechanism proposed above. The overall structure of the sirtuins is highly conserved from bacteria to human, consisting of a large Rossmann-fold domain characterized for NAD binding and a small zinc-binding domain with four invariant cysteine residues. The acetylated substrate peptide binds to the small cleft between the two domains and forms a β-strand with the enzyme mainly via backbone hydrogen bonds, causing the zinc-binding domain to rotate clockwise to

Figure 1.1. The deacetylation mechanism of sirtuins. Some of the oxygen atoms are colored differently to indicate how they originate.


ADP
NH2
OH
NH
S
(His residue in
sirtuin active site)
H3O+
NH3
H+

the Rossmann-fold domain, thereby driving the enzyme from the inactive open state to the active close state. The acetyl lysine inserts into a hydrophobic tunnel and is accessible to the nicotinamide ribose of NAD for the nucleophilic attack of the carbonyl oxygen. Upon the binding of the acetylated substrate peptide, NAD binds to the enzyme in a productive position with the nicotinamide moiety inserting into the proposed C pocket where it is cleaved as the first step of the reaction. The invariant histidine residue forms hydrogen bonds with the 3'-OH of the nicotinamide ribose, supporting the idea that histidine acts as a base to deprotonate the 2'-OH to promote the nucleophilic attack.

Previously, kinetic experiments and mass spectrometry analyses have confirmed the existence of the alkylamidate and the cyclic intermediates (70, 71). However, no direct observation has been obtained until recently when two structures of different sirtuins with the alkylamidate intermediate were solved and shown to support one of the intermediate steps. By using the mechanism-based inhibitor, thioacetyl lysine peptide, the S-alkylamidate intermediate was captured in the Sir2Tm and SIRT3 crystals (72, 73). The intermediate state aligned well with both the acetyl lysine peptide and NAD in a Michaelis-Menten complex, suggesting that only a small rearrangement was required for the removal of nicotinamide and acetyl groups. More sirtuin structures with the oxocarbonium ion or bicyclic intermediate bound are needed to address the other two intermediate steps.

As mentioned earlier, mammalian SIRT1 could deacetylate diverse substrates (20-24, 31, 34, 36, 48), which raises the question whether sirtuins have any substrate specificity. Not surprisingly, human SIRT1 showed no specificity for the amino acid sequence flanking the acetyl lysine of substrate peptides *in vitro* (74). However,

previous kinetic studies reported that yeast Sir2, Hst2, and human SIRT2 showed a few fold differences in either the substrate binding affinity or the catalytic efficiency over several mono-acetylated histone peptides (69). All the sirtuin complex structures with peptide bound showed that the interactions between the acetyl peptide and the enzyme mainly came from the backbone hydrogen bonds, with only a couple of hydrogen bonds from the side chains at the -1 and +2 positions (acetyl lysine at the position 0). Mutation of any individual residue at these two positions had only modest effects on the peptide binding affinity, suggesting that the side chain interactions were not the major contributing factors in sirtuin's substrate recognition (75).

Almost all the previous studies on substrate specificity focused on the primary sequence proximate to the acetylated lysine of substrate peptides. Under physiological conditions, substrate recognition by the sirtuins may be partially determined by their sub-cellular localization and the sequences beyond the catalytic core of sirtuins, as well as by the other parts except the sequences flanking the acetyl lysine of substrate proteins. Recently sirtuins were found to remove propionyl and butyryl groups from lysine (25-29), which implied that sirtuins might exhibit their specificity, not on the sequence near the modified lysine, but on the modification of the lysine residues. This will likely attract more attention from the sirtuin research community with regard to understanding the properties of their substrate binding sites.

Activators and Inhibitors of sirtuins

As described above, sirtuins play important yet complex roles in human health and disease, which has led to increasing attention for their therapeutic potential. For instance, SIRT2 inhibitors have therapeutic potential in the treatment of Parkinson's

disease (60). Efforts have been put forward to identify sirtuin activators and inhibitors. To date, a few small molecules have been reported to be effective in either elevating or inhibiting the activity of the sirtuins (76-80).

Several plant polyphenols were identified by high throughput screening of a small molecule library. Resveratrol, commonly found in red wine, was one of those polyphenols which exhibited a significant activation of SIRT1 using *in vitro* fluorescent assays (76). However, this has been controversial because no SIRT1 activation was detected upon the addition of resveratrol when using non-fluorophore-containing peptides (81, 82).

A number of small molecules have been reported as sirtuin inhibitors, including sirtinol and suramin, but their selectivity among sirtuins requires further optimization (77, 78). As the first product of the sirtuin-catalyzed deacetylation reaction, nicotinamide is the physiological inhibitor which acts as a non-competitive product inhibitor to reverse the reaction and regenerate NAD (79, 80). Recently, the mechanism-based inhibitor, thioacetyl peptide, showed significant potential in inhibiting sirtuins activity by stalling the intermediate at the active site (71).

During my dissertation research, I aimed to investigate those mammalian sirtuins that show low deacetylation activity, such as mammalian SIRT4-7, to figure out their biological functions and physiological substrates. In Chapter 2, I started with the X-ray crystal structures of SIRT5 complexed with different thioacetyl peptides, which showed that the primary sequences flanking the acetylated lysine did not influence the substrate specificity of SIRT5. Interestingly, a SIRT5 structure with a buffer molecule of CHES (N-cyclohexyl-2-aminoethanesulfonic acid) bound showed

that SIRT5 harbored an unusual lysine binding pocket with extra space to accommodate larger modifications and two non-hydrophobic residues, Tyr102 and Arg105, at the bottom of the pocket. These structural features led to the discovery of a novel function for SIRT5, specifically as a NAD-dependent demalonylase and desuccinylase. None of the other six mammalian sirtuins displayed significant demalonylation or desuccinylation activity. In addition, SIRT5 regulated the enzymatic activity of carbamoyl phosphate synthetase 1 (CPS1), a physiological substrate of SIRT5 (83), by desuccinylating CPS1 at lysine 1291. In Chapter 3, I further crystallized and solved a series of structures of SIRT5 bound with a succinyl peptide, or both a succinyl peptide and NAD, or a bicyclic intermediate, delineating the desuccinylation reaction step by step. This is the first time that a bicyclic intermediate was directly observed, which supports the ADPR-peptidyl-imidate mechanism.

In Chapter 4, we investigated a Sir2 homologue from a parasite *Plasmodium falciparum (P. falciparum)* which is the most dangerous species of the *Plasmodium* genus that causes malaria. *P. falciparum* Sir2A (PfSir2A) has been shown to regulate the expression of surface antigens to evade the detection by host immune surveillance. It has been proposed that PfSir2A achieves this by deacetylating histones. However, PfSir2A exhibited weak deacetylation activity in previous studies (84). We demonstrated that PfSir2A preferentially catalyzed the hydrolysis of the medium and long chain fatty acyl modifications from lysine residues *in vitro*. In addition, medium and long chain fatty acyl modifications on lysine residues were shown to exist in the proteome of *P. falciparum*, and these modifications could be reversed by recombinant PfSir2A *in vitro*. The structures of PfSir2A complexed with a myristoyl peptide or

with both a myristoyl peptide and NAD were solved to provide a structural explanation as to how PfSir2A preferentially removes the medium and long chain fatty acyl modifications. The novel function and the crystal structures of PfSir2A may facilitate the development of inhibitors for treating malaria.

1.2. Rho GTPases and Rho GEFs of DOCK180 Family

Rho GTPases: Functions, Posttranslational Modifications, and Regulation

The Rho family of small GTPases, with Cdc42, Rac1, and RhoA as three classical members, have been initially known for their pivotal functions in regulating actin cytoskeleton dynamics (85, 86). For example, the hyperactivation of Cdc42 triggers filopodia formation (87, 88), while activated Rac1 and RhoA induce the formation of lamellipodia and actin stress fibers, respectively (87, 89, 90). With an ever expanding list of downstream targets, Rho GTPases have been demonstrated to be essential for a wide range of cellular processes, including cell polarity, apoptosis, membrane trafficking, phagocytosis, cellular transformation, gene transcription, adhesion, and cell growth (91-93).

Rho GTPase proteins contain a classical guanine nucleotide-binding domain (GTPase or G- domain), followed by a small polybasic region (PBR) and a CAAX motif (where C represents cysteine, A is any aliphatic amino acid, and X is any amino acid) at the C-terminus. To date, a number of X-ray crystal structures of Cdc42, Rac1 and RhoA have been solved to better the understanding of the molecular basis of how they interact with binding partners (94-97). The G-domain, with relatively low intrinsic GTP hydrolytic activity, is responsible for the binding of upstream regulators and downstream effectors. The PBR consists of several positively charged lysine and/or arginine residues, which promote the association of these GTPases with negatively charged membranes (98). The C-terminal CAAX motif is posttranslationally modified by three enzymes. First, the prenyltransferase catalyzes the modification of CAAX motifs and covalently adds a farnesyl or geranylgeranyl

isoprenoid lipid tail to the cysteine residue. Cdc42, Rac1, and RhoA are all geranylgeranylated. Next, the Rce1 (Ras converting enzyme 1) endoprotease catalyzes the cleavage of the –AAX peptide from the C-terminus. Finally, a methyl group is added to the prenylated cysteine by the Icmt (isoprenylcysteine-*O*-carboxyl methyltransferase) (99). The subcellular localizations of Rho GTPases are determined by their posttranslational modification. Cdc42 is widely distributed at the plasma membrane, Golgi, endoplasmic reticulum, and nuclear envelope, whereas Rac1 and RhoA predominantly associate with the plasma membrane (100).

The cellular functions of Rho GTPases largely depend on their subcellular localization. The lipid tails of GTPases insert into cellular membranes where they are able to engage their regulators: GEFs (Guanine nucleotide Exchange Factors), GAPs (GTPase-Activating Proteins), and GDIs (Guanine nucleotide Dissociation Inhibitors). GEFs promote the exchange of GDP for GTP, therefore activating Rho GTPases. GAPs markedly enhance the hydrolysis of GTP and thus inactivate GTPases. Rho GDIs sequester GDP-bound Rho GTPases in the cytosol by inhibiting nucleotide exchange (98, 101, 102). Rho GTPases switch between a GDP-bound inactive form and a GTP-bound active form to function as molecular switches to turn on and off signaling pathways.

The Dbl Family GEFs

Two families of GEFs have been identified for Rho GTPases: the classical Dbl family (103) and the recently discovered DOCK180 family (104-106). Dbl, as the founding member of Dbl family, was first identified as an oncogene in human diffuse B-cell lymphoma. Sequence analysis of Dbl uncovered a region with high homology

to the yeast Cdc24 protein which was shown to serve as a GEF for the yeast Cdc42 protein (107). To date, about 70 members of the Dbl family have been identified, including the Rac1-specific GEF Tiam1 and the Cdc42-specific GEF Intersectin (101). All the Dbl family members contain two conserved domains, the Dbl homology (DH) and the pleckstrin homology (PH) domains. The DH domain is responsible for catalyzing nucleotide exchange, while the PH domain has been demonstrated to interact with the plasma membrane (108-111).

Structural studies, especially the complex structure of Rac1 and DH-PH module of Tiam1, have elucidated the mechanism of how DH-PH domains of Dbl family members promote the nucleotide exchange of their cognate GTPases (96). The DH domain harbors three conserved regions (CR1-CR3) which form a three-helix bundle and provide the interface for Rac1 binding. The highly conserved residue Glu1047 from Tiam1 interacts via hydrogen bonding with three residues (Thr35, Val36, and Tyr32) from the switch I region of Rac1, which induces a conformational change within switch I, and subsequently destabilizes a bound Mg^{2+} ion and disrupts nucleotide binding. The Mg^{2+} ion binding site is further sterically blocked by the movement of Ala59 within the switch II region of Rac1, such that Ala59 forms a hydrogen bond with the highly conserved residue Lys1195 of Tiam1. Those structural features collectively delineate the mechanism of nucleotide exchange by Dbl family GEFs.

The residues 1177-1179 in Tiam1, the most variant positions among other DH domains, contact residues from the β2- β3 strands of Rac1 thus accounting for the specificity of Tiam1 for Rac1 versus Cdc42 and RhoA. Biochemical studies have shown that position 56 of Rac1 or Cdc42 is pivotal for this specificity. Changing

Trp56 of Rac1 to Phe (the corresponding residue in Cdc42) prevented Rac1 from being activated by Tiam1, whereas the Rac1 (W56F) mutant could be activated by the Cdc42-specific GEF, Intersectin (112). Similarly, substitution of Phe at position 56 of Cdc42 for Trp enabled Cdc42 to be activated by Tiam1 but not by Intersectin. These studies demonstrated that position 56 of Rac1 and Cdc42 is necessary and sufficient for the specific recognition by their cognate Dbl family GEFs.

In the Rac1-Tiam1 complex structure, the PH domain orientates away from the DH domain and fails to make contact with Rac1. Functionally, the DH domain exhibits full GEF activity *in vitro*. However, the PH domain is required for the transforming ability of Dbl *in vivo* (113). This suggests that the PH domain is necessary for the biological functions of Dbl family members. It has been proposed that the PH domain of Dbl family members is essential for membrane binding via interactions with phosphoinositides, albeit the binding of phosphoinositides to PH domains is weak and low in specificity (114, 115). Therefore, although the PH domain is essential for the cellular functions of Dbl proteins, the mechanism of how the PH domain regulates the GEF activity awaits further structural and biochemical investigation.

The DOCK180 Family GEFs: Discovery, Functions, and Classification.

The other family of GEFs for Rho GTPases is the atypical DOCK180 family, which shares no sequence similarity with the classical Dbl family of GEFs (116, 117). DOCK180-related proteins have been shown to activate Rac1 and/or Cdc42, but not RhoA. The founding member of the DOCK180 family, Ced-5, was originally identified as a binding partner for the oncogenic c-Crk protein in *Caenorhabditis*

elegans (105, 118). DOCK180 family proteins are evolutionarily conserved from worms, to flies and mammals. Eleven DOCK180-related proteins have been identified in mammals, designated as DOCK1-DOCK11 (DOCK1 is also known as DOCK180) (116, 119, 120) (Fig. 1.2).

DOCK180 family members have been demonstrated to regulate multiple cellular processes, including cell migration and phagocytosis (118, 121-124), tumor suppression (125), and neuronal development and differentiation (126, 127). DOCK180 binds to the scaffold protein ELMO and forms a trimeric complex which activates Rac1. The Rac1 GTPase is a key regulator of the cell cytoskeleton and is essential for cell migration. Recently, RhoG, activated by the GEF TRIO, was found to be functioning upstream of the CrkII/ELMO/DOCK180/Rac1 pathway, which is essential for the engulfment of apoptotic cells (124). DOCK4, indentified as a Rap1 GEF, has been shown to be mutated in some human cancer cells, including prostate and ovarian cancers. Overexpression of wild type DOCK4 in mouse osteosarcoma cells represses cell growth and tumor invasion (125). DOCK180-related proteins play important roles in almost every process of neuronal cell development, including migration, neurite outgrowth, axon guidance, and synapse formation. DOCK1 and DOCK3 have been reported to promote neurite outgrowth induced by the signaling of nerve growth factor (NGF) in PC12 cells (126). DOCK7 has been demonstrated to be activated by the growth factor neuregulin-1 (NRG-1) and to regulate axon formation through Rac1 and Cdc42 activation (127). DOCK4 and DOCK9 have been shown to be enriched in hyppocampal neurons and to be associated with the synapse formation.

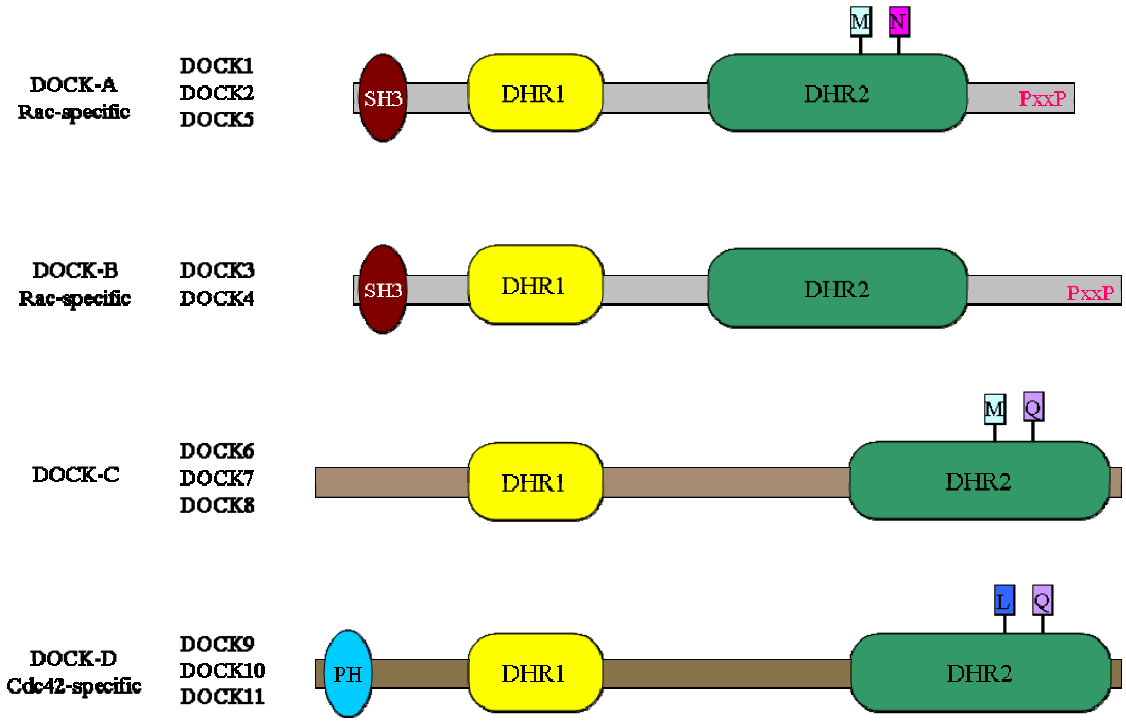
The eleven mammalian DOCK proteins are classified into four subfamilies: DOCK-A to DOCK-D subfamily (117) (Fig. 1.2). The DOCK-A subfamily contains

DOCK1, 2, and 5, while the DOCK-B subfamily consists of DOCK3 and DOCK4. Proteins within these two subfamilies share 50-65% sequence identity. Only members in the subfamilies DOCK-A and DOCK-B possess the proline-rich motif PxxP which mediates the interaction with c-CRK protein. Also, DOCK-A and DOCK-B members have a SH3 domain in the N-terminus. This SH3 domain can interact with the PxxP motif within the C-terminus of the same DOCK protein and mediate autoinhibition of the GEF activity. The binding with other regulatory proteins can release the autoinhibhition by binding to the SH3 domain and enhance the accessibility of the DHR2 domain to its cognate GTPase. DOCK6-8 fall within the DOCK-C subfamily, and DOCK9-11 comprise the DOCK-D subfamily. The DOCK-C subfamily is more related to the DOCK-D subfamily than to the DOCK-A/B subfamily, with about 30% sequence identity with DOCK-D members. The DOCK-D proteins contain an N-terminal PH domain which has been proposed to be important for membrane targeting and protein-protein interactions (128).

Regulation of DOCK180 Family Proteins

To date, little work has been done to study the regulation of DOCK180-related proteins. Recently, a few studies have been conducted on some DOCK180 members, including DOCK180 and DOCK7. Given the primary sequence similarity of the members within the family, it is likely that the regulatory mechanism found for a specific member might be shared with the whole family, although further research is required to confirm this idea.

Figure 1.2. The eleven members of the DOCK180 family in mammals (adapted from reference 128). DOCK180 family was classified into DOCK-A, B, C, D subfamilies. One of the significant characteristics of DOCK180 family is that all the members harbor two DOCK homology domains: DHR1 and DHR2. Two residues: Met and Asn in DOCK-A subfamily, Leu and Gln in DOCK-D subfamily, are important for the specific recognition of Rac and Cdc42, respectively. The corresponding residues in DOCK-C subfamily are Met and Gln.


DOCK-A
Rac-specific
DOCK1
DOCK2
DOCK5
SH3
DHR1
DHR2
PxxP
M
N
DOCK-B
Rac-specific
DOCK3
DOCK4
SH3
DHR1
DHR2
PxxP
DOCK-C
DOCK6
DOCK7
DOCK8
DHR1
DHR2
M
Q
DOCK-D
Cdc42-specific
DOCK9
DOCK10
DOCK11
PH
DHR1
DHR2
L
Q

- *Associating with Phosphoinositides*

The DHR1 domain of DOCK180 has been shown to interact with PtdIns (3,4,5) P3 (PIP3), which enhances the translocation of DOCK180 from the cytosol to the plasma membrane (129). This occurs in response to PIP3 produced by the phosphatidyinositol 3-kinase (PI3K) upon the stimulation of growth factors. The DHR1 domain is highly conserved among members of the DOCK180 family. Thus, the mechanism of associating with phosphoinositides can be applied by other members to translocate to the membrane and to activate downstream targets.

- *Binding to Adaptors*

DOCK180 was originally found to bind to the SH3 domain of CrkII, an adaptor protein which also contains an SH2 domain for binding to phosphorylated tyrosine residues (105). Additionally, the DOCK180-CrkII complex is always found bound to ELMO proteins via the interaction between the SH3 motif of DOCK180 and the PxxP motif of ELMO (116). ELMO1 is essential for the full GEF activity of DOCK180 and for the subsequent activation of Rac1 by DOCK180. All the DOCK-A and DOCK-B subfamily members possess an SH3 motif within their N-terminus. Therefore, interacting with adaptors can be a common approach utilized by DOCK180-related proteins to regulate their cellular functions.

- *Inter- and Intra- Molecular Interactions*

It has been reported that the N-terminus of DOCK9 DHR2 domain mediates the dimerization of the protein, which enhances its GEF activity (120). Additionally, the N-terminal SH3 domain of DOCK180 binds to the C-terminal PxxP motif, leading to the subsequent autoinhibition of its GEF activity (130). This autoinhibition mechanism can be applied to other DOCK-A and DOCK-B subfamily members,

because they all harbor these SH3 and PxxP motives. DOCK11 has been found to interact with the active form of Cdc42, and this binding of active Cdc42 further activates DOCK11, forming a positive feedback activation loop (131).

- *Phosphorylation*

The activities of the classical Dbl family GEFs are modulated by serine/threonine phosphorylation (132, 133). Similarly, DOCK180 was found be phosphorylated on serine/threonine residues (134). Interestingly, DOCK7 has been demonstrated to be directly phosphorylated on residue Tyr1118 by the EGF receptor family member ErbB2, in a heregulin-dependent manner (127). This tyrosine phosphorylation directly activates DOCK7 to promote the nucleotide exchange of the downstream GTPases Rac1 and Cdc42. DOCK7 belongs to the DOCK-C subfamily which is different from the other three subfamilies that have either SH3, PxxP, or PH motifs for intermolecular interactions. Herein, it is possible that DOCK-C subfamily members may act as direct substrates of signaling molecules that regulate their GEF activity.

- *Ubiquitination*

DOCK180 was more stable when overexpressed with ELMO in cells, which suggested that ELMO might block the ubiquitination and subsequent degradation of DOCK180 (135). However, the ubiquitinaiton of DOCK180 has still not been definitively demonstrated.

The DHR1 and DHR2 Domains of DOCK180 Family

One of the major features of the DOCK180 family is that all members possess two DOCK homology regions, DHR1 and DHR2, which do not share sequence

similarity with PH and DH domains of the Dbl family (117). The DHR1 domain is located in the N-terminus of DOCK180 family members and consists of about 250 residues, which share weak homology with the C2 domains found in the DOCK-A and DOCK-B subfamilies, whereas no C2 homology was found in the DHR1 domain of the DOCK-C or DOCK-D subfamily members. The DHR1 domain of DOCK1 has been demonstrated to specifically bind to PIP3-enriched membranes (129). The X-ray crystal structure of the DHR1 domain of DOCK1 has revealed that several highly conserved residues (such as lysine and arginine residues) present on the surface form a highly basic pocket to accommodate the negatively charged PIP3 from cellular membranes (136). The GEF activity of DOCK1 is independent from the DHR1 domain. However, DHR1 has an important functional role as it has been shown to be required for Rac1-mediated cell polarity and cell elongation.

The DHR2 domain is located downstream of the DHR1 domain and consists of about 500 residues which are highly conserved among the DOCK180 family. The DHR2 domain shares no sequence similarity with the DH domain from the Dbl family GEFs. Therefore, the DOCK180 family members are referred to as the atypical Rho GEFs. Previously, the DHR2 domain has been demonstrated to be necessary and sufficient to activate GTPases. However, a recent study on DOCK180 showed that a limited C-terminal portion of the DHR2 domain of DOCK180 displayed full GEF activity (137). The N-terminus of the DOCK9 DHR2 domain is responsible for dimerization which may play a role in promoting its maximal GEF activity (120, 138). This suggested that all the DOCK proteins might exist as homodimers, with dimerization being mediated by the N-terminal portion of the DHR2 domain.

The Specificity of DOCK180 Family on Rac and/or Cdc42

A good deal of biochemical studies have established that the DOCK-A and DOCK-B subfamily members are Rac-specific GEFs, while the DOCK-D proteins can only specifically activate Cdc42 (117, 139). It is proposed that the DOCK-C subfamily exhibits no specificity and can recognize both Rac1 and Cdc42 (127, 140). What is the mechanism by which DOCK180 GEFs activate Rho GTPases? Why do different subfamilies recognize different Rho G-proteins? Structural studies on the DOCK9-Cdc42 and DOCK2-Rac1 complexes have elucidated the activation mechanism and shed light on how the specificity of the DOCK180 proteins toward their substrate GTPases is manifested (138). The overall structure of the DOCK9-DHR2 domain consists of three lobes: lobe A-C, of which lobe A comes from the N-terminal portion of DHR2 and provides critical intermolecular interactions for dimerization. Lobe B and C make contacts with the switch I and switch II domains of Cdc42 and cause conformational changes only in switch I, which is different from Dbl GEFs that cause conformational changes in both switch I and switch II. Upon the binding of DHR2, the residue Phe28 of Cdc42 inserts into the hydrophobic pocket of lobe B and weakens the binding of GDP. The nucleotide sensor of DHR2, the α10 insert, moves to the nucleotide-binding site of Cdc42. Within this α10 insert sensor, an invariant residue from plants to humans, Val1951, helps to exclude Mg^{2+}, therefore further destabilizing GDP binding. For Dbl GEFs, the binding of the GEF also induces a conformational change in switch II, which moves a highly conserved Ala residue into the position normally occupied by Mg^{2+}, thereby releasing Mg^{2+} from the GTPase. The invariant Val residue suggests that all of the DOCK180 proteins share the same mechanism for activating their cognate GTPases. This is further supported by the observation that the

replacement of valine with alanine abolished the GEF activity of DOCK9 and DOCK1.

Biochemical studies have shown that several residues in the switch I and β2-β3 regions of Rac or Cdc42 are important for the specific recognition by their cognate DOCK proteins (137, 139). These residues include Ala27, Gly30, and Trp56 for Rac, and Lys27, Ser30, and Phe56 for Cdc42, through which Rac and Cdc42 undergo interactions with DOCK proteins. Rac-to-Cdc42 substitutions of these residues can enable Rac to at least be partially activated by the Cdc42-specific DOCK9 DHR2, and vice versa. Among those residues, position 56 is the key residue for specific recognition, given that the Rac1 W56F mutant only showed basal activation by the DOCK1 DRH2c. In the X-ray crystal structure of the DOCK9-Cdc42 complex (or the DOCK2-Rac complex) (138, 141), the side chain of Phe56 (or Trp56) interacts with Leu1941 and Gln1944 (or Met1529 and Asn1532) of the DOCK9 (or the DOCK2) DHR2 domain. Replacing Met1524 of the DOCK1 DHR2c with leucine eliminated about 75% of the activation of wild-type Rac1, while restoring about 50% of the activation of the Rac1 W56F mutant. This suggests that Met1524 is crucial for the ability of DOCK1 to distinguish Rac1 from Cdc42. It is reasonable to assume that residue Asn1527 in DOCK1 (Asn1532 in DOCK2 and compartment Gln1944 in DOCK9) should also be important for the specific recognition of Rac1 (Rac2 and Cdc42).

However, little is known regarding why DOCK-C subfamily members (including DOCK6-8) are capable of the dual recognition of Rac and Cdc42, given that DOCK-C members have not exhibited robust GEF activity yet *in vitro*. It is controversial that DOCK7 can activate both Cdc42 and Rac1, due to the conflicting

results from *in vitro* GEF assays (127, 142). Similarly, one study reported that DOCK6 showed no catalytic activity toward either Rac or Cdc42 (140). Therefore, it will be interesting to ultimately characterize the DOCK-C subfamily members.

In chapter 5, I show that the DOCK7-DHR2 domain is a potent GEF for prenylated Cdc42 and Rac1 in a model liposome system, demonstrating that the prenylation and resulting membrane localization of Cdc42 or Rac1 are pivotal for the activation by DOCK7. Additionally, as demonstrated previously for DOCK1 (137), we show that DOCK7 harbors residues that impart GEF specificity and these can be mutated to shift a given DOCK-DHR2 activity profile. Finally, using our liposome reconstitution assays, we show that the DOCK7-DHR2 limit domain (DHR2s) possesses an N-terminal site distinct from the nucleotide exchange active site that binds preferentially to the active forms of Cdc42 and Rac1, which helps to recruit DOCK7 to the membrane. This recruitment results in an increase in the effective concentration of the GEF at the membrane surface that in turn provides for an accelerated rate of nucleotide exchange for prenylated Cdc42 in the presence of liposomes.

References

1. Gottlieb, S. and Esposito, R. E. (1989) *Cell* **56**, 771-776.
2. Michan, S., and Sinclair, D. (2007) *Biochem. J.* **404**, 1-13.
3. Brachmann, C. B., Sherman, J. M., Devine, S. E., Cameron, E. E., Pillus, L. and Boeke, J. D. (1995) *Genes Dev.* **9**, 2888-2902.
4. Frye, R. A. (2000) *Biochem. Biophys. Res. Commun.* **273**, 793-798.
5. Tsukamoto, Y., Kato, J., and Ikeda, H. (1997) *Nature* **388**, 900-903.
6. Michishita, E., Park, J. Y., Burneskis, J. M., Barrett, J. C., and Horikawa, I. (2005). *Mol. Biol. Cell* **16**, 4623-4635
7. Frye, R. A. (1999) *Biochem. Biophys. Res. Commun.* **260**, 273-279.
8. Tanny, J. C., Dowd, G. J., Huang, J., Hilz, H. and Moazed, D. (1999) *Cell* **99**, 735-745.
9. Liszt, G., Ford, E., Kurtev, M., and Guarente, L. *J. Biol. Chem.* **280**, 21313-21320 (2005).
10. Haigis, M. C., Mostoslavsky, R., Haigis, K. M., Fahie, K., Christodoulou, D. C., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Karow, M., Blander, G., Wolberger, C., Prolla, T. A., Weindruch, R., Alt., F.W., and Guarente, L. (2006) *Cell* **126**, 941-954.
11. Kowieski, T. M., Lee, S. and Denu, J. M. (2008) *J. Biol. Chem.* **283**, 5317-5326.
12. Du, J., Jiang, H., and Lin, H. (2009) *Biochemistry* **48**, 2878-2890.
13. Imai, S.-i., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000) *Nature* **403**, 795-800.

14. Landry, J., Sutton, A., Tafrov, S. T., Heller, R. C., Stebbins, J., Pillus, L. and Sternglanz, R. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 5807-5811.

15. Smith, J. S., Brachmann, C. B., Celic, I., Kenna, M. A., Muhammad, S., Starai, V. J., Avalos, J. L., Escalante-Semerena, J. C., Grubmeyer, C., Wolberger, C., and Boeke, J. D. (2000) *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6658-6663.

16. North, B. J., Marshall, B. L., Borra, M. T., Denu, J. M., and Werdin, E. (2003) *Mol. Cell* **11**, 437-444.

17. Schwer, B., Bunkenborg, J., Verdin, R. O., Andersen, J. S. and Verdin, E. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10224-10229.

18. Schlicker, C., Gertz, M., Papatheodorou, P., Kachholz, B., Becher, C. F. W., and Steegborn, C. (2008) *J. Mol. Biol.* **382**, 790-801.

19. Michishita, E., McCord, R. A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T. L. A., Barrett, J. C., Chang, H. Y., Bohr, V. A., Ried, T., Gozani, O., and Chua, K. F. (2008) *Nature* **452**, 492-496.

20. Vaziri, H., Dessain, S. K., Eaton, E. N., Imai, S., Frye, R. A., Pandita, T. K., Guarente, L., and Weinberg, R. A. (2001) *Cell* **107**, 149-159.

21. Bouras, T., Fu, M., Sauve, A. A., Wang, F., Quong, A. A., Perkins, N. D., Hay, R. T., Gu, W., and Pestell, R. G. (2005) SIRT1 deacetylase p300 *J. Biol. Chem.* **280**, 10264-10276.

22. Brunet, A., Sweeney, L. B., Sturgill, J. F., Chua, K. F., Greer, P. L., Lin, Y., Tran, H., Ross, S. E., Mostoslavsky, R., Cohen, H. Y., Hu, L. S., Cheng, H., Jedrychowski, M. P., Gygi, S. P., Sinclair, D. A., Alt, F. W. and Greenberg, M. E. (2004) *Science* **303**, 2011-2015.

23. Yeung, F., Hoberg, J. E., Ramsey, C. S., Keller, M. D., Jones, D. R., Frye, R. A. and Mayo, M. W. (2004) *EMBO* **23**, 2369-2380.

24. Rodgers, J. T., Lerin, C., Haas, W., Gygi, S. P., Spiegelman, R. M. and Puigserver, P. (2005) *Nature* **434**, 113-118.

25. Smith, B. C., and Denu, J. M. (2007) *J. Biol. Chem.* **282**, 37256-37265.

26. Garrity, J., Gardner, J. G., Hawse, W., Wolberger, C., and Escalante-Semerena, J. C. (2007) *J. Biol. Chem.* **282**, 30239-30245.

27. Liu, B., Lin, Y., Darwanto, A., Song, X., Xu, G., and Zhang, K. (2009) *J. Biol. Chem.* **284**, 32288-32295.

28. Cheng, Z., Tang, Y., Chen, Y., Kim, S., Liu, H., Li, S. S. C., Gu, W., and Zhao, Y. (2009) *Mol. Cell. Proteomics.* **8**, 45-52.

29. Bheda, P., Wang, J. T., Escalante-Semerena, J. C., and Wolberger, C. (2010) *Protein Sci.* **20**, 131-139.

30. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., Falck, J. R., Peng, J., Gu, W., and Zhao, Y. (2007) *Mol. Cell. Proteomics.* **6**, 812-819.

31. Vaquero, A., Scher, M., Lee, D., Erdjument-Gromage, H., Tempst, P. and Reinberg, D. (2004) *Mol. Cell* **16**, 93-105.

32. Black, J. C., Mosley, A., Kitada, T., Washburn, M. and Caery, M. (2008) *Mol. Cell* **32**, 449-455.

33. Wang, F. and Tong, Q. (2009) *Mol. Biol. Cell* **20**, 801-808.

34. Yang, Y., Hou, H., Haller, E. M., Nicosia, S. V. and Bai, W. (2005) *EMBO* **24**, 1021-1032.

35. Celic, I., Masumoto, H., Griffith, W. P., Meluh, P., Cotter, R. J., Boeke, J. D. amd Verreault, A. (2006) *Curr. Biol.* **16**, 1280-1289.

36. Jeong, J., Juhn, K., Lee, H., Kim, S., Min, B., Lee, K., Cho, M., Park, G. and Lee, K. (2007) *Exp. Mol. Med.* **39**, 8-13.

37. Mostoslavsky, R., Chua, K. F., Lombard, D. B., Pang, W. W., Fischer, M. R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M. M., Mills, K. D., Patel, P., Hsu, J. T., Hong, A. L., Ford, E., Cheng, H., Kennedy, C., Nunez, N., Bronson, R., Frendewey, D., Auerbach, W., Valenzuela, D., Karow, M., Hottiger, M. O., Hursting, S., Barrett, J. C., Guarente, L., Mulligan, R., Demple, B., Yancopoulos, G. D. and Alt, F. W. (2006) *Cell* **124**, 315-329.

38. Kaidi, A., Weinert, B. T., Choudhary, C. and Jackson, S. P. (2010) *Science* **329**, 1348-1353.

39. Jin, Y., Kim, Y., Kim, D., Baek, K., Kang, B. Y., Yeo, C. and Lee, K. (2008) *Biochem. Biophys. Res. Commun*. **368**, 690-695.

40. Li, S., Banck, M., Mujtaba, S., Zhou M., Sugrue, M. M. and Walsh, M. J. (2010) *PLoS ONE* **5**, e10486.

41. Cheng, H., Mostoslavsky, R., Saito, S. I., Manis, J. P., Gu, Y., Patel, P., Bronson, R., Appella, E., Alt, F. W., and Chua, K. F. (2003) *Proc. Natl. Acad. Sci. U.S.A.* **100**, 10794-10799.

42. Hirschey, M. D., Shimazu, T., Goetzman, E., Jing, E., Schwer, B., Lombard, D. B., Grueter, C. A., Harris, C., Biddinger, S., Ilkayeva, O. R., Stevens, R. D., Li, Y., Saha, A. K., Ruderman, N. B., Bain, J. R., Newgard, C. B., Farese Jr R. V., Alt, F. W., Kahn, C. R. and Verdin, E. (2010) *Nature* **464**, 121-126.

43. Shimazu, T., Hirschey, M. D., Hua, L., dittenhafer-Reed, K. E., Schwer, B., Lombard, D. B., Li, Y., Bunkenborg, J., Alt, F. W., Denu, J. M., Jacobson, M. P., and Verdin, E. (2010) *Cell Metab.* **12**, 654-661.

44. Bao, J., Scott, I., Lu, Z., Pang, L., Dimond, C. C., Gius, D. and Sack, M. N. (2010) *Free Radic Biol. Med.* **49**, 1230-1237.

45. Kim, H., Patel, K., Muldoon-Jacobs, K., Bisht, K. S., Aykin-Burns, N., Pennington, J. D., van der Meer, R., Nguyen, P., Savage, J., Ovens, K. M., Vassiliopolous, O., Ozden, O., Park, H., Singh, K. K., Abdulkadir, S.A., Spitz, D.R., Deng, C. and Gius, D. (2010) *Cancer Cell* **17**, 41-52.

46. Cimen, H., Han, M. J., Yang, Y., Tong, Q., Koc, H. and Koc, E. C. (2010) *Biochemistry* **49**, 304-311.

47. Kendrick, A. A., Choudhury, M., Rahman, S. M., McCurdy, C. E., Friederich, M., Vanhove, J. L., Watson, P. A., Birdsey, N., Bao, J., Gius, D., Sack, M. N., Jing, E., Kahn, C.R., Friedman, J. E. and Jonscher, K. R. (2010) *Biochem. J.* **433**, 505-514.

48. Hirschey, M. D., Shimazu, T., Capra, J. A., Pollard, K. S. and Verdin, E. (2011) *Aging* **3**, 1-7.

49. Yu, T., McIntyre, J. C., Bose, S. C., Hardin, D., Owen, M. C. and McClintock, T. S. (2005) *J. Comp. Neurol.* **483**, 251-262.

50. Qin, W., Yang, T., Ho, L., Zhao, Z., Wang, J., Chen, L., Zhao, W., Thiyagarajan, M., MacGrogan, D., Rodgers, J. T., Puigserver, P., Sodoshima, J., Deng H., Pedrini, S., Gandy, S., Sauve, A. A. and Pasinetti, G. M. (2006) *J. Biol. Chem.* **281**, 21745-21754.

51. Tissenbaum, H. A., and Guarente, L. (2001) *Nature* **410**, 227-230.

52. Anderson, R. M., Bitterman, K. J., Wood, J. G., Medvedik, O., and Sinclair, D. A. (2003) *Nature* **423**, 181-185.

53. Mair, W., Goymer, P., Pletcher, S. D. and Partridge, L. (2003) *Science* **301**, 1731-1733.

54. Austad, S. N. (1989) *Exp. Gerontol.* **24**, 83-92.

55. Weindruch, R. and Walford, R. L. (1982) *Science* **215**, 1415-1418.

56. Yamamoto, H., Schoonjans, K., and Auwerx, J. (2007) *Mol. Endorinol.* **21**, 1745-1755.

57. Schwer, B., Schumacher, B., Lombard, D. B., Xiao, C., Kurtev, M. V., Gao, J., Schneider, J. I., Chai, H., Bronson, R. T., Tsai, L. H., Deng, C. X., and Alt, F. W. (2010) *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21790-21794.

58. Anekonda, T. S., and Reddy, P. H. (2006) *J. of Neurochem.* **96**, 305-313.

59. Green, K. N., Steffan, J. S., Martinez-Coria, H., Sun, X., Schreiber, S. S., Thompson, L. M., and LaFerla, F. M. (2008) *J. of Neurosci.* **28**, 11500-11510.

60. Outeiro, T. F., Kontopoulos, E., Altmann, S. M., Kufareva, I., Strathearn, K. E., Amore, A. M., Volk, C. B., Maxwell, M. M., Rochet, J., McLean, P. J., Young, A. B., Abagyan, R., Feany, M. B., Hyman, B. T., and Kazantsev, A. G. (2007) *Science* **317**, 516-519.

61. Garske, A. L., Smith, B. C., and Denu, J. M. (2007) *ACS Chem. Biol.* **2**, 529-532.

62. Gao, F., Cheng, J., Shi, T., and Yeh, E. T. (2006) *Nat. Cell Biol.* **8**, 1171-1177.

63. Ashraf, N., Zino, S., Macintyre, A., Kingsmore, D., Payne, A. P., Geroge, W. D., and Shiels, P. G. (2006) *Br. J. Cancer* **95**, 1056-1061.

64. De Nigris, F., Cerutti, J., Morelli, C., Califano, D., Chiariotti, L., Viglietto, G., Santelli, G., and Fusco, A. (2002) *Br. J. Cancer* **86**, 917-923.

65. Frye, R. (2002) *Br. J. Cancer* **87**, 1479.

66. Sauve, A. A., Wolberger, C., Schramm, V. L., and Boeke, J. D. (2006) *Annu. Rev. Biochem.* **75**, 435-465.

67. Denu, J. M. (2005) *Curr. Opin. Chem. Biol.* **9**, 431-440.

68. Smith, B. C., Hallows, W. C., and Denu, J. M. (2008) *Chem. Biol.* **15**, 1002-1013.

69. Borra, M. T., Langer, M. R., Slama, J. T., and Denu, J. M. (2004) *Biochemistry* **43**, 9877-9887.

70. Smith, B. C., and Denu, J. M. (2006) *Biochemistry* **45**, 272-282.

71. Smith, B. C., and Denu, J. M. (2007) *Biochemistry* **46**, 14478-14486.

72. Hawse, W. F., Hoff, K. G., Fatkins, D. G., Daines, A., Zubkova, O. V., Schramm, V. L., Zheng, W., and Wolberger, C. (2008) *Structure* **16**, 1368-1377.

73. Jin, L., Wei, W., Jiang, Y., Peng, H., Cai, J., Mao, C, Dai, H., Choy, W., Bemis, J. E., Jirousek, M. R., Milne, J. C., Westphal, C. H., and Perni, R. B. (2009) *J. Biol. Chem.* **284**, 24394-24405.

74. Blander, G., Olejnik, J., Krzymanska-Olejnik, E., McDonagh, T., Haigis, M. Yaffe, M. B. and Guarente, L. (2005) J. *Biol. Chem.* **280**, 9780-9785.

75. Cosgrove, M. S., Bever, K., Avalos, J. L., Muhammad, S., Zhang, X. and Wolberger, C. (2006) *Biochemistry* **45**, 7511-7521.

76. Howitz, K. T., Bitterman, K. J., Cohen, H. Y., Lamming, D. W., Lavu, S., Wood, J. G., Zipkin, R. E., Chung, P., Kisielewski, A., Zhang, L., Scherer, B. and Sinclair. D. A. (2003) *Nature* **425**, 191-196

77. Grozinger, C. M., Chao, E. D., Blackwell, H. E., Moazed, D. and Schreiber, S. L. (2001) *J. Biol. Chem.* **276**, 38837-38843.

78. Trapp, J., Meier, R., Hongwiset, D., Kassack, M. U., Sippl, W. and Jung, M. (2007) *Chem. Med. Chem.* **2**, 1419-1431.

79. Bitterman, K. J., Anderson, R. M., Cohen, H. Y., Latorre-Esteves, M. and Sinclair D. A. (2002) *J. Biol. Chem.* **277**, 45099-45107.

80. Jackson, M. D., Schmidt, M. T., Oppenheimer, N. J. and Denu, J. M. (2003) *J. Biol. Chem.* **278**, 50985-50998.

81. Kaeberlein, M., McDonagh, T., Heltweg, B., Hixon, J., Westman, E. A., Caldwell, S. D., Napper, A., Curtis, R., DiStefano, P. S., Fields, S., Bedalov, A. and Kennedy, B. K. (2005) *J. Biol. Chem.* **280**, 17038-17045.

82. Borra, M. T., Smith, B. C. and Denu, J. M. (2005) *J. Biol. Chem.* **280**, 17187-17195.

83. Nakagawa, T., Lomb, D. J., Haigis, M. C., and Guarente, L. (2009) *Cell* **137**, 560-570.

84. French, J. B., Cen, Y., and Sauve, A. A. (2008) *Biochemistry 47*, 10227-10239.

85. Johnson, D. I. and Pringle, J. R. (1990) *J. Cell Biol.* **111**, 143-152.

86. Shinjo, K., Koland, J. G., Hart, M. J., Narasimhan, V., Johnson, D. I., Evans, T. and Cerione, R. A. (1990) *Proc. Natl. Acad. Sci. U. S. A.* **87**, 9853-9857.

87. Nobes, C. D. and Hall, A. (1995) *Cell* **81**, 53-62.

88. Kozma, R., Ahmed, S., Best, A. and Lim, L. (1995) *Mol. Cell. Biol.* **15**, 1942-1952.

89. Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D. and Hall, A. (1992) *Cell* **70**, 401-410.

90. Ridley, A. J. and Hall, A. (1992) *Cell* **70**, 389-399.

91. Cerione, R. A. (2004) *Trends Cell Biol.* **14**, 127-132.

92. Chardin, P. (1998) *Biochimie* **70**, 865-868.

93. Tapon, N. and Hall, A. (1997) *Curr. Opin. Cell Biol.* **9**, 86-92.

94. Nassar, N., Hoffman, G. R., Manor, D., Clardy, J. C. and Cerione, R. A. (1998) *Nat. Struct. Biol.* **5**, 1047-1052.

95. Hoffman, G. R., Nassar, N. and Cerione, R. A. (2000) *Cell* **100**, 345-356.

96. Worthylake, D. K., Rossman, K. L., and Sondek, J. (2000) *Nature* **408**, 682–688.

97. Wei, Y., Zhang, Y., Derewenda, U., Liu, X., Minor, W., Nakamoto, R. K., Somlyo, A. V., Somlyo, A. P., and Derewenda, Z. S. (1997) *Nat. Struct. Biol.* **4**, 699-703.

98. Williams, C. L. (2003) *Cell. Signal.* **15**, 1071-1080.

99. Sebti, S. M., and Der, C. J. (2003) *Nat. Rev. Cancer* **3**, 945-951.

100. Roberts, P. J., Mitin, N., Keller, P. J., Chenette, E. J., Madigan, J. P., Currin, R. O., Cox, A. D., Wilson, O., Kirschmeier, P., and Der, C. J. (2008) *J. Biol. Chem.* **283**, 25150-25163.

101. Hoffman, G. R., and Cerione, R. A. (2002) *FEBS Lett.* **513**, 85–91.

102. Johnson, J. L., Erickson, J. W., and Cerione, R. A. (2009) *J. Biol. Chem.* **284**, 23860–23871.

103. Hart, M. J., Eva, A., Evans, T., Aaronson, S. A., and Cerione, R. A. (1991) *Nature* **354**, 311–314.

104. Takai, S., Hasegawa, H., Kiyokawa, E., Yamada, K., Kurata, T., and Matsuda, M. (1996) *Genomics* **35**, 403–404.

105. Hasegawa, H., Kiyokawa, E., Tanaka, S., Nagashima, K., Gotoh, N., Shibuya, M., Kurata, T., and Matsuda, M. (1996) *Mol. Cell. Biol.* **16**, 1770–1776.

106. Erickson, J. W., and Cerione, R. A. (2004) *Biochemistry* **43**, 837–842.

107. Srivastava, S. K., Wheelock, R. H., Aaronson, S. A., and Eva, A. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 8868–8872.

108. Worthylake, D. K., Rossman, K. L., and Sondek, J. (2004) *Structure* **12**, 1078–1086.

109. Feng, Q., Albeck, J. G., Cerione, R. A., and Yang, W. (2002) *J. Biol. Chem.* **277**, 5644–5650.

110. Feng, Q., Baird, D., and Cerione, R. A. (2004) *EMBO J.* **23**, 3492–3504.

111. Baird, D., Feng, Q., and Cerione, R. A. (2005) *Curr. Biol.***15**, 1–10.

112. Gao, Y., Xing, J., Streuli, M., Leto, T. L., and Zheng, Y. (2001) *J. Biol. Chem.* **276**, 47530–47541.

113. Zheng, Y., Zangrilli, D., Cerione, R. A., and Eva, A. (1996) *J. Biol. Chem.* **271**, 19017-19020.

114. Lemmon, M. A., and Ferguson, K. M. (2000) *Biochem. J.* **350**, 1-18.

115. Snyder, J. T., Rossman, K. L., Baumeister, M. A., Pruitt, W. M., Siderovski, D. P., Der, C. J., Lemmon, M. A., and Sondek, J. (2001) *J. Biol. Chem.* **276**, 45868-45875.

116. Brugnera, E., Haney, L., Grimsley, C., Lu, M., Walk, S. F., Tosello-Trampont, A. C., Macara, I. G., Madhani, H., Fink, G. R., and Ravichandran, K. S. (2002) *Nat. Cell Biol.* **4**, 574–582.

117. Cote, J. F., and Vuori, K. (2002) *J. Cell Sci.* **115**, 4901–4913.

118. Wu, Y. C., and Horvitz, H. R. (1998) *Nature* **392**, 501–504.

119. Meller, N., Merlot, S., and Guda, C. (2005) *J. Cell Sci.***118**, 4937-4946.

120. Meller, N., Irani-Tehrani, M., Ratnikov, B. I., Paschal, B. M., and Schwartz, M. A. (2004) *J. Boil. Chem.* **279**, 37470-37476.

121. Park, D., Tosello-Trampont, A. C., Elliott, M. R., Lu, M., Haney, L. B., Ma, Z., Klibanov, A. L., Mandell, J. W., and Ravichandran, K. S. (2007) *Nature* **450**, 430–434.

122. Wang, X., Wu, Y. C., Fadok, V. A., Lee, M. C., Gengyo-Ando, K., Cheng, L. C., Ledwich, D., Hsu, P. K., Chen, J. Y., Chou, B. K., Henson, P., Mitani, S., and Xue, D. (2003) *Science* **302**, 1563–1566.

123. Wu, Y. C., Tsai, M. C., Cheng, L. C., Chou, C. J., and Weng, N. Y. (2001) *Dev. Cell* **1**, 491–502.

124. Henson, P. M. (2005) *Curr. Biol.* **15**, R29–R30.

125. Yajnik, V., Paulding, C., Sordella, R., McClatchey, A. I., Saito, M., Wahrer, D. C., Reynolds, P., Bell, D. W., Lake, R., van den Heuvel, S., Settleman, J., and Haber, D. A. (2003) *Cell* **112**, 673–684.

126. Miyamoto, Y., and Yamauchi, J. (2010) *Cell. Signaling* **22**, 175–182.

127. Yamauchi, J., Miyamoto, Y., Chan, J. R., and Tanoue, A. (2008) *J. Cell. Biol.* **181**, 351-365.

128. Cote, J., and Vuori, K. (2007) *Trends Cell Biol.* **17**, 383-393.

129. Cote, J., Motoyama, A. B., Bush, J. A., and Vuori, K. (2005) *Nat. Cell Biol.* **7**, 797-807.

130. Lu, M., Kinchen, J. M., Rossman, K. L. Grimsley, C., Hall, M., Sondek, J., Hengartner, M. O., Yajnik, V., and Ravichandran, S. (2005) *Curr. Biol.* **15**, 371-377.

131. Lin, Q., Yang, W., Baird, D., Feng, Q., and Cerione, R. A. (2006) *J. Biol. Chem.* **281**, 35253-35262.

132. Schmidt, A., and Hall, A. (2002) *Genes Dev.* **16**, 1587-1609.

133. Rossman, K. L., Der, C. J., and Sondek, J. (2005) *Nat. Rev. Mol. Cell Biol.* **6**, 167-180.

134. Kiyokawa, E., Hashimoto, Y., Kurata, T., Sugimura, H., and Matsuda, M. (1998) *J. Biol. Chem.* **273**, 24479-24484.

135. Makino, Y., Tsuda, M., Ichihara, S., Watanabe, T., Sakai, M., Sawa, H., Nagashima, K., Hatakeyama, S., and Tanaka, S. (2006*J. Cell Sci.* **119**, 923-932.

136. Premkumar, L., Bobkov, A. A., Patel, M., Jaroszewski, L., Bankston, L. A., Stec, B., Vuori, K., Cote, J., and Liddington, R. C. (2010) *J. Biol. Chem.* **285**, 13211-13222.

137. Wu, X., Ramachandran, S., Lin, M., Cerione, R. A., and Erickson, J. W. (2011) *Biochemistry* **50**, 1070-1080.

138. Yang, J., Zhang, Z., Roe, S. M., Marshall, C. J., and Barford, D. (2009) *Science* **325**, 1398-1402.

139. Kwofie, M. A., and Skowronski, J. (2008) *J. Biol. Chem.* **283**, 3088–3096.

140. Miyamoto, Y., Yamauchi, J., Sanbe, A., and Tanoue, A. (2007) *Exp. Cell Res.* **313**, 791-804.

141. Kulkarni, K., Yang, J., Zhang, Z., and Barford, D. (2011) *J. Biol. Chem.* **286**, 25341-25351.

142. Watabe-Uchida, M., John, K. A., James, J. A., Newey, S. E., and Van Aelst, L. (2006) *Neuron* **51**, 727–739.

Chapter 2

SIRT5 is an NAD-dependent Demalonylase and Desuccinylase[1]

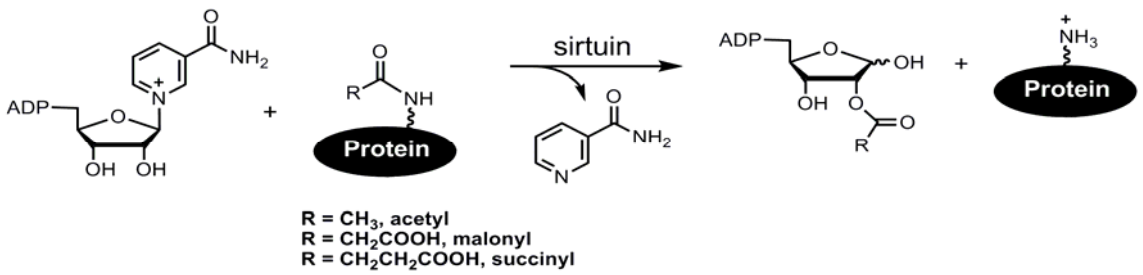
2.1 Introduction

Silent Information Regulator 2 (Sir2) proteins, or sirtuins, are a class of evolutionally conserved enzymes with nicotinamide adenine dinucleotide (NAD)-dependent protein deacetylation activity (Fig. 2.1) and putative ADP-ribosyltransferase activity (1, 2). Since the initial discovery of the deacetylase activity of sirtuins by Guarente and coworkers (3), many important biological functions of sirtuins have been revealed (2), including the regulation of life span (3-8), transcription (9-11), apoptosis and cell survival (12-16), genome stability (17, 18) and metabolism (8,19-25). The NAD-dependent deacetylation mechanism of sirtuins has been studied extensively and the generally accepted mechanism is shown in Fig. 1.1 (26-30). Several crystal structures of sirtuins have been reported which support the proposed enzymatic reaction mechanism (31-33).

There are seven sirtuins in mammals, SIRT1-7. Three of them, SIRT3, SIRT4, and SIRT5, are in mitochondria and possibly regulate metabolic pathways in mitochondria (34). All other sirtuins are either nuclear or cytosolic. Based on sequence similarity, sirtuins from different species can be grouped into different classes (35). Mammalian SIRT1, SIRT2, and SIRT3 belong to Class I, SIRT4 belongs to Class II, SIRT5 belongs to Class III, and SIRT6 and SIRT7 belong to Class IV. Of the seven

[1] This work was published in *Science* **334**, 806-809. AAAS' License to Publish allows me to include this paper in my dissertation. I was a co-first author, responsible for the protein purification, crystallization, X-ray diffraction data collection, and structure determination of all the SIRT5 complexes. Some of the figures have been rearranged for readers' convenience.

Figure 2.1. The sirtuins catalyzed NAD-dependent deacylation reaction. The reaction generates *O*-acyl-ADP-ribose.


ADP
O
NH2
N+
O
OH
OH
+
R
O
NH
Protein
sirtuin
O
NH2
N
ADP
O
OH
OH
O
R
O
+
NH3
Protein
R = CH3, acetyl
R = CH2COOH, malonyl
R = CH2CH2COOH, succinyl

mammalian sirtuins, only Class I sirtuins (SIRT1, SIRT2, and SIRT3) have been shown to have robust deacetylase activity. SIRT4 has no detectable deacetylase activity (23, 34). SIRT5 was reported to have weak deacetylase activity (36) and our own work (see below) further confirmed this. SIRT6 was originally reported to lack deacetylase activity (34, 37), but a recent report from Chua and coworkers suggested that SIRT6 can specifically deacetylate Histone H3 K9 (18). However, the *in vitro* activity is weak (18). No activity has been detected for SIRT7 (34), which localizes in the nucleolus.

What is the biochemical function of sirtuins with little or no deacetylation activity (i.e. Class II, III and IV sirtuins)? It has been suggested that sirtuins lacking deacetylase activity could function as protein ADP-ribosyltransferases (23, 37). We and others have studied the ADP-ribosylation activity of sirtuins and found that it is orders of magnitude slower than the deacetylase activity of Class I sirtuins or the ADP-ribosylation activity of a bacterial ADP-ribosyltransferase, diphtheria toxin (38, 39). Here we present data showing that mammalian SIRT5 is an NAD-dependent demalonylase and desuccinylase *in vitro* (Fig. 2.1) with catalytic efficiencies similar to that of SIRT1-catalyzed deacetylation. Furthermore, we show that protein succinylation and malonylation exist in mammalian proteins. The finding that SIRT5 prefers to hydrolyze succinyl/malonyl lysine residues suggests that other sirtuins that show little or no deacetylase activity may also prefer acyl groups other than acetyl.

2.2 Experimental Procedures

Cloning, Expression and Purification of Human Sirtuins. Human SIRT1-3, 5, and 6 were expressed as previously described (39). Human *SIRT7* coding sequence was PCR-amplified using primers JT072_SIRT7 (1-400) EcoRI5 (5'-AGTCAGGAA TTCATGGCAGCCGGGGGTCT-3') and JT073_SIRT7 (1-400) XhoI3 (5'-AGTCAGCTCGAGT TACGTCACTTTCTTCCTTTTT-3'). Amplified product was digested with EcoRI and XhoI. The digested PCR product was purified and ligated into the similarly digested expression vector pET28a. C-terminal Flag-tagged *SIRT5* (*Flag-SIRT5*) and truncated *SIRT5(34-302)* were cloned using TOPO and Gateway cloning technology (Invitrogen Corp., Carlsbad, CA) into pDEST-F1 for expression. SIRT7, Flag-SIRT5 and SIRT5(34-302) were expressed in *E. coli* and purified as described (39). After purification, Flag-SIRT5 and SIRT5(34-302) were digested by TEV at room temperature for 2 h and purified by HisTrap^{TM} HP Column (GE Healthcare, Piscataway, NJ) and gel filtration on a HiLoad 26/60 Superdex 75 prep grade column (GE Healthcare, Piscataway, NJ). Protein was dialyzed into crystallization buffer (20 mM Tris, pH 8.0, 20 mM NaCl, 5% glycerol), concentrated into 16 mg/mL, flash frozen by liquid nitrogen, and stored at -80 °C for crystallization.

Protein Crystallization. P53K382 acetyl, or thioacetyl peptide, Histone H3K9 acetyl, or thioacetyl, or malonyl, or succinyl, or thiosuccinyl peptide [4- KQTAR (ac, tac, mal, suc**)** STGGKA-15] was used for crystallization. 50 mM peptide stock solutions were made by dissolving peptide powder in double distilled water and neutralized with 8 M NaOH before mixed with SIRT5. SIRT5-peptide mixtures were prepared at a 1:20 protein:peptide molar ratio and incubated for 30~60 min on ice. The final protein concentration was 10 mg/mL. Crystals were grown by the method of hanging drop

vapor diffusion. H3K9 acetyl and malonyl peptide failed to give any crystals under the conditions I've tried. The SIRT5-H3K9 thioacetyl peptide cocrystals were grown in the condition of 20% PEG8000, 0.1 M CHES, pH 9.5 at 18 °C. The SIRT5-H3K9 succinyl peptide co-crystals were grown in the condition of 16% PEG4K, 6% Glycerol at 18 °C, while the SIRT5-H3K9 thiosuccinyl peptide co-crystals were grown in the condition of 30% PEG10k, 0.1 M TRIS, pH8.5 at room temperature.

Data Collection and Structure Determination. 15% Glycerol was added to the mother liquid as the cryoprotectant for SIRT5-H3K9 thioacetyl co-crystals. SIRT5-H3K9 succinyl co-crystals were soaked in the cryoprotectant solution (18% PEG4K, 15% Glycerol) with 10 mM NAD for 2-10 minutes at room temperature immediately before data collection. To obtain the intermediate structure, SIRT5-H3K9 thiosuccinyl co-crystals were soaked in the cryoprotectant solution (30% PEG10k, 0.1 M TRIS, pH8.5, 15% Glycerol) with 10 mM NAD for 2-5 hours at 4 °C, and flash frozen in liquid nitrogen for data collection. All the X-ray diffraction data were collected at CHESS (Cornell High Energy Synchrotron Source) A1 or F1 station. The data were processed using the programs HKL2000 (40). All the structures of SIRT5 complexes were solved by molecular replacement using the program Molrep from the CCP4 suite of programs (41). The SIRT5-ADPR structure (PDB code: 2B4Y) was served as the search model. Refinement and model building were performed with REFMAC5 and COOT from CCP4. The X-ray diffraction data collection and structure refinement statistics are shown in Table 2.1.

Table 2.1. Crystallographic data collection and refinement statistics.

	SIRT5-Thioacetyl H3K9	SIRT5-Succinyl H3K9-NAD
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	52.54, 67.88, 156.75	52.69, 69.41, 156.32
α, β, γ (º)	90, 90, 90	90, 90, 90
Resolution (Å)	50-2.00	50-1.55
R_{sym} or R_{merge} (%)	9.0 (34.4)	6.6 (39.8)
$I/\sigma I$	38.65 (4.23)	42.26 (5.15)
Completeness (%)	99.9 (99.8)	99.7 (98.5)
Redundancy	6.8 (5.3)	7.1 (5.7)
Refinement		
Resolution (Å)	30-2.00	30-1.55
No. reflections	39228	79514
R_{work} / R_{free} (%)	21.52/25.80	14.26/19.09
No. of protein atoms	8170	8170
No. of ligand/ion molecules		
Thioacetyl H3K9	2	---
Succinyl H3K9	---	2
CHES	2	---
NAD	---	2
Zn	2	2
No. of water	187	732
R.m.s deviations		
Bond lengths (Å)	0.0256	0.0230
Bond angles (º)	1.774	2.031

Numbers showed in the parentheses are for the highest resolution shell.

Deacetylation, Demalonylation, and Desuccinylation Activity Assay and Determination of k_{cat} and K_m. The deacylase activity of human SIRT1, SIRT2, SIRT3, SIRT5, SIRT6 and SIRT7 were measured by detecting the deacylated peptide from the acyl peptides using LC-MS. Purified sirtuin was incubated with 0.3 mM acyl peptides, 1.0 mM NAD in 20 mM Tris-HCl buffer (pH 7.5) containing 1 mM DTT in 60 μL reactions for 2 h at 37°C. The reactions were stopped with 60 μL 10% TFA and analyzed by LC-MS.

For determination of k_{cat} and K_m, human SIRT1, SIRT2, SIRT3 and SIRT5 were measured by detecting the deacylated peptide from H3K9 acyl peptides using HPLC. Purified sirtuin was incubated with 1.0 mM of NAD in 20 mM Tris-HCl buffer (pH 7.5) containing acyl peptides (0–750 μM) and 1 mM DTT in 60 μL reactions at 37°C. The reactions were stopped with 100 mM HCl and 160 mM acetic acid, analyzed by HPLC with a reverse phase C18 column (250 × 4.6 mm, 90 A, 10 μm, GraceVydac, Southborough, MA), with a linear gradient of 0% to 20% B for 10 min (1 mL/min). Product quantification was based on the area of absorption monitored at 215 nm, assuming hydrolysis of the acyl group does not affect the absorption. The k_{cat} and K_m values were obtained by curve-fitting the $V_{initial}$/[E] versus [S] plot using KaleidaGraph. For SIRT5 R105M, the observed second order rate constant, k_{obs} (rate/([sirtuin][NAD])) was detected instead of k_{cat} and K_m because of the very weak deacylation activity. The experiments were done in duplicate.

For comparing the deacetylation, demalonylation and desuccinylation activities of SIRT5 on different peptide backbones, histone H3, GDH and ACS2 peptides with two tryptophan residues at the C-terminal were used to allow better detection and quantification on HPLC. The determination of k_{cat} and K_m was carried out essentially

the same as mentioned above with slight modifications. The reactions were quenched with 60 μL 10% TFA. The chromatography gradient was 0% to 50% B for 20 min (1 mL/min). The peptides were detected and quantified on the LC by the absorption at 280 nm.

Purification of O-Ma-ADPR and O-Su-ADPR with HPLC and Analysis by MS. SIRT5 or SIRT1 (1 μM) was incubated with 0.5 mM malonyl or succinyl peptides and 1.0 mM NAD in 20 mM Tris-HCl buffer (pH 7.5) with 1 mM DTT in 60 L reactions for 2 h at 37°C. The reactions were terminated by adding 60 μL 10% TFA. After centrifugation to remove precipitated proteins, the supernatant was analyzed by HPLC using a 50 mM ammonium acetate isocratic system on a Sprite TARGA C18 column (40 × 2.1 mm, 5 μm, Higgins Analytical, Inc.). The product *O*-Ma-ADPR (retention time 1.6 min) and *O*-Su-ADPR (retention time 2.3 min) was collected and the molecular weights were confirmed by MALDI-MS. The ADPR (retention time 1.2 min) and NAD (retention time 5 min) have also been confirmed by MALDI-MS.

O-Su-ADPR generated from bovine liver mitochondrial peptide mixtures was purified as above and analyzed using an Agilent 1100 high-performance liquid chromatographer coupled to an ABI 4000 Q-trap mass spectrometer operating in IDA negative ion mode. Chromatography consisted of an HILIC column (Nest Group, 100 Å, 5 μm, polyhydroxyethyl A, 1 × 150 mm) eluted with a gradient of acetonitrile versus 10 mM ammonium acetate at 0.05 ml/min. This extra LC step was needed because the sample was more complicated than the reactions using only synthetic peptides.

Detection of Succinyl Lysine from Bovine Liver Mitochondrial Proteins Using the ^{32}*P-NAD Assay.* Bovine liver mitochondria was isolated as previously described (42).

Mitochondria from 5 g bovine liver was lysed for 30 min at 4°C in ice-cold lysis buffer (25 mM Tris-HCl pH 8.0, 50 mM NaCl, 0.1% Triton X-100) containing protease cocktail inhibitor (P8340, Sigma). The supernatants were collected and exchanged to 25 mM Tris-HCl (pH 8.0) with 50 mM NaCl using centrifugal filter (MILLIPORE, Billerica, MA) to remove endogenous NAD. The extracts were stored at -80°C. For trypsin digestion, 1.5 mg of the bovine liver mitochondria proteins or BSA (used as the control) was dissolved in 6 M urea, 60 mM Tris-HCl (pH 8.0), 15 mM DTT in a reaction volume of 450 μL. The solution was heated at 95°C for 15 min and then cooled to room temperature. Then 22.5 μL of 1M iodoacetamide (final concentration ~50 mM) was added and the mixture was incubated at room temperature with gentle mixing for 1 h. Then 3.6 mL of 50 mM Tris-HCl (pH 7.4) with 1 mM $CaCl_2$ was added to the reaction mixture to lower the urea concentration to 0.75 M. Finally, 150 μL of 100 μg/mL modified trypsin (Promega Corporation, Madison, WI) were added and the reaction mixture was incubated at 37 °C for 12 h. After quenching the reaction by adding 65 μL 10% TFA to pH 2~3, the digested peptides were desalted by using Sep-Pak C18 cartridge 1cc/50 mg (Waters Corporation, Milford, MA) and lyophilized.

To detect the acyl-ADPR compounds formed in sirtuin-catalyzed deacylation reactions, reactions were performed in 10 μL solutions with 1 μCi ^{32}P-NAD (ARC Inc., ARP 0141, 800Ci/mmol, 0.125μM), 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM DTT. The acyl peptide substrates used were 100 μM H3K9 acetyl, malonyl, or succinyl peptide, 2 μg calf thymus histones (Roche Applied Science, Indianapolis, IN), 20 μg bovine liver mitochondrial peptides, or 20 μg BSA peptides. The reactions

were incubated with 1 μM SIRT5 or SIRT1 at 37°C for 1 h. CD38 catalytic domain was used to generate ADPR as a control. A total of 0.5 μL of each reaction were spotted onto silica gel TLC plates and developed with 7:3 ethanol:ammonium bicarbonate (1 M aqueous solution). After development, the plates were air-dried and exposed to a PhosphorImaging screen (GE Healthcare, Piscataway, NJ). The signal was detected using a STORM860 phosphorimager (GE Healthcare, Piscataway, NJ).

Affinity Purification of Lysine-succinyl Peptides and Protein Identification. Flag-SIRT5 (25 μg) was bound onto 100 μL anti-Flag M2 affinity gel (A2220, Sigma) by incubation at 4°C in NETN buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA, 0.5% NP40) for 2 hr (43). The supernatant was removed and the gel was washed three times with NETN buffer. The tryptic mitochondria peptides (~1 mg) obtained above were resolubilized in 0.5 mL NETN buffer and insoluble particles were removed by centrifugation at 10,000 × g for 10 min. Affinity purification was carried out by incubating the peptides with Flag-SIRT5 bound anti-Flag M2 affinity gel at 4°C for 3 h with gentle shaking. The gel was washed three times with 1 mL of NETN buffer and twice with ETN buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, 1 mM EDTA). The bound peptides were eluted three times with 100 μL of 0.1% TFA. The elutions were combined and lyophilized. The resulting peptides were cleaned using C18 ZipTips (Millipore, Bedford, Massachusetts) according to the manufacture's instructions, prior to LC-MS/MS analysis performed at the Proteomic and MS Facility of Cornell University. Tandem mass spectra were searched against NCBI-nr database with MASCOT search engine (Matrix Science, London, UK) using acetyl, malonyl, and succinyl lysine as modifications.

Detection of Lysine Succinylation on CPS1 Peptides Using the ^{32}P-NAD Assay. The CPS1 band was cut from a SDS-PAGE gel of the bovine liver mitochondria lysate. The protein was in-gel digested with trypsin and extracted and desalted as the following. The gel band was washed in 100 μL water for 5 min, followed by 100 μL 100 mM Ammonium bicarbonate : acetonitrile (1:1) for 10 min and finally 50 μL acetonitrile for 5 min. The acetonitrile was then discarded and the gel band was allowed to dry in the ventilated fume hood for 5-10 min. The gel slice was then rehydrated with 15 μL trypsin solution (10 μg/mL modified trypsin in 1mM HCl) on ice for 30 min. The trypsin solution was topped with 10 μL 50 mM Ammonium bicarbonate with 10% acetonitrile. The digestion reactions were kept at 30°C for 12 h. The resulting solution was acidified with formic acid (1% in final). The trypsin digested peptides were extracted twice with 30 μl of 50% acetonitrile with 0.2% TFA (45 min incubation at room temperature followed by 5 min sonication). The third extraction was done with 30 μl of 90% acetonitrile with 0.2% TFA (5 min). All the extracts were combined and lyophilized. When dried, the peptides were dissolved in 12 μl of 0.1% TFA and desalted by ZipTips (Millipore, Bedford, Massachusetts). The desalted peptides were lyophilized again and reconstituted in water. The GDH peptides from in-gel digestion and the histone peptides from in-solution digestion were prepared as described above. The ^{32}P-NAD assays were carried out as described above. To lower the detection limit, higher concentrations of sirtuins were used. The SIRT5 was used at a final concentration of 52 μM and the Hst2 was used at 24 μM. The sample peptides were used at a concentration of 0.3 μg/μl and the control peptides were at 20 μM.

Generation of SIRT5 Deficient Mouse Line. SIRT5 +/+ and -/- mice were generated at the Institut Clinique de la Souris (Strasbourg, France). Briefly, exon 4 of SIRT5 locus was flanked with loxP sites using standard genetic engineering and gene targeting procedures. The resulting SIRT5 floxed mice were bred with CMV-Cre transgenic mice to generate germline SIRT5 deficient (SIRT5 -/-) mice and control SIRT5 +/+ mice. The absence of SIRT5 mRNA in different tissues of SIRT5 -/- mice was confirmed by Q-RT-PCR analysis and the loss of SIRT5 protein expression was verified by western blot using an anti-SIRT5 antibody (Abcam ab62740).

Detection of Succinylation Level on CPS1 from SIRT5 wt and KO Mouse Livers Using the ^{32}P-NAD Assay. The 29-week-old male SIRT5 +/+ and -/- littermates were fasted overnight (from 6:00 PM to 10:00 AM) and then provided with free access to food for four hours prior to sacrifice. Liver tissues were rapidly removed, snap-frozen with liquid nitrogen, and stored at -80°C for analysis. The liver samples were first broken into small pieces, and then homogenized in 1mL of the lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1% NP-40, 10% glycerol). The crude lysates were incubated at 4°C on shaker for 30 min then centrifuged at 10 000 g, 4°C for 10 min. The concentration of the lysate was determined by Bradford assay.

The lysate with total protein of 200 μg was incubated with 3 μl of the CPS1 antibody (Abcam ab3682) at 4°C for 60min. 40 μl of the Protein A/G agarose beads (Santa Cruz Biotechnology sc-2003) was then added and incubated at 4°C overnight. The beads were washed 3 times in Tris buffer (25 mM Tris-HCl pH 8.0, 50 mM NaCl, 0.1% NP-40). The beads with the immunopurified CPS1 (0.5-1.0 μg) was incubated with 1 μCi ^{32}P-NAD and 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 10 mM DTT. The SIRT5 was used at the concentration of 10 μM and Hst2 was used at 1 μM. The

control peptides were used all at 20 μM concentration. The reaction was incubated at 37°C for 1 h. 1.8 μL sample was loaded onto the TLC plate. Separation by TLC and detection by autoradiography was performed as described above.

Identification of Acyl Lysine Residues from Commercial Proteins by LC-MS/MS. Commercial GDH (Sigma G2626), malate dehydrogenase (Sigma M2634), citrate synthase (Sigma C3260), and pyruvate dehydrogenase (Sigma P7032) were in-solution digested with trypsin. Typically, 1.0 mg of the protein was dissolved in 6 M guanidine hydrochloride, 150 mM Tris-HCl (pH 8.0), 15 mM DTT in a reaction volume of 400 μL. The solution was incubated at room temperature for 60 min. Then 20 μL of 1M iodoacetamide (final concentration ~50 mM) was added and the mixture was incubated at room temperature with gentle mixing for another 60 min. The excess iodoacetamide was quenched with 40mM DTT. Then 3.6 mL of 50 mM Tris-HCl (pH 7.4) with 1 mM $CaCl_2$ was added to the reaction mixture to lower the guanidine hydrochloride concentration to 0.6 M. Finally, 100 μL of 100 μg/mL modified trypsin (Promega Corporation, Madison, WI) were added and the reaction mixture was incubated at 37 °C for 12 h. After quenching the reaction by adding 65 μL 10% TFA to pH 2~3, the digested peptides were desalted by using Sep-Pak C18 cartridge 1cc/50 mg (Waters Corporation, Milford, MA) and lyophilized.

Identification of Protein Acylation by nanoLC/MS/MS Analyses. The tryptic digest was reconstituted in 50 μL of 2% acetonitrile with 0.5% formic acid. About 200 ng of tryptic digest were injected for nanoLC-ESI-MS/MS analysis. The initial analysis was performed in UltiMatePlus nanoLC (Dionex, Sunnyvale, CA) coupled with to a hybrid triple quadrupole linear ion trap mass spectrometer, 4000 Q Trap from ABI/MDS Sciex (Framingham, MA) equipped with Micro Ion Spray Head II ion source.

Additional analysis was carried out using UltiMate3000 nanoLC (Dionex, Sunnyvale, CA) coupled with a LTQ Orbitrap Velos (Thermo-Fisher Scientific, San Jose, CA) mass spectrometer equipped with "Plug and Play" nano ion source device (CorSolutions LLC, Ithaca, NY).

Tryptic peptides were separated on a PepMap C-18 RP nano column (5 μm, 75 μm i.d. × 150 mm, Dionex), eluted in a 60 to 90-minute gradient of 10% to 40% acetonitrile in 0.1% formic acid at 300 nL/min. For 4000 Q Trap analysis, MS data acquisition was acquired using Analyst 1.4.2 software (AB SCIEX, Framingham, MA) in the positive ion mode for information dependant acquisition (IDA) analysis with nanospray voltage at 1.85 kV and heated interface at 150°C used for all experiments. In IDA analysis, after each survey and an enhanced resolution scan, three highest intensity ions with multiple charge states were selected for tandem MS (MS/MS) with rolling collision energy applied for detected ions based on different charge states and *m/z* values. For Orbitrap analysis, each of tryptic digests was acquired at both CID-based parallel data-dependent acquisition (DDA) mode where FT mass analyzer was used for one survey MS scan followed by MS/MS scans on top 7 most intensity peaks with multiple charged ions above a threshold ion count of 5000 in LTQ mass analyzer, and HCD-based DDA mode where one FT survey scan was followed by 5 MS/MS scan in FT analyzer. MS survey scans and MS/MS scans were acquired at a resolution of 60,000 and 7,500 (fwhm at *m*/*z* 400) respectively. The Orbitrap Velos is operated in positive ion mode with nano spray voltage set at 1.5 kV and source temperature at 225 °C. Either internal calibration using the background ion signal at m/z 445.120025 as a lock mass or external calibration using Ultramark 1621 for FT mass analyzer is performed. The normalized collision energy for CID and HCD were set at 35 % and

38% respectively. All data are acquired under Xcalibur 2.1 operation software (Thermo-Fisher Scientific).

The MS/MS data generated from both 4000 Q Trap and LTQ Orbitrap Velos were submitted to Mascot 2.3 (Matrix Science, Boston, MA) for database searching using in-house licensed Mascot local server and the search was performed to query to SwissProt database (taxonomy: mammal) with three missed cleavage sites by trypsin allowed. The peptide tolerance was set to 1.2 Da and MS/MS tolerance was set to 0.6 Da for the data from 4000 Q Trap. For Orbitrap data, the peptide tolerance was set to 10 ppm and MS/MS tolerance was set to 0.6 Da (for MS/MS spectra acquired by LTQ) and 0.05 Da (MS/MS spectra acquired by Orbitrap FT mass analyzer). A fixed carbamidomethyl modification of cysteine and several variable modifications on lysine acetylation, malonylation, succinylation and carbamylation, deamidation of asparagine and glutamine, and methionine oxidation were applied. Only significant scores for the peptides defined by Mascot probability analysis (www.matrixscience.com/help/scoring_help.html#PBM) greater than "identity" were considered for the peptide identification and modification site determinations. All MS/MS spectra for the identified peptides with acylation modifications were manually inspected and validated.

2.3 Results

To check whether the lack of robust deacetylase activity for SIRT4-7 is due to strict requirement for the peptide sequence, the activities of six human sirtuins (all except SIRT4, which could not be expressed in soluble forms in *E. coli*) were monitored using 16 different acetyl peptides derived from physiological substrates of sirtuins. Under the experimental conditions used, SIRT1-3 and 5 showed deacetylase activity, but SIRT6 and SIRT7 did not. All 16 peptides could be deacetylated by SIRT1-3, while only eight could be deacetylated slowly by SIRT5 (data not shown). A histone H3K9 acetyl peptide (residue 4-15, H_2N-KQTAR(Kac)STGGKA-COOH) was one of the best substrates for SIRT1-3 and SIRT5. With this peptide, the k_{cat} and K_m of different sirtuins were determined (Table 2.2). The catalytic efficiency (k_{cat}/K_m) of SIRT5 was about 500-fold lower than that of SIRT1.

To understand why the deacetylase activity of SIRT5 is weak, a crystal structure of SIRT5 in complex with a thioacetyl peptide was obtained (the corresponding acetyl peptide could not be crystallized with SIRT5). A buffer molecule, CHES (*N*-cyclohexyl-2-aminoethanesulfonic acid), was also bound to SIRT5 (Fig. 2.2A). The interactions between the thioacetyl peptide and SIRT5 involved mostly backbone hydrogen bonding (Fig. 2.2B), similar to what was observed for *Thermotoga maritime* Sir2 (Sir2Tm), a sirtuin with robust deacetylase activity (44). Thus, it appears that the exact acetyl peptide sequence may not be crucial for recognition by SIRT5.

When the SIRT5 structure was superimposed with the structure of Sir2Tm in complex with an acetyl peptide and NAD (PDB 2h4f) (31) (Fig. 2.3B), the positions of the thioacetyl lysine in SIRT5 and the acetyl lysine in Sir2Tm were almost identical in

Table 2.2. The kinetic parameters of four human sirtuins on H3 K9 acetyl peptide.

Sirtuins	k_{cat} (s^{-1})	K_m for acetyl peptide (μM)	k_{cat}/K_m ($s^{-1}M^{-1}$)
SIRT1	0.039 ± 0.001	38 ± 4	1.0×10^3
SIRT2	0.030 ± 0.001	190 ± 14	1.6×10^2
SIRT3	0.012 ± 0.001	50 ± 9	2.4×10^2
SIRT5	ND*	ND (>750)*	2.0

* The k_{cat} and K_m values for SIRT5 could not be determined because the V versus [S] plot was linear (K_m was greater than the highest substrate concentration tested). Thus only k_{cat}/K_m value can be obtained.

The k_{cat} and K_m values were obtained by curve-fitting the $V_{initial}/[E]$ versus [S] plot using KaleidaGrap.

Figure 2.2. The complex structure of SIRT5-tacH3K9-CHES. (A) Overall structure of SIRT5 in complex with an H3K9 thioacetyl peptide. The thioacetyl lysine side chain (green) and the CHES molecule (cyan) are shown in stick (cyan). (B) H-bonding interactions (dashed yellow lines) between thioacetyl peptide (green) and SIRT5 (light grey).

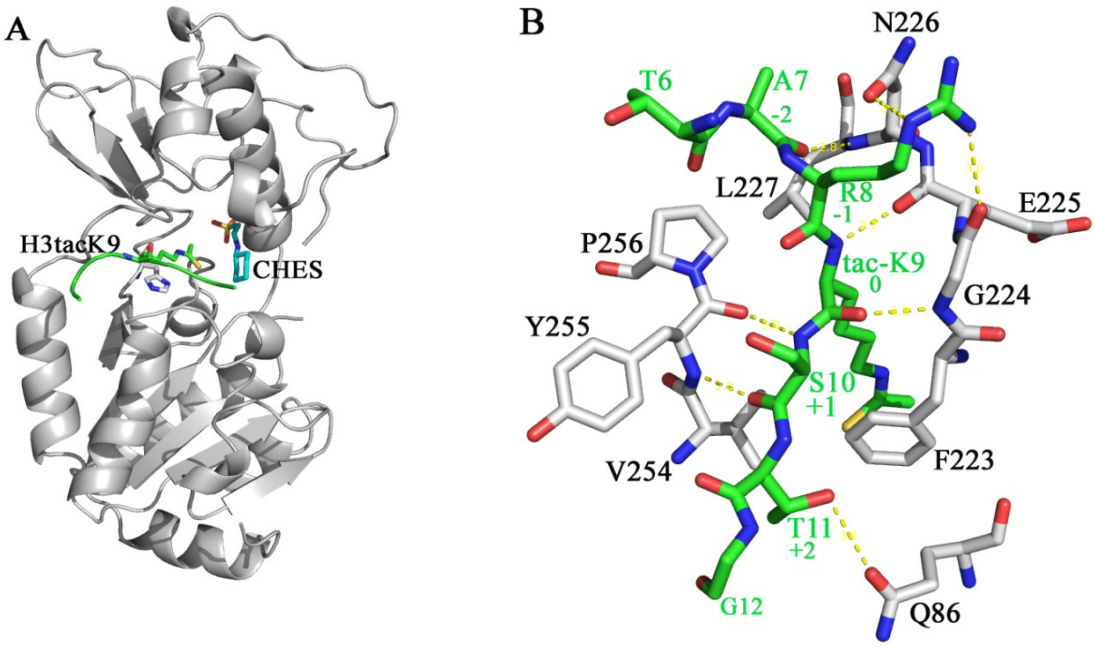
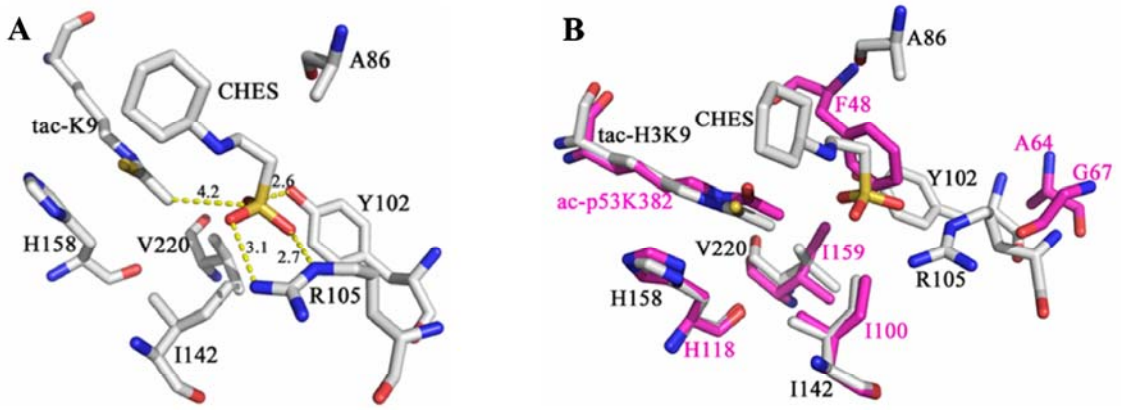
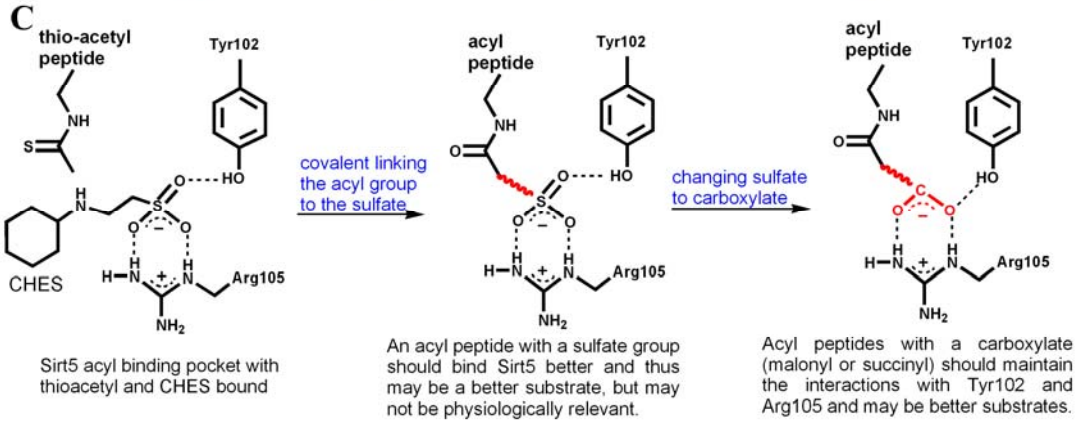

A
H3tacK9
CHES
B
N226
T6
A7
-2
L227
R8
-1
E225
P256
tac-K9
0
G224
Y255
S10
+1
V254
F223
T11
+2
G12
Q86

Figure 2.3. The structure of SIRT5 revealed an unusual acyl pocket. (A) The acyl pocket of SIRT5 was partially occupied by the sulfate from the CHES molecule via interactions with Arg105 and Tyr102. The sulfur was 4.2 Å away from the thioacetyl group. (B) Alignment of SIRT5-thioacetyl peptide structure (grey) and Sir2Tm-acetyl peptide structure (PDB 2h4f, magenta). (C) The rationale for predicting that malonyl/succinyl peptides could be better substrates for SIRT5. (D) SIRT5-succinyl peptide-NAD ternary structure showing that the succinyl group interacted with Tyr102 and Arg105.





Sirt5 acyl binding pocket with thioacetyl and CHES bound

An acyl peptide with a sulfate group should bind Sirt5 better and thus may be a better substrate, but may not be physiologically relevant.

Acyl peptides with a carboxylate (malonyl or succinyl) should maintain the interactions with Tyr102 and Arg105 and may be better substrates.

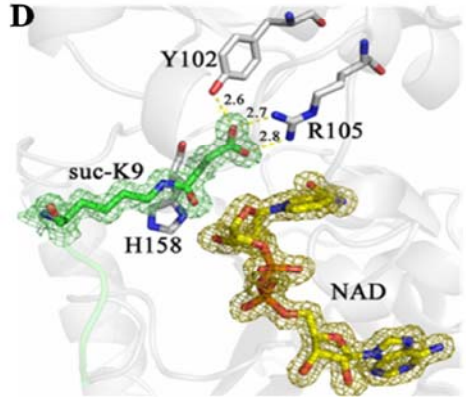
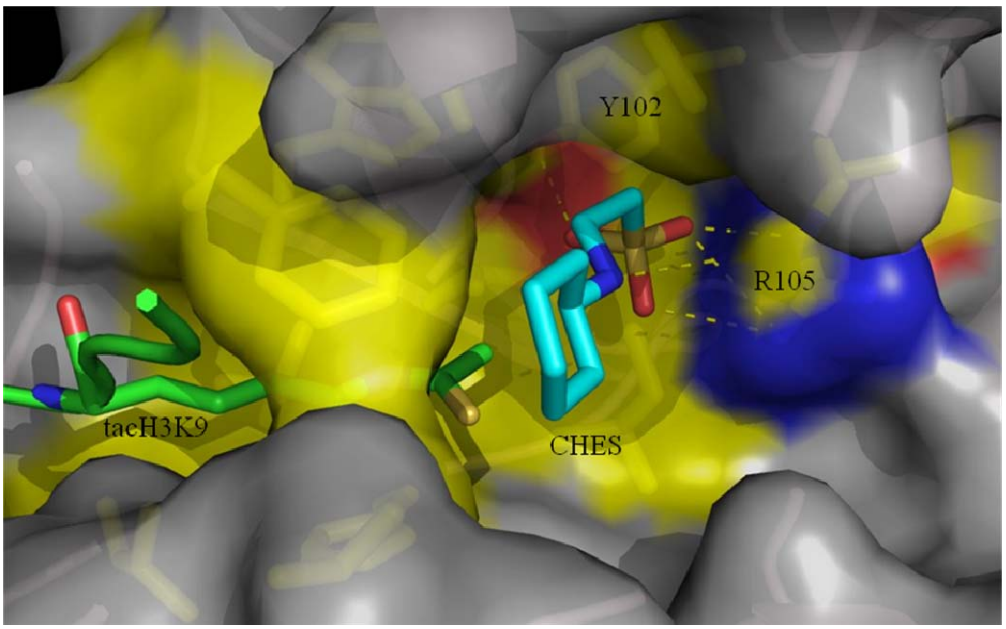


Figure 2.4. The thioacetyl lysine binding tunnel of SIRT5. SIRT5 was bound with H3K9 tac peptide (green) and a buffer molecule CHES (cyan) which interacted with two hydrophilic residues: Y102 (red) and R105 (blue), at the bottom of the hydrophobic lysine binding tunnel (yellow).


Y102
R105
tacH3K9
CHES

the superimposed structures. The acetyl group in the Sir2Tm structure was surrounded by three hydrophobic residues, Phe48, Ile100, and Ile159. In contrast, the thioacetyl lysine in the SIRT5 structure entered into a hydrophobic tunnel with two hydrophilic residues, Y102 and R105, at the bottom (Fig. 2.4). There was extra space between thioacetyl group and the bottom of the tunnel, which was occupied by CHES. CHES formed a hydrogen bond with the polar residue Y102 and salt bridges with positively charged residue R105 (Fig. 2.3A and 2.4). The distance between the thioacetyl group and the sulfur from CHES was about 4.2 Å which equals to the length of 2-3 single carbon-carbon bonds (Fig. 2.3A). All those structural features of SIRT5-H3K9 thioacetyl-CHES complex led us to the rationale presenting in Fig. 2.3C. If the acetyl group was elongated to a propoinyl or butyryl group, the longer chain would occupy the extra space which would make the longer chain bind tighter to SIRT5. Since the sulfate from CHES made contacts with residue Y102 and R105, if the modified lysine contained a sulfate group at the end, it should be a better substrate than acetyl lysine. However, acyl chain bearing a sulfate group at the end was physiologically irrelevant. Substituting of sulfate group with carboxylate will make a propionyl or a butyryl group into a malonyl or succinyl group, respectively, both of which are physiologically relevant. We predicted that malonyl or succinyl lysine peptide would be better substrates for SIRT5.

In cells, the most common acyl-CoA molecules with a carboxylate group are malonyl-CoA and succinyl-CoA (45). Malonyl-CoA, made from acetyl-CoA by acetyl-CoA carboxylase in the cytosol and the mitochondria, is a precursor for fatty acid biosynthesis (46, 47). Succinyl-CoA is an intermediate in the Krebs cycle in the mitochondria. Since acetyl-CoA is used to modify proteins in cells, it is possible that

malonyl- and succinyl-CoA could also be used to modify proteins. Thus, H3K9 malonyl and succinyl peptides were synthesized and tested for hydrolysis by SIRT5.

Liquid chromatography–mass spectrometry (LC-MS) was used to monitor the reactions. SIRT1, 2, and 3 catalyzed the hydrolysis of the acetyl peptide, but not the malonyl or succinyl peptide (Fig. 2.5A-I). In contrast, with SIRT5, little hydrolysis of the acetyl peptide was observed, but the malonyl and succinyl peptides were hydrolyzed significantly (Fig. 2.5J-L). SIRT6 and SIRT7 had no detectable activity on the acyl peptides under the conditions tested (Fig. 2.5M-R). Thus, SIRT5 is a desuccinylase and demalonylase.

What is the mechanism of SIRT5's demalonylation and desuccinylation? If SIRT5 uses the same mechanism to catalyze demalonylation and desuccinylation, *O*-malonyl-ADPR (*O*-Ma-ADPR) or *O*-succinyl-ADPR (*O*-Su-ADPR) should be produced. These products were indeed detected by mass spectroscopy (data not shown). The formation of *O*-Ma-ADPR and *O*-Su-ADPR was also detected using ^{32}P-NAD (Fig. 2.6A). Thus, the mechanism for SIRT5-catalyzed desuccinylation or demalonylation is similar to the deacetylation mechanism of Class I sirtuins.

The k_{cat} and K_m values for SIRT5-catalyzed deacetylation, demalonylation, and desuccinylation were determined with three different peptide sequences (Table 2.3). With all three peptide sequences, the catalytic efficiencies for demalonylation and desuccinylation were much (29 to >1000 fold) higher than that for deacetylation (Table 2.3).

Figure 2.5. Among all human sirtuins tested, only SIRT5 catalyzed lysine demalonylation and desuccinylation *in vitro*. The enzymatic reactions were analyzed by LC-MS. Pink traces showed the ion intensities (10x magnified) for the masses of the acyl peptides (acetyl, m/z 1274.0; malonyl, m/z 1318.0; succinyl, m/z 1332.0) and blue traces showed the ion intensities (10x magnified) for the mass of the deacylated peptide (m/z 1232.0). Black traces showed the ion intensity for all masses from 100-2000 (total ion counts or TIC). With SIRT1, hydrolysis was observed for the acetyl peptide (A), but not malonyl (B) or succinyl (C) peptide. With SIRT5, hydrolysis of the acetyl peptide was barely detectable (J) while hydrolysis of the malonyl (K) and succinyl (L) peptides were obvious. SIRT2 and SIRT3 only displayed deacetylation activity (D, G), while no activity was detected for SIRT6 and SIRT7 (M-R).

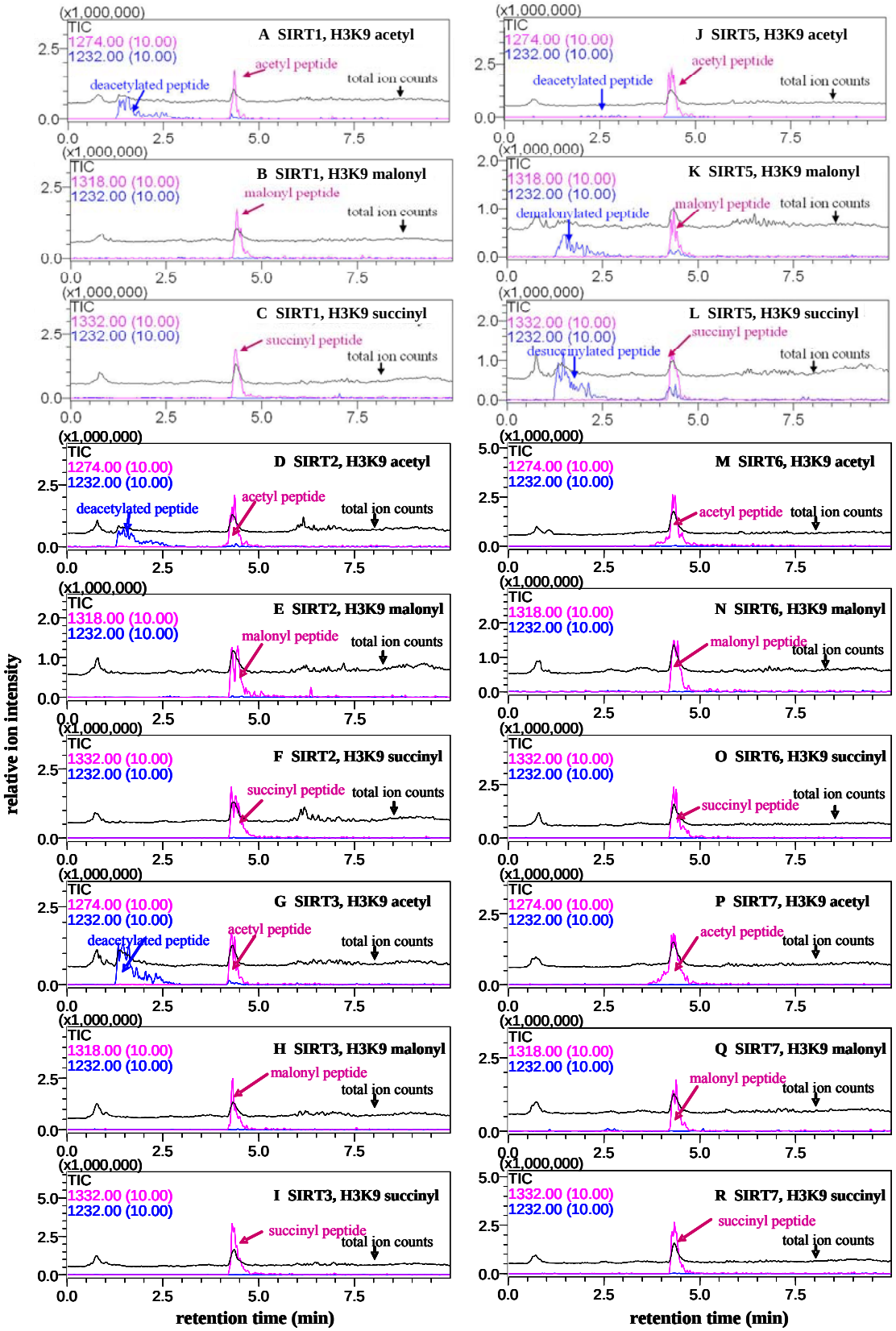

A SIRT1, H3K9 acetyl
B SIRT1, H3K9 malonyl
C SIRT1, H3K9 succinyl
D SIRT2, H3K9 acetyl
E SIRT2, H3K9 malonyl
F SIRT2, H3K9 succinyl
G SIRT3, H3K9 acetyl
H SIRT3, H3K9 malonyl
I SIRT3, H3K9 succinyl
J SIRT5, H3K9 acetyl
K SIRT5, H3K9 malonyl
L SIRT5, H3K9 succinyl
M SIRT6, H3K9 acetyl
N SIRT6, H3K9 malonyl
O SIRT6, H3K9 succinyl
P SIRT7, H3K9 acetyl
Q SIRT7, H3K9 malonyl
R SIRT7, H3K9 succinyl
TIC
1274.00 (10.00)
1318.00 (10.00)
1332.00 (10.00)
1232.00 (10.00)
deacetylated peptide
acetyl peptide
demalonylated peptide
malonyl peptide
desuccinylated peptide
succinyl peptide
total ion counts
(x1,000,000)
relative ion intensity
retention time (min)

Table 2.3. The kinetic parameters of SIRT5 on acetyl, malonyl, and succinyl peptides with different sequences.

Peptide		k_{cat} (s^{-1})	K_m **for peptide** (μM)	k_{cat}/K_m ($s^{-1}M^{-1}$)
H3 K9 (KQTAR**K**STGGWW***)**	deacetylation	ND	ND (>750)	7.8
	demalonylation	0.037 ± 0.003	6.1 ± 2.8	6.1×10^3
	desuccinylation	0.025 ± 0.002	5.8 ± 2.7	4.3×10^3
GDH K503 (SGASE**K**DIVHSGWW***)**	deacetylation	ND	ND (>750)	<2**
	demalonylation	0.014 ± 0.001	8.7 ± 1.3	1.6×10^3
	desuccinylation	0.028 ± 0.002	14 ± 4	2.0×10^3
ACS2 K628 (KTRSGKVMRRWW***)**	deacetylation	ND	ND (>750)	18
	demalonylation	0.079 ± 0.008	150 ± 40	5.2×10^2
	desuccinylation	0.268 ± 0.051	450 ± 150	6.0×10^2

* Two tryptophan residues were added at the C-terminal of the peptide to facilitate the detection by UV-Vis absorption during the HPLC assay.

** No activity. The value was estimated based on the detection limit.

ND: cannot be determined either because no activity was observed or because V versus [S] was linear (even the highest [S] of 750 μM used was much smaller than the K_m).

A crystal structure of SIRT5 in complex with a succinyl peptide and NAD was obtained (Table 2.1). The structure (Fig. 2.3D) showed that the carboxylate from succinyl interacted with Tyr102 and Arg105, consistent with what was predicted based on the structure of SIRT5 with CHES bound (Fig. 2.3C). Changing Arg105 to Met or Tyr102 to Phe significantly increased the K_m for desuccinylation (Table 2.4). Thus, Tyr102 and Arg105 were important for binding succinyl and malonyl groups.

We next determined whether lysine malonylation or succinylation exists *in vivo*. Lysine succinylation was reported to occur on *E. coli* homoserine trans-succinylase (48), but lysine malonylation has not been reported. We thought that detecting SIRT5-catalyzed formation of *O*-Ma-ADPR or *O*-Su-ADPR using ^{32}P-NAD could be a sensitive assay to detect the presence of malonyl or succinyl lysine. H3K9 acetyl, malonyl, and succinyl peptides were incubated with SIRT5 or SIRT1 in the presence of ^{32}P-NAD. The small molecule products were then separated by thin-layer chromatograph (TLC) and detected by autoradiography. With malonyl and succinyl peptides, SIRT5 consumed all the NAD molecules and new spots that corresponded to *O*-Ma-ADPR and *O*-Su-ADPR appeared (Fig. 2.6A). With the acetyl peptide, essentially no NAD was consumed by SIRT5 (Fig. 2.6A). Incubation of the acetyl peptide or calf thymus histones with SIRT1 produced *O*-Ac-ADPR (Fig. 2.6A). The *O*-Su-ADPR spot was separated from the *O*-Ac-ADPR and *O*-Ma-ADPR spots. Thus, SIRT5-catalyzed formation of *O*-Su-ADPR from the hydrolysis of succinyl peptides could be detected using ^{32}P-NAD.

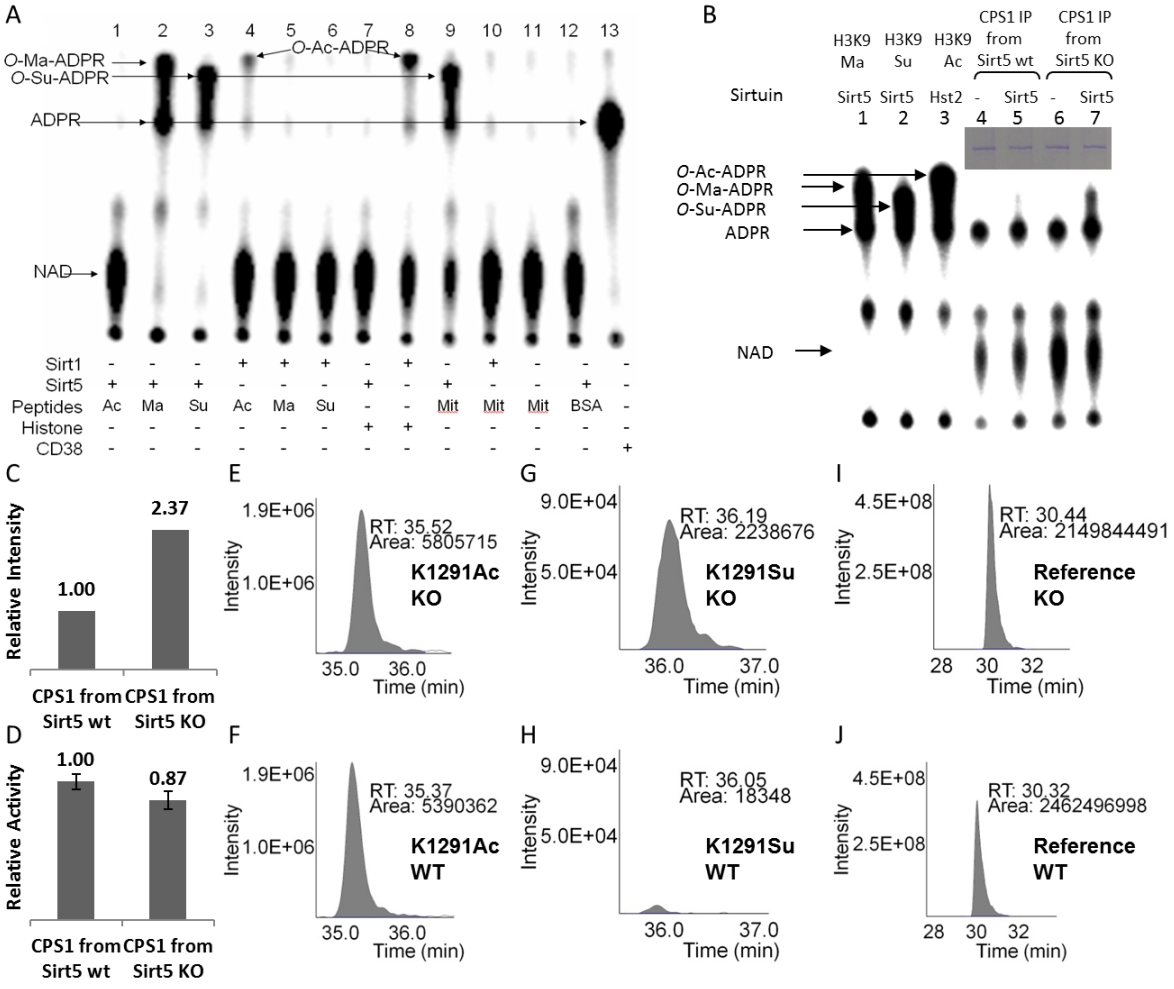
The ^{32}P-NAD assay was then used to detect whether succinyl lysine was present in bovine liver mitochondrial proteins because SIRT5 is localized in mitochondria (7). When bovine liver mitochondrial peptides were treated with SIRT5 and ^{32}P-NAD, the

Table 2.4. The kinetic parameters of mutant SIRT5 on H3K9 acetyl and succinyl peptides.

		k_{cat} (s^{-1})	K_m **for peptide** (μM)	k_{cat}/K_m ($s^{-1}M^{-1}$)
SIRT5	**deacetylation**	ND*	ND (> 750)*	2
	desuccinylation	0.029 ± 0.002	41 ± 11	710
SIRT5 H158Y	**deacetylation**	no activity observed	no activity observed	-
	desuccinylation	ND*	ND (> 750)*	75
SIRT5 Y102F	**deacetylation**	ND*	ND (> 750)*	2
	desuccinylation	ND*	ND (> 750)*	397
SIRT5 R105M	**deacetylation**	ND*	ND (> 750)*	0.5
	desuccinylation	ND*	ND (> 750)*	0.9

The k_{cat} and K_m values cannot be determined because the V versus [S] plot is linear (K_m is much greater than the highest substrate concentration tested, 750 μM). Thus only k_{cat}/K_m value can be obtained.

Figure 2.6. SIRT5 catalyzed lysine desuccinylation *in vivo*. (A) Succinyl lysine was detected in bovine liver mitochondria. SIRT5-catalyzed hydrolysis of malonyl and succinyl peptides could be detected using ^{32}P-NAD, which formed ^{32}P-labeled *O*-Ma-ADPR (lane 2) and *O*-Su-ADPR (lane 3). No reaction occurred with acetyl peptide (lane 1). The formation of *O*-Ac-ADPR catalyzed by SIRT1 was detected (lanes 4 and 8). *O*-Su-ADPR was formed when bovine liver mitochondria peptides were incubated with SIRT5 (lane 9), but not with SIRT1 (lane 10). The control with BSA peptides and SIRT5 did not generate *O*-Su-ADPR (lane 12). CD38-catalzyed hydrolysis of NAD was used to generate the standard ^{32}P-ADPR spot (lane 13). (B) Deletion of SIRT5 increased CPS1 succinylation level in mouse liver. CPS1 was immunoprecipitated from wt and SIRT5 KO livers and the level of succinylation was detected using ^{32}P-NAD. Synthetic acyl peptides were used to generate the reference points *O*-Ac-ADPR, *O*-Ma-ADPR, and *O*-Su-ADPR.The amount of immunoprecipitated CPS1 for lane 4-7 is shown above. (C) The intensities of the *O*-Su-ADPR generated in (B) was quantified and compared. Numbers shown above the bars are normalized. (D) The CPS1 activities were measured using the liver lysates from SIRT5 wt and KO mice. ($n=3$, $p<0.05$) (E) and (F), MS signal intensities for CPS1 K1291 acetyl peptides from SIRT5 KO and wt mice. (G) and (H), MS signal intensities for CPS1 K1291 succinyl peptides from SIRT5 KO and wt mice. (I) and (J), MS signal intensities for a reference peptide from SIRT5 KO and wt mice.


A
1 2 3 4 5 6 7 8 9 10 11 12 13
O-Ma-ADPR
O-Su-ADPR
O-Ac-ADPR
ADPR
NAD
Sirt1 - - - + + + - + - + - - -
Sirt5 + + + - - - + - + - - + -
Peptides Ac Ma Su Ac Ma Su - - Mit Mit Mit BSA -
Histone - - - - - - + + - - - - -
CD38 - - - - - - - - - - - - +
B
H3K9 Ma
H3K9 Su
H3K9 Ac
CPS1 IP from Sirt5 wt
CPS1 IP from Sirt5 KO
Sirtuin
Sirt5 Sirt5 Hst2 - Sirt5 - Sirt5
1 2 3 4 5 6 7
O-Ac-ADPR
O-Ma-ADPR
O-Su-ADPR
ADPR
NAD
C
Relative Intensity
1.00
2.37
CPS1 from Sirt5 wt
CPS1 from Sirt5 KO
D
Relative Activity
1.00
0.87
CPS1 from Sirt5 wt
CPS1 from Sirt5 KO
E
1.9E+06
1.0E+06
Intensity
RT: 35.52
Area: 5805715
K1291Ac
KO
35.0 36.0
Time (min)
F
1.9E+06
1.0E+06
Intensity
RT: 35.37
Area: 5390362
K1291Ac
WT
35.0 36.0
Time (min)
G
9.0E+04
5.0E+04
Intensity
RT: 36.19
Area: 2238676
K1291Su
KO
36.0 37.0
Time (min)
H
9.0E+04
5.0E+04
Intensity
RT: 36.05
Area: 18348
K1291Su
WT
36.0 37.0
Time (min)
I
4.5E+08
2.5E+08
Intensity
RT: 30.44
Area: 2149844491
Reference
KO
28 30 32
Time (min)
J
4.5E+08
2.5E+08
Intensity
RT: 30.32
Area: 2462496998
Reference
WT
28 30 32
Time (min)

formation of *O*-Su-ADPR was detected (Fig. 2.6A), suggesting that bovine liver mitochondrial proteins contained succinyl lysine. Control reactions with SIRT1, without sirtuins, or with BSA peptides did not produce *O*-Su-ADPR or *O*-Ac-ADPR (Fig. 2.6A). The identity of *O*-Su-ADPR was further confirmed by LC-MS/MS (data not shown).

To identify succinylated proteins, succinyl peptides from bovine liver mitochondria were affinity purified using a FLAG-tagged SIRT5 (SIRT5-FLAG) and then identified by LC-MS/MS. Three succinylated proteins were identified: HMG-CoA synthase 2, thiosulfate sulfurtransferase, and aspartate aminotransferase. The sites of succinyl modification were identified by MS/MS (data not shown). Furthermore, LC-MS/MS identified succinyl lysine from three commercial mitochondrial enzymes purified from animal tissues: glutamate dehydrogenase (GDH), malate dehydrogenase, and citrate synthase (data not shown). Thus, lysine succinylation occurs on mammalian mitochondrial proteins.

LC-MS/MS of commercial mitochondrial enzymes identified three and two malonyl lysine residues on GDH and malate dehydrogenase, respectively (data not shown). Thus, protein lysine malonylation exists in mammalian cells.

SIRT5 is known to regulate the activity of carbamoyl phosphate synthase 1 (CPS1) (24). We therefore sought to test whether CPS1 was a desuccinylation target of SIRT5. To confirm that SIRT5's desuccinylase functions *in vivo*, a SIRT5 knockout (KO) mouse strain was generated using standard technology. Consistent with earlier reports (24), CPS1 activity was 15% higher in SIRT5 wt than in SIRT5 KO mice. CPS1 was immunoprecipitated from wt and SIRT5 KO mouse liver and incubated with recombinant SIRT5 and ^{32}P-NAD. More *O*-Su-ADPR was formed with CPS1

from SIRT5 KO mouse than with CPS1 from wt mouse (Fig. 2.6B). As a control, the succinylation levels on immunoprecipitated tubulin were the same in wt and SIRT5 KO mouse. Using MS, we identified three lysine residues of CPS1 that are both acetylated and succinylated: Lys44, Lys287, and Lys1291 (data not shown). For Lys44 and Lys287, the levels of acetylation and succinylation did not change in SIRT5 KO mice (data not shown). For Lys1291, succinylation level increased >20-fold in SIRT5 KO mice compared with the level in wt mice (Fig. 2.6, G and H). In contrast, acetylation levels of Lys1291 did not change in SIRT5 KO compared with wt (Fig. 2.6, E and F). Thus, SIRT5 functions as a desuccinylase *in vivo*.

2.4 Discussion

Here we have demonstrated that SIRT5 is an NAD-dependent demalonylase and desuccinylase. The demalonylase or desuccinylase activity is much higher than its deacetylase activity. The preference for negatively charged acyl groups can be explained by the presence of Tyr102 and Arg105 in the active site of SIRT5, which are conserved in most Class III sirtuins (35). Presumably all Class III sirtuins with the conserved Arg and Tyr should have NAD-dependent desuccinylase and/or demalonylase activity.

We showed that lysine malonylation and succinylation occur on several mammalian proteins. Protein lysine succinylation has been observed recently in bacteria (49). Protein lysine malonylation has not been reported previously.

Acetylation occurs on numerous metabolic enzymes and regulates their activities in mammals and bacteria (50, 51). Given that succinyl-CoA and malonyl-CoA are common metabolites like acetyl-CoA (45), and all the succinylated and malonylated proteins we found are metabolic enzymes, it is likely that protein lysine succinylation and malonylation function similarly to acetylation and regulate metabolism (50, 51).

REFERENCES:

1. Sauve, A. A., Wolberger, C., Schramm, V. L., and Boeke, J. D. (2006) *Annu. Rev. Biochem.* **75**, 435-465.
2. Michan, S., and Sinclair, D. A. (2007) *Biochem. J.* **404**, 1-13.
3. Imai, S.-i., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000) *Nature* **403**, 795-800.
4. Tissenbaum, H. A., and Guarente, L. (2001) *Nature* **410**, 227-230.
5. Anderson, R. M., Bitterman, K. J., Wood, J. G., Medvedik, O., and Sinclair, D. A. (2003) *Nature* **423**, 181-185.
6. Motta, M. C., Divecha, N., Lemieux, M., Kamel, C., Chen, D., Gu, W., Bultsma, Y., McBurney, M., Guarente, L. (2004) *Cell* **116**, 551-563.
7. van der Horst, A., Tertoolen, L. G. J., de Vries-Smits, L. M. M., Frye, R. A., Medema, R. H., and Burgering, B. M. T. (2004) *J. Biol. Chem.* **279**, 28873-28879.
8. Picard, F., Kurtev, M., Chung, N., Topark-Ngarm, A., Senawong, T., Machado de Oliveira, R., Leid, M., McBurney, M. W., and Guarente, L. (2004) *Nature* **429**, 771-776.
9. Fulco, M., Schiltz, R. L., Iezzi, S., King, M. T., Zhao, P., Kashiwaya, Y., Hoffman, E., Veech, R. L., and Sartorelli, V. (2003) *Mol. Cell* **12**, 51-62.
10. Bouras, T., Fu, M., Sauve, A. A., Wang, F., Quong, A. A., Perkins, N. D. Hay, R. T., Gu, W., and Pestell, R. G. (2005) *J. Biol. Chem.* **280**, 10264-10276.
11. Ford, E., Voit, R., Liszt, G., Magin, C., Grummt, I. and Guarente, L. (2006) *Genes Dev.* **20**, 1075-1080.
12. Vaziri, H., Dessain, S. K., Eaton, E. N., Imai, S.-I., Frye, R. A., Pandita, T. K.,

Guarente, L., and Weinberg, R. A. (2001) *Cell* **107**, 149-159.

13. Luo, J., Nikolaev, A. Y., Imai, S.-i., Chen, D., Su, F., Shiloh, A., Guarente, L., and Gu, W. (2001) *Cell* **107**, 137-148.

14. Cheng, H.-L., Mostoslavsky, R., Saito, S. i., Manis, J. P., Gu, Y., Patel, P., Bronson, R., Appella, E., Alt, F. W., and Chua, K. F. (2003) *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10794-10799.

15. Zhao, W., Kruse, J.-P., Tang, Y., Jung, S. Y., Qin, J. and Gu, W. (2008) *Nature* **451**, 587-590.

16. Outeiro, T. F., Kontopoulos, E., Altmann, S. M., Kufareva, I., Strathearn, K. E., Amore, A. M., Volk, C. B., Maxwell, M. M., Rochet, J., McLean, P. J., Young, A. B., Abagyan, R., Feany, M. B., Hyman, B. T., and Kazantsev, A. G. (2007) *Science* **317**, 516-519.

17. Mostoslavsky, R., Chua, K. F., Lombard, D. B., Pang, W. W., Fischer, M. R., Gellon, L., Liu, P., Mostoslavsky, G., Franco, S., Murphy, M. M., Mills, K. D., Patel, P., Hsu, J. T., Hong, A. L., Ford, E., Cheng, H., Kennedy, C., Nunez, N., Bronson, R., Frendewey, D., Auerbach, W., Valenzuela, D., Karow, M., Hottiger, M. O., Hursting, S., Barrett, J. C., Guarente, L., Mulligan, R., Demple, B., Yancopoulos, G. D. and Alt, F. W. (2006) *Cell* **124**, 315-329.

18. Michishita, E., McCord, R. A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T. L. A., Barrett, J. C., Chang, H. Y., Bohr, V. A., Ried, T., Gozani, O., and Chua, K. F. (2008) *Nature* **452**, 492-496.

19. Qiao, L. and Shao, J. (2006) *J. Biol. Chem.* **281**, 39915-39924.

20. Rodgers, J. T., Lerin, C., Haas, W., Gygi, S. P., Spiegelman, B. M. and

Puigserver, P. (2005) *Nature* **434**, 113-118.

21. Schwer, B., Bunkenborg, J., Verdin, R. O., Andersen, J. S. and Verdin, E. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10224-10229.

22. Hallows, W. C., Lee, S. and Denu, J. M. (2006) *Proc. Natl. Acad. Sci. U. S. A.* **103**, 10230-10235.

23. Haigis, M. C., Mostoslavsky, R., Haigis, K. M., Fahie, K., Christodoulou, D. C., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Karow, M., Blander, G., Wolberger, C., Prolla, T. A., Weindruch, R., Alt., F.W., and Guarente, L. (2006) *Cell* **126**, 941-954.

24. Nakagawa, T., Lomb, D. J., Haigis, M. C. and Guarente, L. (2009) *Cell* **137**, 560-570.

25. Lagouge, M., Argmann, C., Gerhart-Hines, Z., Meziane, H., Lerin, C., Daussin, F., Messadeq, N., Milne, J., Lambert, P., Elliott, P., Geny, B., Laakso, M., Puigserver, P., and Auwerx, J. (2006) *Cell* **127**, 1109-1122.

26. Sauve, A. A., Celic, I., Avalos, J., Deng, H., Boeke, J. D. and Schramm, V. L. (2001) *Biochemistry* **40**, 15456-15463.

27. Jackson, M. D. and Denu, J. M. (2002) *J. Biol. Chem.* **277**, 18535-18544.

28. Jackson, M. D., Schmidt, M. T., Oppenheimer, N. J. and Denu, J. M. (2003) *J. Biol. Chem.* **278**, 50985-50998.

29. Smith, B. C. and Denu, J. M. (2006) *Biochemistry* **45**, 272-282.

30. Sauve, A. A. and Schramm, V. L. (2003) *Biochemistry* **42**, 9249-9256.

31. Hoff, K. G., Avalos, J. L., Sens, K. and Wolberger, C. (2006) *Structure* **14**, 1231-1240.

32. Zhao, K., Harshaw, R., Chai, X. and Marmorstein, R. (2004) *Proc. Natl. Acad.*

Sci. U. S. A. **101**, 8563-8568.

33. Hawse, W. F., Hoff, K. G., Fatkins, D. G., Daines, A., Zubkova, O. V., Schramm, V. L., Zheng, W., and Wolberger, C. (2008) *Structure* **16**, 1368-1377.

34. Michishita, E., Park, J. Y., Burneskis, J. M., Barrett, J. C. and Horikawa, I. (2005) *Mol. Biol. Cell* **16**, 4623-4635.

35. Frye, R. A. (2000) *Biochem. Biophys. Res. Commun.* **273**, 793-798.

36. Schuetz, A., Min, J., Antoshenko, T., Wang, C.-L., Allali-Hassani, A., Dong, A., Loppnau, P., Vedadi, M., Bochkarev, A., Sternglanz, R., and Plotnikov, A. N. (2007) *Structure* **15**, 377-389.

37. Liszt, G., Ford, E., Kurtev, M. and Guarente, L. (2005) *J. Biol. Chem.* **280**, 21313-21320.

38. Kowieski, T. M., Lee, S. and Denu, J. M. (2008) *J. Biol. Chem.* **283**, 5317-5326.

39. Du, J., Jiang, H. and Lin, H. (2009) *Biochemistry* **48**, 2878-2890.

40. Otwinowski, Z. and Minor, W. (1997) *Methods Enzymol.* **276**, 472-494.

41. Collaborative. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760-763.

42. Frezza, C., Cipolat, S., and Scorrano, L. (2007) *Nat. Protoc.* **2**, 287-295.

43 Kim, S. C., Sprung, R., Chen, Y., Xu, Y., Ball, H., Pei, J., Cheng, T., Kho, Y., Xiao, H., Xiao, L., Grishin, N. V., White, M., Yang, X. J., and Zhao, Y. (2006) *Mol. Cell* **23**, 607-618.

44. Cosgrove, M. S., Bever, K., Avalos, J. L., Muhammad, S., Zhang, X. and Wolberger, C. (2006) *Biochemistry* **45**, 7511-7521.

45. Gao, L., Chiou, W., Tang, H., Cheng, X., Camp, H. S. and Burns, D. J. (2007)

Journal of Chromatography B **853**, 303-313.

46. Kim, K.-H. (1997) *Ann. Rev. Nutr.* **17**, 77.

47. Saggerson, D. (2008) *Ann. Rev. Nutr.* **28**, 253.

48. Rosen, R., Becher, D., Büttner, K., Biran, D., Hecker, M. and Ron, E. Z. (2004) *FEBS Letters* **577**, 386-392.

49. Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y. and Zhao, Y. (2011) *Nat Chem Biol* **7**, 58-63.

50. Zhao, S., Xu, W., Jiang, W., Yu, W., Lin, Y., Zhang, T., Yao, J., Zhou, L., Zeng, Y., Li, H., Shi, J., An, W., Hancock, S. M., He, F., Qin, L., Chin, J., Yang, P., Chen, X., Lei, Q., Xiong, Y., and Guan, K. L. (2010) *Science* **327**, 1000-1004.

51. Wang, Q., Zhang, Y., Yang, C., Xiong, H., Lin, Y., Yao, J., Li, H., Xie, L., Zhao, W., Yao, Y., Ning, Z., Zeng, R., Xiong, Y., Guan, K. L., Zhao, S., and Zhao, G. P. (2010) *Science* **327**, 1004-1007.

Chapter 3

The Bicyclic Intermediate Structure Provides Insights into the Desuccinylation Mechanism of SIRT5

3.1 Introduction

Protein lysine acetylation is an important and reversible posttranslational modification which regulates protein function. Sirtuins have been widely recognized as a family of NAD-dependent deacetylases which remove acetyl groups from protein lysine residues (1, 2). Sirtuins have been shown to play crucial roles in the regulation of numerous cellular processes, including DNA repair, cell survival and apoptosis, and energy metabolism (reviewed in ref. 2, 3). They have been implicated in human health and diseases, including lifespan extension (4, 5), cancers (6-9), neurodegenerative diseases (10-13), and metabolic diseases (14, 15).

Extensive biochemical and structural studies have uncovered the deacetylation mechanism (2, 16, 17). Upon the binding of both acetylated substrate and NAD, the nicotinamide of NAD was released first; the carbonyl oxygen of acetyl group then attacks the C1' position of nicotinamide ribose, forming the alkylamidate intermediate; the ribose 2'-OH deprotonated by the enzyme attacks the amidate at the carboxyl carbon, generating the 1'-2'-cyclic intermediate, followed by the hydrolysis of a water molecule, producing 2'-*O*-acetyl-ADP-ribose (2'- *O*AADPr) which can be non-enzymatically isomerized into 3'-*O*AADPr. The absolutely conserved histidine residue among the sirtuin family serves as a general base to deprotonate the 2'-OH directly or through the deprotonation of 3'-OH for attacking the 1'-*O*-alkyl-amidate. Kinetic experiments and mass spectrometry have suggested the existence of the

alkylamidate and the cyclic intermediates (18, 19). By using the mechanism-based inhibitor, thioacetyl lysine peptide, the S-alkylamidate intermediate was captured in Sir2Tm and SIRT3 crystals (20, 21). To date, no cyclic intermediate has been directly observed yet.

Among the seven sirtuins in mammals, SIRT1-3 have been demonstrated as robust deacetylases, while SIRT4-7 show little or undetectable deacetylation activity (22-29). Compared to its high deacetylation activity, human SIRT2 has been shown to be a less efficient depropionylase and debutyrylase (30). These findings were coupled by the indentification of protein lysine propionylation and butyrylation (31). However, in contrast to its weak deacetylation activity, SIRT5 was identified as an efficient desuccinylase and demalonylase (32). Furthermore, many mitochondrial proteins were found to contain lysine malonylation and succinylation. Independently, Zhao and coworkers also identified lysine succinylation and malonylation as novel posttranslational modifications (33, 34).

Among the seven human sirtuins, SIRT5 was the only one with demalonylase and desuccinylase activity (32). SIRT5 has been demonstrated to play a pivotal role in ammonium disposal by regulating the enzymatic activity of carbamoyl phosphate synthetase 1 (CPS1) (35). It is of great interest to develop SIRT5-specific inhibitors to expand its cellular function and to explore therapeutic applications. The unique activity of SIRT5 enabled us to develop thiosuccinyl peptide as mechanism-based inhibitor specific for SIRT5 (36).

Here, together with the SIRT5-succinyl peptide-NAD ternary complex structure (PDB code: 3RIY) representing the Michaelis-Menten complex, I am able to delineate the desuccinylation reaction in SIRT5 crystals step by step by providing two

new structures of the SIRT5 complex: one is a complex with the succinyl lysine peptide, representing the enzyme-substrate binding step; the other is a complex with the 1',2'-cyclic intermediate that represents the enzyme-intermediate II step. To our knowledge, this is the first piece of evidence supporting the existence of the 1',2'-bicyclic intermediate. This intermediate structure also suggests that thiosuccinyl peptide iinhibits SIRT5 by forming a stalled bicyclic intermediate at the active site.

3.2 Experimental Procedures

Protein Cloning, Expression, and Purification. Truncated *SIRT5(34-302)* was cloned, using TOPO and Gateway cloning technology (Invitrogen Corp., Carlsbad, CA) into pDEST-F1 for expression, expressed in *E. coli* and purified as previously described (37). Purified Protein was dialyzed into crystallization buffer (20 mM Tris, pH 8.0, 20 mM NaCl, 5% glycerol), concentrated to 16 mg/mL, flash frozen by liquid nitrogen, and stored at -80 °C for crystallization.

Protein Crystallization. Histone H3K9 succinyl or thiosuccinyl peptides [4- KQTAR (suc or tsuK) STGGKA-15] were used for crystallization. SIRT5-peptide mixtures were prepared at a 1:20 protein:peptide molar ratio and incubated for 30~60 min on ice. The final protein concentration was 10 mg/mL. Crystals were grown by the method of hanging drop vapor diffusion. The SIRT5-H3K9 succinyl peptide co-crystals were grown in the condition of 16% PEG 4K, 6% Glycerol at 18 °C, while the SIRT5-H3K9 thiosuccinyl peptide co-crystals were grown in the condition of 30% PEG 10K, 0.1 M TRIS, pH8.5 at room temperature.

Data Collection and Structure Determination. SIRT5-H3K9 succinyl co-crystals were soaked in the cryoprotectant solution (18% PEG4K, 15% Glycerol) at room temperature immediately before data collection. To obtain the intermediate structure, SIRT5-H3K9 thiosuccinyl co-crystals were soaked in the cryoprotectant solution (30% PEG10k, 0.1 M TRIS, pH8.5, 15% Glycerol) with 10 mM NAD for 0.5-16 hours at 4 °C, and flash frozen in liquid nitrogen for data collection. All the X-ray diffraction data were collected at the CHESS (Cornell High Energy Synchrotron Source) A1 or F1 station. The data were processed using the programs HKL2000 (38). The two structures of SIRT5 complexes were solved by molecular replacement using the

program Molrep from the CCP4 suite of programs (39). The SIRT5-H3K9 succinyl peptide-NAD structure (PDB code: 3RIY) served as the search model. Refinement and model building were performed with REFMAC5 and COOT from CCP4. The X-ray diffraction data collection and structure refinement statistics are shown in Table 3.1.

Table 3.1. Crystallographic data collection and refinement statistics.

	SIRT5-sucH3K9	SIRT5-bicyclic intermediate
Data collection		
Space group	$P2_12_12_1$	$P2_12_12_1$
Cell dimensions		
a, *b*, *c* (Å)	52.69, 67.03, 157.63	52.40, 66.76, 156.86
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50-2.00	50-1.70
R_{sym} or R_{merge} (%)	9.4 (48.4)	7.1 (45.8)
$I / \sigma I$	88.6 (3.9)	214.0 (9.6)
Completeness (%)	99.6 (97.4)	98.5 (93.5)
Redundancy	6.8 (6.1)	6.3 (3.3)
Refinement		
Resolution (Å)	50-2.00	50-1.70
No. reflections	39242	61666
R_{work} / R_{free} (%)	22.38 /27.61	19.90/24.65
No. of protein residues	550	550
No. of ligand/ion		
Succinyl	2	2
NAD	--	2
Zn	2	2
No. of water	100	354
R.m.s deviations		
Bond lengths	0.047	0.025
Bond angles (°)	2.10	2.23

Numbers showed in the parentheses are for the highest resolution shell.

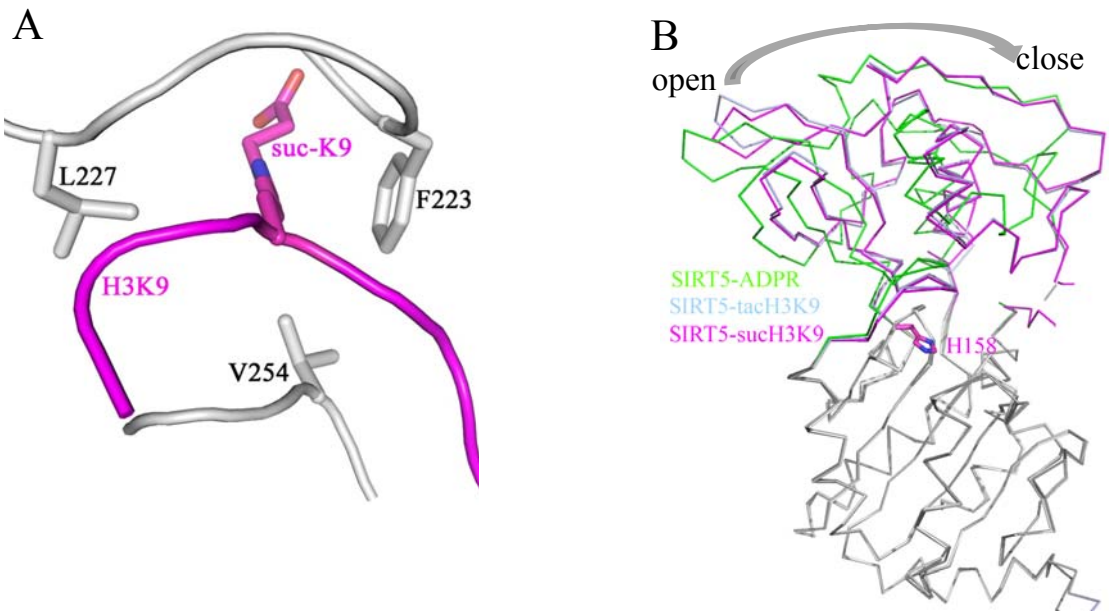
3.3 Results

Overall structure of SIRT5 with H3K9 succinyl peptide

Histones have been reported as physiolocial substrates for several mammalian sirtuins (40-43). Previously, H3K9 peptide was found to be one of the best *in vitro* substrates for all the mammalian sirtuins (32). Biochemical data suggested that SIRT5 shared the mechanism of dessuccinylation similar to that of deacetylation by other sirtuins (32). Herein, we set out to obtain structural insights to elucidate the desuccinylation mechanism of SIRT5 by co-crystallizing SIRT5 with a 12-mer H3K9 peptide (4-KQTAR(Ksuc**)**STGGKA-15) containing succinylated lysine 9. The crystal is in the $P2_12_12$ space group with two SIRT5 molecules in an asymmetric unit. The NAD-stabilizing loop was partially disordered with residues 71-74 (or 65-74 in the other SIRT5 molecule) invisible in the structure.

At least three residues on each side of the succinyl lysine of H3K9 peptide were visible in the structure. The binding pattern of the succinyl H3K9 peptide to SIRT5 was the same as that of acetyl peptides to other sirtuins (44, 45). The succinyl H3K9 peptide formed an anti-parallel β sheet with one loop from the Zn-binding domain and the other loop from the Rossmann fold domain (Fig. 3.1A). This β sheet was stabilized by the main-chain hydrogen bonds from the enzyme and substrate peptide. The structural alignment within the three SIRT5 structures: SIRT5-ADPR, SIRT5-tacH3K9 and SIRT5-sucH3K9, suggested that the interactions within the β sheet drove the Zn-binding domain to rotate clockwise to the Rossmann domain, resulting in SIRT5 having moved from an inactive open state to an active close state upon the substrate binding (Fig. 3.1B). This movement was independent of the interactions made by the acylated lysine side chain, since the zinc-binding domain in

Figure 3.1. The complex structure of SIRT5-sucH3K9. (A) Three hydrophobic residues (Phe223, Leu227 and Val254) of SIRT5 (grey) defined the entrance of the substrate lysine (magenta). (B) Substrate peptide binding caused the Zn-binding domain (colored) to rotate clockwise to the Rossmann fold domain (grey).


A
suc-K9
L227
F223
H3K9
V254
B
open
close
SIRT5-ADPR
SIRT5-tacH3K9
SIRT5-sucH3K9
H158

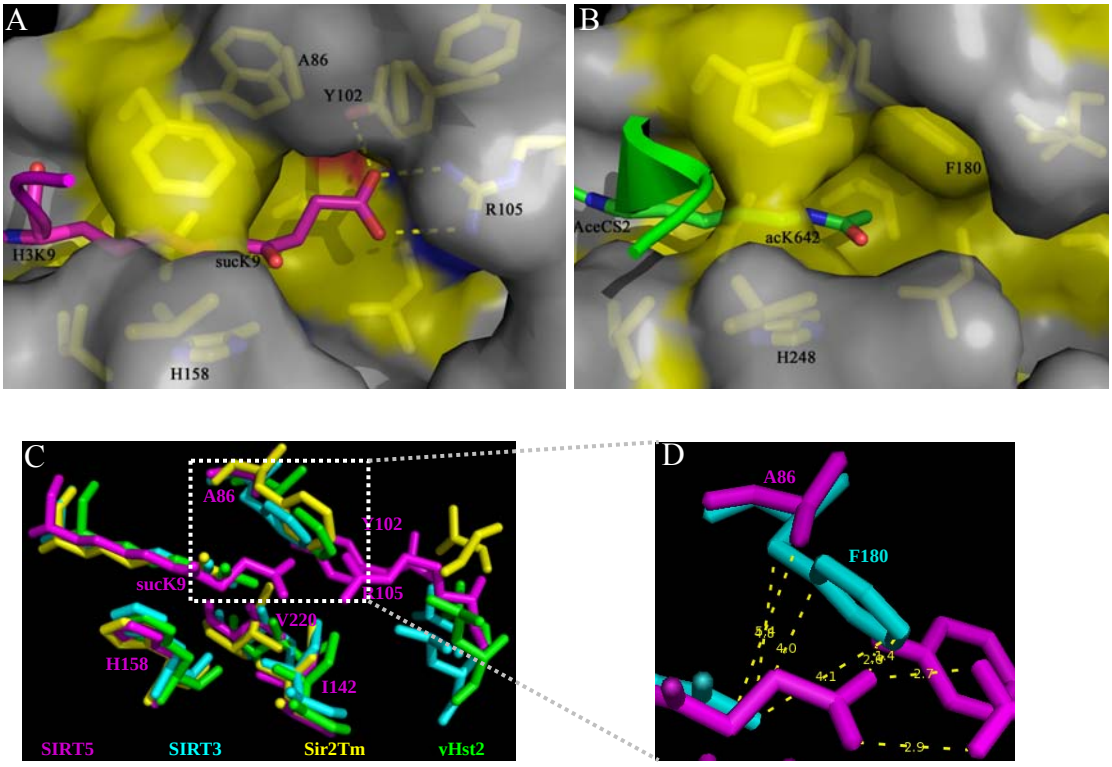
the SIRT5-tacH3K9 structure moved to the same extent as that in the SIRT5-sucH3K9 structure (Fig. 3.1B).

At the entrance of the lysine binding pocket, we found that the lysine residue was surrounded by three hydrophobic residues from SIRT5, F223, L227 and V254 (Fig. 3.1A), which are highly conserved within the sirtuin family. Although these three residues are distant in the primary sequence, they were close enough to form a small triangle structurally, and helped to define the entrance of the acylated lysine of the substrate.

Comparison between SIRT5 with H3K9 succinyl peptide and other sirtuins

Other studies have reported that acetyl lysine was surrounded by all hydrophobic residues (44). In SIRT5, however, two non-hydrophobic residues, Tyr102 and Arg105, are positioned in the deep end of the succinyl lysine binding pocket, and interact with the succinyl group, which implied the specific recognition of the negatively charged acyl group by SIRT5 (Fig. 2A) (32). SIRT5 harbored a larger acyl lysine binding pocket than did SIRT3, due to the replacement of a smaller residue Ala86 of SIRT5 compared to Phe180 of SIRT3 (Fig. 3.2B). In the alignment, Phe180 of SIRT3 was only 1.4 Å away from the succinyl group in SIRT5, which implied that the SIRT5 A86F mutant could hinder succinyl group from interacting with Tyr102 and Arg105 (Fig. 3.2D). Therefore, this mutant should display decreased desuccinylation activity. This should be similar to the de-propionylation of sirtuins (46). The structural alignment also showed that the succinyl lysine peptides bound to the same place as did acetyl lysine peptide in other sirtuins (Fig. 3.2C).

Figure 3.2. The structural features of SIRT5 suggested that SIRT5 was optimized for desuccinylation. (A) Surface representation of SIRT5's succinyl lysine binding pocket (grey), which consisted of several hydrophobic residues (yellow sticks) and two non-hydrophobic residues Tyr102 and Arg105 which interacted and stabilized the succinyl lysine (magenta). (B) Surface representation of SIRT3's acetyl lysine binding pocket (grey). This pocket consisted of all hydrophobic residues (yellow sticks). The residue F180 lies at the upper right bottom of the pocket. (C) Structural alignment among SIRT5 (magenta), SIRT3 (cyan), Sir2Tm (yellow) and yHst2 (green) showed that the succinyl lysine in SIRT5 bound to the same place as acetyl lysine in other sirtuins. (D) Detailed view of SIRT5 and SIRT3 alignment: the smallest distance from succinyl group to Phe180 of SIRT3 was 1.4 Å, indicating that phenylalanine would cause spatial hindrance for succinyl lysine.

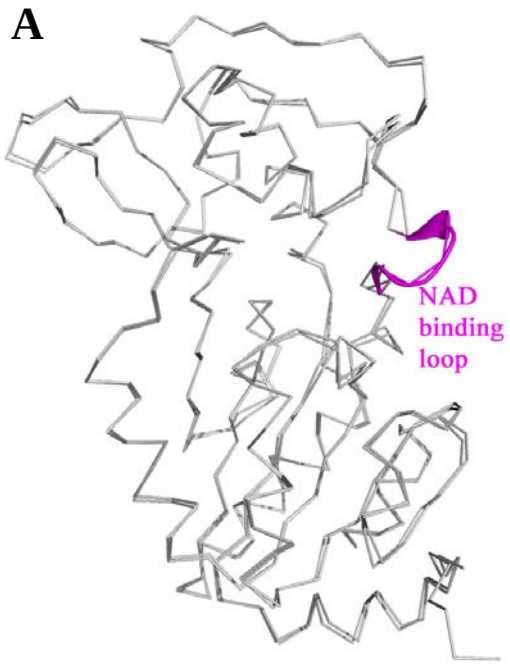

A
A86
Y102
R105
H3K9
sucK9
H158
B
F180
AceCS2
acK642
H248
C
A86
Y102
sucK9
R105
V220
H158
I142
SIRT5
SIRT3
Sir2Tm
vHst2
D
A86
F180

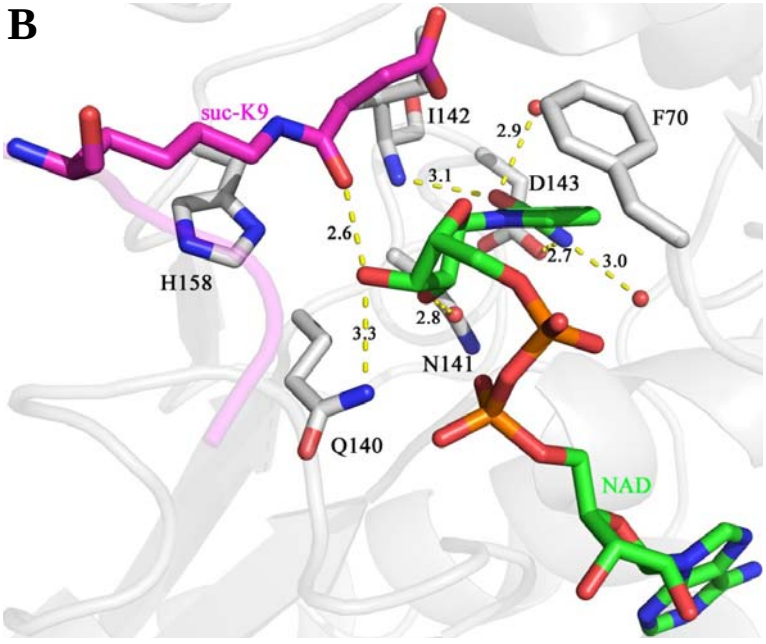
Comparison between SIRT5-sucH3K9 and SIRT5-sucH3K9-NAD

Biochemical studies have demonstrated that the acetyl substrate bound to sirtuins first, followed by NAD binding (47). This ordered binding manner ensured that NAD adopted a productive conformation which led to the completion of the deacetylation reaction (48). Previously, we were able to capture the Michaelis-Menten complex of SIRT5 with both succinyl lysine peptide and NAD bound. In this structure, the NAD binding loop was ordered and stabilized NAD at the active site (Fig. 3.3A). As described above, binding of succinyl lysine substrate caused the movement between the two domains of SIRT5. The alignment of Cα atoms between the succinyl lysine peptide-bound structure and the Michaelis-Menten complex structure of SIRT5 yielded a RMS of 0.357, arguing that NAD did not cause any further movement between the two domains of SIRT5 (Fig. 3.3A).

Similar to NAD in other sirtuins, the nicotinamide group of NAD inserted into the C pocket of SIRT5 where it formed hydrogen bonds with the side chain of the invariant residue Asp143 and the main chain nitrogen of Ile142, as well as two water molecules (Fig. 3.3B). Those interactions caused the carboxylamide of nicotinamide to rotate off the plane of pyridine ring, which activated the cleavage of nicotinamide. Residue Phe70 from the NAD binding loop paralleled roughly to the pyridine ring but almost perpendicular to the ribosyl ring, making room for the nicotinamide release. The other two conserved residues, Gln140 and Asn141, formed hydrogen bonds via their side chains with 3' and 2' –OH of N-ribose, respectively (Fig. 3.3B). The carboxyl oxygen of the succinyl group formed an additional hydrogen bond with 3'-OH of the ribose. Those interactions collectively positioned N-ribose in an orientation which favored the cleavage of nicotinamide. The catalytic residue His158 did not

Figure 3.3. NAD bound to SIRT5 in a productive conformation in the Michaelis-Menten complex. (A) The alignment between SIRT5-H3K9 succinyl peptide and SIRT5-Michaelis complex indicated that NAD binding did not change the overall structure of SIRT5, except that the NAD binding loop (magenta) was ordered in the Michaelis complex structure. (B) The nicotinamide and N-ribose of NAD made extensive interactions with SIRT5 (grey cartoon) and water molecules (red dot) to enable the cleavage of nicotinamide. Residues were shown in grey, hydrogen bonds in yellow, succinyl lysine in magenta, NAD in green. Phe70, almost perpendicular to N-ribose, formed Van der Waals interactions with the nicotinamide ring. His158 did not interact with NAD. The carboxyl oxygen of the succinyl group made a hydrogen bond with N-ribose.


A
NAD
binding
loop

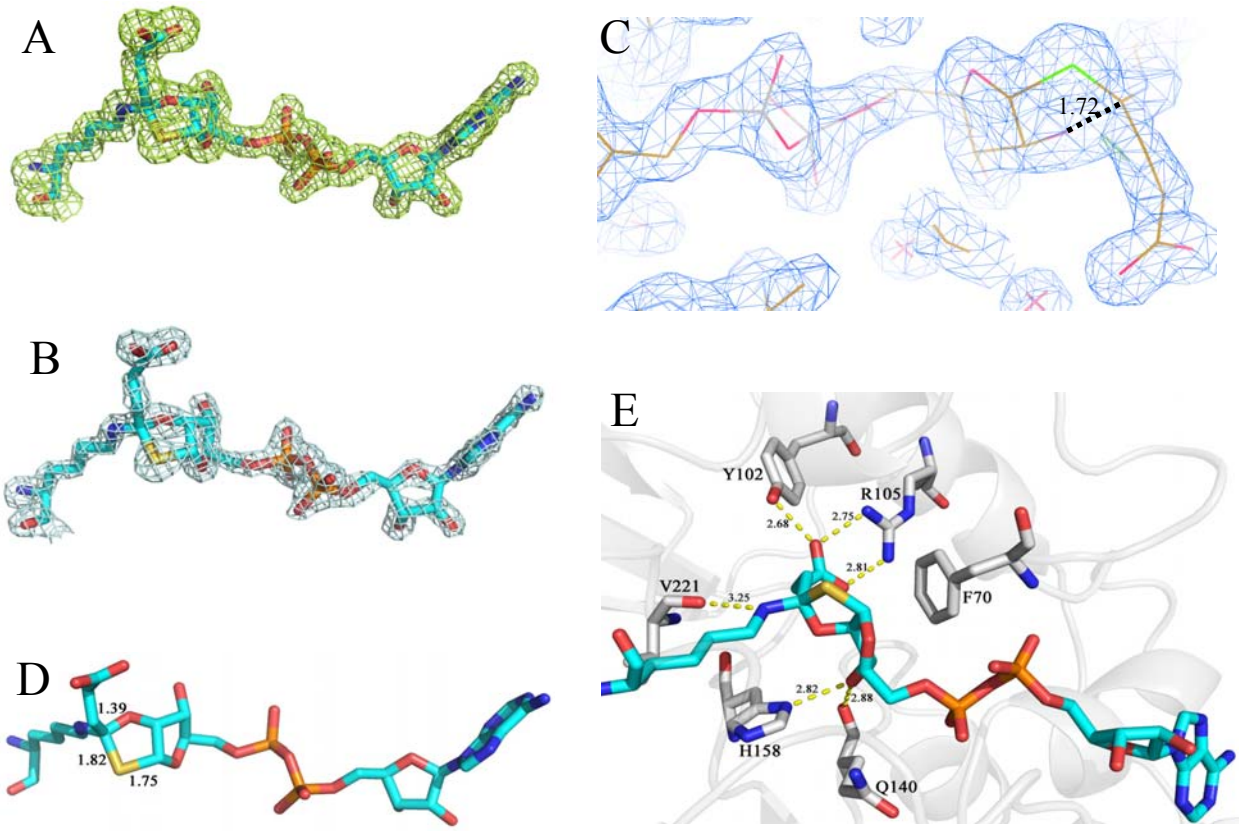

B
suc-K9
I142
F70
2.9
3.1
D143
2.6
2.7
3.0
H158
2.8
3.3
N141
Q140
NAD

contact with NAD at this stage, consistent with the finding that catalytically deficient yeast HST2 mutant had little effect on nicotinamide release rate (49).

Structure of SIRT5-bicyclic intermediate II

Extensive studies have established the mechanism of deacetylation of sirtuins, which includes the formation of two intermediates: the O-alkylamidate intermediate I and the bicyclic intermediate II (2, 16, 17). Thioacetyl lysine peptides have been reported as inhibitors for sirtuins with deacetylase activities, because the corresponding intermediate can be stalled at the active site, resulting in much lower turn-over rate (19). Because SIRT5 is the only mammalian sirtuin that prefers succinyl, we designed a thiosuccinyl lysine peptide as a SIRT5-specific inhibitor. Indeed, it was demonstrated that it inhibited SIRT5 with an IC_{50} of 5 μM while it did not inhibit SIRT1, 2, or 3, even at 100 μM (36). In order to better understand the mechanism of SIRT5's desuccinylation, we co-crystallized SIRT5 with thiosuccinyl H3K9 peptide, then soaked the co-crystals in 10 mM NAD for 0.5-16 hours. The S-alkylamidate intermediate I was unable to be captured in the crystals while all the data sets contained the S-bicyclic intermediate II. The 14 hour NAD-soaking structure was determined to 1.7 Å resolution (Fig. 3.4). In the crystal, the thiosuccinyl lysine peptide reacted with NAD and formed the bicyclic intermediate II. Other groups have captured the S-alkylamidate intermediate I in Sir2Tm and SIRT3 (19, 20) via similar approaches. In our structure, however, when the S-alkylamidate intermediate I was fitted into the density, the 2'-OH group was only 1.7 Å from the succinyl carbon, which is similar to the distance of a carbon-oxygen single bond (about 1.4 Å) (Fig. 3.4C). The cyclic intermediate II fitted the density better when the 2' OH group attacked the carbon of the succinyl group (Fig. 3.4D). An average carbon-sulfur single

Figure 3.4. The complex structure of SIRT5-bicyclic intermediate II. (A) The 2Fo-Fc omit electron density map (1σ) showing the S-cyclic intermediate II. (B) The Fo-Fc map (2σ) showing the same view as in (A). (C) The possible fitting of the S-alkylamidate intermediate I. 2Fo-Fc map (blue) contoured at 1.0 σ and Fo-Fc map (green) contoured at 3 σ. However, the distance between 2'-OH oxygen and carbon of succinyl group was only 1.72Å. (D) The bond length in the S-cyclic intermediate II; (E) the interactions that S-cyclic intermediate II made.

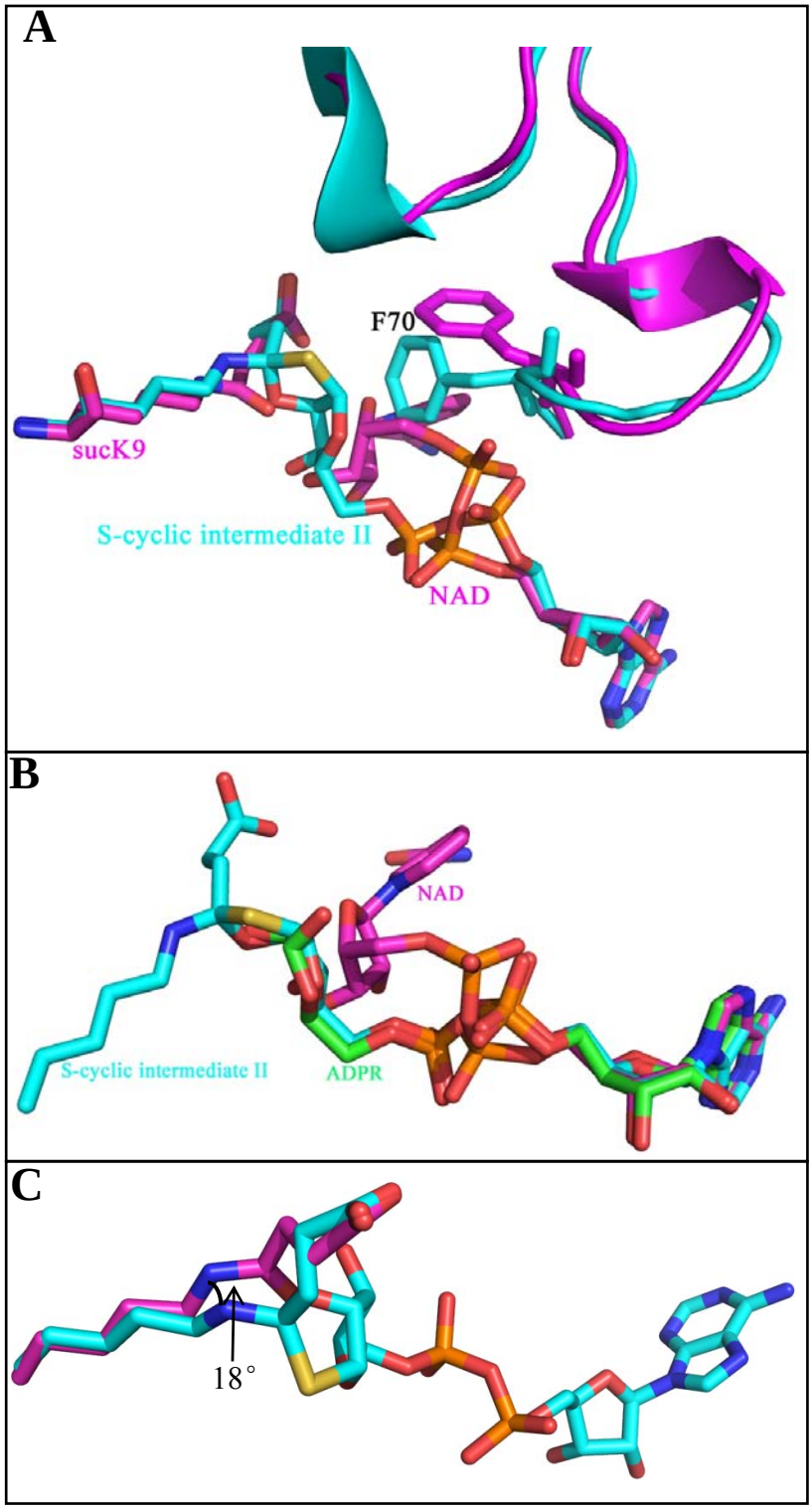

A
B
C
1.72
D
1.39
1.82
1.75
E
Y102
R105
2.68
2.75
2.81
V221
3.25
F70
2.82
2.88
H158
Q140

bond is 1.8 Å, and those bonds in the intermediate II were in great agreement with that (Fig. 3.4D). The cyclic intermediate II was stabilized by extensive hydrogen bonds from both the backbone and the side-chains as well as hydrophobic interactions (Fig. 3.4E). Tyr102 and Arg105 specifically recognized succinyl group. The catalytic residue H158 formed a hydrogen bond with 2'-OH of N-ribose, and deprotonated it to promote the nucleophilic attack of succinyl carbon by 3'-OH. Residue Gln140 was absolutely conserved in all mammalian sirtuins, formed a hydrogen bond with N-ribose via the backbone oxygen, and played a key role in positioning the N-ribose during the reaction. The highly conserved residues, Val221, interacted with and stabilized the lysine side chain via a hydrogen bond of its backbone oxygen with ε-N of lysine. The benzyl ring of Phe70 paralleled with N-ribose ring and formed π-π hydrophobic interactions, thus stabilizing the intermediate at the active site. All those interactions stabilized and positioned the cyclic intermediate at the active site in a conformation that is favorable for the turn-over.

Comparison between SIRT5-bicyclic intermediate and the Michaelis intermediate

Compared to the Michaelis-Menten complex structure of SIRT5, the NAD binding loop of the cyclic intermediate remained almost the same, indicating that the nicotinamide cleavage and intermediate formation did not interfere ADPR binding and stabilization (Fig. 3.5A). However, residue Phe70 on the NAD binding loop adopted two different orientations. In the Michaelis-Menten complex structure, Phe70 was perpendicular to the ribosyl ring of NAD, which was proposed to favor nicotinamide escape (20). In the cyclic intermediate structure, Phe70 paralleled the ribose face, which prevented the base-exchange reaction of nicotinamide from generating NAD (Fig. 3.5A). Compared to the ADPR in the SIRT5-ADPR structure (PDB code:

Figure 3.5. The comparison between the bicyclic intermediate II (cyan) and the Michaelis-Menten complex (magenta) of SIRT5. (A) The NAD binding loop was the same in two structures, except that residue Phe70 was almost perpendicular to each other. In the intermediate structure, Phe70 was parallel to the nicotinamide ribose, while in the Michaelis-Menten structure, Phe70 was perpendicular to the ribose. (B) After the cleavage of nicotinamide, the ribose flipped to the succinyl lysine, repositioning the two phosphates to facilitate the formation of the intermediate. (C) The lysine side chain was rotated about 18º to form the intermediate.

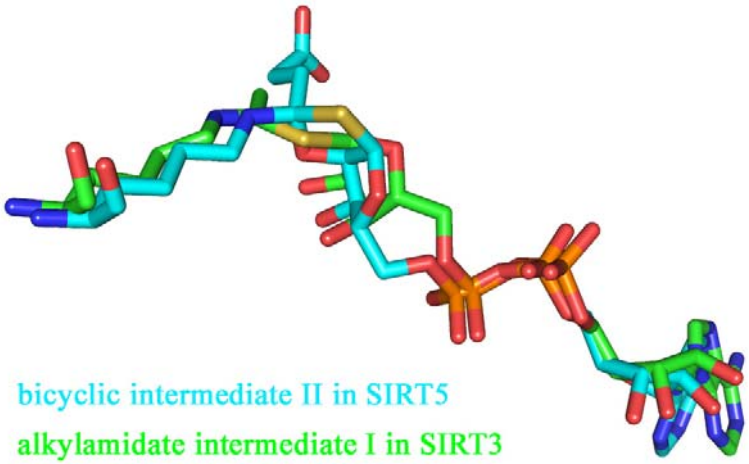

A
F70
sucK9
S-cyclic intermediate II
NAD
B
NAD
S-cyclic intermediate II
ADPR
C
18°

2B4Y), the cyclic intermediate exhibited the same conformation in the ADPR fragment, while NAD in the Michaelis-Menten complex oriented in a different way, especially in the phosphates and N-ribose part (Fig. 3.5B). Those differences suggested that NAD first bound to SIRT5 in the conformation which favored the cleavage of nicotinamide, and then after nicotinamide released, the N-ribose flipped and rotated to some degree which favored the nucleophilic attack of the succinyl group. The lysine side chain was rotated about 18º to form the intermediate, which caused corresponding movement of the succinyl group except that the carboxylate of succinyl group remained the contacts with Tyr102 and Arg105 (Fig. 3.5C).

Comparison between SIRT5-bicyclic intermediate and SIRT3-intermediate I

Previously, kinetic studies and Mass Spectrometry have demonstrated the existence of the alkylamidate and 1', 2'-bicyclic intermediates (18, 19). In 2008, Hawse *et al*. trapped a S-alkylamidate intermediate in Sir2Tm of 2.5 Å resolution using thioacetyl peptide (20). That was the first direct observation of an alkylamidate intermediate. Later, a similar alkylamidate intermediate was obtained in human SIRT3, which aligned well with the intermediate in Sir2Tm (21). We captured a 1', 2'-bicyclic intermediate II in human SIRT5 crystal and solved the structure to 1.7 Å resolution. The ADP-ribose part of the alkylamidate and cyclic intermediates superimposed well except the orientation of N-ribose (Fig. 3.6). Compared to the alkylamidate intermediate I in SIRT3, the ribose plane rotated some degree toward the succinyl group, which favored the nucleophilic attack of 2'-OH to the succinyl carbon. This suggested that the nucleophilic attack of 2'-OH of the ribose caused conformational change in the ADP-ribose.

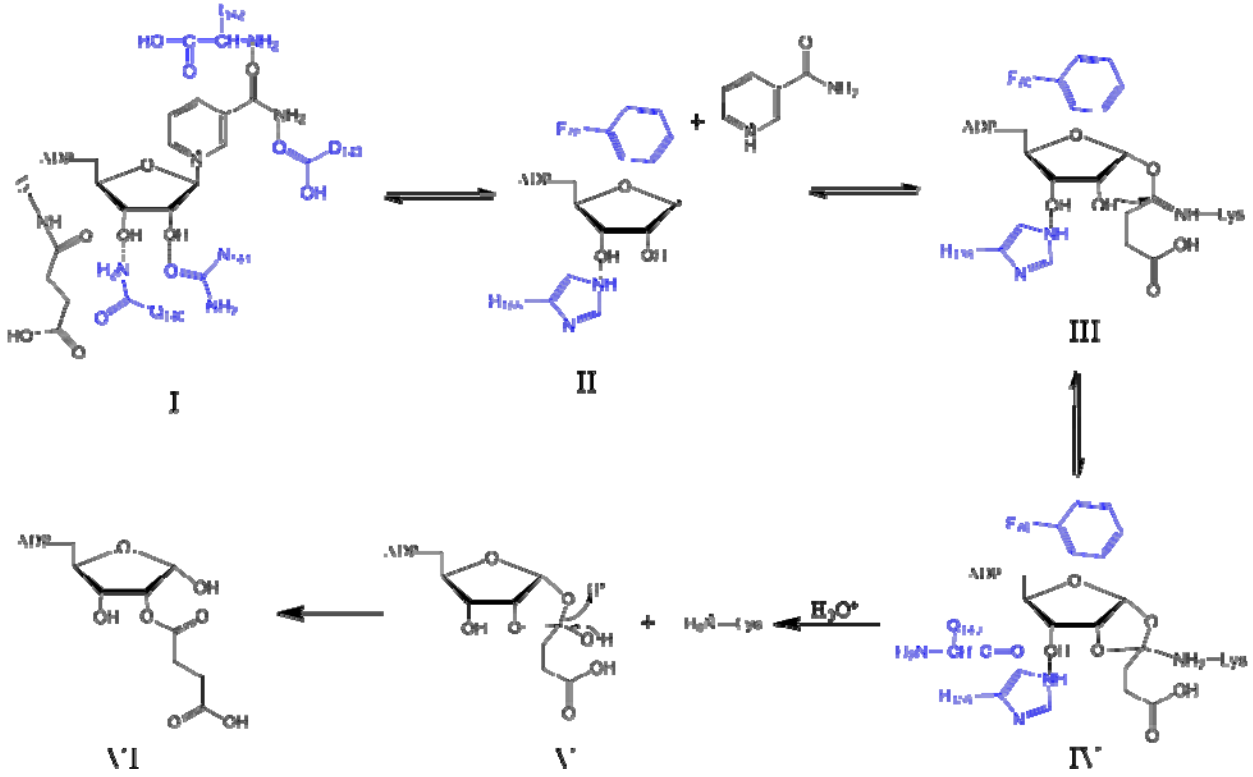
Figure 3.6. The structural alignment between SIRT5-bicyclic intermediate II and SIRT3-alkylamidate intermediate I showed that the N-ribose changed its conformation to facilitate the nucleophilic attack of 2'-OH to the succinyl carbon.


bicyclic intermediate II in SIRT5
alkylamidate intermediate I in SIRT3

3.4 Discussion

SIRT5 was recently identified as a novel desuccinylase and demalonylase (32), which is different from the typical sirtuin function of deacetylase. Biochemical studies suggested that SIRT5 used a mechanism similar to deacetylation to remove a malonyl or succinyl modification from lysine residues. Previously, two alkylamidate intermediates were shown to be trapped in the Sir2Tm and human SIRT3 structures (20, 21), supporting the ADPR-peptidyl-imidate mechanism of deacetylation. Our study demonstrates, for the first time, that a bicyclic intermediate can be directly observed in crystals, providing one more piece of evidence that sirtuins utilize the ADPR-peptidyl-imidate mechanism to remove acyl group from substrate lysine. The desuccinylation mechanism of SIRT5 is summarized in Fig. 3.7. Upon the formation of the Michaelis-Menton complex, the oxygen atom of the carboxyl from succinyl group forms a hydrogen bond with 3'-OH of nicotinamide ribose of NAD (Fig. 3.7 I). The two residues of SIRT5, Gln140 and Asn141, which are invariant among sirtuins, form additional hydrogen bonds with 3'-OH and 2'-OH of the ribose, respectively. The carboxyl amide of nicotinamide interacts with the absolutely conserved Asp143 and the highly conserved Ile142. Those interactions collectively force the nicotinamide into a high energy conformation which subsequently causes the cleavage of nicotinamide and generates the transient ionic intermediate (Fig. 3.7 II). It is likely that this oxocarbonium-ion intermediate is stabilized via hydrogen bond between 3'-OH of the ribose and the catalytic residue His158 and via π-π hydrophobic interactions between the ribose ring and residue Phe70 (Fig. 3.7 II). The release of nicotinamide leads to the rotation of the nicotinamide ribose (Fig. 3.5B), which disrupts the interactions between the ribose and residue Gln140 and Asn41

Figure 3.7. The structure-based desuccinylation mechanism of SIRT5. SIRT5 residues are colored in blue. (I) Upon the binding of a succinylated lysine peptide and an NAD, the extensive interactions between NAD and peptide & enzyme drive NAD to a productive conformation which results in the cleavage of nicotinamide. (II)The release of nicotinamide generates a positively charged oxacarbonium-ion transient intermediate which is stabilized by interactions with His158 and Phe70. (III) Upon the leave of nicotinamide, the ribose rotates to expose 1'-carbon to the carboxyl oxygen of succinyl group for nucleophilic attack, followed by the formation of the O-alkylamidate intermediate. (IV) The 2'-OH of the ribose attacks the carboxyl carbon of succinyl group and generates the 1', 2'-bicyclic intermediate which is additionally stabilized by hydrogen bond between 3'-OH and main chain oxygen of Gln140. (V and VI) The bicyclic intermediate is hydrolyzed into free lysine and succinyl-*O*-ADP ribose.


I
II
III
IV'
V'
VI
H3O+

(Fig. 3.7 II). The new conformation of nicotinamide ribose favors the attack of the carboxyl oxygen of succinyl group, producing the ADPR-peptidyl-imidate intermediate (Fig. 3.7 III), followed by the attack of 2'-OH to the carboxyl carbon of succinyl group to form the bicyclic intermediate (Fig. 3.7 IV). The bicyclic intermediate is further hydrolyzed into a free lysine and a succinyl-*O*-ADP ribose (Fig. 3.7 V and VI).

We were unable to obtain the alkylamidate intermediate and the thiosuccinyl-*O*-ADP-ribose product, even though we soaked co-crystals of SIRT5-thiosuccinyl H3K9 peptide in 10 mM NAD at 4 ºC from 0.5 to 16 hours. Regardless of the NAD soaking time we tested, we could only trap the bicyclic intermediate at the active site. This is quite different from the soaking experiments with Sir2Tm or SIRT3, which could generate the alkylamidate intermediate and thioacetyl-*O*-ADP-ribose product with relatively short and long soaking time in NAD, respectively (20, 21). This suggests that, although sirtuins share similar mechanism to remove acyl groups from modified lysines, different sirtuins may preferentially stabilize different intermediates at the active site. SIRT5 favors the bicyclic intermediate as the more stable specie than the alkylamidate one. Additionally, the type of acyl group may affect the stability of those two intermediates.

Similar to lysine acetylation/deacetylation, lysine malonylation/demalonylation and succinylation/desuccinylation should be important in regulating protein functions. To date, only carbamoyl phosphate synthetase 1 (CPS1) has been identified as a desuccinylation substrate of SIRT5 (32). Further exploration of SIRT5's novel substrates awaits future effort. The thiosuccinyl peptide is a mechanism-based inhibitor for SIRT5, which specifically inhibits SIRT5 with the IC_{50} value of 5 μM

(36). This is about 5-fold lower than that of suramin, generally accepted as an inhibitor for SIRT5. The structural study presented here will facilitate the design of specific inhibitors to fully characterize SIRT5's cellular functions as well as to search for more physiological substrates of SIRT5. Additionally, SIRT5 inhibitors can be potential drugs for therapeutic applications.

References

1. Imai, S.-i., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000) *Nature* **403**, 795-800.
2. Sauve, A. A., Wolberger, C., Schramm, V. L., and Boeke, J. D. (2006) *Annu. Rev. Biochem.* **75**, 435-465.
3. Michan, S., and Sinclair, D. (2007) *Biochem. J.* **404**, 1-13.
4. Tissenbaum, H. A., and Guarente, L. (2001) *Nature* **410**, 227-230.
5. Anderson, R. M., Bitterman, K. J., Wood, J. G., Medvedik, O., and Sinclair, D. A. *Nature* **423**, 181-185 (2003).
6. Gao, F., Cheng, J., Shi, T., and Yeh, E. T. (2006) *Nat. Cell Biol.* **8**, 1171-1177.
7. Ashraf, N., Zino, S., Macintyre, A., Kingsmore, D., Payne, A. P., Geroge, W. D., and Shiels, P. G. (2006) *Br. J. Cancer* **95**, 1056-1061.
8. De Nigris, F., Cerutti, J., Morelli, C., Califano, D., Chiariotti, L., Viglietto, G., Santelli, G., and Fusco, A. (2002) *Br. J. Cancer* **86**, 917-923.
9. Frye, R. (2002) *Br. J. Cancer* **87**, 1479.
10. Outeiro, T. F., Kontopoulos, E., Altmann, S. M., Kufareva, I., Strathearn, K. E., Amore, A. M., Volk, C. B., Maxwell, M. M., Rochet, J., McLean, P. J., Young, A. B., Abagyan, R., Feany, M. B., Hyman, B. T., and Kazantsev, A. G. (2007) *Science* **317**, 516-519.
11. Garske, A. L., Smith, B. C., and Denu, J. M. (2007) *ACS Chem. Biol.* **2**, 529-532.
12. Anekonda, T. S., and Reddy, P. H. (2006) *J. of Neurochem.* **96**, 305-313.
13. Green, K. N., Steffan, J. S., Martinez-Coria, H., Sun, X., Schreiber, S. S., Thompson, L. M., and LaFerla, F. M. (2008) *J. of Neurosci.* **28**, 11500-11510.

14. Yamamoto, H., Schoonjans, K., and Auwerx, J. (2007) *Mol. Endorinol.* **21**, 1745-1755.

15. Schwer, B., Schumacher, B., Lombard, D. B., Xiao, C., Kurtev, M. V., Gao, J., Schneider, J. I., Chai, H., Bronson, R. T., Tsai, L. H., Deng, C. X., and Alt, F. W. (2010) *Proc. Natl. Acad. Sci. U. S. A.* **107**, 21790-21794.

16. Denu, J. M. (2005) *Curr. Opin. Chem. Biol.* **9**, 431-440.

17. Smith, B. C., Hallows, W. C., and Denu, J. M. (2008) *Chem. Biol.* **15**, 1002-1013.

18. Smith, B. C., and Denu, J. M. (2006) *Biochemistry* **45**, 272-282.

19. Smith, B. C., and Denu, J. M. (2007) *Biochemistry* **46**, 14478-14486.

20. Hawse, W. F., Hoff, K. G., Fatkins, D. G., Daines, A., Zubkova, O. V., Schramm, V. L., Zheng, W., and Wolberger, C. (2008) *Structure* **16**, 1368-1377.

21. Jin, L., Wei, W., Jiang, Y., Peng, H., Cai, J., Mao, C, Dai, H., Choy, W., Bemis, J. E., Jirousek, M. R., Milne, J. C., Westphal, C. H., and Perni, R. B. (2009) *J. Biol. Chem.* **284**, 24394-24405.

22. Frye, R. A. (1999) *Biochem. Biophys. Res. Commun.* **260**, 273-279.

23. Frye, R. A. (2000) *Biochem. Biophys. Res. Commun.* **273**, 793-798.

24. North, B. J., Schwer, B., Ahuja, N., Marshall, B., and Verdin, E. (2005) *Methods* **36**, 338-345.

25. Michishita, E., Park, J. Y., Burneskis, J. M., Barrett, J. C., and Horikawa, I. (2005). *Mol. Biol. Cell* **16**, 4623-4635

26. Haigis, M. C., Mostoslavsky, R., Haigis, K. M., Fahie, K., Christodoulou, D. C., Murphy, A. J., Valenzuela, D. M., Yancopoulos, G. D., Karow, M.,

Blander, G., Wolberger, C., Prolla, T. A., Weindruch, R., Alt., F.W., and Guarente, L. (2006) *Cell* **126**, 941-954.

27. Schuetz, A., Min, J., Antoshenko, T., Wang, C., Allali-Hassani, A., Dong, A., Loppnau, P., Vedadi, M., Bochkarev, A., Sternglanz, R., and Plotnikov, A. N. (2007) *Structure* **15**, 377-389.

28. Schlicker, C., Gertz, M., Papatheodorou, P., Kachholz, B., Becher, C. F. W., and Steegborn, C. (2008) *J. Mol. Biol.* **382**, 790-801 (2008).

29. Liszt, G., Ford, E., Kurtev, M., and Guarente, L. *J. Biol. Chem.* **280**, 21313-21320 (2005).

30. Smith, B. C., and Denu, J. M. (2007) *J. Biol. Chem.* **282**, 37256-37265.

31. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., Falck, J. R., Peng, J., Gu, W., and Zhao, Y. (2007) *Mol. Cell. Proteomics.* **6**, 812-819.

32. Du, J., Zhou, Y., Su, X., Yu, J. J., Khan, S., Jiang, H., Kim, J., Woo, J., Kim, J. H., Choi, B. H., He, B., Chen, W., Zhang, S., Cerione, R. A., Auwerx, J., Hao, Q., and Lin, H. (2011) *Science* **334**, 806-809.

33. Zhang, Z., Tan, M., Xie, Z., Dai, L., Chen, Y., and Zhao, Y. (2011) *Nat. Chem. Biol.* **7**, 58-63.

34. Peng, C., Lu, Z., Xie, Z., Cheng, Z., Chen, Y., Tan, M., Luo, H., Zhang, Y., He, W., Yang, K., Zwaans, B. M., Tishkoff, D., Ho, L., Lombard, D., He, T. C., Dai, J., Verdin, E., Ye, Y., and Zhao, Y. (2011) *Mol. Cell. Proteomics.* In press.

35. Nakagawa, T., Lomb, D. J., Haigis, M. C., and Guarente, L. (2009) *Cell* **137**, 560-570.

36. He, B., Du, J., and Lin, H. (2012) *J. Am. Chem. Soc.* Just accepted.

37. Du, J., Jiang, H., and Lin, H. (2009) *Biochemistry* **48**, 2878-2890.

38. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 472.

39. Collaborative (1994) *Acta C rystallogr. D Biol. Crystallogr.* **50**, 760.

40. Vaquero, A., Scher, M. B., Lee, D. H., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2004) *Mol. Cell* **16**, 93-105.

41. Vaquero, A., Scher, M. B., Lee, D. H., Sutton, A., Cheng, H., Alt, F. W., Serrano, L., Sternglanz, R., and Reinberg, D. (2006) *Genes Dev.* **20**, 1256-1261.

42. Michishita, E., McCord, R. A., Berber, E., Kioi, M., Padilla-Nash, H., Damian, M., Cheung, P., Kusumoto, R., Kawahara, T. L. A., Barrett, J. C., Chang, H. Y., Bohr, V. A., Ried, T., Gozani, O., and Chua, K. F. (2008) *Nature* **452**, 492-496.

43. Yang, B., Zwaans, B. M. M., Eckersdorff, M., and Lombard, D. B. (2009) *Cell Cycle* **8**, 2662-2663.

44. Avalos, J. L., Celic, I., Muhammad, S., Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2002) *Mol. Cell* **10**, 523-535.

45. Hoff, K. G., Avalos, J. L., Sens, K., and Wolberger, C. (2006) *Structure* **14**, 1231-1240.

46. Bheda, P., Wang, J. T., Escalante-Semerena, J. C., and Wolberger, C. (2010) *Protein Sci.* **20**, 131-139.

47. Borra, M. T., Langer, M. R., Slama, J. T., and Denu, J. M. (2004) *Biochemistry* **43**, 9877-9887.

48. Avalos, J. L., Boeke, J. D., and Wolberger, C. (2004) *Mol. Cell* **13**, 639-648.

49. Jackson, M. D., Schmidt, M. T., Oppenheimer, N. J., and Denu, J. M. (2003) *J. Biol. Chem.* **278**, 50985-50998.

Chapter 4

Plasmodium falciparum Sir2A Preferentially Hydrolyzes Medium and Long Chain Fatty Acyl Lysine[2]

4.1 Introduction

Sirtuins are a family of enzymes known as nicotinamide adenine dinucleotide (NAD)-dependent deacetylases (1, 2). They regulate a variety of biological processes, including transcription and metabolism (3, 4). *Plasmodium falciparum* (*P. falciparum*) contains two sirtuins, PfSir2A and PfSir2B (5). It was shown that these two sirtuins regulate the expression of surface antigens to evade the detection by host immune surveillance (6, 7). Thus, inhibiting these sirtuins may help fight malaria. It was thought that PfSir2A and PfSir2B achieve this physiological function by deacetylating histones. *In vitro* studies on PfSir2A showed that it has deacetylase activity (8). However, the activity was weak compared to several other sirtuins (9), such as human SIRT1 and yeast Sir2. It was also reported that PfSir2A had ADP-ribosyltransferase activity (8). However, several reports questioned whether the ADP-ribosyltransferase activity of sirtuins was physiologically relevant since the measured activity of several sirtuins was weak (10, 11).

In addition to acetylation, lysine propionylation and butyrylation have been

[2] This work was published in *ACS Chemical Biology* on Oct. 21, 2011 online. Permission of reusing it in this dissertation was granted by ACS chemical Biology, copyright 2011 American Chemical Society. I was a co-first author, responsible for the protein purification, crystallization, and diffraction data collection and processing, structure solving of PfSir2A complexes. Some of the figures have been rearranged for readers' convenience.

reported as posttranslational modifications that occur on proteins, including histones (12-14). Many fatty acyl-CoA molecules exist in cells as metabolic intermediates. If the shorter chain fatty acyl CoA molecules (acetyl-CoA, propionyl-CoA, and butyryl-CoA) are used as acyl donors to modify proteins, it is possible that longer chain fatty acyl-CoA molecules can also be used to modify protein. Given that PfSir2A's deacetylase activity is weak, we set out to investigate whether longer chain fatty acyl lysine can be accepted as better substrates by PfSir2A.

4.2 Experimental Procedures

Cloning, Expression, and Purification of PfSir2A. PfSir2A gene was custom synthesized by Genscript. The sequence was codon optimized for overexpression in *E. coli* and cloned into pET-28a (+) vector with BamHI and XhoI restriction sites. The PfSir2A expression vector was then introduced into an *E. coli* BL21 with pRARE2. Successful transformants were selected by plating the cells on kanamycin (50 mg mL^{-1}) and chloramphenicol (20 mg mL^{-1}) luria broth (LB) plates. Single colonies were selected and grown in LB with kanamycin (50 mg mL^{-1}) and chloramphenicol (20 mg mL^{-1}) overnight at 37°C. On the following day the cells were then subcultured (1:1000 (v/v)) into a 2 L LB with kanamycin (50 mg mL^{-1}) and chloramphenicol (20 mg mL^{-1}). The cells were induced with 500 M of isopropyl β-D-1-thiogalactopyranoside (IPTG) at OD_{600} of 0.6 and grown overnight at 15°C, 200 rpm. The cells were harvested by centrifugation at 6000 rpm for 10 minutes at 4°C (Beckman Coulter Refrigerated Floor Centrifuge) and passed through an EmulsiFlex-C3 cell disruptor (AVESTIN, Inc.) 3 times. Cellular debris was removed by centrifuging at 20000 rpm for 30 minutes at 4°C (Beckman Coulter). The PfSir2A was then purified using gravity flow Ni-affinity chromatography (Sigma) and dialyzed into 25 mM Tris pH 8.0, 50 mM NaCl, 1 mM DTT, 10% (v/v) glycerol. The proteins were then aliquoted and kept frozen at -80°C. For crystallization, the His_6-tag of PfSir2A was removed by overnight incubation at 4°C with 30 unit mL^{-1} of thrombin (Haematologic Technologies Inc.), followed by Ni-affinity column purification to separate the undigested PfSir2A from the digested one. Then, the tag-free PfSir2A was further purified by FPLC with Superdex™ 75 column (Bio-rad), dialyzed into crystallization buffer (20 mM HEPES, pH 7.1, 20 mM NaCl, 5 mM DTT, 3% (v/v) glycerol), concentrated into 10 mg mL^{-1}, flash frozen by liquid

nitrogen, and stored at -80 °C for crystallization.

Synthesis of Acyl Peptides. Solvents were purchased from Fisher Scientific unless otherwise specified and peptide synthesis reagents and Fmoc-protected amino acids and derivatives were purchased from Creosalus Inc.

The H3K9 (NH_2-KQTARK*STGGWW-COOH) backbone was prepared using standard solid phase peptide synthesis (SPPS) at room temperature (RT). Wang resins SS (100-200 mesh, 1% DVB, 10 mmole/g) were placed into a peptide synthesis vessel along with 5 mL of anhydrous dichloromethane (DCM) for 5 hours. The first activated amino acid solution was freshly prepared with 0.32 mmoles of Fmoc-W-OH, 0.32 mmoles of O-benzotriazole-*N,N,N',N'*-tetramethyl-uronium-hexafluoro-phosphate (HBTU), 0.133 mmoles of 4-dimethylaminopyridine (DMAP), 0.64 mmoles of diisopropylethylamine (DIEA, added last) and an appropriate amount of anhydrous *N,N'*-dimethylformamide (DMF). The resin was incubated with the solution overnight at RT. The resins were then washed with DMF (5 times) before incubating with a cocktail of acetic anhydride, pyridine and DMF (2:1:3 ratio (v/v)) for 30 minutes to block remaining amino groups on the resin. Kaiser test was used to test the success of the coupling. Once coupling was confirmed, 20% (v/v) piperidine in DMF was used to remove Fmoc. All subsequent activated amino acid derivatives were freshly prepared with 0.24 mmoles of the amino acid, 0.24 mmole of HBTU, 0.21 mmole of N-hydroxybenzotriazole (HOBT), 0.48 mmole of DIEA in DMF and reaction time was 2 hours at RT.

The lysine to be modified by different acyl groups (K*) was protected with allyl carbamate (Alloc) on the side chain while the N-terminal K was protected by Boc. After all the peptide coupling steps were done on the resin, the Alloc group was

removed by incubating the resin in a cocktail of DCM, morpholine (Sigma, 2.5%) and glacial acetic acid (5%) with a 1:1 weight ratio of the original resin and tetrakis(triphenylphosphine)palladium(0) for 4 hours under nitrogen. Several washes of 0.5% (v/v) DIEA in DCM and 0.02 M of diethyldicarbamate in DMF were carried out to remove the palladium. The resin was then incubated with solutions for putting on different acyl groups. The acylation solutions contained 0.24 mmoles of fatty acids of different chain lengths (acetic anhydride, butyric acid, octanoic acid, or myristic acid), 0.24 mmole of HBTU, 0.21 mmole of HOBT, 0.48 mmole of DIEA and DMF. The resin was then washed with DMF and the peptides were cleaved with a mixture of trifluoroacetic acid (TFA), 5% (v/v) water, 5% (w/v) phenol, 2.5% (v/v) ethanedithiol and 5% (v/v) thioanisole. TFA was removed from the filtered peptide solution and the peptides were precipitated out by the addition of ether and lyophilized. The crude peptides were dissolved in water and purified by HPLC (Beckman Coulter System Gold 125P solvent module and 168 Detector) using a TARGA C18 column (250 x 20 mm, 10 μM, Higgins Analytical, Inc.) with mobile phase A (water with 0.1% (v/v) TFA) and B (acetonitrile with 0.1% (v/v) TFA) at a gradient of 20% B to 100% B in 50 minutes and a flow rate of 10 mL/min. The peptides were monitored at 215 nm and 280 nm and fractions were collected. LCMS (SHIMADZU LCMS-QP8000α with a Sprite TARGA C18 column (40 × 2.1 mm, 5 μm, Higgins Analytical, Inc.) was used to confirm the peptide mass and high purity fractions were lyophilized. The solvents used in LCMS were water with 0.1% (v/v) formic acid and acetonitrile with 0.1% (v/v) formic acid.

HPLC Assay and Kinetics. Activity of PfSir2A was determined using HPLC, by quntifying the modified and unmodified H3K9 peptide. The reaction contained 20 mM

of Tris pH 8.0, 1 mM DTT, 20 μM H3K9 modified peptide, 1 mM of NAD and 1 μM of PfSir2A and was incubated at 37^{o}C for 1 hour. The reaction was quenched with 1 volume of 10% (v/v) TFA and spun down for 10 minutes at 18,000 g (Beckman Coulter Microfuge) to separate the PfSir2A from the reaction. The supernatant was then analyzed by HPLC.

The k_{cat} and K_m values were determined using HPLC to quantify the amount of product formed with varying concentrations of the modified peptide with 1 mM of NAD, 20 mM Tris pH 8.0, 50 mM DTT, 0.5 μM of PfSir2A (butyryl, octanoyl and myristoyl H3K9 peptide), and 1 μM of PfSir2A was used for acetyl H3K9. Peptide concentration used for H3K9 acetyl and butyryl were both 2, 4, 8, 16, 32, 64, 128, and 256 μM with an incubation time of 40 and 20 minutes, respectively. Peptide concentration used for H3K9 octanoyl was 1, 2, 4, 8, 16, 32, 64, and 128 μM with an incubation time of 15 minutes. Peptide concentration used for H3K9 myristoyl was 1, 2, 3, 4, 5, 6, 8, and 16 μM with an incubation time of 1 minute. The stock solutions of the different peptides were made in different solvents. H3K9 acetyl was stored in water while the longer fatty acyl peptides, butyryl and octanoyl were stored in 1:1 (v/v) DMSO:water solutions. The myristoyl peptide is especially hydrophobic and was stored in DMSO. If only water was used to dissolve them, the peptides would stick to the plastic tubes used and led to errors in the peptide concentration. The final concentrations of DMSO in the assays varied from 0 to 10% by volume. The quenched reactions were then analyzed via HPLC using a reverse phase analytical column (Sprite TARGA C18, 40 × 2.1 mm, 5 μm, Higgins Analytical, Inc.) with a 0% to 70% B gradient in 8 minutes at 1 mL/min. The acetyl peptide has a very close retention time to the unmodified peptide and a different column was used to separate the peaks

(Kinetex XB-C18 100A, 75x 4.60 mm, 2.6 um, Phenomenex). The product peak and the substrate peaks were both quantified and converted to initial rates, which were then plotted against the modified peptide concentration and fitted using the Kaleidagraph program.

^{32}P-NAD Assay. The reaction contained 50 mM Tris pH 8.0, 150 mM NaCl, 10 mM DTT, 60 μM H3K9 modified peptide, 0.1 μCi of ^{32}P-NAD (American Radiolabeled Chemicals, ARP 0141-250 μCi) and 1 μM of Pfsir2A and was incubated at 37^{o}C for 1 hour. *P.* falciparum whole cell lysate (100 μL) was first denatured with 6 M of Urea, 15 mM of DTT at 37^{o}C for 15 minutes. Then it was alkylated with 50 mM of iodoacetamide in the dark at RT for 1 hour. The solution was then diluted so that the final concentration of urea was less than 0.75 M and digested with 0.1 μg/μL trypsin and 50 mM Tris pH 7.4, and 1 mM $CaCl_2$ overnight at 37^{o}C. The digest was quenched with 10% TFA to a final pH of 2 to 3 and desalted with a Waters C18 Sep-Pak column. The peptides were eluted 5 times with 1 mL of 90% ACN/0.1%TFA and lyophilized. The peptides were then solubilized in 50 μL of water and 1 μL was used in the ^{32}P-NAD assay under conditions described above. The reaction was incubated at 37^{o}C for 15 minutes and 1 μL of the reaction mixture was spotted onto a polyester backed silica plate (100 μm thick, Waters). After the spots were dried, the plate was run for 6 cm in 30:70 (v/v) 1M ammonium bicarbonate:95% ethanol. The plate was dried and exposed overnight in a PhosphorImaging screen (GE Healthcare). The signal was detected using a STORM860 phosphorimager (GE Healthcare).

Crystallization, Data Dollection, and Structural Refinement. PfSir2A was mixed with H3K9 myristoyl peptide at the protein:peptide molar ratio of 1:10, diluted into 3 mg mL^{-1} with crystallization buffer, and incubated on ice for 30-60 minutes. Crystals were

grown at room temperature with hanging drop vapor diffusion method at the condition of 16% (w/v) PEG 3350, 0.1 M NaF, 7% (v/v) formamide. PfSir2A-H3K9 myristoyl co-crystals were soaked in the cryoprotectant solution (18% (w/v) PEG 3350, 0.1 M NaF, 10% (v/v) formamide, 15% (v/v) glycerol) with 10 mM NAD for 2-10 minutes at room temperature immediately before data collection. All data were collected at CHESS (Cornell High Energy Synchrotron Source) F2 station. The data were processed using the programs HKL2000 (*24*). Using the program Molrep from the CCP4 suite of programs (*25*), the structures were solved by molecular replacement with PfSir2A-AMP structure (PDB code: 3JWP) as the search template. Refinement and model building were performed with REFMAC5 and COOT from CCP4. The X-ray diffraction data collection and structure refinement statistics were shown in Table 4.1.

Atomic coordinates and structure factors were deposited in the Protein Data Bank under accession codes 3U3D and 3U31 for PfSir2A-myrH3K9 and PfSir2A-myrH3K9-NAD, respectively.

Table 4.1. Crystallographic data collection and refinement statistics.

	PfSir2A-myrH3K9	PfSir2A-myrH3K9-NAD
Data collection		
Space group	$P2_12_12$	$P2_12_12$
Cell dimensions		
a, *b*, *c* (Å)	32.51, 103.33, 105.51	32.17, 102.73, 105.18
α, β, γ (º)	90, 90, 90	90, 90, 90
Resolution (Å)	50-2.40	50-2.20
R_{sym} or R_{merge} (%)	14.9 (75.3)	12.9 (74.8)
$I/\sigma I$	23.83 (1.75)	24.7 (2.21)
Completeness (%)	98.5 (97.2)	99.9 (99.9)
Redundancy	8.2 (4.4)	6.7 (5.0)
Refinement		
Resolution (Å)	50-2.40	50-2.20
No. reflections	18972	23688
R_{work} / R_{free} (%)	22.38 /27.61	20.24/24.65
No. of protein residues	263	263
No. of ligand/ion		
Myristoyl	1	1
NAD	--	1
Glycerol	1	1
Zn	1	1
No. of water	25	75
R.m.s deviations		
Bond lengths	0.047	0.024
Bond angles (º)	2.10	2.07

Numbers showed in the parentheses are for the highest resolution shell.

4.3 Results

The PfSir2A gene was de novo synthesized. The protein was expressed in *E. coli* and affinity purified to near homogeneity. For substrates, we synthesized histone H3 peptides bearing acetyl, butyryl, octanoyl, and myristoyl groups on Lys9. To facilitate the detection of the peptides by ultra-violet (UV) light absorption, two Trp residues were added to the C-terminal of the peptides. A high-pressure liquid chromatography (HPLC) assay was used to monitor the activity of PfSir2A on these different acyl peptides. Interestingly, all four acyl peptides could be hydrolyzed (Fig. 4.1). The butyryl, octanoyl and myristoyl peptides could be hydrolyzed more efficiently than the acetyl peptide. The myristoyl peptides appeared to be hydrolyzed most efficiently.

To quantitatively compare the activity of PfSir2A on different acyl peptides, kinetic studies were carried out. The kinetics data (Table 4.2) suggested that acetyl H3K9 peptide was the least efficient substrate among the four acyl peptides tested, with a k_{cat}/K_m of 26 $s^{-1}M^{-1}$. The k_{cat}/K_m value for deacetylation was comparable to that reported by Sauve and coworkers (*9*). The butyryl, octanoyl, and myristoyl peptides were hydrolyzed with much higher catalytic efficiencies. In particular, the catalytic efficiencies for the hydrolysis of myristoyl peptide were more than 300-fold higher than that for the hydrolysis of acetyl peptide. The K_m value for the myristoyl peptide was lower than 1μM (PfSir2A was saturated with 2 μM of the myristoyl peptide. The K_m value could not be accurately determined because of the detection limit at low substrate concentrations). The enzymology data demonstrated that PfSir2A preferentially hydrolyzes medium and long chain fatty acyl lysine.

To understand the structural basis for PfSir2A' preference for longer chain fatty acyl groups, we co-crystallized PfSir2A with an H3K9 myristoyl peptide to generate

Figure 4.1. PfSir2A could hydrolyze medium and long chain fatty acyl lysine more efficiently than acetyl lysine. (A) Overlaid HPLC traces showing PfSir2A-catalyzed hydrolysis of different fatty acyl lysine peptides. Acyl peptides were used at 20 μM, PfSir2A at 1 μM, and NAD at 500 μM. The corresponding synthetic peptide without any acyl lysine modification (H3K9WW unmodified) was used as the control to indicate the position of the hydrolysis product formed.

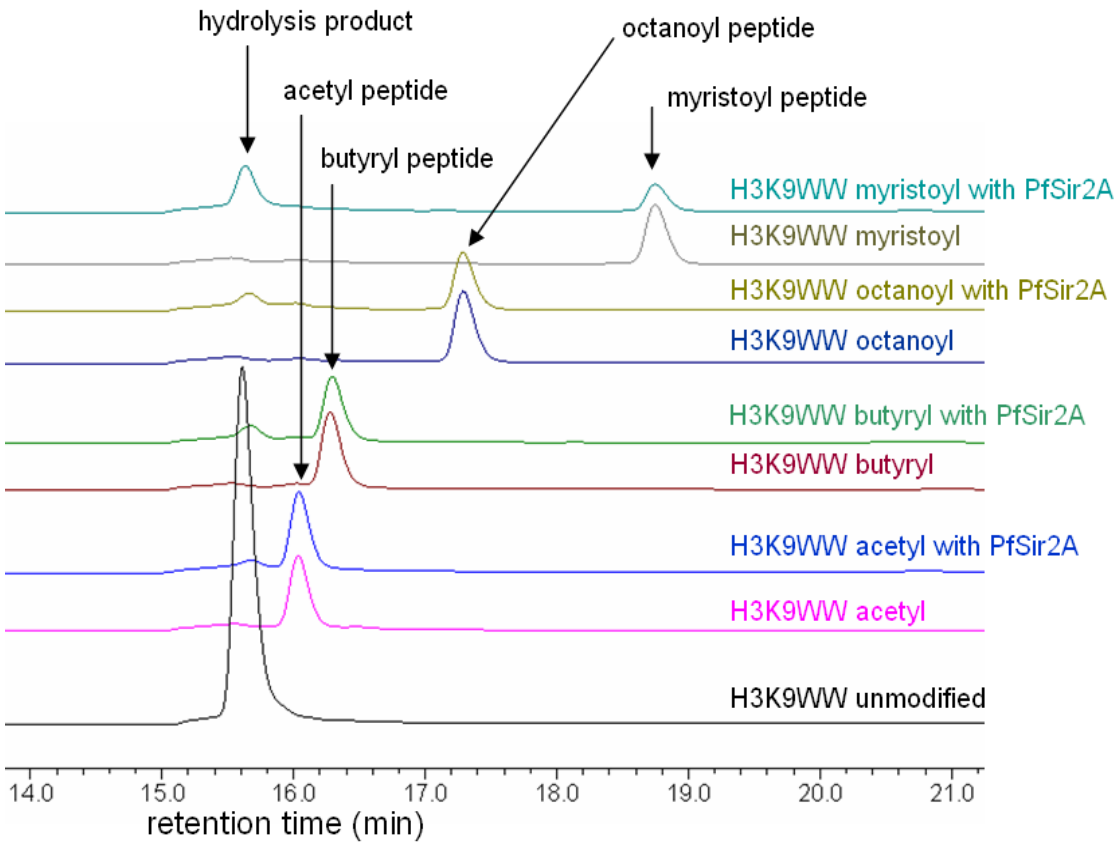

hydrolysis product
acetyl peptide
butyryl peptide
octanoyl peptide
myristoyl peptide
H3K9WW myristoyl with PfSir2A
H3K9WW myristoyl
H3K9WW octanoyl with PfSir2A
H3K9WW octanoyl
H3K9WW butyryl with PfSir2A
H3K9WW butyryl
H3K9WW acetyl with PfSir2A
H3K9WW acetyl
H3K9WW unmodified
14.0
15.0
16.0
17.0
18.0
19.0
20.0
21.0
retention time (min)

Table 4.2. Kinetics data for PfSir2A on different acyl peptides.

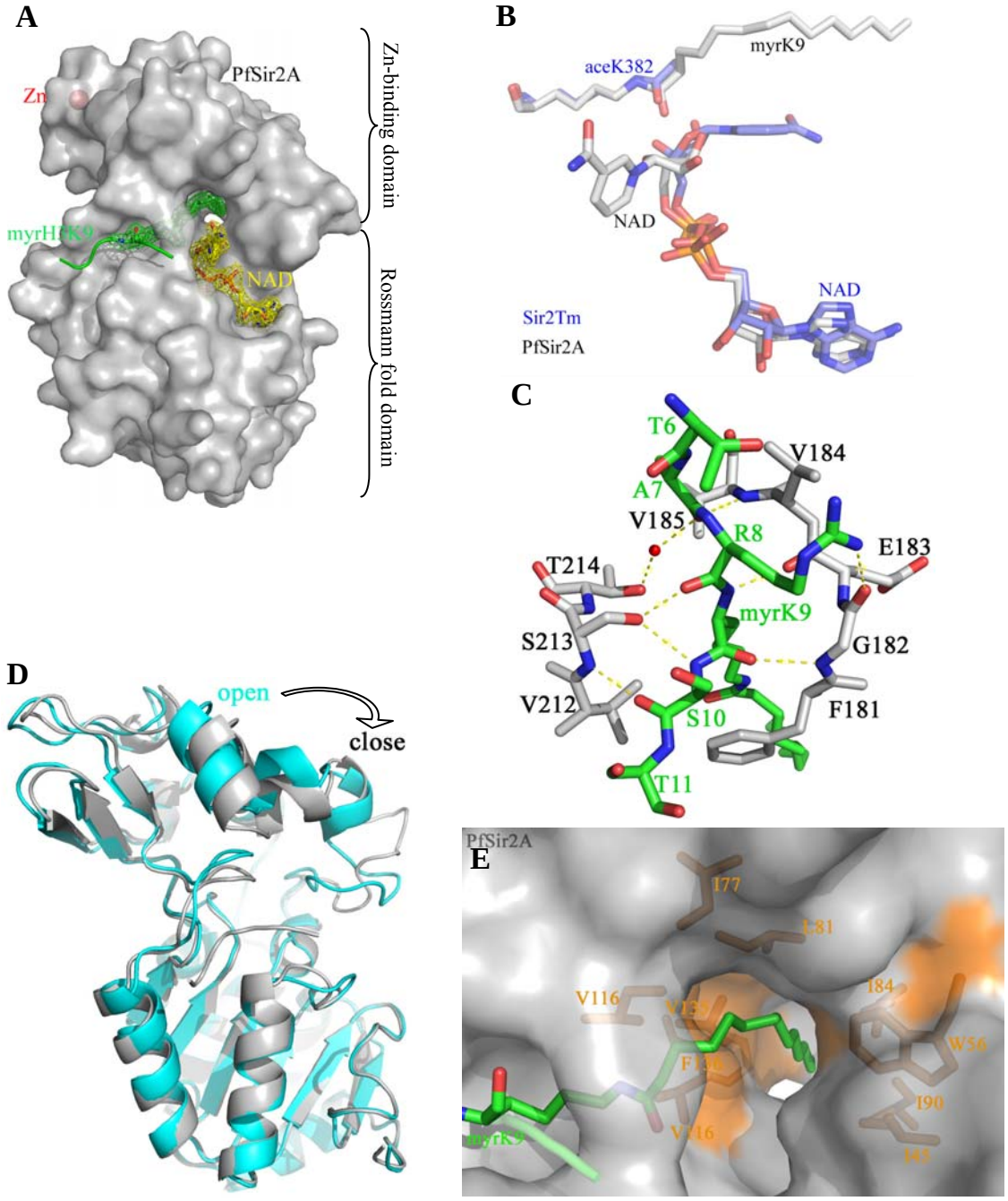
substrate	k_{cat} (s^{-1})	K_m **for peptide** (μM)	k_{cat}/K_m ($s^{-1}M^{-1}$)
H3K9 acetyl	0.001 ± 0.0002	39 ± 9	2.6×10^{1}
H3K9 butyryl	0.001 ± 0.0002	8 ± 1	1.6×10^{2}
H3K9 octanoyl	0.001 ± 0.004	1.2 ± 0. 3	9.2×10^{2}
H3K9 myristoyl	0.01± 0.002	<1.0[a]	$>1.0 \times 10^{4}$

[a]. The K_m value cannot be accurately determined due to the detection limit when substrate concentration was lower than 1 μM.

the PfSir2A-H3K9 myristoyl complex. The co-crystal was then soaked briefly in an NAD solution to obtain a ternary complex of PfSir2A with H3K9 myristoyl peptide and NAD. The structures were solved using molecular replacement with the deposited PfSir2A structure PDB 3JWP as the search model. The overall structure of PfSir2A was similar to other sirtuins, containing a small Zn-binding domain and a large Rossmann fold domain (Fig. 4.2A) (15-18). The two substrates, H3K9 myristoyl peptide and NAD, bound to the clefts between the two domains. This binding mode of the two substrates was similar with the reported ternary complex structures of other sirtuins. For instance, the peptide substrates of the *Thermotoga maritima* Sir2 (Sir2Tm, PDB 2H4F, one of the first sirtuin ternary complex structureswith both NAD and acetyl peptide bound) and PfSir2A superimposed well (Fig. 4.2B) (19). The interactions between PfSir2A and H3K9 myristoyl peptide mainly came from main chain hydrogen bonds (Fig. 4.2C), in agreement with other studies of sirtuins (19). Compared with the structure without any acyl peptide bound (PDB 3JWP), the binding of H3K9 myristoyl peptide to PfSir2A caused the Zn-binding domain to rotate clockwise to the Rossmann fold domain, so that PfSir2A moved from an open state to a close state which is similar to that observed in Sir2Tm (Fig. 4.2D) (19). However, different from Sir2Tm, PfSir2A had a long open hydrophobic tunnel that accommodated the myristoyl group (Fig. 4.2E). The hydrophobic tunnel was surrounded by several hydrophobic residues (Ile45, Trp56, Ile77, Ile80, Ile84, Ile90, Val116, Val135, Phe136, Ile178, and Leu181). This structure feature suggested that PfSir2A was optimized for recognizing fatty acyl groups.

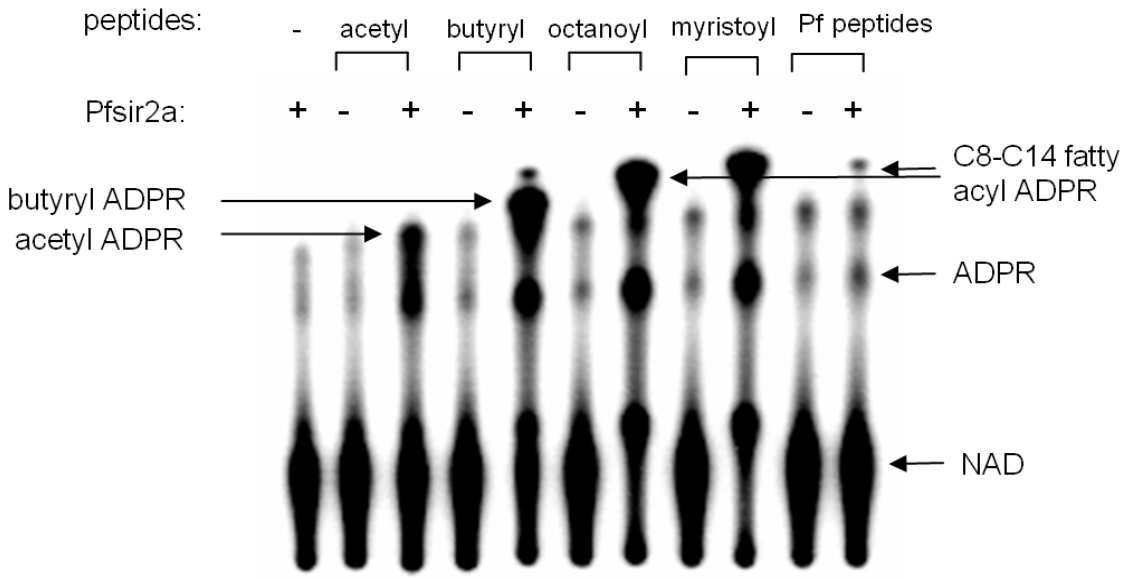
The enzymology and structural data led to the hypothesis that PfSir2A may function to remove medium and long chain fatty acyl groups in malaria parasite.

Figure 4.2. Structural basis for the recognition of myristoyl lysine by PfSir2A. (A) Overall structure of PfSir2A. The Fo-Fc map at 1.6σ shows the H3K9 myristoyl peptide (green) and NAD (yellow) at the active site. (B) The structural alignment between Sir2Tm (blue) and PfSir2A (grey). The positions of H3K9 myristoyl peptide and NAD in PfSir2A were similar to the positions of acetyl peptide and NAD in Sir2Tm. (C) Hydrogen bonding interactions between the H3K9 myristoyl peptide (green) and PfSir2A (grey). (D) Structural alignment between PfSir2A-AMP (cyan, PDB code: 3JWP) and PfSir2A-myrH3K9 showed that the binding of the substrate peptide myrH3K9 drove PfSir2A from an inactive open state to an active close state. (E) PfSir2A had a long open hydrophobic tunnel which accommodated fatty acyl groups. PfSir2A surface representation: grey; myristoyl lysine: green; hydrophobic residues: orange.


A
PfSir2A
Zn
myrH3K9
NAD
Zn-binding domain
Rossmann fold domain
B
myrK9
aceK382
NAD
NAD
Sir2Tm
PfSir2A
C
T6
V184
A7
V185
R8
E183
T214
myrK9
S213
G182
V212
S10
F181
T11
D
open
close
E
PfSir2A
I77
L81
V116
I84
W56
I90
I43
myrK9

Protein lysine myristoylation has been reported to occur on several mammalian proteins (20, 21). To test whether malaria parasite proteins have medium or long chain fatty acyl modifications on lysine residues, a sensitive assay was developed to detect the presence of fatty acyl lysine in malaria parasites. With the use of ^{32}P-NAD, the formation of fatty acylated ADP-ribose (ADPR) in PfSir2A-catalyzed defatty acylation of synthetic acyl peptides can be detected after thin-layer chromatography (TLC) separation and autoradiography. Longer chain fatty acyl ADPR species were more hydrophobic and thus moved faster than shorter chain fatty acyl ADPR species (Fig. 4.3). Notably, most NAD molecules were consumed when octanoyl and myristoyl peptides were incubated with PfSir2A, while there were still NAD molecules left when acetyl and butyryl peptides were incubated with PfSir2A. This observation confirmed the kinetic studies that octanoyl and myristoyl peptides were more efficient substrate for PfSir2A. When total protein extracts of malaria parasites were incubated with ^{32}P-NAD and PfSir2A, the formation of a longer chain fatty acyl ADPR was detected. The position of the fatty acyl ADPR was similar to that formed the reactions with synthetic octanoyl and myristoyl peptides, suggesting that fatty acyl group on *P. falciparum* proteins should have a similar chain length. The intensity of the fatty acyl ADPR spot was weak, suggesting that the concentration of the fatty acyl peptide in our *P. falciparum* protein extract was low. However, we could repeatedly detect this spot using the ^{32}P-NAD assay. In addition, compared with the negative control without PfSir2A, the intensity for the acetyl ADPR spot did not increase. Therefore, PfSir2A's deacetylase activity could not be detected using *P. falciparum* protein extract, but the activity of removing longer fatty acyl groups could be detected.

Figure 4.3 ^{32}P-NAD assay could detect the presence of medium or long chain fatty acyl lysine modifications on *P. falciparum* proteins. PfSir2A were incubated with ^{32}P-NAD and synthetic peptides bearing different acyl modifications. Negative controls were reactions without PfSir2A or peptides. The reactions were resolved by TLC and detected by autoradiography. With *P. falciparum* peptides (last two lanes), the acyl ADPR spot formed was similar to the C8-C14 acyl ADPR, suggesting that such fatty acyl groups were present and could be removed by PfSir2A.


peptides:
-
acetyl
butyryl
octanoyl
myristoyl
Pf peptides
Pfsir2a:
+ - + - + - + - + - +
butyryl ADPR
acetyl ADPR
C8-C14 fatty
acyl ADPR
ADPR
NAD

4.4 Discussion

In summary, our enzymology and structural data demonstrated that PfSir2A was more efficient at removing medium and long chain fatty acyl groups than acetyl groups from peptides. Although it is known that other sirtuins can also hydrolyze propionyl and butyryl lysine, but the activity is typically weaker than the hydrolysis of acetyl lysine (22, 23). Therefore, our work demonstrates for the first time that longer fatty acyl lysines can be the preferred substrate for a sirtuin. The biochemical data suggest that *P. falciparum* proteins contain such fatty acyl lysine modifications, which can be removed by PfSir2A *in vitro*. The data imply that the biological function of PfSir2A may be achieved by its activity of removing medium and long chain fatty acyl groups. The detailed structures of the fatty acyl groups and the abundance of such modifications in comparison to the well-known acetyl lysine modification await future studies. The finding that PfSir2A could remove longer fatty acyl groups suggests that other sirtuins, especially those that have weak or no deacetylase activity, may also have this activity. sirtuins should therefore be called “NAD-dependent deacylases”, instead of “NAD-dependent deacetylases”. The discovery of a robust activity for PfSir2A will also facility the development of PfSir2 inhibitors, which may have therapeutic value in treating malaria.

References

1. Imai, S.-i., Armstrong, C. M., Kaeberlein, M., and Guarente, L. (2000) *Nature* **403**, 795-800.
2. Sauve, A. A., Wolberger, C., Schramm, V. L., and Boeke, J. D. (2006) *Annu. Rev. Biochem.* **75**, 435-465.
3. Imai, S.-i., and Guarente, L. (2010) *Trends in Pharmacological Sciences* **31**, 212-220.
4. Haigis, M. C., and Sinclair, D. A. (2010) *Annu. Rev. Pathol.* **5**, 253-295.
5. Frye, R. A. (2000) *Biochem. Biophys. Res. Commun.* **273**, 793-798.
6. Freitas-Junior, L. H., Hernandez-Rivas, R., Ralph, S. A., Montiel-Condado, D., Ruvalcaba-Salazar, O. K., Rojas-Meza, A. P., Mâncio-Silva, L., Leal-Silvestre, R. J., Gontijo, A. M., Shorte, S., and Scherf, A. (2005) *Cell* **121**, 25-36.
7. Tonkin, C. J., Carret, C. l. K., Duraisingh, M. T., Voss, T. S., Ralph, S. A., Hommel, M., Duffy, M. F., Silva, L. M. d., Scherf, A., Ivens, A., Speed, T. P., Beeson, J. G., and Cowman, A. F. (2009) *PLoS Biol.* **7**, e1000084.
8. Merrick, C. J., and Duraisingh, M. T. (2007) *Eukaryotic Cell* **6**, 2081-2091.
9. French, J. B., Cen, Y., and Sauve, A. A. (2008) *Biochemistry* **47**, 10227-10239.
10. Kowieski, T. M., Lee, S., and Denu, J. M. (2008) *J. Biol. Chem.* **283**, 5317-5326.
11. Du, J., Jiang, H., and Lin, H. (2009) *Biochemistry* **48**, 2878-2890.
12. Chen, Y., Sprung, R., Tang, Y., Ball, H., Sangras, B., Kim, S. C., Falck, J. R., Peng, J., Gu, W., and Zhao, Y. (2007) *Mol. Cell. Proteomics* **6**, 812-819.
13. Garrity, J., Gardner, J. G., Hawse, W., Wolberger, C., and Escalante-Semerena, J. C. (2007) *J. Biol. Chem.* **282**, 30239-30245.

14. Liu, B., Lin, Y., Darwanto, A., Song, X., Xu, G., and Zhang, K. (2009) *J. Biol. Chem.* **284**, 32288-32295.

15. Finnin, M. S., Donigian, J. R., and Pavletich, N. P. (2001) *Nat. Struct. Mol. Biol.* **8**, 621-625.

16. Zhao, K., Chai, X., Clements, A., and Marmorstein, R. (2003) *Nat Struct Mol Biol* **10**, 864-871.

17. Avalos, J. L., Celic, I., Muhammad, S., Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2002) *Mol. Cell* **10**, 523-535.

18. Min, J., Landry, J., Sternglanz, R., and Xu, R.-M. (2001) *Cell* **105**, 269-279.

19. Hoff, K. G., Avalos, J. L., Sens, K., and Wolberger, C. (2006) *Structure* **14**, 1231-1240.

20. Stevenson, F. T., Bursten, S. L., Locksley, R. M., and Lovett, D. H. (1992) *J. Exp. Med.* **176**, 1053-1062.

21. Stevenson, F. T., Bursten, S. L., Fanton, C., Locksley, R. M., and Lovett, D. H. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7245-7249.

22. Smith, B. C., and Denu, J. M. (2007) *J. Biol. Chem.* **282**, 37256-37265.

23. Bheda, P., Wang, J. T., Escalante-Semerena, J. C., and Wolberger, C. (2011) *Protein Science* **20**, 131-139.

24. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 472-494.

25. Collaborative. (1994) *Acta Crystallogr. D Biol. Crystallogr.* **50**, 760-763.

Chapter 5

Prenylation and Membrane Localization of Cdc42 are Essential for Activation by DOCK7

5.1 Introduction

Rho family of GTPases plays pivotal roles in many cellular processes, including cell polarity, motility, vesicular trafficking, cell-cycle progression, and gene expression (1-4). Their functions largely depend on their cellular localization, which are determined by the posttranslational modification of the C-terminal CAAX motif (where C represents cysteine, A is any aliphatic amino acid, and X is any amino acid) (5, 6). The prenyltransferase catalyzes the modification of CAAX motifs and adds a farnesyl or geranylgeranyl isoprenoid lipid tail to the cysteine residue (7). For instance, both Cdc42 and Rac1 are geranylgeranylated. The attached lipid tail of the GTPase inserts into the cellular membrane and brings the GTPase to the subcellular compartments where it engages its regulators: GEFs (Guanine nucleotide Exchange Factors), GAPs (GTPase Activating Proteins), and GDIs (Guanine nucleotide Dissociation Inhibitors). GEFs promote the exchange of GDP for GTP, therefore activating Rho GTPases. GAPs enhance the hydrolysis of GTP and inactivate GTPases. Rho GDIs sequester GDP-bound Rho GTPases in the cytosol by inhibiting nucleotide exchange (8-10).

Two families of GEFs have been identified for Rho GTPases: the classical Dbl family (11) and the recently discovered DOCK180 family (12-14). The Dbl family consists of ~70 members, all of which contain two conserved domains, the Dbl homology (DH) and the pleckstrin homology (PH) domain (15-18). The DH domain is the catalytic domain for nucleotide exchange, while the PH domain has been

demonstrated to interact with the plasma membrane (19, 20). The mammalian DOCK180 family contains 11 members, DOCK1-11 (with DOCK180 as DOCK1) (21, 22). One of the major features of the DOCK180 family is that all members possess two DOCK homology regions, DHR1 and DHR2, which share no sequence similarity with the PH and DH domains of the Dbl family, respectively. The DHR1 domain shares low homology with the C2 motif and associates with the cellular membrane (23). The DHR2 domain is necessary and sufficient to exchange GDP for GTP on GTPases (22). Recently, our laboratory has identified a limit C-terminal portion of DHR2, designated DHR2c, which exhibited full GEF activity (24).

DOCK180 family members have been shown to activate Rac1 and/or Cdc42, but not RhoA. These are classified into four subfamilies: DOCK-A to DOCK-D subfamilies. The DOCK-A subfamily contains DOCK1, 2, and 5, while the DOCK-B subfamily consists of DOCK3 and DOCK4, all of which are Rac1-specific GEFs. DOCK9-11 comprise the DOCK-D subfamily and are Cdc42 specific GEFs. DOCK6-8 fall within the DOCK-C subfamily, which are capable of activating both Rac1 and Cdc42 (25). The Barford group has solved two complex structures: Cdc42 with DOCK9-DHR2 and Rac1 with DOCK2-DHR2, which not only elucidated the mechanism by which DOCK180 family members activate Rho GTPases but also shed light on the specificity of DOCK180 proteins and their points of contact (26, 27). Two residues at positions 27 and 56 of Rac1 and Cdc42 are key determinants of specificity for nucleotide exchange by DOCK180 proteins. Wu *et al*. have shown that residues from the α10 helix of DOCK180 proteins play a key role in the selectivity of Rac1 or Cdc42 (24).

DOCK-C subfamily members have been proposed to activate both Rac1 and

Cdc42. However, using traditional *in vitro* GEF assays, no robust GEF activity has been observed for DOCK-C members, such as DOCK6 and DOCK7 (27-30). Therefore it is of interest to investigate whether DOCK7 is a Rac1-specific or Cdc42-specific GEF. *In vivo* the DOCK-C members are important in the regulation of neuronal cell development. DOCK6 has been reported to regulate the early development of neurite formation of neuron cells (28, 31). In addition, DOCK7 has been demonstrated to be involved in axon formation by activating Rac1 and Cdc42 (29, 30). Also, DOCK7 was shown to interact with the TSC1/TSC2 complex, which implies that DOCK7 may act as a regulator of the mTOR complex (32).

To better understand the molecular basis of DOCK7's function, I set out to investigate the biochemical characteristics of DOCK7. Here I present data to show that DOCK7 showed undetectable GEF activity toward non-prenylated Rac1 or Cdc42 in solution, but exhibited robust activation of prenylated Rac1 and Cdc42 in a model liposome system. In addition, I identify two key residues in DOCK7 which convey GTPase exchange specificity and the mutations of these two residues shift the DOCK7-DHR2 activity profile. Additionally, I show that DOCK7 possesses a distal site that is distinct from the nucleotide exchange site and preferentially binds to the active forms of Rac1 and Cdc42, which helps to recruit DHR2 to the membrane surface to enhance the rate of GTPase nucleotide exchange.

5.2 Experimental Procedures

Plasmid Constructs. The coding region of DOCK7 (accession No. from EMBL: DQ118679) was ligated into the Bac-to-Bac baculovirus expression vector pFastBac^{TM}HT C (Invitrogen). The cDNAs encoding the DOCK truncation constructs: DHR2 (amino acid 1433-1992), DHR2s (amino acid 1530-1970), and DHR2c (amino acid 1688-1992), were inserted into pFastBac^{TM}HT C vector for insect cell expression. Also, DHR2 and DHR2s were cloned into ppSUMO vector for *E.coli* expression. DHR2c was inserted into pET28b for *E.coli* expression. DOCK7-DHR2c mutants were generated with the QuickChange site-directed mutagenesis kit (Stratagene). The full-length Cdc42 and Rac1 were cloned into the pFastBac^{TM}HT C, pGEX-4T-1, and pET28a vectors.

Preparation of Insect Cell Expressed Proteins. All the target proteins cloned into pFastBac^{TM}HT C vector were expressed in *Spodoptera frugiperda* (Sf21) cells. The expression of Cdc42 and Rac1 were carried out at Kinnakeet Biotechnology (Midlothian, VA), and the purification was performed as described previously (10). All the DOCK7 constructs were expressed in 0.5 L Sf21 cell suspension culture for 3 days after virus infection. The cells were pelleted down by centrifuging at 500g at 4°C for 5 minutes and stored at -80°C for purification. The cell pellet was suspended in 20 mL lysis buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 2% Triton X-100) with protease inhibitor cocktail (Roche) and disrupted by Dounce homogenization. The lysate was spun down at 9000g for 20 minutes at 4 °C. The supernatant was incubated with Ni-NTA agarose beads (Qiagen) for 30 minutes at 4°C. The beads were washed with 20 volumes of washing buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 0.1% CHAPS, 30 mM imidazole), and protein was eluted with 10 mL elution buffer (50 mM Tris, pH

8.0, 150 mM NaCl, 0.1% CHAPS, 200 mM imidazole). The fraction containing the target protein was concentrated to 3 mL.

Preparation of E.coli Expressed Proteins. All the DOCK7 constructs which were cloned either into the pET28 or ppSUMO vector, and the Cdc42 and Rac1 constructs which were cloned into either pGEX or pET28 vector, were expressed in *Escherichia coli*. A single colony of *E.coli* BL21(DE3) or Rosetta 2 strain containing the target plasmid was inoculated in 5 mL LB medium with 50 μg/mL kanamycin or 100 μg/mL carbenicillin (RPI) overnight at 37°C. The overnight culture was subsequently transferred to 1 L 2xYT medium with antibiotic and inoculated at 37 °C until OD600 reached 0.6, followed by induction with 0.1 mM isopropyl 1-thio-β-D-galactopyranoside (IPTG) (RPI) at room temperature for overnight. Bacteria were harvested by centrifugation at 4000g for 10 minutes at 4 °C and stored at -80 °C for future purification. The His_6-tagged DOCK7 pellets were suspended in lysis buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 20 mM imidazole) with protease inhibitor cocktail (Roche), and lysed by sonication on ice. The lysates were centrifuged at 20000 g for 30 minutes at 4 °C, and the resulting supernatants were incubated with Ni-NTA agarose beads for 30 minutes at 4°C. The beads were extensively washed with washing buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 40 mM imidazole), and the proteins were eluted with 20 mL elution buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 200 mM imidazole). The proteins were then concentrated in 50 mM Tris, pH 8.0, 50 mM NaCl. For His_6-tagged Cdc42 and Rac1, bacterial pellets were suspended in 25 mL lysis buffer (20 mM Tris, pH 8.0, 500 mM NaCl, 5 mM $MgCl_2$, 100 μM GDP, 20 mM imidazole) with protease inhibitor cocktail (Roche), and lysed for 1 hour incubation on ice with 150 mg lysozyme (Sigma), 45 mg deoxycholic acid (Sigma),

and DNase I powder (Roche). The following purification procedure was the same as for His_6-tagged DOCK7 proteins, except that the buffers contained 5 mM $MgCl_2$. For GST-tagged Cdc42 and Rac1, bacterial pellet was suspended in 25 mL lysis buffer (50 mM Tris, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 100 μM GDP, 1 mM DTT) with protease inhibitor cocktail, and lysed for 1 hour as described above. The lysate was spun down and the supernatant was incubated with glutathione sepharose beads (Amersham Bioscience) at 4 °C for 1 hour. The beads were washed with washing buffer (50 mM Tris, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT). The protein was eluted with 20 mL elution buffer (50 mM Tris, pH 8.0, 10 mM reduced glutathione, 5 mM $MgCl_2$). Both His_6- and GST- tagged Cdc42 and Rac1 were changed to storage buffer (50 mM Tris, pH 7.5, 50 mM NaCl, 2 mM $MgCl_2$, 0.5 mM DTT), concentrated to 200 μM, flash froze by liquid nitrogen, and stored at -80 °C for assays.

Preparation of Liposomes. The lipids used in this study were purchased from Avanti Polar Lipids. The model liposomes contained 35% phosphatidylethanolamine, 35% cholesterol, 25% phosphatidylserine, and 5% phosphatidylinositol. Two approaches were used to prepare liposomes for different purposes. Rapid solvent exchange was used to make large liposomes that can be pelleted by low speed centrifugation (33). For *in vitro* GEF assays, small lipid vesicles with 1 μm diameter were prepared by extrusion using Avanti mini-extruder.

Nucleotide-loading of Cdc42 and Rac1. Cdc42-MantGDP, Rac1-MantGDP, Cdc42-GTPγS, and Rac1-GTPγS were prepared by incubating insect cell-expressed prenylated Cdc42 or Rac1 with 40-fold excess MantGDP or 10-fold excess GTPγS in the buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.1% CHAPS) with 10

mM EDTA for 1.5 hour on ice, followed by the addition of 20 mM $MgCl_2$ and a 0.5 hour incubation to quench the excess EDTA. The excess nucleotide was removed by Ni-column purification as described above. The eluted nucleotide-loaded Cdc42 and Rac1 were changed to the buffer (20 mM Tris, pH 8.0, 150 mM NaCl, 5 mM $MgCl_2$, 0.1% CHAPS) and concentrated to 20 μM.

In vitro GEF Assays. Nucleotide exchange on Cdc42 or Rac1 was monitored by the fluorescence change of Mant nucleotide with an exciting wavelength of 360 nm and an emission wavelength of 440 nm. Using a Varian Clary Eclipse fluorimeter, all samples were continuously stirred at 25 °C in TBSM buffer (50 mM Tris, pH 7.5, 50 mM NaCl, 5 mM $MgCl_2$) with 1 μM Mant nucleotide. For non-prenylated Cdc42 or Rac1, GST-Cdc42 or GST-Rac1 was added to the cuvette followed by the addition of DOCK7-DHR2s or DHR2c. For prenylated Cdc42 or Rac1, His_6-tagged Cdc42 or Rac1 was first incubated with liposomes for 15 minutes at room temperature, then added to a cuvette to make a final liposome concentration of 20 μM. DOCK7-DHR2s or DHR2c was added to initiate the exchange. To check whether the active form of Cdc42 or Rac1 had an effect on the activation of Cdc42 or Rac1 by DOCK7, DHR2s or DHR2c was incubated with Cdc42-GTPγS or Rac1-GTPγS for 15 minutes at room temperature before it was added to the reaction.

Liposome Centrifugation Assays. To assay the binding preference of DOCK7 to the three different forms of prenylated Cdc42 or Rac1 (nucleotide-free, GDP bound, GTP bound), 0.2 nmol insect cell expressed Cdc42 or Rac1 was incubated with 100 μL of 1 μM lipids prepared by rapid solvent exchange for 10 minutes at room temperature, loaded with no nucleotide, GDP, or GTPγS, and centrifuged at a maximum speed on a table top centrifuge for 10 minutes. The supernatant was removed and lipid pellet was

resuspended in 40 μL TBSM buffer, followed by a 30 minutes incubation at room temperature with 160 nmol DOCK7-DHR2s or DHR2c. The mixture was spun down, and the supernatant and lipid pellet were examined by SDS-PAGE to check the partitioning of DOCK7 between supernatant and pellet. The negative control only contained lipids and either DOCK7-DHR2s or DHR2c (without Cdc42 or Rac1).

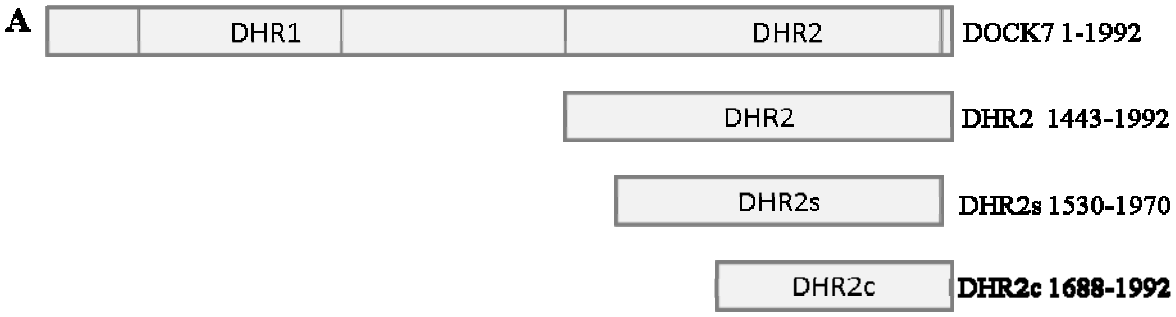
GST-Cdc42/Rac1 Pull-down Assays. To check the binding of DOCK7 to non-prenylated Cdc42 or Rac1, 0.4 nmol *E.coli* expressed GST-Cdc42 or GST-Rac1 was prebound to 15 μL glutathione sepharose beads in the presence of 10 mM EDTA. The negative control only contained beads and either DOCK7-DHR2s or DHR2c (without GST-Cdc42 or GST-Rac1). To prepare GST-Cdc42 or GST-Rac1 preloaded with GDP or GTPγS, excess GDP or GTPγS was added to the samples and incubated for 15 minutes at room temperature, followed by the addition of 20 mM $MgCl_2$ and 320 nmol DOCK7-DHR2s or DHR2c, and incubated for 1 hour at 4 °C. The beads were spun down, washed with TBSM for 3 times, and examined by SDS-PAGE.

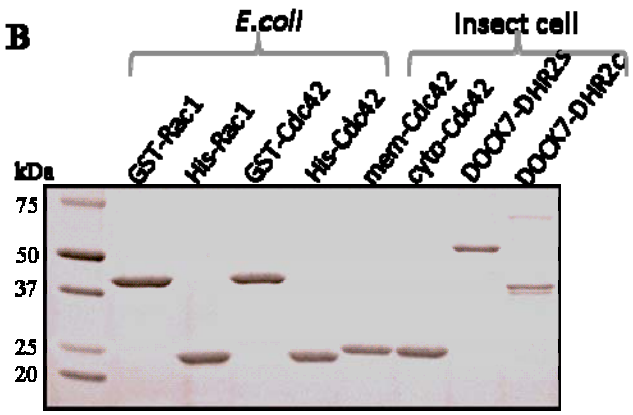
5.3 Results

All the DOCK180 family members have a C-terminal domain of approximately 500 amino acids, termed the DHR2 domain, which is necessary and sufficient for GEF activity (22). Other studies have demonstrated that the recombinant protein of the full length DHR2 domain of DOCK180 or DOCK9 was not stable in solution, whereas a slightly shorter version of DHR2 was sufficiently stable for GEF assays and crystallization (24, 26). Therefore, we designed four different constructs of DOCK7 for overexpression in insect cells to characterize the biochemical features of DOCK7 (Fig. 5.1A). The full length DOCK7 contains 1992 residues that includes 22 residues beyond the DHR2 domain. These C-terminal 22 residues were included in the construct designated DHR2. DHR2s lacked the N-terminal 87 residues of DHR2 as well as the C-terminal 22 residues. DHR2c was designed to test whether it represted a limit functional domain as in the case of DOCK180. The DHR2s and DHR2c constructs were successfully expressed and purified in insect cells (Fig. 5.1B). As previously described, GST-tagged Rac1 and Cdc42 recombinant proteins were used for the GEF assays (Fig. 5.1B). We also expressed His_6-tagged Rac1 and Cdc42 in *E. coli* as the counterparts of His_6-tagged Rac1 and Cdc42 expressed in insect cells (Fig. 5.1B). The Cdc42 and Rac1 present in the membrane fraction of insect cells are prenylated whereas the GTPases present in the cytosolic fractions are not prenylated (Fig. 5.1B). The prenylated recombinant Cdc42 has been shown to bind to both insect cell membranes and reconstituted model liposomes (10).

We first tested whether DOCK7-DHR2s or DHR2c could activate Cdc42 or Rac1. By monitoring the fluorescence changes that accompany the binding of Mant nucleotide to the GTPase, no detectable activation of GST-Cdc42 or GST-Rac1 was

Figure 5.1. DOCK7 and GTPase proteins (A) Schematic representation of the DOCK7 and DHR2 constructs that were examined in this study. (B) SDS-PAGE and Colloidal Blue staining of wild type proteins used in this study.

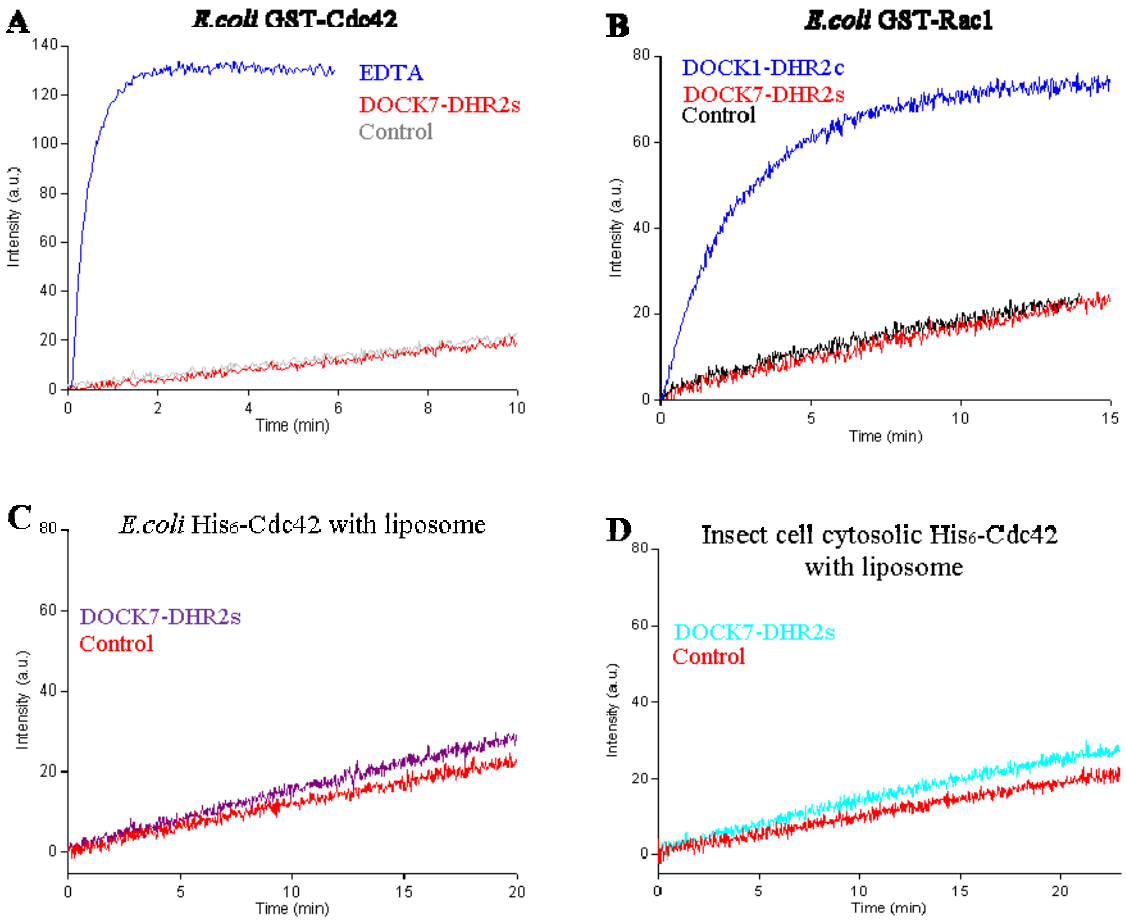

A
DHR1
DHR2
DOCK7 1-1992
DHR2
DHR2 1443-1992
DHR2s
DHR2s 1530-1970
DHR2c
DHR2c 1688-1992


B
E.coli
Insect cell
GST-Rac1
His-Rac1
GST-Cdc42
His-Cdc42
mem-Cdc42
cyto-Cdc42
DOCK7-DHR2s
DOCK7-DHR2c
kDa
75
50
37
25
20

observed in the presence of DOCK7-DHR2s in solution (Fig. 5.2). DHR2c could not initiate noticeable activation of either GST- Cdc42 or GST-Rac1 (data not shown). However, the addition of EDTA caused a significant nucleotide exchange in GST-Cdc42 (Fig. 5.2A), and the Rac1-specific GEF DOCK1-DHR2c could promote the Mant-GDP loading of GST-Rac1 (Fig. 5.2B). To rule out that the GST tag might present an obstacle for the interaction between DOCK7-DHR2s or DHR2c and Cdc42 or Rac1, the replacement of GST-Cdc42 and GST-Rac1 by His_6-Cdc42 and His_6-Rac1, respectively, still resulted in no detectable activation by DOCK7-DHR2s or DHR2c (data not shown). These data were consistent with the findings of another DOCK-C member, DOCK6-DHR2, which showed no activation of Cdc42 or Rac1 using the conventional fluorescence assay (27). However, a modest activation of Rac1 and Cdc42 by DOCK7-DHR2 was detected by using a more sensitive radioactive assay for GDP-GTP exchange (29, 30). Those observations were inconclusive as to whether DOCK-C members were the GEFs for both Cdc42 and Rac1.

Rho GTPase proteins consist of a large GTPase domain in the N-terminus and a polybasic region (PBR) followed by the CAAX motif in the C-terminus. Both PBR and CAAX motif are important for Rho GTPases' cellular functions (5, 6). The positively charged Arg and Lys residues of PBR have been proposed to facilitate the binding of Rho GTPases to cellular membranes. The attached prenylation tail to the cysteine residue of the CAAX motif can insert into membranes where GTPases are able to engage their regulators and effectors. Since DOCK7-DHR2s or DHR2c could not promote significant activation of either Cdc42 or Rac1 in solution, I further tested the prenylated Cdc42 and Rac1 in a model liposome system.

Figure 5.2. DOCK7-DHR2s/c did not activate non-prenylated GST-Cdc42/Rac1. (A) DHR2s (160 nM) (red) showed no ability to stimulate the GDP-to-MantGDP exchange of *E.coli* expressed GST-Cdc42 (200 nM). 10 mM EDTA (blue) was added as the positive control. No GEF was added as a negative control. (B) In contrast to DOCK1-DHR2c (200 nM) (blue), which stimulated the GDP-to-MantGDP exchange of *E.coli* expressed GST-Rac1 (200 nM), DOCK7-DHR2s (160 nM) (red) showed no activation on GST-Rac1. Even in the presence of 20 μM liposome DOCK7-DHR2s (160 nM) (dark purple) showed no activation toward the *E.coli*-expressed His_6-Cdc42 (C), nor toward the cytosolic fraction (non-prenylated) of Cdc42 expressed in insect cells (D).


A
E.coli GST-Cdc42
EDTA
DOCK7-DHR2s
Control
Intensity (a.u.)
Time (min)
B
E.coli GST-Rac1
DOCK1-DHR2c
DOCK7-DHR2s
Control
Intensity (a.u.)
Time (min)
C
E.coli His6-Cdc42 with liposome
DOCK7-DHR2s
Control
Intensity (a.u.)
Time (min)
D
Insect cell cytosolic His6-Cdc42
with liposome
DOCK7-DHR2s
Control
Intensity (a.u.)
Time (min)

The insect cell-expressed geranylgeranylated Cdc42 was incubated with liposomes at room temperature to allow for membrane binding by Cdc42. Cdc42 was observed to bind Mant-GDP upon the addition of DOCK7-DHR2s (Fig. 5.3A). Similarly, insect cell expressed Rac1 was activated upon the addition of DOCK7-DHR2s (Fig. 5.3B). However, non-prenylated Cdc42 suspended in the same liposome mixture did not undergo nucleotide exchange in the presence of DOCK7-DHR2s (Fig. 5.2C). Furthermore, the cytosolic fraction (non-prenylated) of insect cell expressed Cdc42 did not exhibit any significant nucleotide exchange by DOCK7-DHR2s (Fig. 5.2D). These data demonstrate that the prenylation and membrane localization of Cdc42 and Rac1 were essential for the activation by DOCK7. Additionally, DOCK7-DHR2c could activate both Cdc42 and Rac1 to the same extent as DHR2s, supporting a model whereby DOCK7-DHR2 also contained a limit C-terminus which is necessary and sufficient to activate Rho GTPases (Fig. 5.3A). Finally, the activation of Cdc42 by DHR2s was dose-dependent (Fig. 5.3C).

The recently solved complex structures of Cdc42-DOCK9 and Rac1-DOCK2 shed some light regarding the DOCK-GTPase specificity (26, 27). Phe56 of Cdc42, and Trp56 of Rac1, are the key specificity residues for the recognition by DOCK proteins. Substitution of Phe56 of Cdc42 with tryptophan displayed a dramatic decrease in the activation by DOCK9, while the W56F mutant of Rac1 showed much lower efficiency in coupling to DOCK180. Phe56 of Cdc42 interacted with residues Leu1941, Gln1949, and Gly1950 of DOCK9 (Fig. 5.4A). Among them, Leu1941 is conserved in the DOCK-D (Cdc42-specific) subfamily; Gln1949 is invariant in both the DOCK-C and DOCK-D subfamilies; while Gly1950 is absolutely conserved in all the DOCK180 family members. Methionine in the DOCK-A/B (Rac-specific) and

Figure 5.3. DOCK7-DHR2s/c activated prenylated Cdc42 and Rac1 suspended in the model liposomes. In the presence of 20 μM liposomes, either DHR2s (50 nM) (red) or DHR2c (50 nM) (blue) was added to initiate the GDP-to-MantGDP exchange of insect cell-expressed prenylated Cdc42 (200 nM) (A) or Rac1 (200 nM) (B). No GEF was added in the negative control. (C) The activation of Cdc42 by DOCK7-DHR2s displayed a dose-dependent behavior.

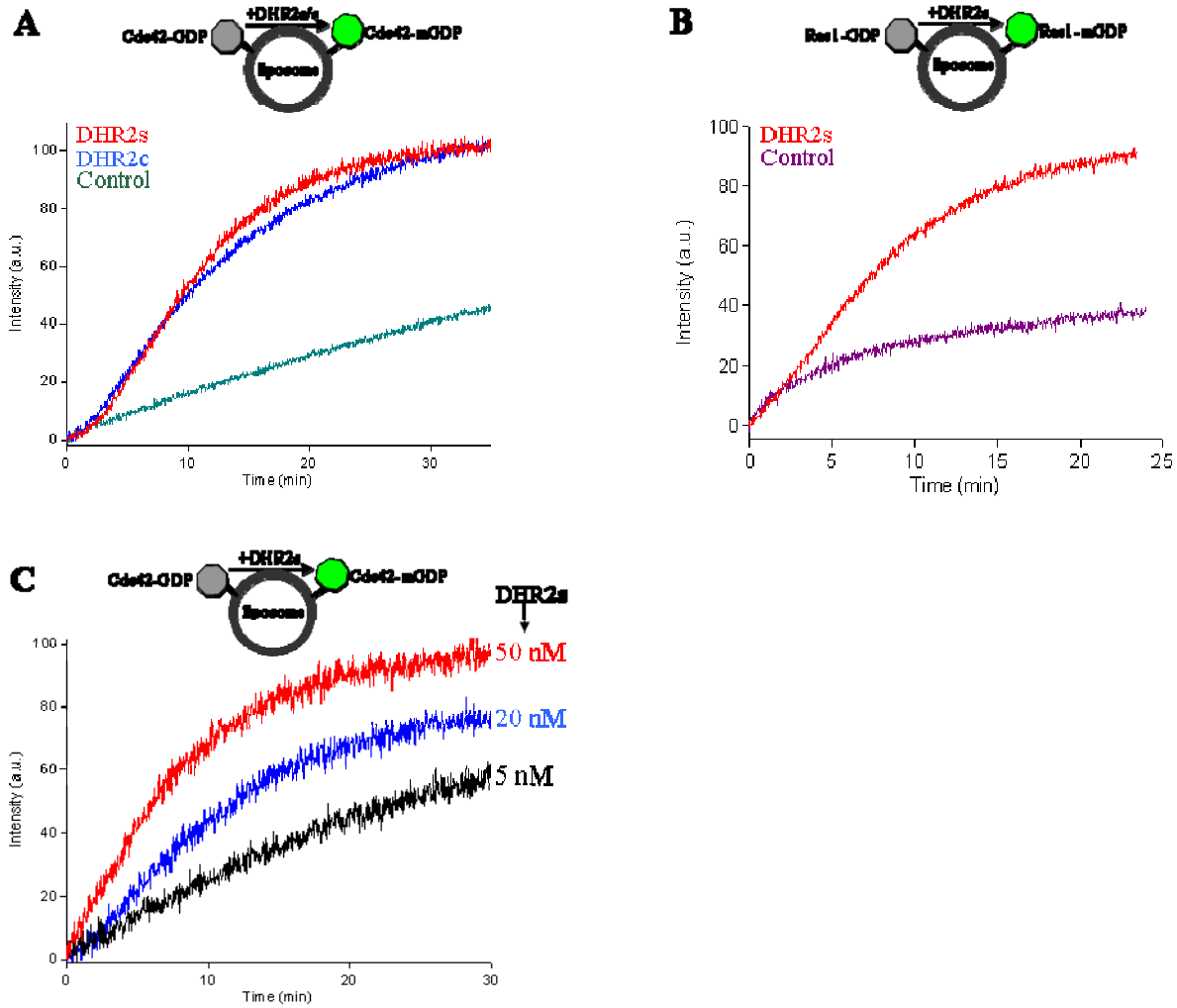
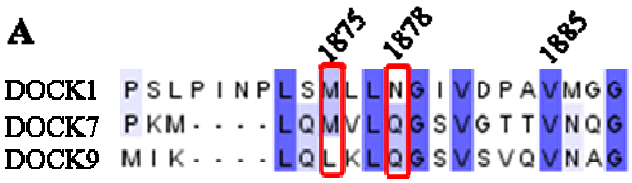
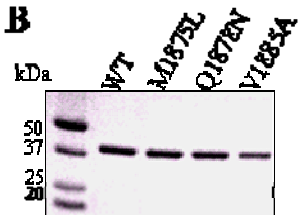
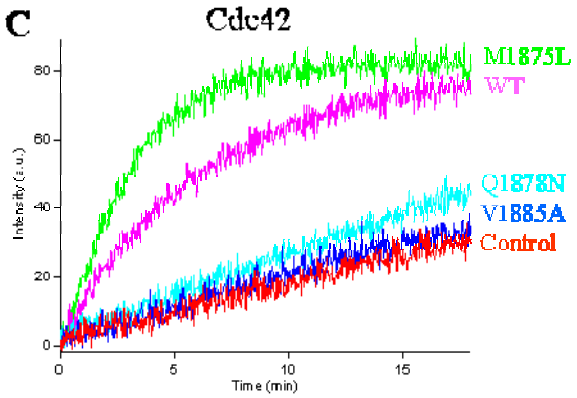

A
Cdc42-GDP
+DHR2s/c
Cdc42-mGDP
liposome
DHR2s
DHR2c
Control
Intensity (a.u.)
Time (min)
B
Rac1-GDP
+DHR2s
Rac1-mGDP
liposome
DHR2s
Control
Intensity (a.u.)
Time (min)
C
Cdc42-GDP
+DHR2s
Cdc42-mGDP
liposome
DHR2s
50 nM
20 nM
5 nM
Intensity (a.u.)
Time (min)

Figure 5.4. Specificity switches of DOCK7 between Cdc42 and Rac1. (A) Sequence comparison between DOCK subfamilies A, C, and D members: DOCK1, DOCK7, and DOCK9, respectively. Numbering is for DOCK7 sequence. (B)SDS-PAGE and Colloidal Blue staining of DOCK7-DHR2c wild type and mutant protein tested in this figure. (C) Mutation of Methionine 1875 of DHR2c to Leucine increases its GEF activity toward Cdc42; while mutation of Glutamine 1878 of DHR2c to Asparagine decreases its GEF activity toward Cdc42.


A
1875
1878
1885
DOCK1 P S L P I N P L S M L L N G I V D P A V M G G
DOCK7 P K M - - - - L Q M V L Q G S V G T T V N Q G
DOCK9 M I K - - - - L Q L K L Q G S V S V Q V N A G

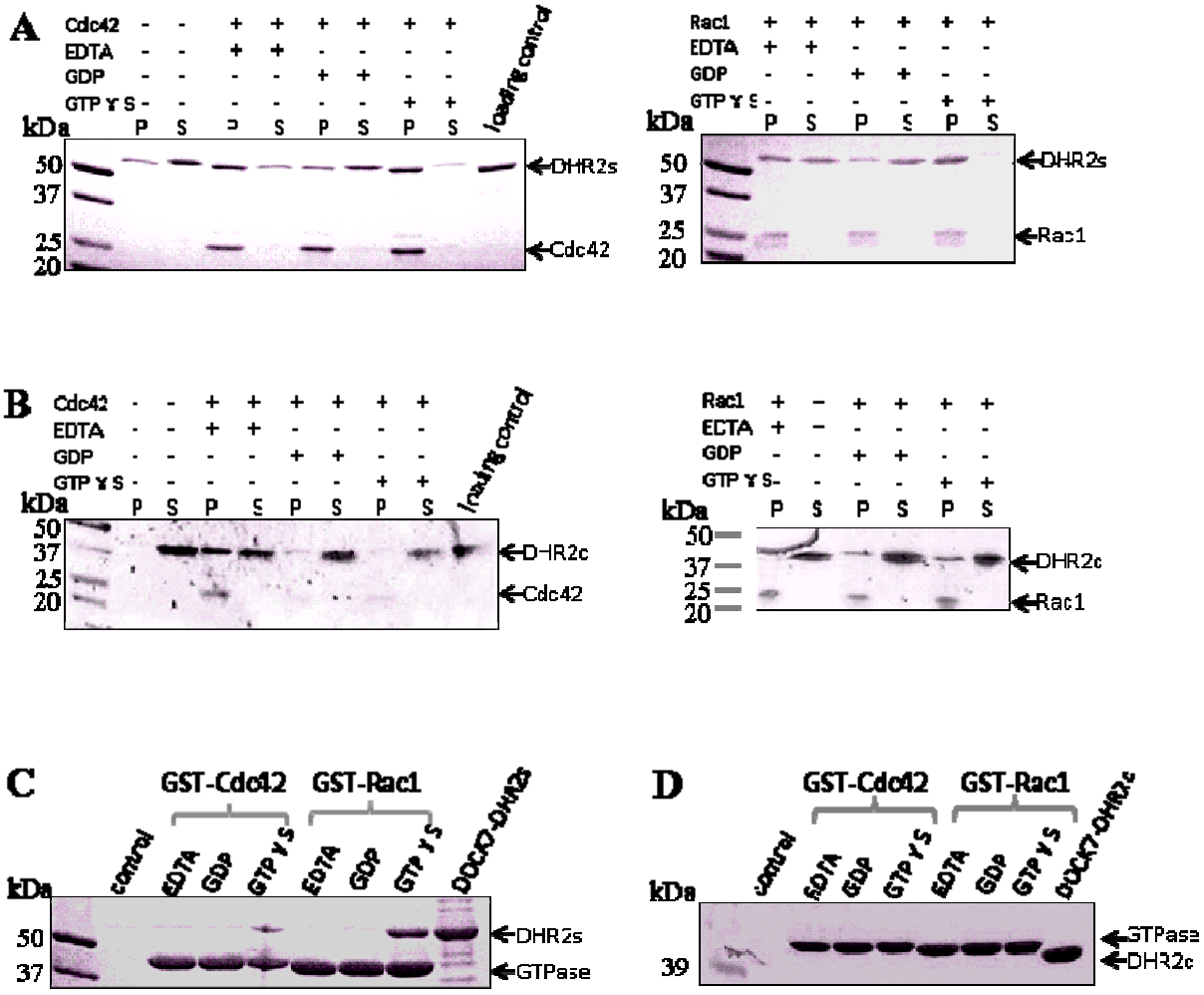

B
kDa
WT
M1875L
Q1878N
V1885A
50
37
25
20


C
Cdc42
M1875L
WT
Q1878N
V1885A
Control
Intensity (a.u.)
Time (min)
0
20
40
60
80
0
5
10
15

DOCK-C subfamilies corresponds to residue Leu1941 in DOCK9; whereas asparagine in the DOCK-A/B (Rac-specific) subfamily corresponds to residue Gln1949 in DOCK9. These alignments collectively suggested that the combination of leucine and glutamine in the DOCK-D subfamily accounted for the specificity for Cdc42, and that the combination of methionine and asparagine in the DOCK-A/B subfamily was responsible for the specificity for Rac. Thereby, the combination of methionine and asparagine in DOCK7 offered an explanation for its dual GEF function for both Cdc42 and Rac1. The M1875L mutant of DOCK7 was proposed to shift it to a Cdc42-specific GEF while the N1878Q mutant was expected to exhibit Rac-specific GEF activity. The DOCK7-DHR2c M1875L and N1878Q mutants were expressed and purified (Fig. 5.4B). As predicted, M1785L exhibited a more robust GEF activity toward Cdc42 than wild-type DHR2c. Conversely, the N1878Q mutant showed almost no GEF activity toward Cdc42. The V1885A mutant of DHR2c was used as a negative control as it has been shown to be catalytically defective, and as predicted, it showed no GEF activity (Fig. 5.4C).

DOCK7-DHR2s and DHR2c showed similar GEF activity, but differed in their abilities to bind to the different states of GTPases in terms of nucleotide association. When prenylated Cdc42 was bound to larger model liposomes which can be pelleted at a low speed, and treated with EDTA, GDP, or GTPγS, DHR2s bound to nucleotide-free and Cdc42-GTPγS with higher affinity than to Cdc42-GDP (Fig. 5.5A). Additionally, compared to nucleotide-free Cdc42, more DHR2s was pelleted by Cdc42-GTPγS, suggesting that DHR2s exhibited a slight preference for GTPγS bound Cdc42 (Fig. 5.5A). Similar behavior was observed with geranylgeranylated Rac1 (Fig. 5.5A). However, DHR2c acted as a typical GEF and bound preferentially to the

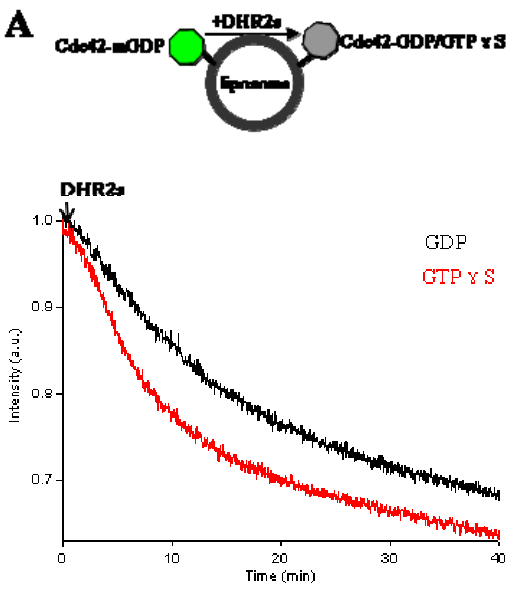
Figure 5.5. DOCK7-DHR2s, but not DHR2c, preferentially bound to the active forms of Cdc42 and Rac1. (A) DHR2s was pelleted with EDTA-treated and GTPγS-treated, prenylated Cdc42 (left panel) or Rac1 (right panel) that was accociated with liposomes. (B) DHR2c was only pelleted with EDTA-treated, prenylated Cdc42 (left panel) or Rac1 (right panel). (C) DHR2s was pulled down with *E.coli*-expressed GTPγS-treated GST-Cdc42 and GST-Rac1, while DHR2c was not able to bind to either of these GTPγS-loaded proteins (D).

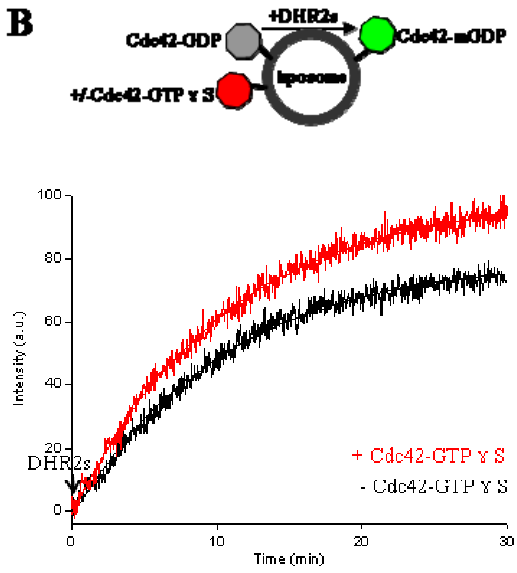

A
Cdc42
EDTA
GDP
GTP γ S
kDa
P S P S P S P S
loading control
50
37
25
20
DHR2s
Cdc42
Rac1
EDTA
GDP
GTP γ S
kDa
P S P S P S
50
37
25
20
DHR2s
Rac1
B
Cdc42
EDTA
GDP
GTP γ S
kDa
P S P S P S P S
loading control
50
37
25
20
DHR2c
Cdc42
Rac1
EDTA
GDP
GTP γ S
kDa
P S P S P S
50
37
25
20
DHR2c
Rac1
C
GST-Cdc42
GST-Rac1
control
EDTA
GDP
GTP γ S
EDTA
GDP
GTP γ S
DOCK7-DHR2c
kDa
50
37
DHR2s
GTPase
D
GST-Cdc42
GST-Rac1
control
EDTA
GDP
GTP γ S
EDTA
GDP
GTP γ S
DOCK7-DHR2c
kDa
39
GTPase
DHR2c

nucleotide-free forms of its cognate GTPases (Fig. 5.5B). These data argue that the limit DHR2c domain was responsible for binding the inactive forms of the GTPases for nucleotide exchange, while the extra N-terminal portion of DHR2s was involved in the recruitment of the active forms of these proteins. These findings were further supported by the results of the GST-Cdc42/Rac1 pull-down assays. When non-prenylated *E. coli*-expressed GST-Cdc42/Rac1 was bound to glutathione beads, and treated by EDTA, GDP, or GTPγS, DHR2s was only pelleted with Cdc42/Rac1-GTPγS (Fig. 5.5C). No detectable amount of DHR2c was pulled-down by any form of GST-Cdc42 or Rac1 (Fig. 5.5D). These data were also in good agreement in that no detectable GEF activity of DHR2s or DHR2c was observed on GST-tagged Cdc42 or Rac1 (Fig. 5.2).

SOS (Son of sevenless), a Ras-specific GEF, has been shown to possess a second distal GTPase binding site, which is distinct from the catalytic site. This distal site is responsible for binding the active form of Ras (i.e. GTP bound Ras) and regulating GEF exchange rates resulting in a positive feedback activation of Ras (34). The binding of Cdc42/Rac1-GTP by DHR2s implied that DHR2s might have a similar distal regulatory site that interacts with Cdc42/Rac1-GTP, representing a potential positive feedback loop in the activation of Cdc42/Rac1. Consistent with this, Cdc42, preloaded with Mand-GDP, showed accelerated binding of GTPγS upon the addition of DHR2s to the model liposome system (Fig. 5.6A). It appears that the Cdc42-GTPγS enhanced the nucleotide exchange as the reaction proceeded. The addition of the active Cdc42-GTPγS to the inactive Cdc42 caused an increase in Mant fluorescence, suggesting that the exchange rate may be influenced by the recruitment of DOCK7-DHR2s to the liposomes via the membrane-associated Cdc42- GTPγS (Fig. 5.6B).

Figure 5.6. DHR2s seems to activate Cdc42 via a positive feedback loop. (A) 200 nM Cdc42 preloaded with MantGDP was assayed for nucleotide exchange in the presence of 160 nM DHR2s, 50 μM GDP or GTPγS. (B) 250 nM Cdc42 was assayed for nucleotide exchange in the presence of 16 nM DHR2s, 1 μM MantGDP, with or without the addition of 100 nM Cdc42-GTPγS.


A
+DHR2s
Cdc42-mGDP
Cdc42-GDP/GTP γ S
Liposome
DHR2s
GDP
GTP γ S
Intensity (a.u.)
Time (min)

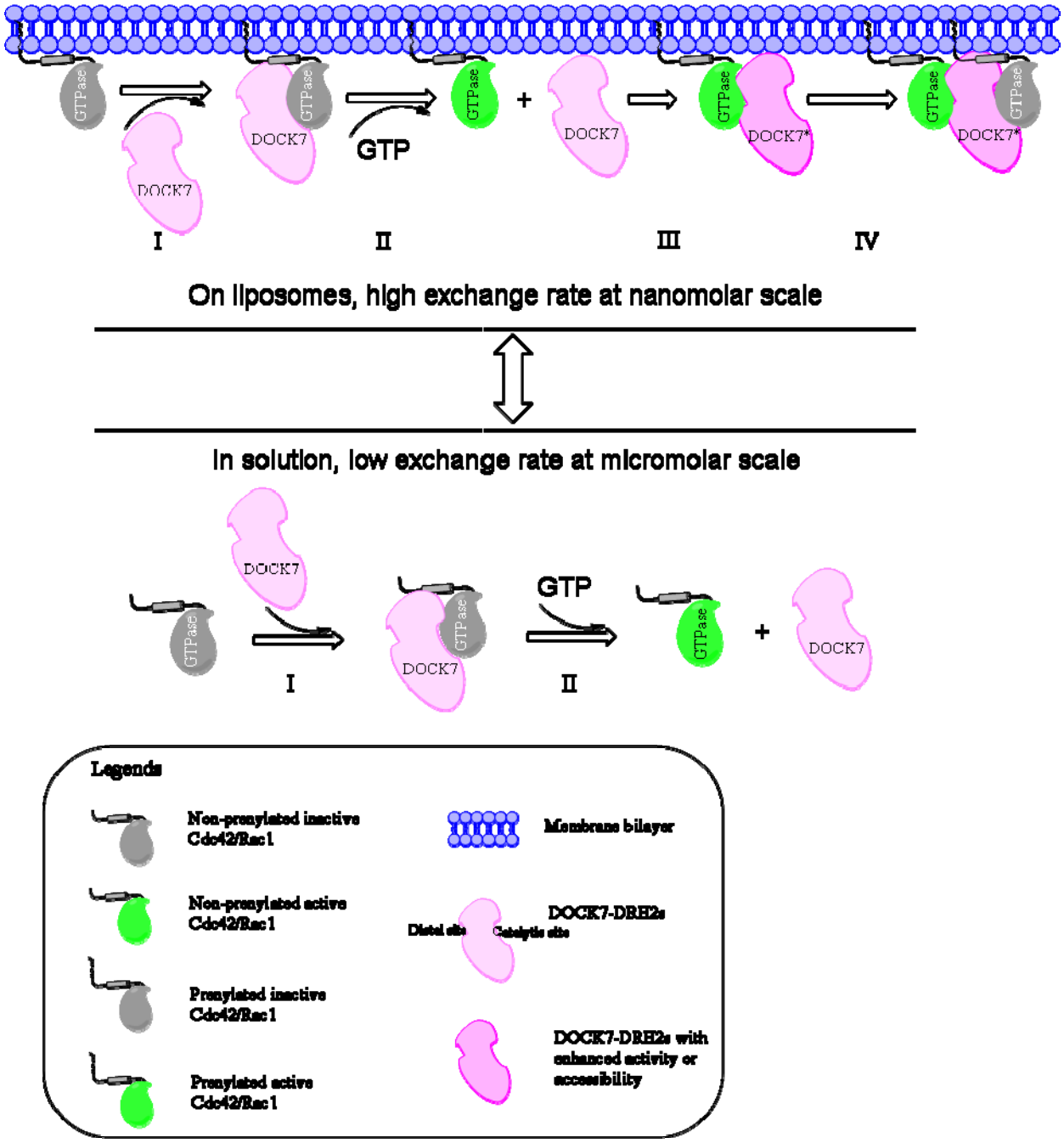

B
+DHR2s
Cdc42-GDP
Cdc42-mGDP
+/-Cdc42-GTP γ S
liposome
DHR2s
+ Cdc42-GTP γ S
- Cdc42-GTP γ S
Intensity (a.u.)
Time (min)

5.4 Discussion

The DOCK180 family members have been demonstrated to regulate multiple cellular processes, including cell migration, cell phagocytosis, cytoskeleton reorganization (35-40), and neuronal differentiation (29-31). They function as upstream activators (GEFs) of Cdc42 and/or Rac1 from the Rho family of small GTPases which are critical players in actin dynamics. The recent structural and functional studies on the DOCK9-Cdc42 and DOCK2-Rac1 complexes have elucidated the mechanism of how DOCK180 members activate their cognate GTPases, and provided insights to the specificity of how DOCK180 proteins recognize GTPases. Both the Rac1-specific DOCK-A member and the Cdc42-specific DOCK-D member shares the same mechanism to activate Rac1 and Cdc42, respectively. It is reasonable to propose that DOCK-C subfamily members utilize the same mechanism, given that replacing the catalytic residue Val1885 of DOCK7-DHR2c with alanine abolished its GEF activity (Fig. 5.4C).

The GEF activity and specificity of the DOCK-A and the DOCK-D subfamilies have been well studied (22, 41). However, little is known about the biochemical characteristics of the DOCK-C subfamily members, such as DOCK7, mainly because DOCK7 displayed little GEF activity toward Cdc42 or Rac1 in solution using standard *in vitro* GEF assays. Interestingly, I found that the DHR2 domain (i.e. the DHR2c and DHR2s constructs) exhibited robust GEF activity toward prenylated Cdc42 or Rac1 in the presence of model liposomes by monitoring changes in the fluorescence of Mant nucleotide upon their binding to the GTPases (Fig. 5.2). DOCK7, at nanomolar concentrations, stimulated the activation of GTPases on liposomes, whereas it showed no GEF activity even when assayed at micromolar concentrations in solution. This

Figure 5.7. Theoretical model of Cdc42/Rac1 activation by DOCK7 via a possible positive feedback loop. In solution, DOCK7-DHR2s could not stimulate the activation of Cdc42 or Rac1, whereas when assayed on liposomes, DOCK7-DHR2s initiated nucleotide exchange on Cdc42 or Rac1. The addition of Cdc42-GTPγS accelerated the rate of nucleotide exchange.


GTPase
DOCK7
GTP
DOCK7*
I
II
III
IV
On liposomes, high exchange rate at nanomolar scale
In solution, low exchange rate at micromolar scale
Legends
Non-prenylated inactive Cdc42/Rac1
Non-prenylated active Cdc42/Rac1
Prenylated inactive Cdc42/Rac1
Prenylated active Cdc42/Rac1
Membrane bilayer
DOCK7-DRH2s
Distal site
Catalytic site
DOCK7-DRH2s with enhanced activity or accessibility

suggests that DOCK7 is sensitive to the lipidation and membrane localization of GTPases (Fig. 5.7). Our observations were mirrored by studies on the Ras GEF, SOS, which showed that the farnesylation of Ras was necessary to further promote SOS-mediated nucleotide exchange (42, 43).

The position 56 of Cdc42 and Rac1 (phenylalanine and tryptophan, respectively) has been identified as the key residue that is responsible for the specific recognition by DOCK180 proteins (24, 27). The side chain of residue 56 of Cdc42 or Rac1 inserts into a binding pocket present within DOCK180-related proteins and makes contacts with the side chains of residues in the pocket, including Leu1941 and Gln1944 of DOCK9 for Cdc42, or Met1529 and Asn1532 of DOCK2 for Rac1. The replacement of Met1524 of DOCK1 can restore about 25% activation of Rac1 W56F mutant (24). The smaller side chain of Phe56 of Cdc42 favors the bulkier side chains of Leu1941 and Gln1944 of DOCK9, whereas the larger side chain of Try56 of Rac1 prefers the smaller side chains of Met1529 and Asn1532 of DOCK2 (or Met1524 and Asn1527 of DOCK1). Therefore, substitution of Phe56 of Cdc42 with tryptophan should cause steric hindrance in the pocket of DOCK9, while substitution of Try56 of Rac1 with phenylalanine should make less contact with DOCK2 or DOCK1. How can DOCK7 recognize both Cdc42 and Rac1? DOCK7 combines Met1875 from the DOCK-A subfamily and Gln1878 from the DOCK-D subfamily at the corresponding positions, which still make enough contact with phenylalanine of Cdc42, but can not cause significant steric hindrance with tryptophan of Rac1. By this “compromise” strategy, DOCK7 is capable of the dual recognition of both Cdc42 and Rac1 at the cost of specificity. Therefore, the M1875L and Q1978N mutants of DOCK7 exhibit specificity toward Cdc42 and Rac1, respectively (Fig. 5.4C).

Previously, Wu *et al*. defined a limit C-terminal region (designated as DHR2c) of the DHR2 domain of DOCK1, which showed full GEF activity (24). Here I also showed that DOCK7 possessed a similar DHR2c portion which exhibited the same catalytic activity as DHR2s (Fig. 5.3A). Although the N-terminus of DOCK9 DHR2 contributed slightly, it appears that all the DOCK180-related members have this smaller domain which is necessary and sufficient to activate its cognate GTPases. However, in contrast to DHR2c, the longer DHR2s of DOCK7 preferentially binds to the active forms of GTPases, which suggests the potential for positive cooperativity through the ability of activated GTPases to recuit the GEF to the liposome surface (Fig. 5.7). Similarly, the PH domain of PDZ-RhoGEF has been demonstrated to bind the active form of RhoA at the effector binding site. (44). Additionally, the DH-PH domain of PDZ-RhoGEF can form a ternary complex with the active form of RhoA throught the PH domain, while the inactive form of RhoA binds to the DH domain. However, activated RhoA does not appear to affect the GEF activity of PDZ-RhoGEF *in vitro*. It is possible that, in order to enhance the accessibility of GEFs, cells take this regulatory strategy to recruit GEFs to membranes. For example, the PH domain of Arno, an Arf1 GEF, binds to myristoylated Arf6 on the liposomes *in vitro*, which in turn increases the GEF activity of Arno, forming a positive feedback loop (45). Lin *et al*. found that the active form of Cdc42 interacted with the N-terminus and release the autoinhibition of the full length DOCK11, leading to the positive feedback activation of Cdc42 (46). Therefore, it is possible that DOCK180 family members recruit the active form of their coupling GTPases to modulate the spatial and temporal regulation of their GEF activity.

References

1. Cerione, R. A. (2004) *Trends Cell Biol.* **14**, 127-132.
2. Chardin, P. (1998) *Biochimie* **70**, 865-868.
3. Ridley, A. J., Paterson, H. F., Johnston, C. L., Diekmann, D., and Hall, A. (1992) *Cell* **70**, 401-410.
4. Tapon, N., and Hall, A. (1997) *Curr. Opin. Cell Biol.* **9**, 86-92
5. Williams, C. L. (2003) *Cell. Signal.* **15**, 1071-1080.
6. Roberts, P. J., Mitin, N., Keller, P. J., Chenette, E. J., Madigan, J. P., Currin, R. O., Cox, A. D., Wilson, O., Kirschmeier, P., and Der, C. J. (2008) *J. Biol. Chem.* **283**, 25150-25163.
7. Sebti, S. M., and Der, C. J. (2003) *Nat. Rev. Cancer* **3**, 945-951.
8. Hoffman, G. R., and Cerione, R. A. (2002) *FEBS Lett.* **513**, 85–91.
9. Nassar, N., Hoffman, G. R., Manor, D., Clardy, J. C., and Cerione, R. A. (1998) *Nat. Struct. Biol.* **5**, 1047-1052.
10. Johnson, J. L., Erickson, J. W., and Cerione, R. A. (2009) *J. Biol. Chem.* **284**, 23860–23871.
11. Hart, M. J., Eva, A., Evans, T., Aaronson, S. A., and Cerione, R. A. (1991) *Nature* **354**, 311–314.
12. Takai, S., Hasegawa, H., Kiyokawa, E., Yamada, K., Kurata, T., and Matsuda, M. (1996) *Genomics* **35**, 403–404.
13. Hasegawa, H., Kiyokawa, E., Tanaka, S., Nagashima, K., Gotoh, N., Shibuya, M., Kurata, T., and Matsuda, M. (1996) *Mol. Cell. Biol.* **16**, 1770–1776.
14. Erickson, J. W., and Cerione, R. A. (2004) *Biochemistry* **43**, 837–842.

15. Worthylake, D. K., Rossman, K. L., and Sondek, J. (2004) *Structure* **12**, 1078–1086.

16. Feng, Q., Albeck, J. G., Cerione, R. A., and Yang, W. (2002) *J. Biol. Chem.* **277**, 5644–5650.

17. Feng, Q., Baird, D., and Cerione, R. A. (2004) *EMBO J.* **23**, 3492–3504.

18. Baird, D., Feng, Q., and Cerione, R. A. (2005) *Curr. Biol.***15**, 1–10.

19. Lemmon, M. A., and Ferguson, K. M. (2000) *Biochem. J.* **350**, 1-18.

20. Snyder, J. T., Rossman, K. L., Baumeister, M. A., Pruitt, W. M., Siderovski, D. P., Der, C. J., Lemmon, M. A., and Sondek, J. (2001) *J. Biol. Chem.* **276**, 45868-45875.

21. Brugnera, E., Haney, L., Grimsley, C., Lu, M., Walk, S. F., Tosello-Trampont, A. C., Macara, I. G., Madhani, H., Fink, G. R., and Ravichandran, K. S. (2002) *Nat. Cell Biol.* **4**, 574–582.

22. Cote, J. F., and Vuori, K. (2002) *J. Cell Sci.* **115**, 4901–4913.

23. Cote, J., Motoyama, A. B., Bush, J. A., and Vuori, K. (2005) *Nat. Cell Biol.* **7**, 797-807.

24. Wu, X., Ramachandran, S., Lin, M., Cerione, R. A., and Erickson, J. W. (2011) *Biochemistry* **50**, 1070-1080.

25. Cote, J., and Vuori, K. (2007) *Trends Cell Biol.* **17**, 383-393.

26. Yang, J., Zhang, Z., Roe, S. M., Marshall, C. J., and Barford, D. (2009) *Science* **325**, 1398-1402.

27. Kulkarni, K., Yang, J., Zhang, Z., and Barford, D. (2011) *J. Biol. Chem.* **286**, 25341-25351.

28. Miyamoto, Y., Yamauchi, J., Sanbe, A., and Tanoue, A. (2007) *Exp. Cell Res.* **313**, 791-804.

29. Watabe-Uchida, M., John,K. A., James, J. A., Newey, S. E., and Van Aelst, L. (2006) *Neuron* **51**, 727–739.

30. Yamauchi, J., Miyamoto, Y., Chan, J. R., and Tanoue, A. (2008) *J. Cell. Biol.* **181**, 351-365.

31. Miyamoto, Y., and Yamauchi, J. (2010) *Signaling* **22**, 175–182.

32. Nellist, M., Burgers, P. C., van den Ouweland, A. M. W., Halley, D. J. J., and Luider, T. M. (2005) *Biochem. Biophys. Res. Commun.* **333**, 818-826.

33. Buboltz, J. T., and Feigenson, G.W. (1999) *Biochim, biophys. Acta* **1417**, 232-245.

34. Margarit, S. M., Sondermann, H., Hall, B. E., Nagar, B., Hoelz, A., Pirruccello, M., Bar-Sagi, D., and Kuriyan, J. (2003) *Cell* **112**, 685-695.

35. Wu, Y. C., and Horvitz, H. R. (1998) *Nature* **392**, 501–504.

36. Park, D., Tosello-Trampont, A. C., Elliott, M. R., Lu, M., Haney, L. B., Ma, Z., Klibanov, A. L., Mandell, J. W., and Ravichandran, K. S. (2007) *Nature* **450**, 430–434.

37. Wang, X., Wu, Y. C., Fadok, V. A., Lee, M. C., Gengyo-Ando, K., Cheng, L. C., Ledwich, D., Hsu, P. K., Chen, J. Y., Chou, B. K., Henson, P., Mitani, S., and Xue, D. (2003) *Science* **302**, 1563–1566.

38. Wu, Y. C., Tsai, M. C., Cheng, L. C., Chou, C. J., and Weng, N. Y. (2001 *Dev. Cell* **1**, 491–502.

39. Henson, P. M. (2005) *Curr. Biol.* **15**, R29–R30.

40. Yajnik, V., Paulding, C., Sordella, R., McClatchey, A. I., Saito, M., Wahrer, D. C., Reynolds, P., Bell, D. W., Lake, R., van den Heuvel, S., Settleman, J., and Haber, D. A. (2003) *Cell* **112**, 673–684.

41. Kwofie, M. A., and Skowronski, J. (2008) *J. Biol. Chem.* **283**, 3088–3096.

42. Porfiri, E., Evans, T., Chardin, P., and Hancock, J. F. (1994*J. Biol. Chem.* **269**, 22672-22677.

43. Pechlivanis, M., Ringel, R., Popkirova, B., and Kuhlmann, J. (2007) *Biochemistry* **46**, 5341-5348.

44. Chen, Z., Medina, F., Liu, M., Thomas, C., Sprang, S. R., and Sternweis, P. C. (2010) *J. Biol. Chem.* **285**, 21070-21081.

45. Stalder, D., Barelli, H., Gautier, R., Macia, E., Jackson, C. L., and Antonny, B. (2011) *J. Biol. Chem.* **286**, 3873-3883.

46. Lin, Q., Yang, W., Baird, D., Feng, Q., and Cerione, R. A. (2006) *J. Biol. Chem.* **281**, 35253-35262.

Chapter 6

Conclusions and Future Directions

6.1 Conclusions

My thesis research has involved structure-function studies of members of the sirtuin family, and biochemical characterization of the Cdc42/Rac-activator DOCK7. Below, I will briefly summarize the main highlights from each of these lines of study and then put forward some interesting considerations for future work.

Human SIRT5 is an NAD-dependent demalonylase and desuccinylase. SIRT5 has been reported to have weak deacetylase activity (1). Co-crystallization of SIRT5 with acetylated peptides failed to produce any crystals amenable for X-ray diffraction. However, the thioacetylated peptides were successfully co-crystallized with SIRT5. The X-ray crystal structure of SIRT5 in complex with thioacetyl H3K9 peptide and a buffer molecule CHES showed that SIRT5 harbored a larger acyl binding pocket with two non-hydrophobic residues, Tyr102 and Arg105, at the bottom of the pocket. The unique features of SIRT5's acyl binding pocket and the binding of CHES inspired us and led to the discovery of the novel functions of SIRT5. Now we believe that SIRT5 is a demalonylase and desuccinylase with robust activity, comparable to the deacetylase activity of SIRT1. As expected, the succinyl group of the H3K9 peptide interacts with Tyr102 and Arg105 of SIRT5, such that mutating either of these two residues dramatically decreases the enzymatic activity. Furthermore, using mass spectrometry, a number of mitochondrial proteins have been identified to be malonylated and succinylated. Some of the lysine residues of proteins can be both acetylated and succinylated. Finally, SIRT5 modulates the enzymatic activity of carbamoyl phosphate synthetase 1 (CPS1) by catalyzing its desuccinylation at lysine

1291.

The mechanism of SIRT5's desuccinylation is similar to that used by other sirtuins to catalyze deacetylation. To delineate the mechanism of desuccinylation, I have solved a series of SIRT5 complex structures, including the complexes of SIRT5-succinyl H3K9 peptide, SIRT5-succinyl H3K9 peptide-NAD, and SIRT5-*S*-bicyclic intermediate, representing the substrate binding step, the Michaelis-Menten complex step, and the intermediate II step, respectively. The 1',2'-bicyclic intermediate complex structure supports the peptidyl-ADPR-amidate mechanism. This is the first time that a bicyclic intermediate was directly observed in an X-ray crystal structure. This intermediate was obtained by co-crystallizing SIRT5 with its specific inhibitor, the thiosuccinyl H3K9 peptide, followed by soaking in NAD in the time range of 0.5-16 hours. I was unable to obtain the alkylamidate intermediate I, which suggested that the inhibition mechanism of thiosuccinyl peptide may be that the *S*-1',2'-bicyclic intermediate was stalled at the active site of SIRT5.

Plasmodium falciparum Sir2A (PfSir2A) preferentially removes the medium and long chain fatty acyl modifications from lysine residues. PfSir2A has been shown to have weak deacetylase activity (2). We found that PfSir2A was more efficient in the hydrolysis of the medium and long chain (C8-C14) fatty acyl groups attached to lysine. The crystal structures of PfSir2A in complex with the myristoyl H3K9 peptide or with both peptide and NAD uncovered the structural features of PfSir2A which explained PfSir2A's capability of catalyzing the deacylation of medium and long chain fatty acyl groups from lysine. PfSir2A possesses a long open hydrophobic acyl binding tunnel which accommodates the myristoyl group. Additionally, modifications leading to the incorporation of the medium and long chain fatty acids were reported to

occur in the parasite *P. falciparum*.

Prenylation and membrane localization of Cdc42/Rac1 are essential for the activation by DOCK7. No robust GEF activity of DOCK7 has been reported towards either Cdc42 or Rac1 (3, 4). In Chapter 5 of my thesis, I show that DOCK7 is unable to stimulate the activation of non-prenylated Cdc42/Rac1 in solution. However, the activation of prenylated Cdc42/Rac1 can be catalyzed by DOCK7 in model liposomes.

The combination of methionine and glutamine in the α10 helix of DOCK7 enables it to dually recongnize Cdc42 and Rac1. Position 56 of Cdc42 and Rac1 (i.e. Phe56 in Cdc42 and Trp56 in Rac1) has been shown to be the key site for the specific recognition by their cognate GEFs (5, 6). The X-ray crystal structures of two complexes (Cdc42-DOCK9 and Rac1-DOCK2 complexes) have revealed that two residues in the α10 helix of DOCK proteins, leucine and glutamine in DOCK9, and methionine and asparagine in DOCK2, interact with Phe56 of Cdc42 and Trp56 of Rac1, respectively (6, 7). DOCK7 combines methionine and glutamine at the corresponding positions, allowing it to exhibit dual GEF activity on Cdc42 and Rac1. Replacing methionine with leucine, or glutamine with asparagine, in DOCK7 can cause it to become a Cdc42-specific or Rac1-specific GEF, respectively.

The N-terminus of DOCK7-DHR2s possesses a distal site for the binding of the active forms of Cdc42/Rac1. The results of liposome spin-down assays indicate that DOCK7 possesses an allosteric site which preferentially binds to the active forms of Cdc42/Rac1. This atypical GTPase-GEF binding interaction allows for a possible positive feedback activation of these GTPases by DOCK7.

6.2 Future Directions

Studies of the sirtuins- SIRT5, together with SIRT3 and SIRT4, are the three mammalian sirtuins which predominantly reside in the mitochondria (8). SIRT3 has been demonstrated to regulate cellular metabolism by deacetylating a number of metabolic enzymes, including Acetyl-CoA synthetase 2 (ACS2) (9), 3-hydroxyl-3-methylglutary CoA synthetase 2 (HMGCS2) (10), long-chain acyl CoA dehydrogenase (LCAD) (11), glutamate dehydrogenase (GDH), and isocitrate dehydrogenase 2 (ICDH2) (12). Very few proteins have been identified as substrates of SIRT5. Recently, SIRT5 has been shown to modulate the enzymatic activity of CPS1 to regulate the disposal of ammonium which is a product of glutamine metabolism, suggesting an important role of SIRT5 in energy metabolism (13). Therefore, future identification of the physiological substrates of SIRT5 should provide interesting insights into its cellular functions.

SIRT5 is the only mammalian sirtuin with demalonylse and desuccinylase activities. This unique feature enables us to design SIRT5-specific inhibitors, including thiosuccinyl lysine peptide with an IC50 in the micromolar scale (Bin He, unpublished data). The thiosuccinyl peptide forms a bicyclic intermediate which is stalled at the active site of SIRT5, thereby decreasing the turnover rate of the product. The structure of SIRT5 bound to this intermediate will provide insights into the optimization of SIRT5-specific inhibitors. Because SIRT5 has been shown in our laboratory to be essential for the growth of different cancer cell lines (Bo Li, unpublished data), the design of specific SIRT5 inhibitors must have therapeutic applications.

Given that metabolic changes are key features of cancer cells (14), it will be of

interest to further investigate the role that SIRT5 plays in tumorigenesis, and in particular, to identify the relevant substrates for SIRT5 in cancer cells as well as possible regulators of its enzymatic activity.

Similar to the case of SIRT5, the complex structures of PfSir2A will facilitate the development of PfSir2A-specific inhibitors, which show great potential in treating malaria. PfSir2A has been reported to control the exclusive expression of cell surface antigens to avoid the clean-up by the host immune surveillance, and *Plasmodium falciparum* is an exceedingly dangerous species that is responsible for 90% of the deaths arising from malaria (15, 16).

Studies of DOCK7- DOCK7 binds to both inactive and active forms of prenylated Cdc42/Rac1. However, an important question concerns the identity of the regions in DOCK7 that interact with the inactive and active forms of the GTPases. Two alternative methods can be used to answer this question. Biochemically, a series of truncation constructs of the limit functional domain of DOCK7 (i.e. DHR2s) can be cloned, expressed, and purified to test their ability to bind to the GTPases. This method should allow the mapping of the binding site for the inactive and acive forms of the GTPases. Structurally, the Cdc42·GTPγS-DHR2s-Cdc42·GDP (or Rac1·GTPγS-DHR2s-Rac1·GDP) complex can be obtained by mixing the purified GTPase and DHR2s recombinant proteins, and further purified for crystallization. The crystal structure would provide important information regarding how DOCK7-DHR2s can bind both forms of the GTPase.

All of the characterizations of DOCK7 in Chapter 5 of my thesis were performed using either DHR2s or DHR2c. Considering that the full length DOCK7 harbors a DHR1 domain in the N-terminus as well as the linker between DHR1 and

DHR2, it is likely that the other parts of the protein beyond DHR2 will affect the GEF activity of DOCK7. For example, the N-terminus of DOCK11 mediates a positive feedback in activating Cdc42. Thus, it should be interesting to perform biochemical characterizations of the GEF activity for the full length DOCK7, both using *in vitro* liposome system, as well as ultimately in cells.

References

1. Schuetz, A., Min, J., Antoshenko, T., Wang, C., Allali-Hassani, A., Dong, A., Loppnau, P., Vedadi, M., Bochkarev, A., Sternglanz, R., and Plotnikov, A. N. (2007) *Structure* **15**, 377-389.
2. French, J. B., Cen, Y., and Sauve, A. A. (2008) *Biochemistry 47*, 10227-10239.
3. Watabe-Uchida, M., John, K. A., James, J. A., Newey, S. E., and Van Aelst, L. (2006) *Neuron* **51**, 727–739.
4. Yamauchi, J., Miyamoto, Y., Chan, J. R., and Tanoue, A. (2008) *J. Cell. Biol.* **181**, 351-365.
5. Gao, Y., Xing, J., Streuli, M., Leto, T. L., and Zheng, Y. (2001) *J. Biol. Chem.* **276**, 47530-47541.
6. Kulkarni, K., Yang, J., Zhang, Z., and Barford, D. (2011) *J. Biol. Chem.* **286**, 25341-25351.
7. Yang, J., Zhang, Z., Roe, S. M., Marshall, C. J., and Barford, D. (2009) *Science* **325**, 1398-1402.
8. Michishita, E., Park, J. Y., Burneskis, J. M., Barrett, J. C. and Horikawa, I. (2005) *Mol. Biol. Cell* **16**, 4623-4635.
9. Schwer, B., Bunkenborg, J., Verdin, R. O., Andersen, J. S. and Verdin, E. (2006) *Proc. Natl. Acad. Sci. U.S.A.* **103**, 10224-10229.
10. Shimazu, T., Hirschey, M. D., Hua, L., dittenhafer-Reed, K. E., Schwer, B., Lombard, D. B., Li, Y., Bunkenborg, J., Alt, F. W., Denu, J. M., Jacobson, M. P., and Verdin, E. (2010) *Cell Metab.* **12**, 654-661.
11. Hirschey, M. D., Shimazu, T., Goetzman, E., Jing, E., Schwer, B., Lombard, D. B., Grueter, C. A., Harris, C., Biddinger, S., Ilkayeva, O. R., Stevens, R. D.,

Li, Y., Saha, A. K., Ruderman, N. B., Bain, J. R., Newgard, C. B., Farese Jr R. V., Alt, F. W., Kahn, C. R. and Verdin, E. (2010) *Nature* **464**, 121-126.

12. Schlicker, C., Gertz, M., Papatheodorou, P., Kachholz, B., Becher, C. F. W., and Steegborn, C. (2008) *J. Mol. Biol.* **382**, 790-801.

13. Nakagawa, T., Lomb, D. J., Haigis, M. C. and Guarente, L. (2009) *Cell* **137**, 560-570.

14. Hsu, P. P., and Sabatini, D. M. (2008) *Cell* **134**, 703-707.

15. Freitas-Junior, L. H., Hernandez-Rivas, R., Ralph, S. A., Montiel-Condado, D., Ruvalcaba-Salazar, O. K., Rojas-Meza, A. P., Mâncio-Silva, L., Leal-Silvestre, R. J., Gontijo, A. M., Shorte, S., and Scherf, A. (2005) *Cell* **121**, 25-36.

16. Tonkin, C. J., Carret, C. l. K., Duraisingh, M. T., Voss, T. S., Ralph, S. A., Hommel, M., Duffy, M. F., Silva, L. M. d., Scherf, A., Ivens, A., Speed, T. P., Beeson, J. G., and Cowman, A. F. (2009) *PLoS Biol.* **7**, e1000084.

17. Lin, Q., Yang, W., Baird, D., Feng, Q., and Cerione, R. A. (2006) *J. Biol. Chem.* **281**, 35253-35262.