STRUCTURAL BASIS FOR ENZYME PROMISCUITY AND SPECIFICITY

Insights from human cytosolic sulfotransferase (SULT) and sirtuin (SIRT) families

by

Wang Pan

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy

Department of Medical Biophysics
University of Toronto

STRUCTURAL BASIS FOR ENZYME PROMISCUITY AND SPECIFICITY
Insights from human cytosolic sulfotransferase (SULT) and sirtuin (SIRT) families

Wang Pan

Doctor of Philosophy, 2011

Department of Medical Biophysics
University of Toronto

ABSTRACT

Understanding the structural basis of specificity and promiscuity of paralogous enzymes is important for deciphering molecular mechanisms and is a necessary step towards designing enzyme-specific modulators. The main objective of this thesis is to provide structural insights that relate protein local sequences to their observed binding and activity profiles through the study of two human protein families – cytosolic sulfotransferases (SULTs) and sirtuins (SIRTs). This was achieved by comparing the family-wide ligand binding fingerprints of these two enzyme families with the structural details of their corresponding enzyme-ligand co-crystal structures.

The hSULT enzyme family was profiled against a focused library through binding and activity assays. This suggested a number of novel compounds that bind to the less well-characterized SULT members (SULT1C3 and SULT4A1), and revealed additional broad-spectrum hSULT inhibitors. Based on the profiling data, three enzyme/co-factor/ligand complex structures were solved using X-ray crystallography. The structure of SULT1C2·PAP(3'- phosphoadenosine 5'-phosphate)·pentacholorphenol(PCP) provided a rationale for a novel SULTs inhibition mechanism that depends on substrate acidity. The SULT1B1·PAP·resveratrol structure suggested that the hydrogen-bonding coordination of the 5-OH group on resveratrol is the structural determinant for the observed substrate preference towards resveratrol. SULT2A1·PAP·lithocholic acid(LCA) ternary complex structure confirms that the specificity of SULT2A1 for lithocholic acid derives

from its high hydrophobicity in the substrate binding pocket.

The same approach was used to interrogate the interaction of the sirtuins with their peptide substrates. The binding and enzymatic assays for human sirtuins have suggested that SIRT1 and SIRT2 are generally less discriminate against substrates while class IV sirtuins - SIRT6 and SIRT7 might be highly specific enzymes. Three different biochemical and kinetic assays showed that SIRT6-dependent histone deacetylation is ~1,000 times slower than for other highly active sirtuins. To understand the molecular basis for the specificity and low activity of SIRT6, I determined the first set of crystal structures for SIRT6 in complex with ADPr (ADP ribose) and the non-hydrolyzable analog of *O*AADPr (2'-*O*-acetyl-ADP ribose) – *N*AADPr (2'-*N*-acetyl-ADP ribose). The structures revealed human SIRT6 has unique structural features including a splayed zinc-binding domain, lacks a helix bundle and the conserved, highly flexible, NAD^+-binding loop, which contribute to its observed biochemical behavior.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my supervisor Dr. Aled Edwards for giving me an awesome yet challenging project and granting me great freedom in carrying out the research. I appreciate the opportunity to purse my graduate degree under his guidance. I would also like to thank Drs. Cheryl Arrowsmith and Denis Grant, members of my supervisory committee, for their useful suggestions, insightful discussions, and their enthusiasm for my doctoral research. I am also grateful to my fellow students in the department of medical biophysics for having been inspiring to study and work with.

I have had the pleasure of working with the scientists in the Structural Genomics Consortium (SGC) over the course of my research, and their contribution to my scientific journey has been immense. I would like to extend my gratitude to all past and present members of the epigenetics, biophysics, and crystallography groups. I would also like to express my sincere thanks to John Denu's lab in University of Wisconsin, Madison, for hosting me during my visit, for teaching me enzymology, and for their contributions toward the sirtuin project.

I could not have completed this Ph.D. program without the support of my family. So thank you, Mom and Dad, for your support, love and guidance throughout my life, esp. the phone calls and video chats. I would also like to thank Andy for the late night adventures, the games, the drinks, the food and all the time we spent together, and for being a good listener. Finally, I would like to express my gratitude to my friends, in particular, Wincy and Paul, for their encouragement and constant reminder that there is a great life waiting for me beyond graduate school!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADPr, ADP ribose

ADT, androsterone

AMC, 7-amino-4-methylcoumarin

AMP-PNP, adenosine 5′-(β,γ-imido) triphosphate

CCD, charge-coupled device

COX-1/COX-2, cyclooxygenase-1/cyclooxygenase-2

CPS1, carbamoyl phosphate synthetase 1

CtIP, C-terminal binding protein (CtBP)-interacting protein

CYP, cytochromes P450

DBHD, 3,5-dibromo-4-hydroxy-benzoic acid (6,8-dichloro-4-*oxo*-4H-chromen-3-ylmethylene) hydrazide

DBHM, 3,5-dibromo-4-hydroxyl-benzoic acid (6-chloro-4-*oxo*-4H-chromen-3-ylmethylene) hydrazide

DCNP, 2,6-dichloro-4-nitrophenol

DHEA, dehydroepiandrosterone

ELISA, enzyme-linked immunosorbent assays

FRET, fluorescence-resonance-energy-transfer

GCIP, Grap2 Cyclin-D interacting protein

hSIRT(s), human sirtuin(s)

hSULT(s), human cytosolic sulfotransferase(s)

ITC, isothermal titration calorimetry

LCA, lithocholic acid

*N*AADPr, 2'-*N*-acetyl-ADP ribose

*O*AADPr, 2'-*O*-acetyl-ADP ribose

OH-PCB, hydroxylated polychlorinated biphenol

PAP, 3'-phosphoadenosine 5'-phosphate

PAPS, 3'-phosphoadenosine 5'-phosphosulfate

PAP(S) = PAPS or PAP

PCP, pentachlorophenol

PLP, pyridoxal 5-phosphate

PCB, polychlorinated bipheny

Sir2, silent information regulator 2

T4, 3,3',5,5'-tetraiodo-L-thyronine

TFA, trifluoroacetic acid

Chapter 1

Introduction and Thesis Overview

1.1 Overview

1.1.1 Structural biology advances lead optimization and drug design

It has long been recognized that knowledge of the three dimensional structures of proteins has the potential to accelerate drug discovery. The flu drug Relenza was designed using the crystal structure of neuraminidase (Varghese 1999) and AIDS drugs, such as Agenerase and Viracept, were developed using the crystal structure of HIV protease (Lapatto, Blundell et al. 1989; Miller, Schneider et al. 1989).

The use of high-throughput crystallography for defining structures of all gene products in an organism is known as structural genomics, and there are several worldwide initiatives to define crystal structures of representative protein family members in several genomes (Lesley, Kuhn et al. 2002; Service 2002; Heinemann, Bussow et al. 2003; Rupp 2003). Structures solved in such initiatives provide understanding of macromolecular function at the atomic-level and are potentially useful as a basis for ligand design.

Arguably, the structural biology of the human kinase family, a protein family comprising over 500 sequences in the genome, has provided more information for drug discovery than for many other protein target classes. Due to high degree of structural conversation in the ATP-binding site, targeting the ATP-binding site of protein kinases for developing sufficiently selective agents for therapeutic use was once considered extremely challenging. However, within the last decade there has been a large number of low molecular weight, potent, ATP-site binders reported, many

of which show good selectivity against a broad range of kinases (Fabian, Biggs et al. 2005). The epidermal growth factor receptor (EGFR) kinase inhibitors, OSI774 (Tarceva) and GW572016 (Lapatinib) (Wood, Truesdale et al. 2004) provide excellent examples. Comparison of these two molecules in the ATP-binding site of the tyrosine kinase EGFR revealed differential binding mode (Lapatinib: inactive kinase structure mode;. Tarceva: active kinase mode) (Wood, Truesdale et al. 2004).

Although an understanding of the structural biology of how these inhibitors bind in the ATP-binding site had a considerable impact on developing efficacious and relatively selective drugs, due to the large number of kinase enzymes in the genome, it is unlikely that highly selective compounds for a kinase target of choice can ever be developed if the highly conserved ATP-binding site is targeted.

1.1.2 Two small protein families for studying protein family substrate preferences

For protein families that transfer atoms or chemical groups from a common co-factor to various substrates, it is reasonable to assume targeting the substrate binding site instead of the conserved co-factor binding site will result in more selective modulators. A structure based understanding of the substrate preference within a protein family will not only help us understand the structural basis for enzymatic activity, but will also provide us with valuable insight into modulator design. To obtain such understanding, both a family-wide biophysical/biochemical ligand profile and representative protein-ligand complexes are crucial.

To test the feasibility of identifying ligand preference and to demonstrate a proof of concept for family-wide ligand screen approach, two small human protein families – the human cytosolic sulfotransferases (SULTs) and the human sirtuins (SIRTs), were subjected to ligand profiling and structural biology approaches. While both enzyme families are much smaller in size than the human kinase family, they each have unique properties make them attractive targets. SULT enzymes have relatively high sequence identity but are known to demonstrate substrate preferences; SIRT enzymes have relatively low sequence identity but their substrate preferences have not been well

characterized. In addition, SULTs belong to a metabolic/detoxification pathway recognizing small molecules as substrates while SIRTs exert epigenetic and post-transcriptional control on protein substrates. The hypothesis is that by comparing and contrasting ligand fingerprints of these two enzyme families and by examining their detailed structural information from co-crystal structures, insights into the specificity and promiscuity of ligand binding for these two protein family can be obtained.

1.2 Introduction to SULT and SIRT protein families

1.2.1 The human cytosolic sulfotransferases (SULTs)

The sulfonation process was first characterized in the late 1870s by Eugen Baumann, who was able to isolate the sulfonate conjugated form of phenol from the urine sample of a patient treated with carbolic acid as an antiseptic. Sulfonation converts xenobiotics into less or non-toxic metabolites, and facilitating their excretion in the urine or bile by making them more water soluble, and therefore is characterized as a detoxification process. In addition to its important role in xenobiotic metabolism in the body, sulfonation has many other functions including biosynthesis, mode of action, and homeostasis of many important endogenous chemicals including steroid hormone, iodothyronines, and catecholamines (Visser 1996; Coughtrie, Sharp et al. 1998; Eisenhofer, Coughtrie et al. 1999).

There are 12 unique human cytosolic sulfotransferases (SULTs) encoded by genes of the SULT superfamily (Figure 1.1) (Freimuth, Wiepert et al. 2004). They are responsible for carrying out sulfonation of small molecule metabolites and xenobiotics (another class of sulfotransferases: membrane-associated sulfotransferases sulfonate larger biomolecules, such as carbohydrates and proteins). The SULTs transfer a sulfonate group from the universal sulfonate donor 3'-phosphoadenosine 5'-phosphosulfate (PAPS), to an acceptor molecule which typically contains at least one hydroxyl group (sulfonation of amine groups has been observed (Yamazoe, Nagata

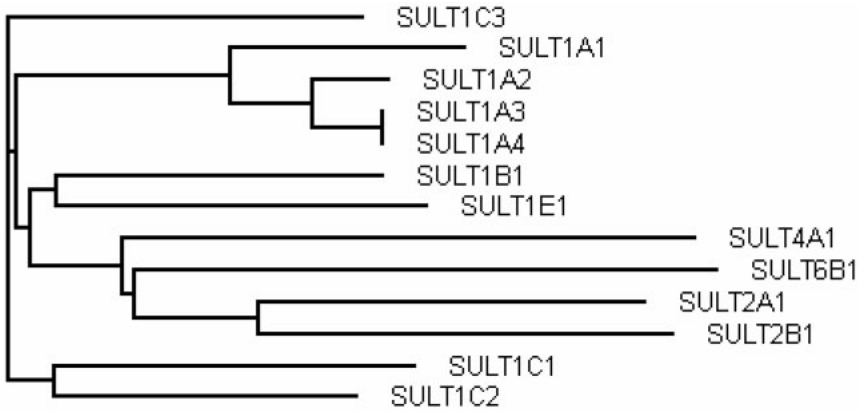


Figure 1.1 **Human SULT family phylogenetic tree.** Human cytosolic sulfotransferase (SULT) protein classification based on primary amino acid sequence.

et al. 1999)), yielding 3'-phosphoadenosine 5'-phosphate (PAP), and the sulfonated product (a reaction scheme is presented in Chapter 2, Figure 2.1). In doing so, the activities of a large array of small endogenous and foreign chemicals including drugs, toxic compounds, steroid hormones and neurotransmitters are modulated.

1.2.2 Biological functions of the human SULT family members

The 12 unique human SULT enzymes partition into four families (Blanchard, Freimuth et al. 2004; Freimuth, Wiepert et al. 2004) - SULT1, SULT2, SULT4 and SULT6. The SULT1 and SULT2 families are the largest, and probably the most important for xenobiotic and endobiotic metabolism. The members in the same family share at least 45% amino acid sequence identity. A SULT family is further divided into subfamilies that include enzymes that share at least 60% amino acid sequence identity (Weinshilboum, Otterness et al. 1997; Blanchard, Freimuth et al. 2004).

The SULT1 family has 4 subfamilies – A, B, C, E, and comprises nine members (1A1, 1A2, 1A3, 1A4, 1C1, 1C2, 1C3, 1B1 and 1E1). The SULT1A3 and SULT1A4 genes appear to have arisen by segmental duplication and encode the same protein (Hildebrandt, Salavaggione et al. 2004).

The three unique SULT1A subfamily members catalyze the sulfonation of many phenolic molecules. SULT1A1 has been studied in the context of conjugation of phenolic drugs such as acetaminophen, minoxidil, and 17α-ethinylestradiol. It also sulfonates endogenous compounds such as 17β-estradiol and iodothyronines as well as environmental xenobiotics such as the

isoflavones. SULT1A2 is very closely related to SULT1A1 at the amino acid level, but exhibits very different affinity for phenolic compounds (Ozawa, Nagata et al. 1995; Her, Raftogianis et al. 1996; Zhu, Veronese et al. 1996; Raftogianis, Wood et al. 1999). Although SULT1A2 cDNAs have been isolated from human liver and colon cDNA libraries, it is not clear that this isoform is expressed in any tissue. SULT1A3 plays the major role in the sulfonation of catecholamines such as dopamine and norepinephrine, a function that is probably specific to humans (and possibly other primates). SULT1A3 is expressed at high levels in fetal liver but hepatic expression is essentially absent in the adult. The gastrointestinal tract is the major site for adults (Richard, Hume et al. 2001), which correlates with the dopaminergic function of the gut, where the majority of dopamine sulfate is produced (Eisenhofer, Coughtrie et al. 1999).

The human SULT1B subfamily has one isoform - SULT1B1. It appears to be involved in the sulfonation of thyroid hormones (Wang, Falany et al. 1998). It also catalyzes the sulfonation of prototypic phenolic substrates such as 2-naphthol and dopamine but not steroid hormones. It appears to show substrate-binding specificity towards thyroid hormones (Wang, Falany et al. 1998) and prototypic phenolic substrates such as 2-naphthol and dopamine but not steroid hormones.

The function of the three SULT1C subfamily members is not yet fully understood. SULT1C1 and 1C2 appear to be most highly expressed in fetal tissues, although there are suggestions that the adult stomach and kidney (SULT1C1) and ovary (SULT1C2) may also be sites of expression (Her, Kaur et al. 1997; Sakakibara, Yanagisawa et al. 1998). SULT1C2 has been shown to catalyze the sulfonation of the procarcinogen *N*-hydoxy-2-acetylaminofluorene as well as *p*-nitrophenol, the prototypic phenolic SULT substrates. Although both SULT1C1 and 1C2 are capable of sulfonating many simple phenols, the function of these enzymes in human is not clear. SULT1C3 gene was identified in a comprehensive human genome-wide search for novel SULT genes (Freimuth, Wiepert et al. 2004). The gene sequence does not contain any obvious inactivating mutation, however, SULT1C3 protein was not detected in 20 different human tissues tested (Freimuth, Wiepert et al. 2004).

SULT1E1 is the only isoform in the human SULT1E subfamily. It is among the most widely

studied of the SULT enzymes. It catalyzes the sulfonation of the 3β-hydroxy group of endogenous (e.g. estrone and estradiol) and xenobiotic (e.g. 17α-ethinylestradiol) estrogen, generally with very high affinity (Falany, Wheeler et al. 1994; Falany 1997). Sulfonation of estrogens inhibits their action at the estrogen receptor, and SULT1E1 has been implicated as important enzyme in estrogen homeostasis, and may be disrupted in breast cancer. The reexpression of recombinant SULT1E1 in the MCF-7 estrogen-responsive breast cancer cell line has been shown to inhibit cell growth (Falany, Macrina et al. 2002). Although the SULT1A1, 2A family also has such activity though with lower affinity, the SULT1E1 isoform is expressed in several steroid hormone-responsive tissues such as the endometrium. It has been shown that SULT1E1 is exquisitely regulated during the menstrual cycle (Buirchell and Hahnel 1975; Rubin, Harrold et al. 1999), where it is postulated to moderate estrogenic stimulation of the endometrium around the time of implantation. A further insight into the importance of estrogen sulfonation has come from the identification that a number of hydroxylated metabolites of the ubiquitous environmental pollutants polychlorinated biphenyls are extraordinarily potent inhibitors (K_i in the pM range) of SULT1E1 (Kester, Bulduk et al. 2000), suggesting a mechanism whereby these chemicals exert their well-known endocrine disrupting effects (Safe 1994). Using standard targeting vector methodology, the first SULT knockout experiment generated SULT1E1 deficient mice in 2001 (Qian, Sun et al. 2001). The phenotype of the knockout mice was observed in males, where age-dependent Leydig cell hypertrophy/ hyperplasia coupled with seminiferous tubule damage occurred. These lesions resulted in abnormal sperm function. In young SULT1E1 knockout mice, a reduction in fertility was observed in the females.

Three proteins (SULT2A1, SULT2B1a and SULT2B1b) are encoded by the two SULT2 family genes. They catalyze sulfonation of hydroxyl groups of steroids, such as androsterone, allopregnanolone and dehydroepiandrosterone. In addition, they also sulfonate a range of xenobiotics, including the bioactiviation of benzylic alcohols of polycyclic aromatic hydrocarbons (Glatt, Engelke et al. 2000). A major difference between the activities of the SULT2A and SULT2B enzymes appears to be that only the SULT2A1 is capable of catalyzing the sulfonation of the phenolic hydroxyl group at the 3 position of estrogens while the SULT2B1 isoforms are more

selective for 3β-hydroxysteriods. SULT2A1 is expressed in adrenal gland, liver, and brain (Falany 1997) and SULT2B1 isoform is expressed in prostate, placenta, and trachea.

SULT2B1 has two variants – SULT2B1a and SULT2B1b. They are the results of alternative transcription initiation, and differ in the amino terminal residues. This difference imparts apparent biochemical distinction such that SULT2B1a preferentially catalyzes the sulfonation of pregnenolone while SULT2B1b sulfonates both cholesterol and pregnenolone (Fuda, Lee et al. 2002).

SULT4A1 is the only member of the SULT4 family and is expressed exclusively in the brain. Although there are speculations suggesting that SULT4A1 may have an important function, no activity or function has been identified for this protein (Falany, Xie et al. 2000).

A series of alternatively spliced SULT6B1 transcripts were identified in the testis, confirming the expression in human, but neither the protein nor its enzymatic activity have been characterized (Freimuth, Wiepert et al. 2004).

1.2.3 The human sirtuins (SIRTs)

The human sirtuins comprise a family of NAD^+-dependent deacetylases conserved evolutionarily from bacteria to humans. Their ortholog in lower organisms – *Sir2* (silent information regulator 2) protein deacetylase has emerged as an important regulator in extending the life spans of *S. cerevisiae*, *C. elegans*, and *D. melanogaster* (Guarente and Picard 2005; Longo and Kennedy 2006). Although the life extending effects of sirtuins in mammals have yet to be demonstrated, the ubiquitously expressed human sirtuins are critical regulators of many cellular pathways including insulin secretion, the cell cycle, and apoptosis, and are associated with a variety of age-associated diseases such as Alzheimer's disease and cancer (Kim, Nguyen et al. 2007; Firestein, Blander et al. 2008; Luo and Altieri 2008; Wang, Sengupta et al. 2008). Sirtuins are classified as the Class III histone deacetylases (HDACs). In contrast to the "classical" Class I and II histone deacetylases (HDACs), which use an eletrophilic Zn^{2+} ion to directly hydrolyse the amide bond

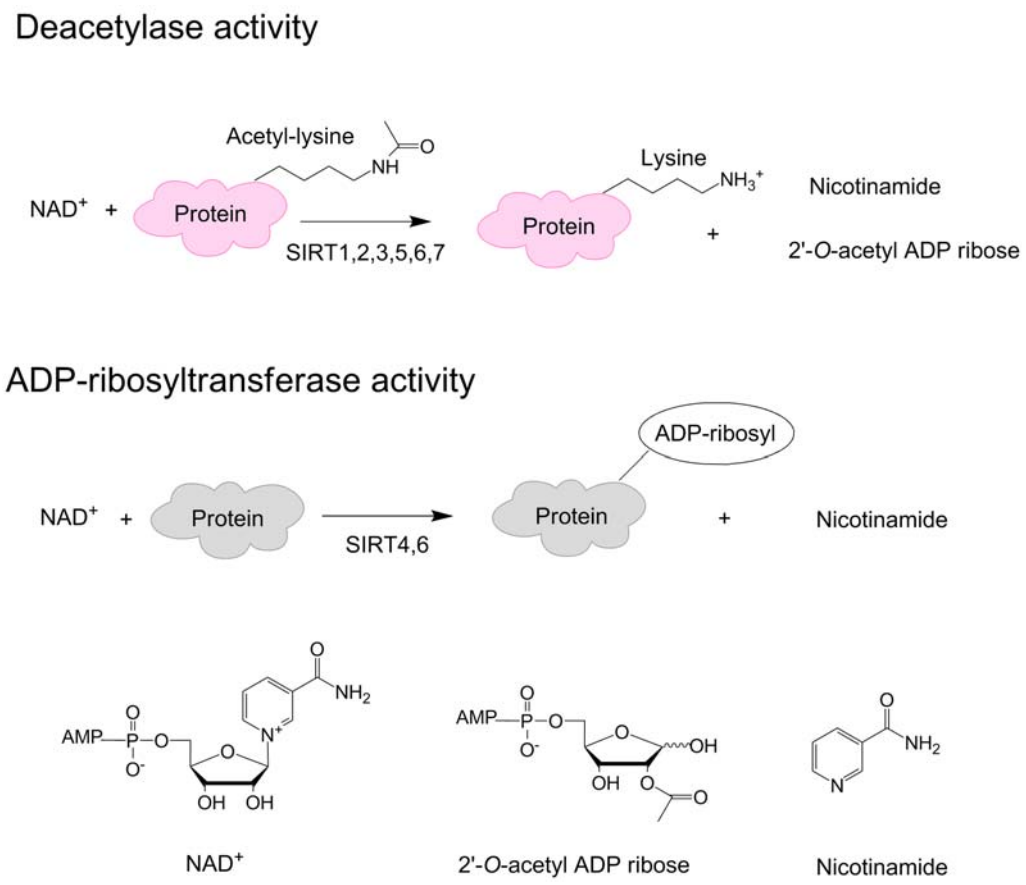


Figure 1.2 **Deacetylation/ADP-ribosyl transfer reaction scheme of SIRTs.**

with water, sirtuins transfer an acetyl group from the lysine side chains of a protein substrate to the co-substrate NAD^+ generating nicotinamide and *O*AADPr (2'-*O*-acetyl-ADP ribose) (Denu 2005) as shown in Figure 1.2. A detailed deacetylation mechanism involving the generation of a peptidyl-imidate intermediate in the first chemical step has been proposed (Sauve, Celic et al. 2001). The requirement of NAD^+ for activity suggested sirtuins may function as cellular energy sensors and serve as a link between cellular metabolism and reverse acetylation mediated cellular pathways. Since all classes of HDACs are capable of deacetylate many non-histone substrates, it is more appropriate to use protein deacetylase as a general designation for these enzymes.

1.2.4 Biological functions of the human SIRT family members

There are seven members in the human sirtuin family, SIRT1-7 (Figure 1.3). Each of them has

distinct cellular targets and diverse cellular localizations (Michishita, Park et al. 2005). SIRT1 has been the most studied, as it shares the highest sequence identity with the founding member of the Sir2 family from yeast. SIRT1 is localized either to the nucleus or the cytoplasm depending on tissue and cell type (Tanno, Sakamoto et al. 2007). Sirt1 transgenic mice display beneficial phenotypes similar to mice on a calorie-restricted diet – they are leaner and more metabolically active (Bordone, Cohen et al. 2007), supporting an anti-aging role of SIRT1. SIRT1 is also linked to oncogenesis. It promoted deacetylation of p53 and p73 as well as E2F1 which represses the expression of target genes inhibiting apoptosis; it also deacetylates BCL-6, which increases its oncogenic activity (Vaziri, Dessain et al. 2001; Bereshchenko, Gu et al. 2002). SIRT1 is also upregulated in human lung cancer, prostate cancer, and leukemia (Yeung, Hoberg et al. 2004; Bradbury, Khanim et al. 2005; Kuzmichev, Margueron et al. 2005).

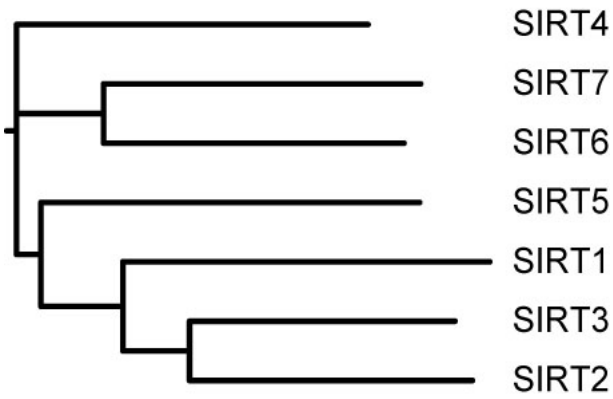


Figure 1.3 **Human SIRT family phylogenetic tree.** Human sirtuin protein classification based on primary amino acid sequence.

SIRT2 is thought to be the only cytoplasmic sirtuin. It has been implicated in the process of cell division via α-tubulin deaceylation (North, Marshall et al. 2003) and histone H4 lysine 16 deaceylation during mitosis (Vaquero, Scher et al. 2006). SIRT1 and SIRT2 cytoplasmic have both been demonstrated to undergo nucleo-cytoplasmic shuttling (North and Verdin 2007).

SIRT3, SIRT4, and SIRT5 are located in mitochondria but the comparison of *Sirt3*$^{-/-}$ and *Sirt3*$^{+/+}$ mice has provided compelling evidence that endogenous Sirt3 is responsible for the majority of protein deacetylation in mitochondria (Lombard, Alt et al. 2007). SIRT4 has exhibited no deacetylase activity to date. However, SIRT4 is reported to inhibit glutamate dehydrogenase through ADP-ribosylation and it plays a role in insulin secretion (Haigis, Mostoslavsky et al. 2006). Although the mouse knockout of *Sirt5* does not affect the bulk acetylation state of mitochondrial proteins (Lombard, Alt et al. 2007), SIRT5 deacetylates and activates carbamoyl phosphate

synthetase 1 (CPS1), which mediates the first step in the urea cycle (Nakagawa, Lomb et al. 2009).

SIRT6 and SIRT7 are chromatin-associated sirtuins. SIRT6 is linked with heterochromatic regions and SIRT7 is found in nucleoli (Michishita, Park et al. 2005). SIRT6 has both ADP-ribosyl transferase and deacetylase activity, and modulates telomeric chromatin (Michishita, McCord et al. 2008) and NF-κB-dependent gene expression through histone H3 Lysine 9 deacetylation (Kawahara, Michishita et al. 2009). Loss of *Sirt6* leads to a shortened lifespan and premature aging (Mostoslavsky, Chua et al. 2006). Neural-specific deletion of *Sirt6* in mice leads to postnatal growth retardation, but the animals reach normal size and ultimately become obese over time (Schwer, Schumacher et al. 2010). SIRT7 is involved in the activation of RNA polymerase I transcription (Ford, Voit et al. 2006). *Sirt7*-knockout mice have shortened life spans with enhanced inflammatory cardiomyopathy (Vakhrusheva, Smolka et al. 2008).

1.3 Structural information governing specificity enables rational design of specific molecular modulators

1.3.1 The interest in human SULTs molecular modulators

In the field of drug discovery, sulfonation pathways have become a promising intervention points due to the many adverse drug reactions that are directly related to genetic and/or environmental effects on the cytochromes P450 (CYP) enzymes. The consequence of this is that many drugs coming to the market in the future will not be metabolized via the CYP system (Hodgson 2001), but rather by conjugating pathways, such as, the SULT system. It will therefore be necessary to fully understand the structure, function and substrate specificities of these SULT enzymes in both human and experimental animal species in order to eventually predict the fate of chemical entities metabolized by them. The anti-diabetic drug troglitazone, serves an example to illustrate this point. This drug was withdrawn from the US market in 2000 only after three years of its launch due to the inhibition of bile salt transport by troglitazone sulfate, which has been implicated in

the severe idiosyncratic hepatoxicity of troglitazone (Funk, Ponelle et al. 2001). In addition, for an increasingly large number of xenobiotics, including environmental pollutants and drugs such as tamoxifen and acetaminophen, or endogenous neurotransmitter such as dopamine, for which sulfonation results in increased biological activity (Glatt 1997). However, for xenobiotics such as *N*-hydroxy arylamines, *N*-hydroxy heterocyclic amines, and hydroxymethyl polycyclic aromatic hydrocarbons, sulfonation is a metabolic activation process leading to highly reactive electrophiles that are both mutagenic and carcinogenic (Falany 1997; Weinshilboum, Otterness et al. 1997).

Designing small molecule inhibitors has long been an interest in the SULT field in order to exploit their potential as drug targets. The main categories of these inhibitors are: cofactor (PAPS) analogues (Horwitz, Misra et al. 1978), a variety of small molecules including 2,6-dichloro-4-nitrophenol (DCNP) and pentachlorophenol (PCP) (Mulder and Scholtens 1977; Kester, Bulduk et al. 2000), various aryl carboxylic acids (Rao and Duffel 1991), dietary compounds and food additives (Bamforth, Jones et al. 1993; Walle, Eaton et al. 1995; Coughtrie and Johnston 2001), common drugs (Bamforth, Dalgliesh et al. 1992), amines (Matsui, Takahashi et al. 1995), and synthetic bisusbstrate analogues (Iyer, Butler et al. 1983).

1.3.2 The highly variable substrate binding site hinders understanding of SULTs specificity

Despite the discovery of many broad spectrum SULT inhibitors, the design of potent and specific inhibitors for each of the human SULT enzymes remains a challenge due to the fact that SULT enzymes are promiscuous with only some degree of substrate specificity.

The binding site of the cofactor PAPS is well characterized due to the availability of many SULT crystal structures complexed with the cofactor product PAP (Pedersen, Petrotchenko et al. 2000; Lee, Fuda et al. 2003; Shevtsov, Petrotchenko et al. 2003; Dombrovski, Dong et al. 2006) and human SULT1E1 complexed with the cofactor PAPS (Pedersen, Petrotchenko et al. 2002). However, only a few structures have been determined with both cofactor product PAP and substrate bound (Gamage, Duggleby et al. 2003; Lee, Fuda et al. 2003; Gamage, Tsvetanov et al.

2005; Lu, Li et al. 2005). The active sites of apo form enzymes exhibit great mobility. Accordingly, this limited knowledge of the flexible substrate binding site has hindered elucidation of structural principles underlying the recognition and utilization of a given substrate or inhibitor.

1.3.3 The interest in human SIRTs molecular modulators

Human sirtuin activity can be modulated by potentially therapeutic small molecules (Howitz, Bitterman et al. 2003; Grubisha, Smith et al. 2005; Porcu and Chiarugi 2005). A number of nonspecific and specific inhibitors of human sirtuins have been discovered by testing natural products, by biochemical- or cell-based screening of small molecule libraries using a forward chemical genetics approach, or by chemical synthesis. Since NAD^+ functions as a cosubstrate for sirutin activity, compounds that are structurally related to NAD^+ are naturally the first inhibitors to be identified. A non-hydrolysable NAD^+ analogue, carba-NAD^+, and nicotinamide have been identified as non-competitive inhibitors of sirtuins (Denu 2005; Grubisha, Smith et al. 2005). For the similar reason, ADPr is also predicted to a broad-spectrum inhibitor of the sirutins. However, due to the presence of a large number of other NAD^+-dependent enzymes, such non-specific inhibitors, might evoke unexpected side effects (Grubisha, Smith et al. 2005; Belenky, Bogan et al. 2007).

Specific inhibitors have been reported for SIRT1 and SIRT2. The 2-hydroxynaphthaldehyde derivative sirtinol has been used extensively in the literature as SIRT1 inhibitor (SIRT1 IC_{50}: 130 μM) (Ota, Tokunaga et al. 2006). Other inhibitors include the coumarin derivative splitomicin (SIRT1 IC_{50}: 60 μM) (Bedalov, Gatbonton et al. 2001), the splitomicin derivative HR-73 (SIRT1 IC_{50}: 5 μM) (Hirao, Posakony et al. 2003), and the indole derivative EX-527 (SIRT1 IC_{50}: 90 nM) (Napper, Hixon et al. 2005). The only sirtuin inhibitors that have been tested in animal models of cancer are cambinol (SIRT1 IC_{50}: 56 μM; SIRT2 IC_{50}: 59 μM) (Heltweg, Gatbonton et al. 2006) and the tenovins (tenovin-6, SIRT1 IC_{50}: 21μM; SIRT2 IC_{50}: 20 μM; SIRT3 IC_{50}: 67 μM) (Lain, Hollick et al. 2008). Both of them are reported to selectively inhibit SIRT1 and SIRT2. AC-93253 (SIRT1 IC_{50}: 45.3 μM; SIRT2 IC_{50}: 6.0 μM; SIRT3 IC_{50}: 24.6 μM) (Zhang, Au et al. 2009)

and vinylnitrile AGK2 (SIRT2 IC_{50}: 3.5 μM) (Outeiro, Kontopoulos et al. 2007) showed selective inhibition of SIRT2 and both inhibitors have lower potency for two closely related SIRT family members, SIRT1 and SIRT3. The substitution of thioacetyl-lysine for acetyl-lysine into sirtuin-specific substrates is yet another strategy. A thioacetyl-lysine-based SIRT2 inhibitor based on the peptide sequence of α-tubulin that was 10- and 39-fold selective over SIRT1 and SIRT3 has been reported (Fatkins and Zheng 2008). Recently, an acetylated lysine-ADP ribose conjugate was reported to inhibit SIRT1 highly selectively over SIRT1 and SIRT2 (Asaba, Suzuki et al. 2009).

Activation of SIRT1 and SIRT3 has been suggested for the treatment of metabolic diseases, whereas SIRT2 and SIRT4 inhibition may be useful in certain cancers and neurodegeneration. Although the data is limited for SIRT5-7, it can be speculated that the activation of sirtuins could lead to a new platform opportunity for drug discovery.

1.3.4 Limited knowledge of ligand and protein interaction for human sirtuins

To date, three (SIRT2, 3, and 5) out of the seven human sirtuins have been structurally characterized. The SIRT2 apo structure was reported in 2001 (Finnin, Donigian et al. 2001) and three SIRT3 complex structures with acetyl-coenzyme A synthetase 2-like peptides were reported recently (Jin, Wei et al. 2009). The only sirutin crystal structure solved with an inhibitor molecule is the SIRT5 structure in complex with suramin, a small molecule identified through our thermo-denaturation screen for SIRT5 stabilizing compounds (Schuetz, Min et al. 2007). This structural study provided a view of a synthetic broad spectrum inhibitory compound in a sirtuin active site revealing that suramin binds into the NAD^+, and the substrate-binding sites. Suramin inhibits SIRT5 NAD^+-dependent deacetylase activity with an IC_{50} value of 22 μM.

Due to the limited number of available sirtuin crystal structures, the low sequence identities of sirtuin core domains, and the even more divergent N and C terminal flanking domains within the sirtuin family (Michishita, Park et al. 2005), it is difficult to understand the mechanisms behind inhibition or activation despite the availability of biochemically identified inhibitors and activators.

1.4 Objective and overall experimental design

The overall objective of my thesis is to gain structural based insights of the two protein families – human cytosolic sulfotransferase (SULTs) and sirtuins (SIRTs). My goal was to understand the structural basis underlying the specificity and promiscuity of these enzymes, in the hopes of using such information to design specific molecular modulators. My first aim was to determine biophysical and biochemical profiles of the protein families. My second aim was to solve X-ray crystal structures of proteins with their substrates or with compounds that have been identified in the biophysical and biochemical studies. My third aim was to relate the binding and assay data with the structural data to explain in part the protein family member specificity.

For the SULTs and SIRTs families, focused potential ligand libraries were constructed according to the known properties of each of the protein families. Then the focused library was screened against purified members of the protein family using a binding assay to identify potential interacting partners. Ligands identified in the binding assay were then subjected to a secondary enzymatic assay to confirm hits and to rule out any false ones. Co-crystal trials were then set up for proteins and their corresponding "hits". When co-crystal structures were solved, the conformation of the active site residues was closely examined along with the interaction of these residues and ligands. Such structural information and the biochemical/biophysical profiles were compared and contrasted within the same protein family to advance the understanding of the ligand preference.

1.5 Summary of thesis chapters

The next two chapters of this thesis focus on the human SULTs family. Chapter 2 is a stand alone story describing a family wide chemical and structural profiling of the entire human SULTs family. It has been previously published (Allali-Hassani, Pan et al. 2007). Chapter 3 is closely connected to the chemical profiling results in Chapter 2. It provides two structural case studies of SULTs with compounds - resveratrol and lithocholic acid. I was also involved in a combined analysis

(not presented in this thesis) to explore the binding site similarities among members of the human SULTs and correlate the predicted binding profile with and the experimental data (Najmanovich, Allali-Hassani et al. 2007). For Chapter 4 and 5, I switch to the human SIRTs family. Chapter 4, a systematic histone peptide binding array followed by focused secondary enzymatic assay, serves as a prequel to to Chapter 5. The results suggested the structurally uncharacterized SIRT6, a class IV sirtuin, is highly specific towards certain peptide sequence. In Chapter 5, the low activity of SIRT6 has been confirmed and a structural based explanation was provided for this unexpected property. This chapter is itself a complete story and has been published (Pan, Feldman et al. 2011). In the final Chapter, I recapitulate the major findings of this thesis followed by their future perspectives.

During the course of my PhD studies, I also had the pleasure to participate in several epigenetic projects in addtion to my thesis work. These projects included: the importance of the human histone acetyltransferase MOF (MYST1) and histone mark - H4K16Ac in DNA damage repair (Li, Corsa et al. 2010); the regulation of activity and specificity of polycomb repressive complex 2 (PRC2) through the binding of different histone marks (Xu, Bian et al. 2010); and structural studies of two malignant brain tumor (MBT) repeat proteins - MBT domain 1 (MBTD1) (Eryilmaz, Pan et al. 2009) and lethal(3)malignant brain tumor-like 2 (L3MBTL2) (Guo, Nady et al. 2009).

Chapter 2

Structural and Chemical Profiling of the Human Cytosolic Sulfotransferases

This Chapter is adapted from:

Allali-Hassani A, Pan PW, Dombrovski L, Najmanovich R, Tempel W, et al. 2007 Structural and Chemical Profiling of the Human Cytosolic Sulfotransferases. PLoS Biol 5(5): e97. doi:10.1371/journal.pbio.0050097

My contributions towards this chapter:

I performed many of the activity assays, purified the proteins for activity and binding assays and crystallized SULT1C2 with PAP and PCP. AA-H performed all binding assays, developed the SULT activity assay and performed some of the activity assays. LD purified and crystallized all other proteins whose structures are reported here. RN performed structure, binding and activity profile analysis. WT solved the structure of SULT1C2 with PAP and PCP, and performed structural analysis. AD helped with x-ray data collection and crystallization and solved the structures of SULTs 1B1, 1C1, 1C2 and 4A1. PL generated all the expression clones. FM performed computational analysis of SULT4A1 and SULT1C2 with PAP and PCP, and PCBs. JT contributed to data analysis and the manuscript. AB led the crystallographic part of the project and contributed to data analysis and the manuscript. ANP led the structural genomics part of the project and contributed to structure analysis and the manuscript. MV led the screening part of the project and contributed to data analysis and the manuscript. AME contributed to overall project design and the manuscript. CHA contributed to overall project design, data analysis and lead the preparation of the manuscript.

2.1 Introduction

Cytosolic sulfotransferases (SULTs) comprise a family of enzymes that catalyze the transfer of a sulfonate group from 3'-phosphoadenosine 5'-phosphosulfate (PAPS) to an acceptor group of the substrate (Figure 2.1). In doing so, they modulate the activities of a large array of small endogenous and foreign chemicals including drugs, toxic compounds, steroid hormones and neurotransmitters. Since sulfonated molecules are highly soluble in water and easily excreted from the organism, SULTs are often referred to as enzymes of chemical defense. In some cases, however, SULTs activate certain compounds from food and the environment into mutagenic and carcinogenic metabolites(Glatt 2000).

To date, 13 human cytosolic sulfotransferase (hSULT) genes have been identified; they partition into four families (Blanchard, Freimuth et al. 2004; Freimuth, Wiepert et al. 2004) - SULT1, SULT2, SULT4 and SULT6. Although the family members share considerable sequence and structural similarity, they appear to have different biological functions. The SULT1 family comprises 9 members divided in four subfamilies (1A1, 1A2, 1A3, 1A4, 1C1, 1C2, 1C3, 1B1 and 1E1). The SULT1A3 and SULT1A4 genes appear to have arisen from a segmental duplication and encode the same protein (Hildebrandt, Salavaggione et al. 2004). Members of the SULT1 family have been shown to sulfonate simple phenols, estradiol and thyroid hormones, as well as environmental xenobiotics and drugs. The SULT2 family has two genes, encoding three proteins (SULT2A1, SULT2B1a and SULT2B1b), which catalyze sulfonation of hydroxyl groups of steroids, such as androsterone, allopregnanolone and dehydroepiandrosterone. SULT4A1 is the only member of the SULT4 family. The fact that it is highly conserved and expressed primarily in the brain suggests an important function, however, no activity or function has been identified for

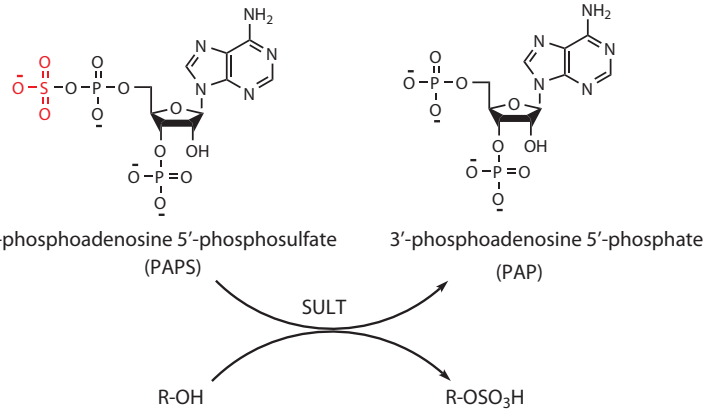


Figure 2.1 **Schematic of the reaction catalyzed by SULTs**.

this gene (Falany, Xie et al. 2000). Finally the SULT6B1 gene is expressed in the testis of primates, but neither the protein nor its enzymatic activity have been characterized (Freimuth, Wiepert et al. 2004).

Recent progress in the structural biology and characterization of the catalytic mechanism of hSULTs has established that many family members have distinct but overlapping substrate specificities and that the enzymes have a sequential catalytic mechanism that is susceptible to substrate inhibition (Chapman, Best et al. 2004; Gamage, Tsvetanov et al. 2005). Nevertheless, only a few of the human enzymes have been subjected to detailed structural and mechanistic studies (Bidwell, McManus et al. 1999; Pedersen, Petrotchenko et al. 2000; Pedersen, Petrotchenko et al. 2002; Rehse, Zhou et al. 2002; Gamage, Duggleby et al. 2003; Lee, Fuda et al. 2003; Shevtsov, Petrotchenko et al. 2003; Chang, Shi et al. 2004; Gamage, Tsvetanov et al. 2005; Lu, Li et al. 2005) and there are no reports of a systematic comparison among all the hSULTs. Understanding the structural and mechanistic basis for specificity among hSULTs is essential to elucidate their role in the metabolism of regulatory hormones, drugs and carcinogens, and may assist in chemical risk assessment and the design of more effective therapeutics. Here we report the crystal structures of five of the 12 structurally unique hSULTs. These structures, combined with those previously reported for 6 other hSULTs allowed a comprehensive comparison of both global and local structural features. We further screened 9 hSULTs for binding activity toward a set of 90 potential substrates and inhibitors, and 8 hSULTs for enzymatic activity towards 31 potential substrates in order to better understand the relationship between binding specificity, activity and structure within the hSULT family. These data combined with detailed structural analysis of substrate binding sites reveal relationships between family members not previously apparent from sequence analysis. “Chemical fingerprints” of the spectrum of small molecules that bind in the presence and absence of the cofactor product, 3’-phosphoadenosine 5’-phosphate (PAP), demonstrate a marked change in the small molecule binding profile upon PAP binding. This result combined with the structural data suggests PAPS has a strong influence on which compounds may bind in the substrate binding site and raises the possibility that the enzymes might be inhibited by chemically related compounds that are not productive substrates. The binding studies also provide insight into potential functions

of the under-characterized SULT1C subfamily and of SULT4A1, an orphan member of the SULT family expressed primarily in the brain.

2.2 Methods

2.2.1 Protein purification and crystallization

The SULT1B1, SULT1C1, SULT1C2, SULT1C3 and SULT4A1 genes were amplified by PCR from the Mammalian Gene Collection clones (accession codes gi:29550928, gi:4507305, gi:28830308, gi:56847626, gi:7657633 for SULT1B1, SULT1C1, SULT1C2, SULT1C3, SULT4A1, respectively) and subcloned into a modified pET28a-LIC vector. Expression and purification of recombinant proteins was as described by Dombrovski *et al* (Dombrovski, Dong et al. 2006). Purified recombinant proteins contained an additional Gly-Ser dipeptide at the N-terminus. Additional details are provided at http://www.sgc.utoronto.ca in the structure gallery for each protein.

Purified SULT1B1, SULT1C1, and SULT1C3 were crystallized in the presence of 2 mM PAP using the hanging drop method at 20 °C by mixing: for SULT1B1 - 2 µl of the protein solution with 2 µl of the reservoir solution containing 0.1 M Bis-Tris, pH 6.5, 0.2 M ammonium sulfate and 16-20% polyethylene glycol 4000; for SULT1C1 - 2 µl of the protein solution with 2 µl of the reservoir solution containing 0.1 M K_2HPO_4 and 12-16% polyethylene glycol 3350; for SULT1C3 - 2 µl of the protein solution with 2 µl of the reservoir solution containing 18% polyethylene glycol 3350, 0.2 M ammonium formate, 0.1 M Bis-Tris, pH 6.5. To obtain crystals of SULT1C2•PAP•pentachlorphenol ternary complex, 10 mg/ml of purified SULT1C2 was mixed with 2 mM PAP and 2 mM PCP in 20 mM MES-NaOH buffer, pH 6.5, and incubated on ice for 30 minutes. SULT1C2•PAP•PCP complex was crystallized using the sitting drop method at 20°C by mixing 0.8 µl of the protein-cofactor-inhibitor mix with 0.8 µl of the reservoir solution containing 25% polyethylene glycol 3350, 0.2 M lithium sulfate, 0.1 M Bis-Tris pH 6.5. SULT4A1

and SULT1C2 crystals were obtained by using the hanging drop method at 20 °C by mixing 2 µl of the protein solution with 2 µl of the reservoir solution containing 20% polyethylene glycol 4000, 0.2 M ammonium tartrate, and 14-20% polyethylene glycol 3350, 0.2 M lithium citrate, 0.1 M sodium citrate, pH 4.6, respectively.

2.2.2 Chemical library preparation

A library of 90 compounds was created for screening sulfotransferases. These compounds were known substrates, products and inhibitors of sulfotransferases, their analogues, and compounds with high similarity to known inhibitors identified from the literature and public databases (www.rcsb.org and www.brenda.uni-koeln.de). Certain substrates such as controlled substances were not included, and some additional compounds were selected through chemical similarity to known SULT substrates and inhibitors using the ChemNavigator search engine (http://www.chemnavigator.com/). The compounds were dissolved in 100% DMSO at 100 mM concentration and subsequently diluted stepwise to 10 mM and 1 mM in Hepes buffer (100 mM Hepes, 150 mM NaCl pH 7.5). The full list of compounds in the library is included in supporting material (Table S2.2).

2.2.3 Ligand binding screens

Screening for ligand binding was performed in 50 µl volume with a final concentration of 1 mM of compound per well, in 384 well plates. The concentration of protein was the same for all wells at 0.4 mg/ml. Ligand binding was detected by monitoring the increase in thermostability of proteins in the presence of ligands. Protein thermostability at pH 7.5 was studied using StarGazer technology that monitors protein stability by its aggregation properties (Senisterra, Markin et al. 2006; Vedadi, Niesen et al. 2006). Protein samples at 0.4 mg/ml were heated from 27 to 80 °C at the rate of 0.5 °C per min in clear-bottom 384-well plates (Nunc) in 50 µl of 100 mM Hepes pH 7.5 and 150 mM NaCl. Protein aggregation was monitored by capturing images of scattered light every 30s with

a CCD (Charge-Coupled Device) camera. The pixel intensities in a pre-selected region of each well were integrated to generate a value representative of the total amount of scattered light in that region. These total intensities were then plotted against temperature for each sample well and fit to the Boltzman equation by nonlinear regression. The point of inflection of each "denaturation" curve was identified as T_{agg} (aggregation temperature). The increase in stability of the protein in the presence of a ligand is shown as ΔT_{agg}.

2.2.4 Sulfotransferase activity screens

Enzyme assays were performed using a HPLC based method which we developed for sulfotransferase activity assay by modifying the protocol previously used to monitor ADP production and ATP hydrolysis by a purified bacterial ATPase (Allali-Hassani, Campbell et al. 2004). SULTs at 1-5 μM were assayed in the presence of 0.1 to 0.5 mM PAPS and different concentrations of each substrate in 100 mM Hepes pH 7.5 by incubating the reaction at 37°C for a period of time from 15 to 120 minutes depending on how fast PAPS was converted to PAP. The K_m values for characterized sulfotransferases are in the range of nM to mM concentrations (Brix, Barnett et al. 1999; Schrag, Cui et al. 2004), with a significant variation in catalytic efficiency and substrate specificity. Based on these observations and considering possible substrate inhibition (Barnett, Tsvetanov et al. 2004; Gamage, Barnett et al. 2006), we tested all sulfotransferases at substrate concentrations of 10, 25 and 100 μM. The reactions were stopped by adding 2 volumes of urea (final concentration of 5.3 M) and the mixture was filtered through a 5 kDa MW cutoff Amicon Ultrafree-MC filter (Millipore, Bedford, MA, U.S.A.) to remove the protein. The ratio of PAP and PAPS was determined after separating them on HPLC using a 4.5 mm x 50 mm WP QUAT, a strong ion-exchange column (J. T. Baker, Phillipsburg, NJ, U.S.A.), using a gradient of triethylamine bicarbonate from 20 to 500 mM applied at 2 ml/min for 7 minutes. The progress of the reaction was monitored by reading the absorbance at 259 nm and the amounts of PAP produced and the remaining PAPS were determined by intergrating the areas underneath the resolved peaks corresponding to PAP(S) using the HPLC software (Waters). All values in Table 2.2 were corrected for the rate of conversion of PAPS to PAP

in the presence of enzyme, but no substrate. This background activity is reported for each SULT in the last row of Table 2.2 and the values are the average of three independent measurements.

2.2.5 Sequence, structure and data clustering analysis

Sequences used to generate Figures 2.3 and 2.8 are as follows: SULT1A1 - NP_803880, SULT1A2 - NP_001045, SULT1A3 - NP_003157, SULT1A4 - NP_001017389, SULT1B1 - NP_055280, SULT1C1 - NP_789795, SULT1C2 - NP_006579, SULT1C3 - NP_001008743, SULT1E1 - NP_005411, SULT2A1 - NP_003158, SULT2B1 - NP_814444, SULT4A1 - NP_055166, SULT6B1 - NP_001027549. We created a multiple sequence alignment of the above mentioned sequences using HMMer (Eddy 1998), and the pfam (Bateman, Coin et al. 2004) Sulfotransferase_1 (PF00685.15) Hidden Markov Model. Sequence similarity is measured using the Tanimoto coefficient of residues in common in the HMM-based alignment. The calculation of local sequence similarities involves the detection of binding site residues (Laskowski 1995), and their subsequent mapping onto the HMM-based alignment. Cofactor (PAP/PAPS) binding residues were also mapped onto the alignment and excluded from all pairwise comparisons and similarity calculations. Substrate binding site structural similarities were detected using a two-stage graph-matching process providing a one-to-one chemical and spatial correspondence between atoms in clefts. The method considers all non-hydrogen atoms and can use large sets of atoms as input allowing larger, over-predicted and apo-form binding sites to be analyzed. In the first stage, the two clefts are superimposed (Arun, Huang et al. 1987) via the detection of the largest clique (Bron and Kerbosch 1973) in a C_{α} association graph corresponding to the largest subset of identical residues in equivalent spatial locations. The first stage is used to constrain the construction of the second stage all-atom association graph. The second stage graph matching results in the detection of the largest subset of heavy atoms of equivalent atoms types (Sobolev, Sorokine et al. 1999), and spatial positions. Pairwise local structural similarity was calculated as a Tanimoto coefficient based on the size of the largest clique in the second graph matching stage. Dissimilarity matrices were derived

from the similarity measures described above. Pairwise experimental catalytic and binding profile dissimilarity matrices were calculated as the L_2 distance of the vectors with the corresponding experimental measurements. Hierarchical clustering was used to create the clustering trees shown in Figure 2.8. The correlation between the cophenetic matrix and the original dissimilarity matrix was used to choose the linkage method that results in the most accurate representation of the original data (Everitt 2001). Average linkage was found to be the clustering method of choice in all instances.

2.2.6 pK_a calculations of polyphenol hydroxyl moiety

To assess the extent to which the acidity of the hydroxyl moiety of hydroxylated polychlorinated biphenols is related to their inhibitory strength, we computed the pK_a values of the 4-hydroxyl group for a series of hydroxylated PCB (polychlorinated bipheny) analogs. This group of compounds and their inhibitory effect on SULT1E1 were previously reported (Kester, Bulduk et al. 2000). To this end we used the pKa calculator in the PC stand alone version of the ACD (www.acdlabs.com) suite of programs. Two clusters of compounds, namely 4-OH-(2,3,4,5,6)Cl and 4-OH-(3,5)Cl are identified as outliers upon computing a linear regression on the relationship ($R^2 = .57$).

2.3 Results

2.3.1 Completion of the structural coverage of hSULTs

The crystal structures of SULT1C3 bound to PAP, apo SULT1C2, a ternary complex of SULT1C2 bound to PAP and the environmental toxin, pentachlorophenol (PCP), and SULT4A1 were solved at 3.2, 2.0, 1.8 and 2.2 Å respectively (Table S2.1 and Figure S2.1). We also recently reported the structures of SULT1B1 and SULT1C1 bound to PAP at 2.1 and 1.8 Å respectively (Dombrovski, Dong et al. 2006). The structures of a single subunit of each of these normally dimeric proteins are

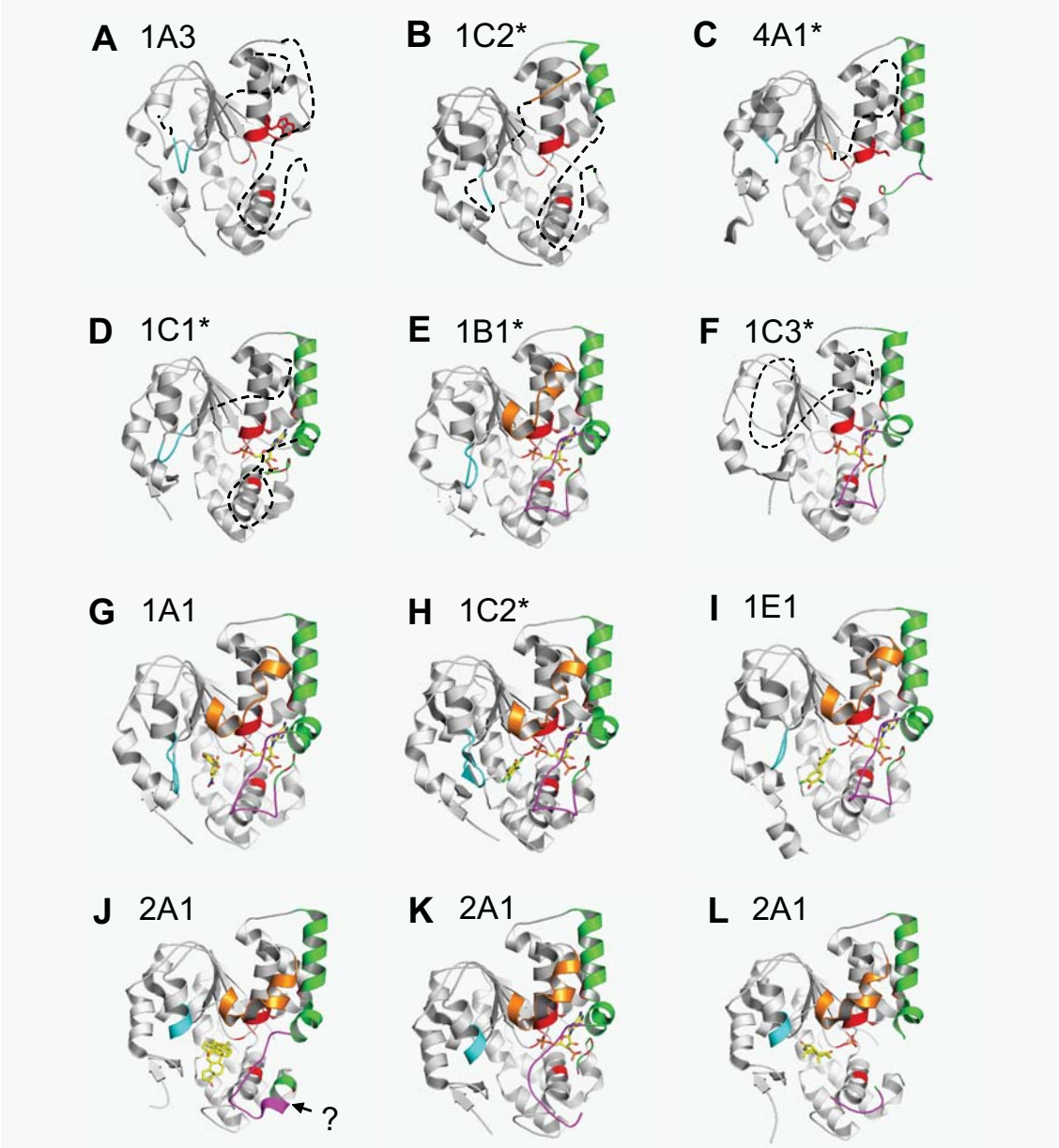


Figure 2.2 **Representative structures for each of the hSULTs grouped according to similar global structural features** (Structures we solved are labeled with *). Apo proteins with considerable disorder in several loops: (A) SULT1A3 (1CJM) (B) SULT1C2 (2AD1), and (C) SULT4A1 (1ZD1). Structures with increased order due to binding of PAP (green helices, purple loop): (D) SULT1C1 (2ETG), (E) SULT1B1 (1XV1), and (F) SULT1C3 (2H8K). Ternary complexes display further ordering of substrate binding loops: (G) SULT1A1 with PAP and *p*-nitrophenol (1LS6), (H) SULT1C2 with PAP and PCP (2GWH), (I) 1E1 with PAP and 3,5,3',5'-tetrachloro-biphenyl-4,4'-diol (1G3M). Structures with unusual features that likely reflect catalytically unproductive proteins: (J) SULT2A1 bound to DHEA but without PAP (1J99); compare to (K) structure of the same protein with PAP (1EFH), and (L) with androsterone (1OV4). The question mark brings the attention to a helix formation, which leads to a non-productive conformation. The PDB code for each structure is shown in parentheses. The proteins are represented as ribbon model, PAP and substrate as stick models colored per-element (carbon in yellow, oxygen in red, nitrogen in blue, and phosphate in magenta). The loops are colored as discussed in the text (gold, cyan, green, purple).

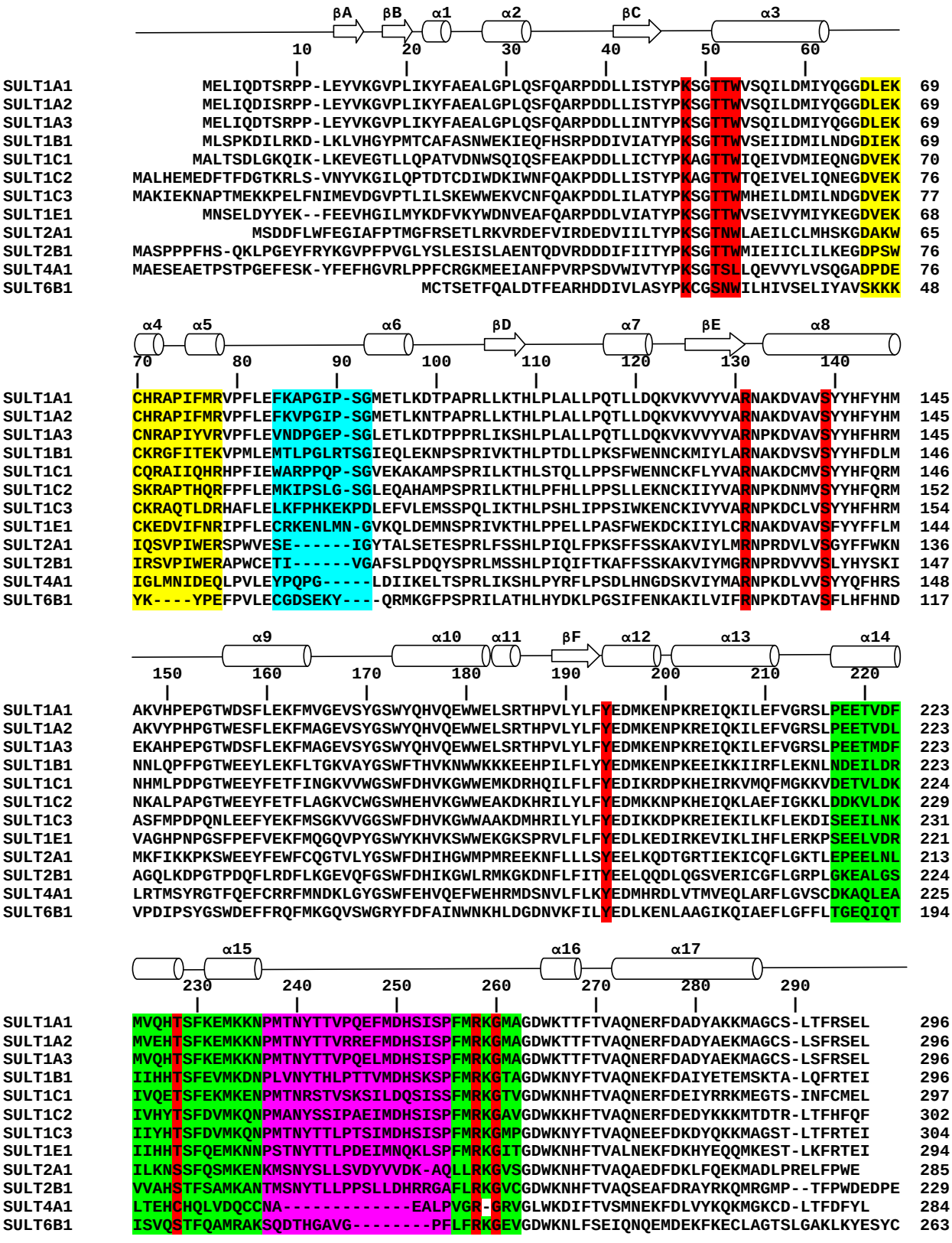


Figure 2.3 **Sequence Alignment of hSULTs showing structural features with color coding corresponding to that of the structures in Figure 2.2**. Highly conserved residues contributing to the PAP(S) binding pocket are highlighted in red.

presented in Figure 2.2 along with a representative structure of each of five other SULT family members previously reported in the literature (Bidwell, McManus et al. 1999; Pedersen, Petrotchenko et al. 2000; Pedersen, Petrotchenko et al. 2002; Rehse, Zhou et al. 2002; Gamage, Duggleby et al. 2003; Shevtsov, Petrotchenko et al. 2003; Chang, Shi et al. 2004). Six additional SULT structures, which are available in the Protein Data Bank (PDB) are presented in Figure S2.2. As expected, all SULTs share the same basic fold: a central four-stranded parallel β-sheet surrounded by α-helices and three loops that are often disordered (dashed lines) in the absence of PAP and/or substrate. These disordered segments comprise a 13 residue loop (shown in gold), a 4-10 residue loop (cyan), and a large 32-46 residue loop (green-and-magenta). These loops have been mapped onto the aligned protein sequences in Figure 2.3 using the same coloring scheme. The degree of disorder and the exact conformation of these loops vary considerably across the family, but in general, the presence of ligands (cofactor and/or substrate) is coupled with increased order, namely the formation of helices α4-α5 (gold), and α14-α15 (green). In some cases partial stabilization can be attributed to molecular packing in the crystal, for example the stabilization of α14 (green) in apo SULT1C2. The binding site for PAP(S) is nearly identical in all structures bound to these ligands, with highly conserved residues contributing to the binding pocket (highlighted in red in Figures 2.2 and 2.3). It is interesting to note that the SULT6B1 sequence in the protein databases (NP_001027549) lacks the N-terminal region, which encodes a β-sheet thought to be an important structural component of the SULT fold (Figure 2.3). We note that the recombinant SULT6B1 did not express in our attempts to purify it from bacteria.

2.3.2 Structural comparison supports a role for PAPS in priming the conformation of substrate binding loops

It is generally agreed that sulfonation takes place via a sequential mechanism in which a ternary enzyme complex is first formed, followed by reaction and release of products (Chapman, Best et al. 2004). However, both random and ordered binding of the substrate and cofactor molecules have been reported and the detailed kinetic mechanism (or mechanisms) of the sulfonate transfer

reaction is the subject of continuing research (reviewed in Ref. (Chapman, Best et al. 2004)). Comparison of all the available structural data provides insight into the order of substrate and cofactor binding. The structures provide evidence for both binary complexes (enzyme/substrate and enzyme/cofactor) consistent with a random bi-bi mechanism and ruling out an ordered mechanism in which binding of substrate requires binding of cofactor (or vice versa). This is in agreement with a detailed kinetic analysis for SULT1E1 (Hoff, Czyryca et al. 2006). However, a closer inspection of the structures also suggests that binding of substrates may not be completely uncoupled from binding of the cofactor. In all the structures with the co-factor product, PAP, α14-α15 and the C-terminal segment of the largest flexible loop (green in Figure 2.2) are ordered. This region contributes three absolutely conserved residues necessary for PAPS binding, T228, R258, and G260 (SULT1A1 numbering and red in Figure 2.2). Importantly, although the other loops (cyan, gold, and magenta) do not contribute directly to PAPS binding, they are more likely to be partially ordered in the presence of PAP(S). The PAP(S)-induced ordering of α14-α15 and residues 256-262 (green and red) may also restrict the conformations available to the intervening substrate-binding magenta loop when PAP is bound. Thus, the structural data suggest that PAPS binding tends to prime the cyan, gold and magenta loops for binding to the substrate.

On the other hand, the structure of SULT2A1 bound to androsterone (Chang, Shi et al. 2004) (Figure 2.2L) hints that binding of substrates does not prime the PAPS-binding loops. In this structure, the substrate-binding cyan and gold loops are ordered, but the magenta loop and adjacent PAPS-binding residues (green and red portion of the loop) are disordered. Thus, while in this case substrate and PAP(S) molecules can each bind independently to the enzyme as in a random bi-bi mechanism, there may be some degree of cooperativity between substrate binding and prior cofactor binding, but not vice versa. Given that the estimated cellular concentration of PAPS is well above that of most substrates, this may be relevant to the catalytic mechanism.

The family-wide structural comparison also suggests an additional or alternative explanation for the well documented substrate inhibitory effect. Previously reported cases of substrate inhibition have been attributed either to two substrate molecules occupying the active site at the same time

(Gamage, Duggleby et al. 2003; Barnett, Tsvetanov et al. 2004) or to the ability of substrates to bind in unproductive orientations at higher concentrations (Rehse, Zhou et al. 2002; Gamage, Tsvetanov et al. 2005). Examination of the structures in Figure 2.2 suggests a third or alternative mechanism; at high concentrations, substrates may bind in a mode in which the binding loops are incompatible with PAPS binding. This case is exemplified by the structure of SULT2A1 with dehydroepiandrosterone (DHEA) (Rehse, Zhou et al. 2002). As shown in Figure 2.2J, this structure can accommodate two substrate orientations at roughly 30° to one another. Comparison with other hSULT structures strongly suggests that SULT2A1 in this structure adopts a non-productive conformation. A portion of the green-and-magenta loop that contributes two residues for PAP(S) binding is folded into a helix, orienting the crucial PAPS binding residues away from the cofactor binding pocket. This helix conformation is not an intrinsic feature of SULT2A1 because in the SULT2A1-PAP complex this region adopts a conformation similar to that in other SULT-PAP structures (compare Figures 2.2K and J with this region highlighted by a question mark). Thus, it appears that the structure adopted by SULT2A1 with two molecules of DHEA is incompatible with PAP(S) binding and that this conformation is induced by the substrate. This is further evidence of "communication" between the substrate binding site and the PAPS binding site.

2.3.3 Ligand Binding and activity profiles reveal enzyme-specific chemical fingerprints

In order to predict and understand the fate of xenobiotics and drug candidates in humans, it is essential to better understand the selectivity and specificity of binding and activity within the hSULT family. While detailed analyses of individual structures have been very informative in this regard (Bidwell, McManus et al. 1999; Pedersen, Petrotchenko et al. 2002; Gamage, Duggleby et al. 2003; Lee, Fuda et al. 2003; Shevtsov, Petrotchenko et al. 2003; Gamage, Tsvetanov et al. 2005; Lu, Li et al. 2005), we sought to compare all active sites relative to the spectrum of small molecules that can bind to each site. However, several of the proteins whose structures were solved in this study have not been previously characterized, and it was difficult to directly compare data from the

Compound	1C1		1C2		1C3		4A1		1B1		1A1		1A3		2A1		1 E1	
	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP	- PAP	+ PAP
T_{agg} of protein (^{0}C)	48.4	57	44.5	52.5	39.2	48.3	63.4	62.9	47.8	58.2	49.3	59	47.1	55.6	48.9	59.4	46.8	56.8
PAP	4.5	NA	6.4	NA	6.5	NA	0	NA	7.8	NA	6.7	NA	6.6	NA	7.4	NA	6.4	NA
inhibitors																		
DBHD	10.8	2.3	0	0	8	0	6	8.3	11.4	0	5.4	4.1	3.5	2.5	8.4	3	3.9	0
DBHBM	9.3	0	0	0	2.8	0	3	0	7.4	0	3.6	5.6	0	2.2	4.9	2.9	2.4	0
pyridoxal 5-phosphate	8.7	0	4.5	0	5.7	0	6	6.3	5.4	0	4.3	0	6.6	0	9	0	7.9	0
AMP-PNP	4.5	0	5.2	0	6.4	0	2.8	4	4.5	0	4.4	0	5.6	0	3.4	0	6.1	0
quercetin dihydrate	31	23	4.6	0	13.5	2.3	18	18	8.4	0	6.5	0	33	0	9.1	20	0	3.9
Catecholamines																		
dopamine	0	0	0	0	0	0	0	0	0	0	0	0	0	2.3	0	0	0	0
epinephrine	0	0	0	0	6.9	0	16	5.9	4.6	0	0	0	0	0	0	0	0	0
norepinephrine	0	0	0	0	5.1	0	2	7	4.5	0	0	0	0	0	0	0	0	0
isoprenaline	0	0	0	0	6.4	0	18	18	5	0	0	0	0	0	0	0	0	0
Steroid-/large/fused ring																		
2-hydroxyestradiol	0	0	0	0	2	0	9	18	0	0	0	0	0	0	3	0	3.1	0
T4 (3,3',5,5'-tetraiodo-L-thyronine)	0	0	0	0	0	0	0	2.3	0	0	0	0	0	0	0	3.5	5.2	0
17α-ethinylestradiol	0	0	0	0	0	0	0	0	0	0	0	3.4	0	0	0	4.2	0	4.3
lithocholic acid	0	0	0	0	0	0	0	0	0	0	0	0	0	0	4.4	0	2.4	0
dehydroisoandrosterone 3-sulfate	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.8	3.1	2.2	0
α-zearalenol	0	0	0	0	0	2.8	0	0	0	0	0	0	0	0	0	3.5	0	2.6
adenosyl cobalamine	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0
estrone	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2.2
apomorphine	0	0	0	3.2	17.6	0	18	18	12.6	0	0	0	0	0	0	0	0	0
Phenols																		
p-cresol	0	0	0	0	7.2	0	7	0	0	0		0	0	0	0	0	0	0
4-aminophenol	0	0	0	0	8.3	4	0	0	3.3	0	0	0	0	0	0	0	0	0
1-naphthol	0	0	0	2.9	0	2.2	0	0	0	2.5	0	4.5	0	0	0	0	0	0
2-naphthol	0	0	0	4.1	0	0	0	0	0	0	0	5	0	0	0	0	0	0
2-ethylphenol	0	0	0	0	0	2.1	0	0	0	0	0	3	0	0	0	0	0	0
2-n-propylphenol	0	0	0	0	0	2.6	0	0	0	0	0	0	0	0	0	0	0	0
2-sec-butylphenol	0	0	0	2	0	2.2	0	0	0	0	0	0	0	0	0	0	0	0
vanillin	0	0	0	2.9	0	0	0	0	0	0	0	4.7	0	0	0	0	0	0
pentachlorophenol	0	0	0	5.1	0	0	0	0	3.7	8.9	0	10.2	0	0	0	0	0	0
2, 6-dichloro-4-nitrophenol	0	0	0	3.5	0	0	0	0	0	3	0	8.7	0	0	0	0	0	2.8
resveratrol	6.6	0	0	0	7.4	0	18	18	8.6	0	0	0	0	0	3.8	0	4.5	0.0
Acids																		
salicylic acid	0	0	0	2.2	0	0	0	0	0	2.4	0	2.5	0	0	0	0	0	0
mefenamic acid	0	0	0	3.6	0	0	0	0	0	2.9	0	9.6	0	0	0	0	0	4.7

Table 2.1 **Thermostability screen for ligand binding to hSULTs.**

The thermostability of each protein (0.4 mg/ml) in the presence and absence of potential ligands (1mM) were measured as described in material and methods. Effect of PAP was also assessed at 1 mM concentration (second row). Except for the numbers in the first row, all the values in this table are ΔT_{agg} (T_{agg} of the protein in the presence of the compound – Tagg in the absence of the compound (Vedadi, Niesen et al. 2006). A ΔT_{agg} value below 2 °C is considered within the error limit of the technique and is shown as zero (0). The values in the first row are for Tagg of each protein in the presence or absence of 10 mM PAP. Over a concentration range of 0.1 to 5 mM, both PAPS and PAP had equivalent stabilizing affects on all SULTs except SULT4A1. Thus, PAP was used as a substitute for PAPS in considering the affect of cofactor upon substrate binding. The ΔTagg values which are less than 4 °C, between 4 and 10 °C and above 10 °C are highlighted by light green, blue and pink respectively. Abbreviations are: AMP-PNP, adenosine 5′-(β,γ-imido) triphosphate; DBHD, 3,5-dibromo-4-hydroxy-benzoic acid (6,8-dichloro-4-oxo-4H-chromen-3-ylmethylene) hydrazide; DBHM, 3,5-dibromo-4-hydroxy-benzoic acid (6-chloro-4-oxo-4H-chromen-3-ylmethylene)-hydrazide; PAP, 3'-phosphoadenosine 5'-phosphate, and NA, not applicable.

Substrate	Sulfotransferase activity (nmol/min/mg)																							
	SULT1C1			SULT1C2			SULT1C3			SULT1B1			SULT1A1			SULT1A3			SULT2A1			SULT1E1		
Substrate concentration (μM)	10	25	100	10	25	100	10	25	100	10	25	100	10	25	100	10	25	100	10	25	100	10	25	100
Catecholamines																								
dopamine	N	N	N	N	N	N	N	N	N	N	N	0.6	2.3	5.2	13.4	7.4	25.7	74.4	N	N	N	N	N	0.7
epinephrine	N	0.1	0.1	N	N	0.8	N	N	N	N	N	1.2	1.8	2.9	8.0	N	0.5	6.0	N	N	N	N	N	N
norepinephrine	N	0.1	0.1	N	N	N	N	N	N	N	N	0.5	1.7	1.7	4.2	0.7	2.7	16.9	N	N	N	N	N	0.7
isoprenaline	N	N	N	N	N	0.8	N	N	N	N	0.6	1.1	2.7	6.1	13.6	5.1	16.6	54.0	N	N	N	N	N	0.8
Steroid-/large/fused ring																								
17α-ethinylestradiol	N	N	N	N	N	N	N	N	N	N	N	N	2.5	7.2	N	N	N	1.4	1.4	1.6	1.3	1.2	4.3	13.3
2-hydroxyestradiol	N	N	0.1	N	N	N	N	N	N	N	N	0.5	1.0	1.8	3.1	N	N	N	1.0	1.5	3.1	N	N	N
lithocholic acid	N	N	N	N	N	N	0.6	0.8	1.1	N	N	N	N	N	N	N	N	N	2.7	8.3	11.4	N	N	N
α-zearalenol	0.3	0.6	1.5	2.0	8.1	9.3	1.1	2.7	4.1	0.8	1.5	1.1	6.3	16.9	25.9	0.8	2.9	5.5	N	N	N	2.5	7.1	13.0
T2 (3,5-diiodo-L-thyronine)	N	N	0.1	N	N	N	N	N	N	N	N	N	1.0	N	N	N	N	N	0.3	0.6	1.2	N	N	N
T3 (3,3',5-triiodo-L-thyronine)	0.4	0.6	1.1	N	N	2.9	N	N	N	N	N	0.9	1.1	1.2	N	N	0.5	1.5	N	N	0.3	2.0	2.2	4.4
T4 (3,3',5,5'-tetraiodo-L-thyronine)	0.3	0.5	0.8	N	N	N	N	N	N	N	N	2.5	N	N	N	N	N	0.6	N	N	1.6	N	1.2	2.1
cholesterol	N	N	N	N	N	0.7	N	N	0.4	N	N	N	0.8	N	N	N	N	N	N	N	N	N	N	N
estrone	0.1	N	N	N	N	0.8	N	N	N	N	N	N	5.2	4.9	5.0	N	N	N	0.3	1.1	1.1	N	0.8	3.4
minoxidil	N	N	N	N	N	N	N	N	N	N	N	1.6	1.6	N	N	N	N	N	N	N	N	N	0.9	1.9
salbutamol	N	N	N	N	N	N	N	N	N	N	N	N	0.9	1.3	N	1.2	2.7	12.7	N	0.3	N	N	N	N
Phenols																								
4-aminophenol	N	N	N	1.3	1.2	2.6	N	N	N	N	0.7	1.7	9.2	16.4	4.7	N	N	N	N	N	N	N	N	1.2
1-naphthol	0.1	0.1	0.6	4.6	11.7	31.5	0.7	1.2	1.8	2.4	5.9	15.3	8.8	6.7	4.1	9.5	24.7	68.5	N	0.4	1.0	5.0	9.6	26.0
2-naphthol	0.2	0.3	1.1	3.3	16.2	11.7	N	N	N	2.9	8.9	16.8	11.1	16.8	10.0	10.0	20.3	42.6	N	0.2	0.8	16.6	41.5	57.9
2-ethylphenol	N	N	N	1.4	4.8	25.8	0.9	1.4	2.2	2.1	5.5	16.5	10.3	10.4	4.9	1.2	4.7	15.2	N	N	0.2	2.5	5.3	15.7
4-ethylphenol	N	N	0.1	1.8	9.9	39.5	N	N	N	1.7	4.0	9.9	4.8	63.0	25.3	0.5	1.3	4.3	N	N	N	9.4	21.9	38.2
2-n-propylphenol	N	N	0.1	0.8	3.4	15.5	N	0.7	0.9	2.1	5.2	14.2	10.0	15.4	8.2	1.0	2.8	11.0	N	N	N	1.8	4.6	12.7
2-sec-butylphenol	0.3	N	N	1.2	N	N	N	1.0	2.1	N	N	1.1	1.0	1.2	1.2	N	1.4	1.2	N	N	N	N	N	N
4-n-amylphenol	0.1	0.1	0.3	N	6.4	17.2	N	N	0.3	0.7	2.2	6.2	5.4	12.5	24.3	1.5	2.4	4.7	N	0.4	1.3	2.6	5.7	20.0
4-n-heptylphenol	N	N	N	N	N	N	N	N	N	N	N	0.5	0.7	2.6	1.6	N	N	0.8	N	N	0.2	N	N	1.0
4-nitrophenol	0.3	0.5	1.9	N	8.5	33.4	N	N	0.6	1.6	5.3	13.2	4.3	21.4	7.2	N	1.7	4.1	N	0.2	0.3	2.4	5.3	16.8
vanillin	0.3	0.8	2.5	5.2	15.8	23.0	N	N	0.8	2.6	8.2	19.0	N	N	N	11.7	29.4	64.0	N	0.3	0.5	5.4	12.2	25.1
1,2,3,4-tetrahydro-1-napthol	N	N	N	N	N	N	N	0.3	0.8	0.9	1.6	3.4	1.3	1.1	N	N	N	N	N	N	0.9	N	N	N
p-cresol	N	N	0.1	4.9	17.7	40.7	N	N	N	3.4	9.2	16.3	11.3	28.5	16.7	1.3	3.4	9.5	N	N	0.3	4.0	9.2	26.8
acetaminophen (Paracetamol)	N	N	0.1	4.8	8.6	22.6	N	N	N	N	0.6	1.1	3.0	5.4	14.1	N	0.7	3.8	N	N	N	1.5	3.6	11.6
tyramine	N	N	N	3.1	9.0	38.2	N	N	N	2.1	5.7	13.0	0.9	0.9	1.2	1.0	1.9	7.5	N	N	N	1.4	4.5	12.2
resveratrol	0.3	0.5	0.3	3.4	15.0	5.2	N	N	N	2.9	6.9	18.4	7.4	3.6	1.0	5.4	16.6	26.1	N	N	N	15.8	36.6	10.3
Background	0.25 ± 0.01			6.2 ± 0.6			0.9 ± 0.2			1.3 ± 0.4			5.4 ± 0.3			6.8 ± 0.4			2.1 ± 0.1			1.7 ± 0.5		

Table 2.2 **Sulfotransferase activity screen of hSULTs.**

Enzymatic activities of the 8 hSULTs were assayed at concentrations of 10, 25, and 100 μM of each substrate as described in material and methods. The first row indicates the concentration of each substrate in μM in each set of experiments. The last row indicates the background rate of conversion of PAPS to PAP in the presence of enzyme, but absence of any substrate. The error in the background rate is the standard deviation of three measurements. All other numbers in the table are the activity in nmol/min/mg of protein after subtracting the background rate. For a given enzyme/substrate pair the substrate concentration that exhibited the highest activity is highlighted in blue. Activity values between one and two standard deviations of the background rate are highlighted in red. Those above 2 standard deviations are in green. N stands for "not detected".

literature due to differences in experimental conditions. Therefore, in order to evaluate specificity and selectivity in a consistent manner, 9 purified, recombinant hSULTs were screened for binding to a library of 90 small molecules (Table S2.2) that comprised known substrates, inhibitors, related hormones, bioamines, drugs (Senisterra, Markin et al. 2006; Vedadi, Niesen et al. 2006). In order to profile the entire hSULT family, we made use of the well-known fact that equilibrium binding of a ligand increases the thermal stability of a protein in a manner proportional to the concentration and binding affinity of the ligand (Senisterra, Markin et al. 2006; Vedadi, Niesen et al. 2006). In a multi-well format, the thermal stability of each hSULT was monitored as a function of temperature and in the presence or absence of compounds (Figure 2.4). In the absence of compounds, well-behaved, sigmoidal thermo-denaturation/aggregation profiles were obtained for hSULTs 1A1, 1A3, 1B1, 1E1, 1C1, 1C2, 1C3, 2A1, and 4A1. SULT2B1b did not denature within the range of temperature used for this type of analysis (up to 80°C). In this screening format, compounds that stabilize a protein by more than 2 degrees are scored as positives (Table 2.1). It was not possible to assay binding to ligands in the presence of the sulfonated co-factor, PAPS, because the sulfonate transfer reaction would have taken place. However, except for SULT4A1, PAPS and PAP had equivalent stabilizing affects on all hSULTs. Thus, PAP was used as a substitute for PAPS in considering the affect of cofactor upon substrate binding, and screens for the binding of ligands were performed in the absence and presence of a saturating amount of PAP.

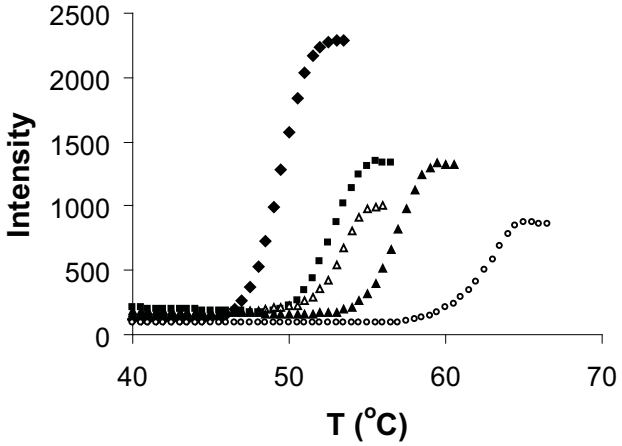


Figure 2.4 **Thermodenaturation/aggregation profile of SULT2A1.** The intensity of scattered light is plotted as a function of temperature for SULT2A1 in the absence of any ligand (♦), in the presence of AMP-PNP (■), lithocholic acid (Δ), PAP(▲) and both PAP and dehydroisoandrosterone 3-sulfate(○). The inflection point of the denaturation curve is taken as T_{agg} reported in Table 2.1.

Based on these binding results, a set of 20 compounds that bound to at least one hSULT plus 11 additional related compounds or known substrates were used as a pool of potential substrates

for enzymatic activity of hSULTs 1A1, 1A3, 1B1, 1C1, 1C2, 1C3, 1E1 and 2A1 (Table 2.2). We monitored the conversion of PAPS to PAP by HPLC as a tractable method of screening multiple proteins against multiple substrates (8 proteins and 31 substrates in this study). An obvious weakness of this assay is that it does not detect the sulfonate group which is being transferred during the reaction. However, the results obtained in this HPLC based assay can serve as a convenient first approximation of enzymatic activities. We note that due to the relatively low sensitivity of this method, we were not able to reliably assay substrates at nM concentrations, and therefore some of the results may be complicated by substrate mediated inhibition. Indeed, inspection of Table 2.2 shows that in some cases the highest levels of activity were observed at lower substrate concentrations. This is especially true of SULT1A1, an enzyme for which significant substrate inhibition has been noted previously (Gamage, Duggleby et al. 2003).

The combined ligand binding and activity screens revealed a unique "chemical fingerprint" for each hSULT (Tables 2.1 & 2.2). First, as expected from previous studies, there was considerable overlap in the substrate specificity for enzymatic activity. For example, all hSULTs assayed here were able to sulfonate a number of phenolic compounds such as naphthols, and/or alkylphenols. However, within each substrate profile there were also elements of specificity. For example, SULT1A1 and SULT1A3 were the only two hSULTs that showed significant activity toward catecholamines compared to other substrates, with SULT1A3 being more specific for dopamine as expected from previous studies (Coughtrie 1998; Dooley 1998; Brix, Barnett et al. 1999; Dajani, Cleasby et al. 1999; Hempel, Barnet et al. 2005). SULT1A3 was also the only protein to bind dopamine in the binding assays, consistent with its designation as human dopamine sulfotransferase (Dajani, Cleasby et al. 1999). It is interesting to note that in the past SULT1A1 and 1A3 have been distinguished from one another in tissue fractions by the higher sensitivity of SULT1A1 to inhibition by 2,6-dichloro-4-nitrophenol (DCNP) (Veronese, Burgess et al. 1994). Although we did not measure inhibition by this compound, we note that SULT1A1 bound DCNP in the presence of PAP, while SULT1A3 did not (Table 2.1).

Six hSULTs (1C1, 1C2, 1E1, 1B1 1A1 and 1A3) had enzymatic activity towards resveratrol,

a polyphenolic compound present in grapes and wine with possible anticancer and cardioprotective activities(Baur and Sinclair 2006). The activity profiles for resveratrol also displayed evidence of substrate inhibition by this compound for SULTs 1C1, 1C2, 1A1 and 1E1. Acetaminophen was a substrate for SULTs 1A1, 1E1, 1A3, 1C2 and 1B1 but not SULT1C3 and SULT2A1. Substrates for SULT1C3 have not been reported previously. Our data indicate that this recently identified member of the hSULT family is able to sulfonate *p*-nitrophenol, 1-naphthol, 2-ethylphenol, 2-*n*-propylphenol and 2-*sec*-butylphenol, as well as the steroid-related compounds, α-zearalenol and lithocholic acid. SULT1C3 appeared to be most active with α-zearalenol (4.1 nmol/min/mg) and 2-ethylphenol (2.2 nmol/min/mg). These data suggest SULT1C3 may contribute to the metabolism of steroid and phenolic compounds. Finally, SULTs 2A1 and 1E1, which are reported to metabolize steroids (Comer, Falany et al. 1993; Falany, Comer et al. 1995; Falany, Krasnykh et al. 1995; Adjei and Weinshilboum 2002), both bound to and sulfonated multiple steroid and steroid-like compounds with different apparent specificities. These data show that despite the limitations of our rapid screening method, the enzymatic activity data in Table 2.2 reflects, to a first approximation, the expected relative substrate activities reported in the literature and reveal new activities toward pharmacologically important compounds. The binding data, on the other hand, suggest a more complicated situation.

2.3.4 PAP is able to alter ligand binding profiles

The ligand binding profiles were remarkably different in the presence and absence of PAP (with the exception of SULT4A1 which will be discussed separately below). Some known substrates for the well characterized SULTs appeared to bind only in the presence of PAP; for example, dopamine for SULT1A3 (Brix, Barnett et al. 1999; Dajani, Cleasby et al. 1999), and 1-naphthol for SULT1B1 (Wang, Falany et al. 1998), while many previously unreported compounds bound to these and other family members in the absence of PAP. It is interesting to note that not all known substrates, nor all of those with reactivity in our activity screens were found to *bind* to the enzyme in the presence of PAP. For example, resveratrol was a substrate for SULT1A3 and SULT1C2 but

did not stabilize either proteins in the binding assays either in the presence or absence of PAP. There are several reasons that may account for these observations. First, the enzymatic screen is likely more sensitive than the ligand-binding assay where compounds with K_m or K_d values in the high μM range are less likely to be detected (Senisterra, Markin et al. 2006; Vedadi, Niesen et al. 2006). Second, some ternary complexes may not be significantly stabilized relative to the PAP-enzyme complex (especially at elevated temperatures). Finally, the presence of the sulfonate group of PAPS may also contribute to binding of substrates and these cases may not be detected in our binding assay. Interestingly, binding of SULT1C1, 1B1 and 1E1 to resveratrol was only observed in the absence of PAP but all are active toward this substrate. Nevertheless, the radically different binding profiles observed in the presence and absence of PAP are consistent with the structure-based mechanisms proposed above. Specifically, PAP (and presumably PAPS) appears to prime the substrate binding loops for subsequent binding to certain substrates, while in the absence of cofactor, the loops are free to bind alternative ligands (perhaps only at high concentrations), or non-productive ligand-bound conformations may exist. This priming of the substrate binding loops is likely made possible by flexibility of the binding loops observed in the structure. This structural plasticity may allow a reconfiguration of substrate binding loops in the absence of PAP(S) in order to bind a different chemical class of compound. For example, both SULT1C3 and SULT1B1 were stabilized by catecholamines in the absence of PAP, but neither showed significant activity toward this class of compounds. This raises the possibility that certain endogenous and/or exogenous compounds, such as those that bind in the absence of PAP, may act as competitive inhibitors of SULTs by occupying the substrate binding pocket and preventing a productive PAPS binding conformation, as for SULT2A1 (Figure 2.2J).

2.3.5 Screening and structural analysis reveals novel mechanism of inhibition

Excluding SULT4A1, three compounds bound to all hSULTs: Adenosine 5′-(β,γ-imido) triphosphate (AMP-PNP), a non-hydrolysable ATP analogue, pyridoxal 5-phosphate (PLP) a competitive inhibitor for sulfotransferases (Kamio, Honke et al. 1995) and quercetin, a potent inhibitor of

SULT1A1 and SULT1E1 (Pacifici 2004). These compounds had been known to inhibit one or more sulfotransferases, but our data suggest that they may be universal SULT inhibitors. AMP-PNP and PLP bound only in the absence of PAP, suggesting that they occupy the PAP binding site, as might be expected from their structural similarity to PAP. Quercetin, found in many fresh fruits and vegetables, is a flavonoid with anti tumor and anti inflammatory activities. It is possible that some of its favorable physiological effects may be related to inhibition of hSULT activity. Additional inhibitors bound to only a subset of the hSULTs, including 3, 5-dibromo-4-hydroxy-benzoic acid (6,8-dichloro-4-oxo-4*H*-chromen-3-ylmethylene) hydrazide (DBHD), 3,5-dibromo-4-hydroxy-benzoic acid (6-chloro-4-oxo-4*H*-chromen-3-ylmethylene) -hydrazide (DBHM), and pentachlorophenol (PCP).

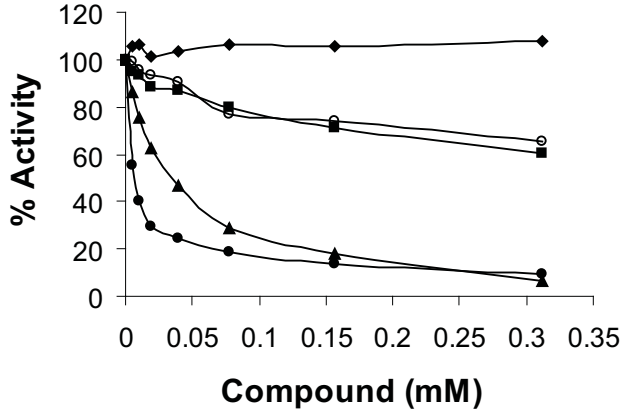


Figure 2.5 **Inhibitory effect of isoprenaline (♦), PLP (○), quercetin (■), DBHD (▲), and PCP (●) on sulfotransferase activity of SULT1B1 with 1-naphthol as substrate.** Sulfotransferase activity of SULT1B1 at different concentrations of each compound (5 to 312 μM) has been assayed as described in material and methods.

The binding profiles in Table 2.1 raise the possibility that compounds that bind to hSULTs only in the absence of PAP may inhibit hSULT activity. In order to investigate this possibility we assayed the activity of SULT1B1 in the presence of increasing concentrations of five compounds found in our screens. These five include known inhibitors (quercetin, DBHD, PLP, PCP) as well as isoprenaline, which binds to SULT1B1 only in the absence of PAP and is a poor substrate for this enzyme (Tables 2.1 and 2.2). As shown in Figure 2.5, PCP and DBHD were strong inhibitors of SULT1B1 while PLP and quercetin had intermediate effects. Isoprenaline however had no inhibitory effect on the activity of SULT1B1 with 1-naphthol, and it was sulfonated by SULT1B1 (Table 2.2), indicating that not all compounds that bind in the absence of PAP are necessarily inhibitors.

PCP is a significant environmental toxin due to its common use as a wood preservative

and in the pulp and paper industry. Its chemical structure is related to hydroxylated metabolites of polychlorinated biphenols (OH-PCBs) whose endocrine disruptive properties may be related to the inhibition of estradiol sulfonation by SULT1E1 (Kester, Bulduk et al. 2000). The mechanism of inhibition of SULT1E1 by OH-PCBs and related compounds has been proposed to take place via both allosteric (Kester, Bulduk et al. 2000; Kester, Bulduk et al. 2002) and competitive (Kester, Bulduk et al. 2002; Shevtsov, Petrotchenko et al. 2003) mechanisms. Our binding data suggest that PCP may be a competitive inhibitor of SULTs 1B1, 1C2, and 1A1 (Table 2.1). For 1C2, and 1A1, PAP was required for binding and for 1B1, PCP binds much better in the presence of PAP suggesting that PCP binds in a substrate-like conformation facilitated by PAP(S). In order to better understand the mechanism of inhibition by PCP, we determined the structure of SULT1C2 bound to PAP and PCP (Figure S2.3). The structure reveals that the protein undergoes a disorder-order transition upon PCP and PAP binding. Helices α4, α5 and α15, loops α5-α6 and α15-α16 are ordered only in the ternary complex but not in the apo SULT1C2 structure (Figures 2.2B and H). PCP is found in the substrate–binding pocket and therefore appears to be a competitive inhibitor, consistent with crystallographic analysis of SULT1E1 bound to PAP and 3,5,3',5'-tetrachlorobiphenol (Shevtsov, Petrotchenko et al. 2003). Comparison of the SULT1C2·PAP·PCP and SULT1E1·PAP·estradiol structures (Figure 2.6) revealed two structural features that may be relevant in explaining the mechanism of PCP inhibition. In the co-crystal structures, the phenol moieties of PCP and estradiol share the same relative position and orientation, positioning the phenolic OH within hydrogen bond distance of the catalytic histidine. This histidine is thought to be deprotonated and made catalytically competent to accept the phenolic hydrogen from estradiol, facilitating nucleophilic attack on the sulfonate of PAPS.

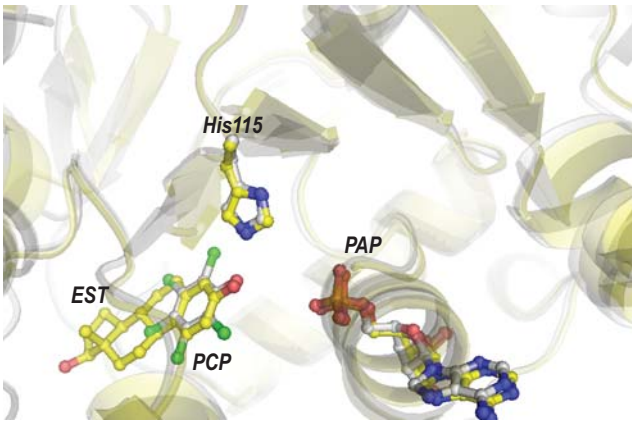


Figure 2.6 **The superimposition of the active sites of the SULT1E1·PAP·17β-estradiol complex in yellow (1AQU) and the SULT1C2·PAP·PCP complex in grey (2GWH).** The PCP (chlorine atoms in green) and 17β-estradiol molecules are highlighted along with the catalytic His115 and PAP. See text for discussion.

Our structure shows that PCP appears to be in a catalytically competent conformation. However, PCP and estradiol differ dramatically in the acidity of their hydroxyl groups; the estimated pK_a for estradiol is approximately 15 while that for PCP is 4.5. Thus, although PCP appears to bind in a catalytically competent conformation, the phenolic oxygen of PCP may be too weak a nucleophile to attack the sulfonate of PAPS.

The inhibitory effects of OH-PCBs and PCP have been interpreted previously in terms of variations in the bound conformation of the halogenated compounds relative to that of estradiol(Kester, Bulduk et al. 2000; Shevtsov, Petrotchenko et al. 2003). Although steric and conformational factors clearly play a role, our structure of SULT1C2 with PCP and PAP suggests a key role for the electronic nature of the halogenated phenols. We examined the calculated pK_a values for the series of 21 4-hydroxyl substituted PCBs for which Kester *et al* reported IC_{50} values (Kester, Bulduk et al. 2000). The results show a strong correlation between the calculated acidity of these phenols with IC_{50} values, with the most acidic compounds having the strongest inhibitory effect (Figure S2.4).

2.3.6 SULT4A1 is an atypical SULT with an atypical structure

Under the conditions of our screens, PAP bound to all hSULTs except for SULT4A1. In order to rule out the possibility that SULT4A1 simply has a much weaker affinity for PAP, we performed titration experiments for several of the proteins with increasing concentrations of PAP ranging from 90 μM to 90 mM (Figure 2.7A). SULT1C1, SULT1C2, SULT1C3 and SULT1B1 showed similar saturation binding curves, reaching saturation at about 100 μM. However, PAP when added at concentrations as high as 90 mM did not stabilize SULT4A1. SULT4A1 is one of the hSULTs that is most divergent in sequence and examination of the binding pocket revealed two significant differences that are predicted to affect PAP binding. First, a Trp in α3 that is conserved in all other hSULTs and stacks with the adenine ring of PAPS (Trp53 in SULT1A1) is replaced with a Leu in SULT4A1 (Figure 2.3). Second, the magenta PAP-binding loop is much shorter in SULT4A1 than in the other hSULTs and lacks the conserved Lys residue that separates the key PAPS binding

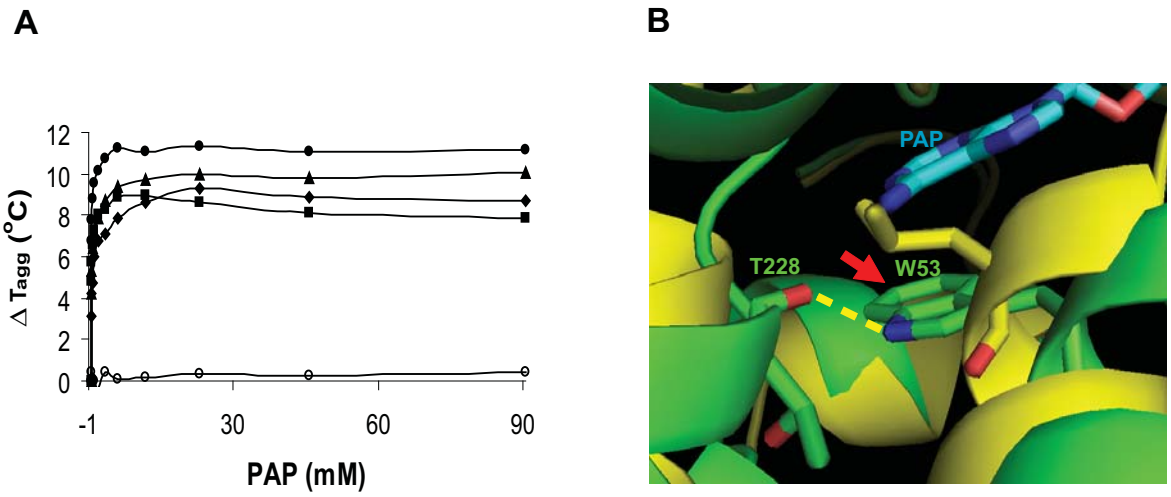


Figure 2.7 **SULT4A1 does not bind PAP.** *A*, thermal aggregation temperatures for SULT1B1(●), SULT1C3(▲), SULT1C1(♦), SULT1C2(■), SULT4A1(○) as a function of PAP concentration. PAP provides no stabilizing effect to SULT4A1. *B*, superimposition of the PAP binding site of SULT1C1 (green) and SULT4A1 (yellow). A conserved Trp residue (W53 of SULT1C1; green ring in the middle of figure, indicated by a red arrow) normally forms π-π stacking interactions with the adenine ring of PAP (blue) and hydrogen bonds to a neighboring conserved Thr (T228 in SULT1C1). SULT4A1 is the only SULT to have a Leu instead of Trp at this position. The Leu residue cannot make stabilizing interactions with Thr or the adenine ring, and its position would cause a severe steric clash with PAP and PAPS.

residues Arg258 and Gly260 (SULT1A1 numbering), which results in these residues being out of register. Taken together, SULT4A1 has a slightly smaller PAPS binding pocket that is predicted to be unable to accommodate the cofactor. Interestingly, some residual electron density was observed in the PAP binding pocket of recombinant SULT4A1. Presumably this derives from a bound small molecule that was co-crystallized, although we were not able to identify it either by modeling atoms into the electron density or using mass spectrometry.

To test the possibility that SULT4A1 might use an alternate sulfonate donor, we tested several potential alternates such as adenosine phosphosulfate, 4-nitrocatechol sulfate, 4-acetylphenyl sulfate, estrone 3-sulfate, indoxyl sulfate and 4-methylumbelliferyl sulfate. None of these compounds stabilized SULT4A1 against thermal aggregation (data not shown). These data strongly suggest that SULT4A1 may not have significant catalytic activity *in vivo*. Indeed, although very weak activity has been reported in one case (Sakakibara, Suiko et al. 2002), other groups have failed to observe activity for human SULT4A1 (Falany, Xie et al. 2000; Glatt and Meinl 2005).

Significantly, this protein, which is expressed primarily in the brain, binds to 2-hydroxylestradiol, thyroid hormone, T4 (3,3',5,5'-tetraiodo-L-thyronine), and the catecholamines, norepinephrine, epinephrine and isoprenaline (but not dopamine), suggesting that SULT4A1 may modulate the bioactivity of these compounds via a mechanism distinct from sulfonation. Of note, SULT4A1 did not bind any simple phenolic compounds under the conditions tested here.

2.3.7 Binding profiles suggest alternative classifications of hSULTs

Examination of the chemical fingerprints reflected in Tables 2.1 and 2.2 suggests that subsets of the hSULTs can be clustered based on the chemical properties of the compounds that they bind – relationships that are not evident from global sequence comparison. To explore alternative activity or structure-based classifications of hSULTs in more detail, we performed average hierarchical clustering on the experimental data in an attempt to identify correlations between the *local* sequence or structural features of the substrate binding pockets and activity profiles among the hSULTs. Figure 2.8 shows the clustering of similarity matrices for each parameter, viewed using trees. Considering only global sequence similarity, the hSULTs cluster according to their nomenclature and phylogenetic relationships (Figure 2.8A). Considering only the 9 proteins for which we have binding or enzymatic data, SULTs 1A1 and 1A3 are most closely related, with a global sequence identity of 95%. The three SULT1C proteins cluster with average sequence identities close to 55%, as do SULT1E1 and 1B1. SULT 2, 4 and 6 subfamilies are relative outliers with sequence identities to all other SULTs considered here around 35% or less. It is well known that related enzymes with sequence identities below 40% have often evolved to have different substrate specificities (Todd, Orengo et al. 2001) and therefore we would expect most of the SULTs to sulfonate different substrates, except perhaps 1A1 and 1A3. In keeping with this concept, the closely related SULTs 1A1 and 1A3 are clearly the most similar within the substrate binding site, as measured by both local sequence and structure comparisons (Figures 2.8B and C). However, the clustering of the other more distantly related SULTs is sufficiently different at the level of local sequence and structure, such that the SULT1C proteins are no longer clustered together and the outliers are different. These comparisons show that the *local* sequence and structures of the

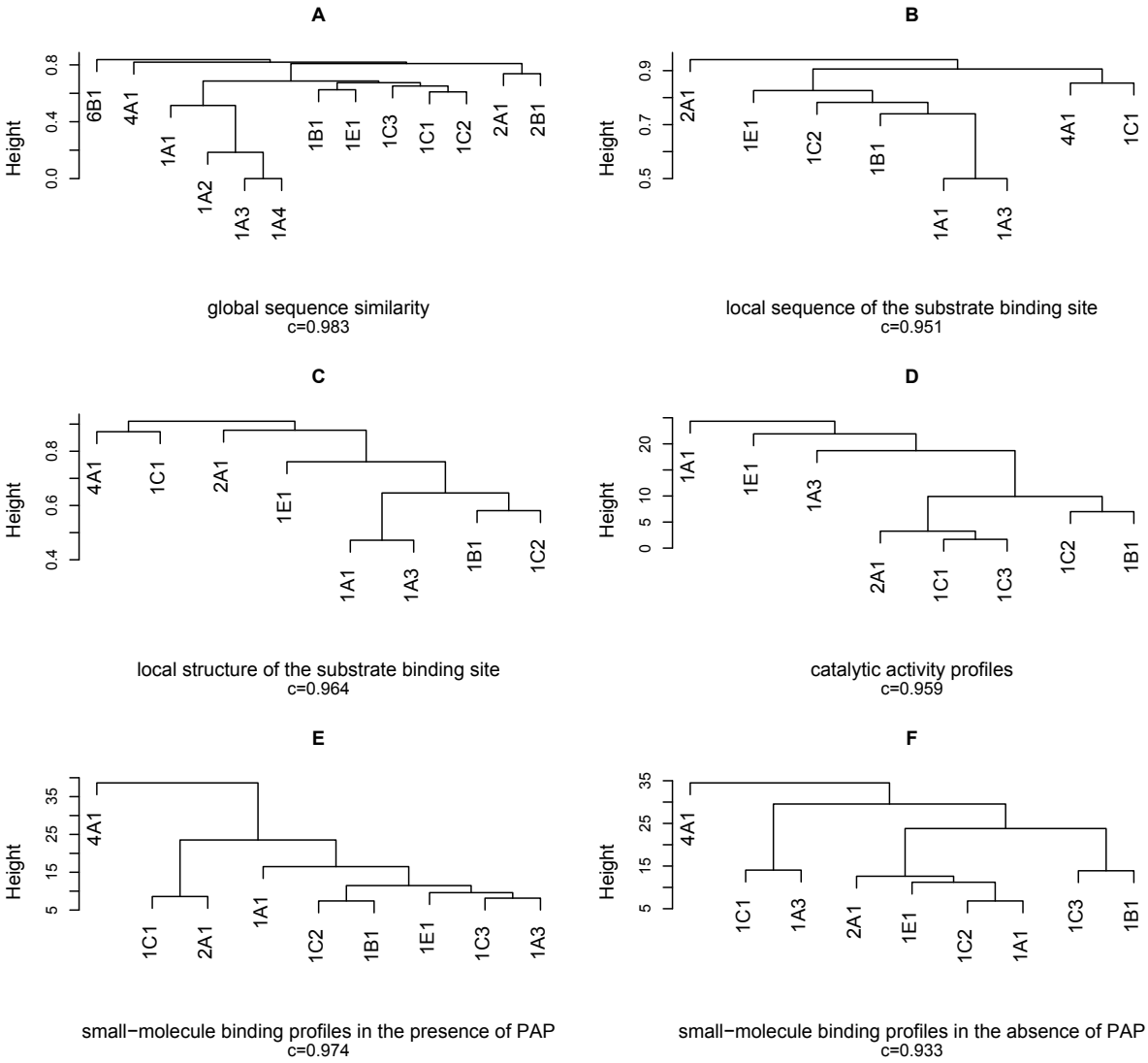


Figure 2.8 **Average linkage hierarchical clustering of SULTs** according to *A*, global sequence similarity; *B*, local sequence of the substrate binding site; *C*, local structure of the substrate binding site; *D*, catalytic activity profiles; *E*, small-molecule binding profiles in the presence of PAP; *F*, small-molecule binding profiles in the absence of PAP. The value under each clustering is the correlation between the original data matrix and the cophenetic matrix - a matrix whose elements represent the height at which these elements first meet in the tree. The higher this correlation, the more accurate the tree represents the original data. See Methods for further details.

substrate binding sites do not correspond to the global sequence relationships.

These results also illustrate why members of the same subfamily do not metabolize the same classes of compounds. While the clustering of hSULTs presented here is likely influenced by the limited subset of compounds used in our binding and activity assays, the results provide an initial view of the family-wide activity based classifications. The trees that cluster the 8/9 proteins according to their binding and activity profiles (Figures 2.8D, E, and F) show a remarkably

different clustering from the sequence and structure-based trees. Firstly, 1A3 and 1A1 no longer cluster together, despite their strong global sequence similarity (95% identity). Inspection of Table 2.1 shows that 1A1 strongly binds most of the phenols and acidic compounds, once PAP is bound,

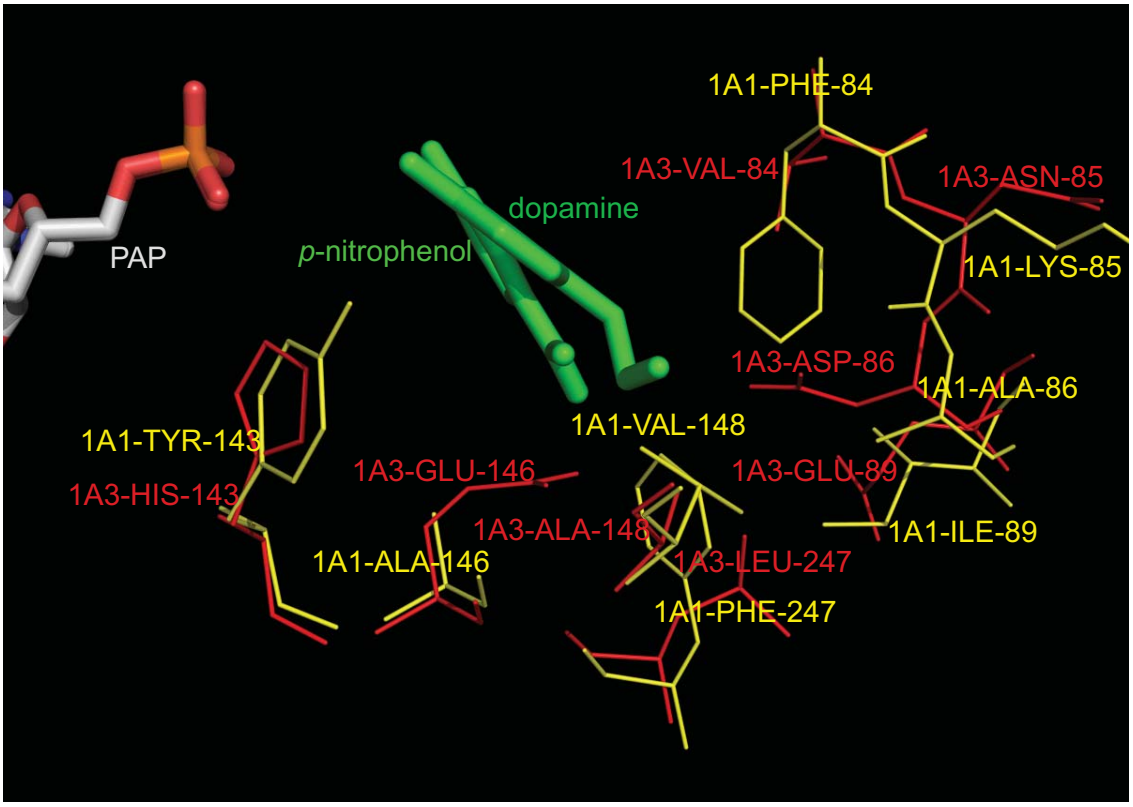


Figure 2.9 **Binding site comparison of SULT1A1 bound to PAP and p-nitrophenol (1LS6) and SULT1A3 bound to PAP and dopamine (2A3R).** The side chains of the 8 binding site residues that differ between these two proteins are shown in yellow (SULT1A1) and red (SULT1A3).

whilst 1A3 shows absolutely no binding of these substrates. These SULTs are also distinguished by their differential reactivity toward catecholamines, with SULT1A3 able to bind dopamine and having higher relative activity toward catecholamines compared to phenols as previously noted (Liu, Suiko et al. 2000). Comparison of the residues in the substrate binding loops of SULT1A1 and SULT1A3 revealed that all residues are identical (and in identical positions in the structures) except for the 8 residues shown in Figure 2.9. These changes map to two loops (residues 84-89, and 143-148) and residue 247 (SULT1A1 numbering). Importantly, SULT1A3, which does not bind acidic compounds, has acidic instead of hydrophobic residues at three of these positions. The net result is a much more negatively charged pocket for SULT1A3 compared to SULT1A1. This

difference would disfavor binding of compounds with a net negative charge to SULT1A3 and favor interactions with the amino group of catecholamines. Thus, the strong local sequence and structure similarities in 1A1 and 1A3 are manifested in their similar ability to bind similar inhibitors and sulfonate catecholamines (as a class), but small local sequence changes in the substrate binding site have enhanced the ability of SULT1A3 to bind catecholamines such as dopamine (Liu, Suiko et al. 2000), and completely changed its ability to bind acidic compounds. The influence of residue 146 on the specificity of SULT1A1 compared to SULT1A3 (Ala *vs* Glu) has been previously noted (Liu, Suiko et al. 2000), however, the results presented here suggest that additional differences in the binding loops also contribute to specificity.

SULT1B1 and 1C2 cluster together in the trees of local structure, activity, and binding in the presence of PAP. These groupings reflect their common ability to bind acidic compounds (and the acidic phenols, PCP and 2,6-dichloro-4-nitrophenol), with promiscuous activity profiles towards phenols. Thus, SULT1B1 and 1C2 appear to be much more closely related in terms structure and activity than their global sequences would indicate.

Finally, Figure 2.8F shows that the clustering of hSULTs again differs when considering the binding of compounds in the absence of PAP. In this case, none of the previously noted similarities is evident. Many of the compounds in Table 2.1 are inhibitors of hSULT activity. The differential clustering in the absence of PAP may reflect possible unrelated configurations of binding site residues when bound to these compounds or the tendency for more disorder in the absence of PAP. The clustering of both binding profiles (Figures 2.8E and F) places SULT4A1 as the furthest outlier, analogous to its position in the global sequence comparison. While this is consistent with the inability of SULT4A1 to bind PAP and catalyze sulfonation, it is interesting to note that this is not due to a radical difference in *local* binding site sequence or structure, as SULT4A1 clusters with SULT1C1 when considering these local factors (Figures 2.8B and C). As noted above, apparently small differences of 1-2 residues in the PAP binding site are likely responsible for this behavior.

2.4 Discussion

An important challenge in structural biology, chemical biology and drug discovery is to relate changes in local sequence and structure to the binding and activity profiles of homologous enzymes in order to better predict or explain substrate and inhibitor specificity within an enzyme family. In the case of the hSULTs, this is particularly desirable in order to predict the fate of xenobiotics, hormones and drug candidates in humans. Like other Phase II detoxification enzymes, hSULTs are known to have broad and overlapping substrate specificities. We and others have shown that this promiscuity derives from the considerable flexibility or plasticity of the hSULT binding sites and therefore, a full understanding of specificity will require multiple 3D structures for each hSULT in complex with substrates and inhibitors, as well as knowledge of the full spectrum of small molecules that bind in both productive and nonproductive conformations. Our structural and chemical profiling data prepare the foundation for such detailed studies.

Here we also reveal a previously unrecognized structural role for the cofactor, PAP(S); namely priming of the often disordered substrate binding loops for interaction with substrates. The "magenta loop" (Figure 2.2), which contributes to both PAPS and substrate binding pockets, can also, in some instances, adopt an inactive conformation and may explain substrate induced inhibition and/or inhibition by compounds that bind in the absence of PAPS as well as known inhibitors. The flexibility of the substrate-binding loops in hSULTs likely contributes to the wide repertoire of compounds that can be accommodated in the substrate binding pocket, only a subset of which lead to productive sulfonation.

Our results provide insight into mechanisms of inhibition of hSULTs. In addition to structural mechanisms of substrate inhibition (above), we identified 3 compounds (PLP, AMP-PNP, and quercetin) that appear to be broad-spectrum hSULT inhibitors, and may also inhibit other non-cytosolic sulfotransferases. We have also provided insight into how PCP and possibly other polychlorinated phenolic compounds can inhibit hSULTs. These compounds, which are known endocrine disruptors, appear to bind in a manner very similar to other productive substrates, but

are unreactive, at least in part, due to their weakly acidic properties. It is predicted that acidic compounds can form salt bridges with adjacent histidines if the pH is below the *pKa* of the histidine. Our experiments were performed at pH 7.5 (above the *pKa* of the histidine) and therefore such predicted ion-pairs were not considered.

Our analyses have also provided insight into the functions of the less well-characterized SULT1C subfamily and SULT4A1. We have identified a number of novel substrates, inhibitors and compounds that bind to these SULTs in the absence of PAP. This data combined with activity assays revealed that SULT1C3 can bind catecholamines and phenolic compounds, but only the latter are substrates. SULT4A1 is inactive as an enzyme in our hands, likely due to its inability to bind PAPS or other sulfonate donors. This orphan SULT likely has an important function in the brain nevertheless, since it is highly conserved and binds well to neurotransmitters, epinephrine and norepinephrine.

The approach outlined here in which simple, medium throughput binding and activity screens can be used to profile properties of purified enzymes has proven extremely useful for identification of novel substrates and analysis of specificity across a human protein family. We demonstrate that the relationship between sequence/structure and function within this small family is remarkably complex and differences in activity can reflect just a few amino acid changes at critical locations within the protein's active site. Global sequence/structure comparisons provide good clues for broad functional classification, but cannot simply define an enzyme's cognate substrate or class of substrates. For the sulfotransferases, which have a large and flexible binding site, it is clearly necessary to perform much more detailed studies to understand both the binding and catalytic activities in terms of local structure. The actual cellular activity of hSULTs will depend on the spectrum of compounds available to a given enzyme and their relative concentrations. The tissue-specific and developmental variation in both hSULT expression and the cellular milieu of small molecules complicates further attempts to *predict* activities. Ultimately, detailed enzymatic characterization of all purified hSULTs as well as cellular assays will be needed to fully understand this family. The data presented here form a basis for further detailed biochemical and structural

studies of both active and inactive enzyme-small molecule complexes in order to fully understand the role of hSULTs in the metabolic fate of endogenous substrates as well as drugs and toxic compounds.

Accession Numbers

The NCBI (http://www.ncbi.nlm.nih.gov) Reference Sequence (RefSeq) protein accession numbers referred in this paper are: SULT1A1 (NP_803880), SULT1A2 (NP_001045), SULT1A3 (NP_003157), SULT1A4 (NP_001017389), SULT1B1 (NP_055280), SULT1C1 (NP_789795), SULT1C2 (NP_006579), SULT1C3 (NP_001008743), SULT1E1 (NP_005411), SULT2A1 (NP_003158), SULT2B1 (NP_814444), SULT4A1 (NP_055166), and SULT6B1 (NP_001027549).

The Mammalian Gene Collection (http://mgc.nci.nih.gov) clones discussed in this paper are: SULT1B1 (gi: 29550928), SULT1C1 (gi: 4507305), SULT1C2 (gi: 28830308), SULT1C3 (gi: 56847626), and SULT4A1 (gi: 7657633)

The Protein Data Bank (http://www.pdb.org) ID discussed in this paper are: SULT1A3 (1CJM), SULT1C2 (2AD1), SULT4A1 (1ZD1), SULT1C1 with PAP (2ETG), SULT1B1 with PAP (1XV1), SULT1C3 with PAP (2H8K), SULT1A1 with PAP and *p*-nitrophenol (1LS6), SULT2B1b with PAP and pregnenolone (1Q20), SULT1E1 with PAP and 3,5,3',5'-tetrachloro-biphenyl-4,4'-diol (1G3M), SULT2A1 with DHEA (1J99), SULT2A1 with PAP (1EFH), SULT1C2 with PAP and PCP (2GWH), SULT1A3 with PAP and dopamine (2A3R), SULT1A1 with PAP and estradiol (2D06), SULT2A1 with androsterone (1OV4), SULT2B1a with PAP and 2-[*N*-cyclohexylamino] ethane sulfonic acid (1Q1Q), and SULT2B1b with PAP and DHEA (1Q22).

Chapter 3

Ternary Complex Structures of Human Cytosolic Sulfotransferases 1B1 and 2A1

3.1 Introduction

A vital feature of the SULT enzyme family is the degree of substrate specificity demonstrated by the individual isozymes. It is a challenging problem to determine what controls the substrate specificity of these enzymes, yet this information is important for predicting the metabolic fate of drugs metabolized by sulfonation, and assisting the future design of small molecule modulators.

3.1.1 Structural features governing SULT specificity

Despite the fact that all SULTs have essentially superimposable crystal structures, some subtle, but important differences in the active sites of the SULT enzymes exist. For example, the human SULT1A3 enzyme displays high selectivity for endogenous and xenobiotic catecholamines (particularly dopamine); this is not shared by its closely related SULT1A1 (93% identical). In a comparison study of SULT1A1 and 1A3, two variable regions (region I and II) were found to be responsible for the observed specificity (Sakakibara, Takami et al. 1998). A chimeric construct of SULT1A1 with region I replaced with region II exhibited specifity that resembles that of SULT1A3 and *vice versa* (Sakakibara, Takami et al. 1998). In addition, by mutating E86A/E89I/E146A in SULT1A3, the catalytic efficiency decreased with dopamine as the substrate and the catalytic efficiency increased to the level of SULT1A1 with *p*-nitrophenol as the substrate. It is postulated

that these negatively charged amino acids interact with the amine group of dopamine (or other monoamine substrates) to stabilize the charge and orient the substrate correctly (Liu, Suiko et al. 2000).

Another study on SULT selectivity suggests that the specificity of SULT1E1 for 17β-estradiol over dehydroepiandrosterone (DHEA) arises from a "gating' phenomenon. For SULT1E1, DHEA is not a substrate due to the fact that the 3β-hydroxy group of DHEA is excluded from the binding pocket because of steric interactions of Y81 with the methyl group at C19 on DHEA. The mutation Y81L allowed this enzyme to bind DHEA (Petrotchenko, Doerflein et al. 1999).

The two alternatively transcribed SULT2B1 gene products, SULT2B1a and SULT2B1b have differently selectivity (2B1a for pregenolone, 2B1b for both pregenolone and cholesterol). The difference between SULT2B1a and SULT2B1b lies in an extra peptide sequence located at the N-terminus of SULT2B1b. A deletion study showed that the N-terminal peptide of 23 amino acids of SULT2B1b is crucial for cholesterol sulfonation, whereas the deletion of 53 amino acids from the C terminus does not affect the activity (Fuda, Lee et al. 2002).

3.1.2 Chemical profiling provides further structural biology cases

Upon detailed analysis of binding and activity data, the human SULTs family-wide profiling studies in Chapter 2 provided a basis for further structural analysis for SULT enzyme and small molecule interactions and therefore, assisting in defining the molecular basis governing the human SULTs substrate specificity.

As stated in the general introduction (Chapter 1) and in the last Chapter on SULTs family-wide characterization, there is at least one crystal structure available for each of the SULT isozymes, but the lack of ternary complexes (i.e. enzyme with both cofactor product PAP and substrate bound) has hindered elucidation of structural features underlying the recognition of small molecules of interest, due to the great mobility of active site residues in the apo forms.

I have chosen two specific human SULT and small molecule pairs for structural studies

described in this chapter – SULT1B1·resveratrol and SULT2A1·lithocholic acid. Neither SULT1B1 nor SULT2A1 had ternary complex structures prior to my work. Both resveratrol and lithocholic acid (LCA) have biological relevance which are discussed in detail in the following sections.

3.1.3 Human SULT1B1 and resveratrol

Resveratrol (*trans*-3,5,4'-trihydroxystilbene, Figure 3.1A), is a polyphenolic phytoalexin found in grapes, red wine, peanuts, and cranberries. It may contribute to the "French paradox" (Kopp 1998), which states the phenomenon that the French population has a relatively low incidence of heart disease despite a diet high in fat content. Over the years, a variety of biological and pharmacological activities of resveratrol including cardioprotective (Constant 1997; Wu, Wang et al. 2001), antioxidative (Leonard, Xia et al. 2003; Murias, Jager et al. 2005), antiplatelet effects (Bertelli, Giovannini et al. 1995), and non-selective inhibitory effects on cyclooxygenase-1 (COX-1) and cyclooxygenase-2 (COX-2) (Likhitwitayawuid, Sawasdee et al. 2002; Murias, Handler et al. 2004) have been proposed. Furthermore, chemopreventive effects of resveratrol have been reported in various experimentally induced tumor models (Jang, Cai et al. 1997; Horvath, Henshall et al. 2005; Kundu and Surh 2008) and such effects have yet to be demonstrated in humans. The treatment of resveratrol in mice has been demonstrated to significantly increased their aerobic capacity, protected them against diet-induced-obesity and insulin resistance (Lagouge, Argmann et al. 2006). The antiaging effects of resveratrol were also postulated based on the observation that mice fed with a high-calorie diet after chronic administration of resveratrol have a significant improvement in health and lifespan (Baur, Pearson et al. 2006). This result was later challenged by a conflicting report that suggested resveratrol did not reduce blood sugar in mice on a high fat diet (Pacholec, Bleasdale et al. 2010).

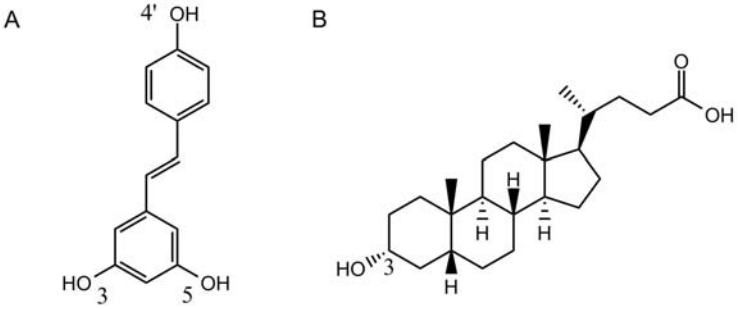


Figure 3.1 **Chemical structures of compounds of interest.** *A*, resveratrol and *B*, lithocholic acid.

If resveratrol indeed possesses any beneficial properties, its purported efficacy as a health supplement or administrated drug would be predicted to be largely dependent on its absorption, and metabolism . It has been demonstrated in humans that resveratrol is well-absorbed when taken orally (Walle, Hsieh et al. 2004). However, due to its rapid metabolism and elimination through phase II conjugation to sulfates and glucuronides (no evidence of phase I functionalization), resveratrol oral bioavailability in its free form is relatively low (Walle, Hsieh et al. 2004). Resveratrol 3-*O*-sulfate is the major sulfonated metabolite of resveratrol in humans (De Santi, Pietrabissa et al. 2000; De Santi, Pietrabissa et al. 2000; Walle, Hsieh et al. 2004; Wenzel and Somoza 2005). In the gut and liver, the mono-sulfonated resveratrol formation is extremely rapid and constitutes the rate-limiting step in resveratrol bioavailability (Walle, Hsieh et al. 2004; Urpi-Sarda, Jauregui et al. 2005). Compared to their parent molecule, sulfonated resveratrol metabolites have poor cytotoxicity in human malignant and nonmalignant breast cancer cell lines (Murias, Miksits et al. 2008; Miksits, Wlcek et al. 2009).

In my chemical profiling of the human SULTs family, resveratrol bound SULT1B1 in the absence of cofactor product PAP in the thermal aggregation binding assay (Table 2.1). The SULT1B1 sulfonation activity is also positively correlated with the three resveratrol concentrations tested, with no substrate inhibition observed (Table 2.2).

The SULT1B1 isozyme is the only SULT1B subfamily member. It is expressed in human liver, colon, small intestine, and blood leukocytes, and it is the major SULT enzyme in small intestine (Riches, Stanley et al. 2009). Obtaining a SULT1B1·PAP·resveratrol ternary complex structure would provide detailed information of the SULT1B1 active site and the structural basis for SULT enzyme recognition of resveratrol. With such knowledge, it may be possible to design potent resveratrol derivatives that are less likely to be metabolized in the phase II sulfonation metabolism.

3.1.4 Human SULT2A1 and lithocholic Acid

Lithocholic acid (Figure 3.1B) is a naturally occurring toxic secondary bile acid. Biosynthetically,

it is a dehydroxylated product of the primary bile acid - chenodeoxycholic acid, generated through the actions of enterobacteria. Lithocholic acid has been reported to be carcinogenic in the intestine and cholestatic in the liver of experimental animals and humans (Javitt 1966; Miyai, Price et al. 1971; Narisawa, Magadia et al. 1974; Kozoni, Tsioulias et al. 2000). Due to its structural similarity to vitamin D, lithocholic acid binds to the vitamin D receptor with low affinity and increase expression of cytochrome P450 (CYP) enzymes in the intestine and the liver of humans (Araya and Wikvall 1999; Makishima, Lu et al. 2002; Deo and Bandiera 2008) and thus can initiate self-induced phase I detoxification.

The conjugation of lithocholic acid is a major pathway for protection against lithocholic acid induced liver damage (Kitada, Miyata et al. 2003). Cytosolic sulfotransferases and UDP-glucuronosyltransferases are responsible for making the major conjugated forms of lithocholic acids - 3-sulfates and 3-glucuronides, respectively (Stiehl 1974; Dionne, Tuchweber et al. 1994; Hofmann 2004). These conjugated forms are less toxic than their parent molecule (Takikawa, Tomita et al. 1991). The increased hydrophilicity due to conjugation facilitates the elimination of the bile acids in the feces and urine (Stiehl, Earnest et al. 1975; Carey, Wu et al. 1979; Oelberg, Chari et al. 1984; Kuipers, Heslinga et al. 1986).

Lithocholic acid was identified as a binding hit for SULT2A1 in the absence of the cofactor product PAP (Table 2.1). In my activity profile (Table 2.2), SULT2A1 displayed preference towards lithocholic acid of the eight human SULTs tested (SULT1C3 had almost negligible level of activity, while the other six SULTs has no observed activity).

Human SULT2A1 is expressed in adrenal gland, liver, and brain (Falany 1997). It was originally termed as dehydroepiandrosterone (DHEA) sulfotransferases, because it has long been known as an enzyme that catalyzes the sulfate conjugation of various hydroxysteroids including DHEA, androsterone (ADT), testosterone, estradiol, and many other endogenous steroids (Chen, Banoglu et al. 1996; Falany 1997; Kakuta, Pedersen et al. 1998). The bile acid sulfonation activity of SULT2A1 has also been documented (Nagata and Yamazoe 2000).

Human SULT2A1·DHEA (PDB: 1J99), SULT2A1·PAP (PDB: 1EFH), and SULT2A1·ADT

(PDB: 1OV4) binary complex structures are already available (Pedersen, Petrotchenko et al. 2000; Rehse, Zhou et al. 2002; Chang, Shi et al. 2004). Using structural alignment, two amino acids - Tyr238 and Met137, were identified as residues responsible in regulating substrate inhibition (Lu, Hsieh et al. 2008). Although these binary complexes provided important information regarding the molecular interactions between steroids, PAP, and SULT2A1, there is currently no ternary complex crystal structure exist to capture protein, cofactor, and substrate together. Considering the preference towards lithocholic acid of SULT2A1 observed in the binding and activity assays, obtaining a SULT2A1·PAP·lithocholic ternary complex structure would give details regarding the SULT2A1 substrate preferences and in the meantime, reveal how the cofactor product PAP and the substrate orient themselves relatively to each other in the SULT2A1 active site.

3.2 Methods

3.2.1 Protein cloning, expression, and purification

Human SULT1B1 and SULT2A1 full-length DNA sequences were amplified by PCR and sub-cloned into the pET28a-LIC vector downstream of the poly-histidine coding region. The plasmids containing the two proteins were then introduced into *E. coli* BL21(DE3) codon plus RIL (Stragagen). The cells were grown in Terrific Broth (TB) in the presence of 50 μg/mL of kanamycin at 37 ºC to an OD_{600} of 0.8. Cells were then induced by the addition of isopropyl-1-thio-D-galactopyranoside (IPTG), with a final concentration 0.5 mM and incubated overnight at 15 ºC before harvesting.

Cells were harvested by centrifugation at 7,000 rpm. The cell pellets were frozen in liquid nitrogen and stored at -80 °C. For purification the cell pastes were thawed and resuspended in lysis buffer (phosphate buffer saline, PBS, pH 7.5, 0.5 M NaCl, 5% glycerol) with protease inhibitor (0.1μM phenylmethyl sulfonyl fluoride, PMSF). The cells were lysed by passing through

a microfluidizer (Microfluidics Corp.). The cell lysates were clarified by centrifugation using a Beckman JLA-16.250 rotor at 16,000 rpms for 1 hour at 4 °C.

The clarified lysates were loaded onto 5 mL HiTrap Chelating column (Amersham Biosciences), charged with Ni^{2+}. The column was washed with 20 mM Tris-HCl buffer, pH 8.0, containing 500 mM NaCl and 50 mM imidazole (10 column volumes), and the protein was eluted with elution buffer (20 mM Tris-HCl, pH 8.0, 500 mM NaCl, 250 mM imidazole). The purified proteins were dialyzed against buffer 20 mM Tris-HCl, pH 8.0, 150 mM NaCl and treated with thrombin (Sigma) overnight at 4 ºC. The proteins were further purified to homogeneity by ion-exchange chromatography on Source 30Q column (1.0x10 cm) (Amersham Biosciences), equilibrated with buffer 20 mM Tris-HCl, pH 8.0, and eluted with a linear gradient of NaCl up to 500 mM concentration (30 colume volumes). Purification yield was 26 mg of SULT1B1 per 1L of culture and 2mg/L for SULT2A1.

3.2.2 Protein crystallization

PAP, resveratrol, and lithocholic acid were purchased from Sigma. PAP was dissolved in the distilled water at 100 mM concentration as a stock solution. Resveratrol and lithocholic acid were both dissolved in DMSO with a stock solution concentration of 100 mM. Compound stock solutions were stored at -20 °C.

Purified 10 mg/mL SULT1B1 was pre-incubated with PAP and resveratrol at 1:10:30 molar ratio. The protein crystal containing both PAP and resveratrol was crystallized using the sitting drop vapor diffusion method at 20 ºC by mixing 2 µL of the protein-compound mixture with 2 µL of the reservoir solution containing 0.1 M Bis-Tris, pH 6.0, 0.001M DTT, 0.1 M ammonium sulfate and 17% polyethylene glycol 3350. Crystals appeared within 24 hrs.

The protein mix contains 10 mg/mL (about 0.3 mM) purified SULT2A1 mixed with 1 mM PAP and 1 mM of lithocholic acid in 20 mM Tris-HCl buffer, pH 8.0. Crystals were obtained using hanging drop vapor diffusion method at 20°C by mixing 2 µl of the protein mix with 2 µl of the

reservoir solution containing 0.05 M Calcium Chloride, 18% PEG 3350, 0.1 M Hepes buffer, pH 7.0, and 10% ethylene glycol. Crystals appeared within 2 weeks.

SULT1B1 crystals were briefly soaked in the corresponding mother liquor supplemented with 16% glycerol as cryoprotectant before freezing in liquid nitrogen. SULT2A1 crystals were frozen instantly in liquid nitrogen, followed by a brief period of contact with paratone-*N* before data collection.

3.2.3 Data collection and processing

SULT1B1 X-ray diffraction data were collected at 100 K at beamline 23ID and SULT2A1 data were collected at 19ID of Advanced Photon Source (APS) at Argonne National Laboratory. Both data sets were processed using the HKL-2000 software suite (Minor 1997).

3.2.4 Structure determination and refinement

SULT1B1·PAP·resveratrol (PDB:3CKL) and SULT2A1·PAP·lithocholic acid (PDB: 3F3Y) were solved by molecular replacement using the program Phaser (McCoy, Grosse-Kunstleve et al. 2005). The model of apo SULT1B1 (PDB:2Z5F) served as a template for the SULT1B1 ternary complex structure. The model of SULT2A1·androsterone binary complex (PDB: 1OV4) was used as a searching template for SULT2A1 ternary complex structure. Refinements were carried out with Refmac (Murshudov, Vagin et al. 1997) for both structures. Graphics program COOT (Emsley and Cowtan 2004) was used for real space model building and visualization. Data statistics are reported in Table 3.1.

3.2.5 Structural analysis

Structure comparisons in this chapter were carried out using the program “The PyMOL Molecular Graphics System” (DeLano Scientific, San Carlos, CA, USA). Sequence alignments were

performed using ClustalW (Thompson, Higgins et al. 1994). Figures were generated with PyMol program (DeLano Scientific, San Carlos, CA, USA). Docking of resveratrol into SULT isozymes were performed using ICM's (Molsoft L.L.C., San Diego, CA USA) built-in local minimization function.

3.3 Results and Discussion

3.3.1 Overall structure of SULT1B1 ternary complex

To investigate how the polyphenolic compound resveratrol interacts with the human SULTs, I solved the ternary complex crystal structure of human SULT1B1 in complex with PAP and *trans*-resveratrol at 2.0 Å (data statistics shown in Table 3.1). The entire amino acid sequence of SULT1B1 is visible in this ternary complex. A monomer structure of the complex is presented in Figure 3.2A. The overall architecture is consistent with the previous structural studies of SULTs. It has a classical SULT fold with a four-stranded parallel β-sheet surrounded by α-helices. The cofactor (PAP/PAPS) binding site, and the binding mode and conformation of PAP are well conserved. Electron density corresponding to resveratrol is observed in the predicted binding pocket of SULT1B1 (Figure 3.2B).

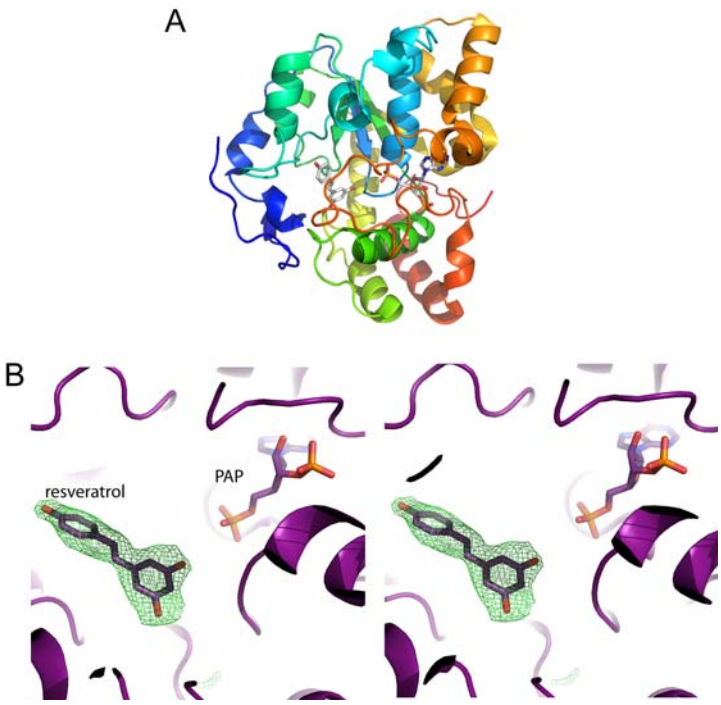


Figure 3.2 **SULT1B1 in complex with PAP and resveratrol.** *A*, ribbon represenation of SULT1B1 complex; *B*, stereo diagram showing PAP and resveratrol in the active site. The F_o-F_c omit density map of resveratrol is shown in green and contoured at 2σ.

Table 3.1 **Refinement Statistics for SULT1B1 and SULT2A1 ternary complex structures.**

	SULT1B1 complex	SULT2A1 complex
Data Collection		
Wavelength (Å)	0.97962	0.97942
Space Group	$P2_1$	$P2_12_12_1$
Unit Cell Parameters		
a, *b*, *c* (Å)	78.8, 47.8, 92.7	79.4, 96.2, 159.3
α, β, γ (°)	90, 93.95, 90	90, 90, 90
Resolution (Å)	50.0-1.79(1.85-1.79)[a]	30.0-2.00 (2.07-2.00)
Unique Reflections	55786(1432)	80907(6973)
Completeness (%)	85.8(22.3)	97.6(85.6)
Redundancy	3.4(1.5)	5.1(3.6)
I/σ(I)	20.0(1.1)	12.9(2.2)
R_{sym} (I)	0.077(0.545)	0.141(0.527)
Refinement		
Resolution (Å)	19.74-2.00	30.0-2.20
Reflections		
Number	44063	60784
Completeness (%)	99.2(94.2)	99.1(97.5)
Test Set (%)	5.1	2.0
R_{work}(%)	18.1(20.4)	23.6(33.9)
R_{free}(%)	24.1(27.7)	29.9(46.4)
E.S.U. (Å)[b]	0.188	0.361
Contents of A.U.[c]		
Protein Molecules/Atoms	2/4932	4/9176
Solvent	671	538
Mean B-Factors (Å^2)		
Protein	28.29	32.9
Ramachandran Plot (%)		
Preferred	98.1	96.5
Allowed	1.9	3.5
Outlier	0.0	0.0
RMSD from Target Geometry		
Bond Lengths (Å)	0.013	0.018
Bond Angles (°)	1.454	1.416
PDB ID	3CKL	3F3Y

[a]Data for the highest resolution shell in parenthesis.
[b]E.S.U. – estimated overall coordinate error based on maximum likelihood.
[c]A.U. – asymmetric unit.

3.3.2 Comparison of SULT1B1 ternary and binary structures

The structure of the binary complex of SULT1B1·PAP was reported previously (Dombrovski, Dong et al. 2006). Upon superimposing the resveratrol complexed structure onto the binary structure (Figure 3.3), the root-mean-square deviation (RMSD) was calculated to be 0.29Å over all atoms, suggesting virtually identical structures. The only visible difference is the conformation of the N-terminal flexible loop. The resveratrol complexed structure has a complete SULT1B1 sequence coverage while the binary complex model has a sequence starting from Pro4. Therefore, the resveratrol complexed structure contains three more ordered amino acid residues at the extreme terminal end.

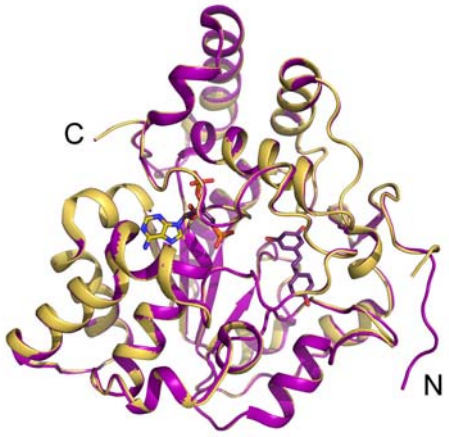


Figure 3.3 **Structural overlays of SULT1B1 binary and ternary complexes.** Yellow: SULT1B1·PAP (PDB: 2Z5F); Pink: SULT1B1·PAP ·resveratrol (PDB: 3CKL).

Only one residue, Arg90, in the substrate binding site shows marked difference between the two structures. The loops centered on this residue in SULT1C1·PAP and SULT1B1·PAP structures were shown to have great conformational variability and the Cα atoms RMSD of Gln90 (SULT1C1) and Arg90 (SULT1B1) was calculated to be 5.0Å (Dombrovski, Dong et al. 2006). In the SULT1B1·PAP binary structure (Figure 3.4A), Arg90 side chain appears to partially cover the hydrophobic substrate binding site and assumes a "closed" conformation in the absence of a ligand. Upon the binding of the substrate - resveratrol (Figure 3.4A), the side chain Arg90 shifts towards an "open" conformation. This "open" conformation captured by the ternary complex structure revealed a larger opening on the protein surface which presumably would allow easy entry and exit of the substrate during the catalysis. Residue Arg90 is unique to the SULT1B1 isozyme (Figure 3.4C). Based on the observed conformational change between the SULT1B1 complexes with and without resveratrol of Arg90, it is likely that this residue is acting as a gating residue for the substrate binding site for SULT1B1.

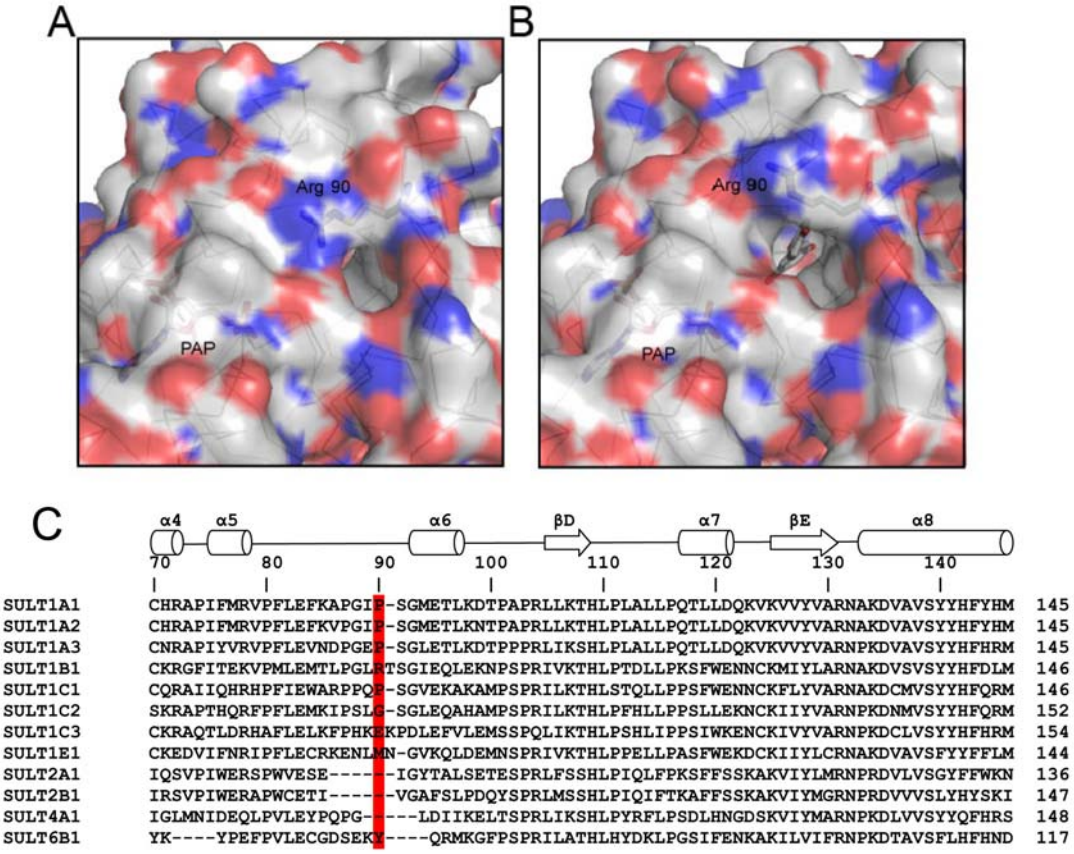


Figure 3.4 **Arg 90 may function as a gating residue guarding substrate entry for SULT1B1.** *A*, SULT1B1·PAP complex (PDB: 2Z5F); *B*, SULT1B1·PAP·resveratrol complex (PDB: 3CKL). *C*, SULT1B1 Arg 90 is unique residue according to SULT family wide sequence alignment.

3.3.3 Structural determinants for resveratrol preference

Resveratrol (Figure 3.1A) has three hydroxyl groups, which are all potential sulfonyl conjugation sites. Two hydroxyl groups (3 and 5-OH) are in *meta* position to each other and another hydroxy group (4'-OH) is located on the other end of the molecule. The binding and activity assays we used to assess the human SULTs chemical profile did not provide information regarding which hydroxyl group or groups on resveratrol would be sulfonated. Resveratrol 3-sulfate is the major reported sulfated metabolite of resveratrol in humans, and other sulfate conjugates such as resveratrol 4'-sulfate, 3,5-disulfate, 3,4'-disulfate, and 3,4',5-trisulfate were identified in *in vivo* studies as

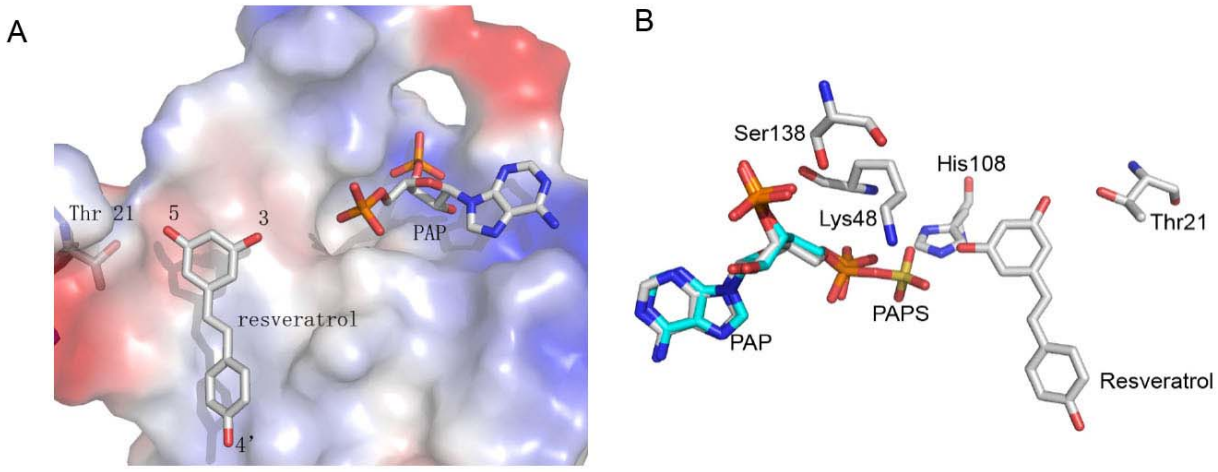


Figure 3.5 **Resverstrol in the SULT1B1 active site**. *A*, resveratrol binds to a generally hydrophobic pocket except the area near the 5-OH side of the compound. *B*, residues surrounding 3 and 5-OH of resveratrol. PAPS from a SULT1E1 structure (PDB:1HY3) was superimposed onto PAP to approximate the reaction center.

well (De Santi, Pietrabissa et al. 2000; De Santi, Pietrabissa et al. 2000; Walle, Hsieh et al. 2004; Wenzel and Somoza 2005).

In the SULT1B1·PAP·resveratrol ternary complex structure (Figure 3.5A), resveratrol was found to be sandwiched in the hydrophobic binding pocket with 3-OH facing the 5-phosphate of PAP and 5-OH forming a hydrogen bond with Thr21. In Figure 3.5B, the actual cofactor PAPS (from SULT1E1 structure PDB:1HY3) was superimposed on to the SULT1B1 ternary complex to mimic the reaction center along with the catalytic His108 and other coordinating residues Ser138 and Lys48. This orientation of the resveratrol provides a structural based explanation for the preference of sulfate conjugation at 3-OH site and thus the observed resveratrol-3-sulfate metabolite. Since resveratrol is the first molecule with multiple hydroxy groups co-crystallized in the human SULT active site, the hydrogen bonding information observed here can also be applied to other stilbenes which have similar structural features as resveratrol.

From the human SULT enzyme profile (Table 2.2), all of the SULT1 family members are capable of catalyzing the sulfonation of resveratrol while none of the SULT2 family members exhibit any activities. To investigate the structural basis for this observation, resveratrol was docked

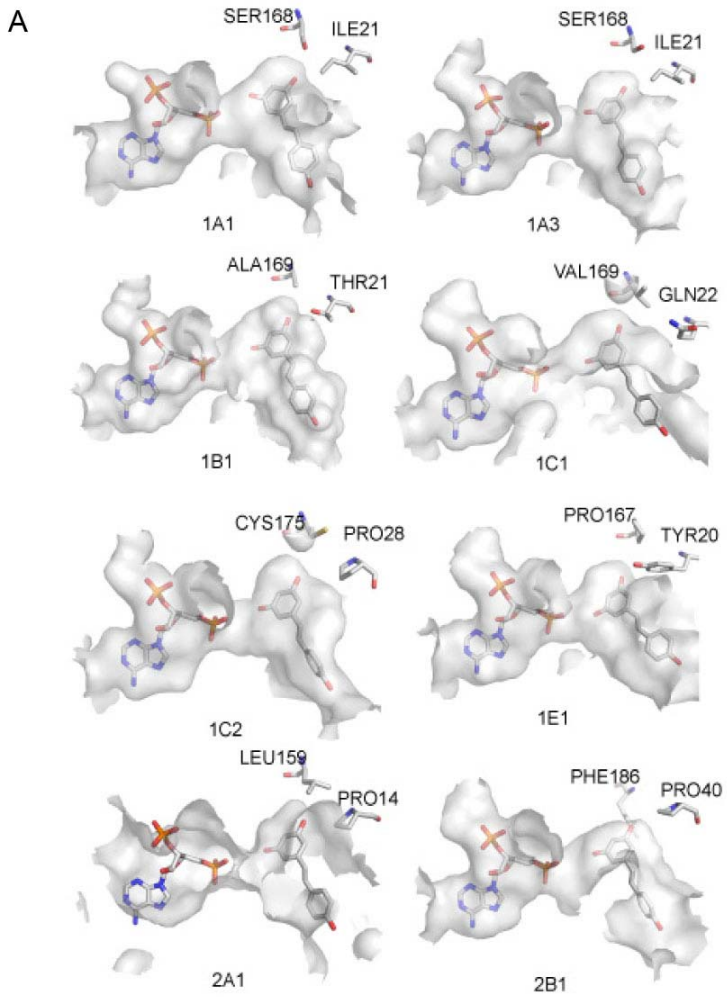
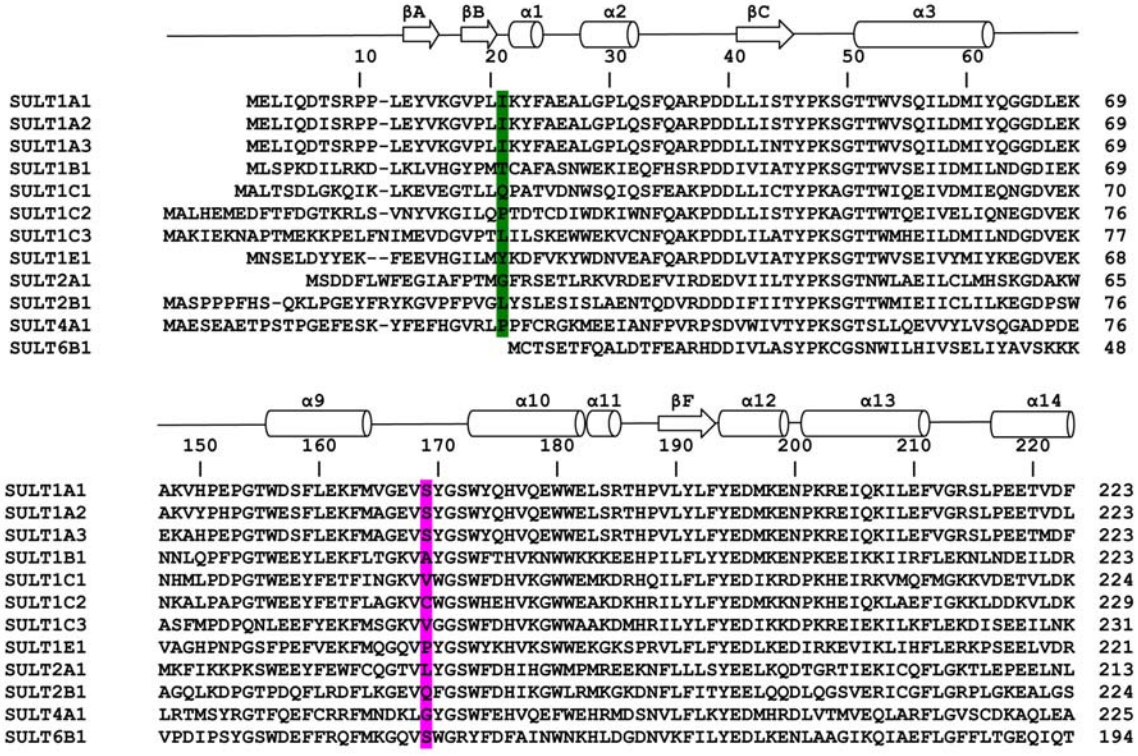


Figure 3.6 **Resveratrol binding across the SULT family is governed by a hydrophilic residue coordinating the 5-OH on resveratrol.** *A*, superimposition of resveratrol into the binding site of human SULTs which have defined substrate binding sites in the crystal structures except for SULT1B1 which is the actual ternary complex [PDB codes used for each isozyme: 1LS6(1A1), 2A3R(1A3), 2ETG(1C1), 2GWH(1C2), 1AQU(1E1), 1J99(2A1), 1Q20(2B1)]; *B*, the two possible locations of hydrophilic residues highlighted in the human SULT family-wide sequence alignment.

B



into the active sites of members of the SULT1 and 2 families (Figure 3.6A). It appears that the presence of residues with polar side chains facing the 5-OH is the key determinant of the observed specificity. The two possible positions of these residues are labeled in green and pink in the SULT family sequence alignment (Figure 3.6B). One residue position is in line with the SULT1B1 Thr21 which resides in the loop between β2 and α1. SULT1C1 and SULT1E1 belong to this type. They have Gln22 and Tyr20 respectively, facing the 5-OH of resveratrol and therefore stabilizing the molecule. The other location of the polar residue is between α9 and α10. SULT1A1, 1A3, and 1C2 belong to this type. The SULT1A members have Ser168 and SULT1C2 has Cys175 within hydrogen bonding distance of the 5-OH of resveratrol. Looking at the two SULT2 family members which do not have activity towards resveratrol, both sites consist of hydrophobic residues that are unable to accommodate the 5-OH.

In summary, by comparing the activity profile towards resveratrol and structural features according to the solved SULT1B1 ternary complex structure, it can be speculated that for SULT isozymes which are capable of interacting through hydrogen bonding with the hydroxy groups of resveratrol can sulfonate the molecule. The 3-OH sulfonation is the preferred reaction site due to the orienting of 5-OH. To make a resveratrol analogue that is less likely to be sulfonated, the 5-OH site can be modified into a bulky hydrophobic group and this will destablize the analogue binding in the substrate binding site. In addition, future modulator designs can target the coordinating residues (highlighted in Figure 3.6B) which have been identified in this structural study.

3.3.4 Overall structure of SULT2A1 ternary complex

To obtain a more complete picture of the SULT2A1 enzyme active site and to examine the specificity of SULT2A1 towards lithocholic acid observed in the family wide activity profile (Table 2.2), I solved the first ternary complex crystal structure of SULT2A1 with PAP and lithocholic acid at 2.2 Å (data statistics shown in Table 3.1). In accordance with all of the human SULT structures solved so far, the SULT2A1 ternary complex has a classical SULT fold with a four-stranded parallel β-sheet surrounded by α-helices. A monomer of the complex is shown in ribbon diagram in Figure

3.7A. Electron densities corresponding to both lithocholic acid and cofactor product PAP are shown in the active site (Figure 3.7B).

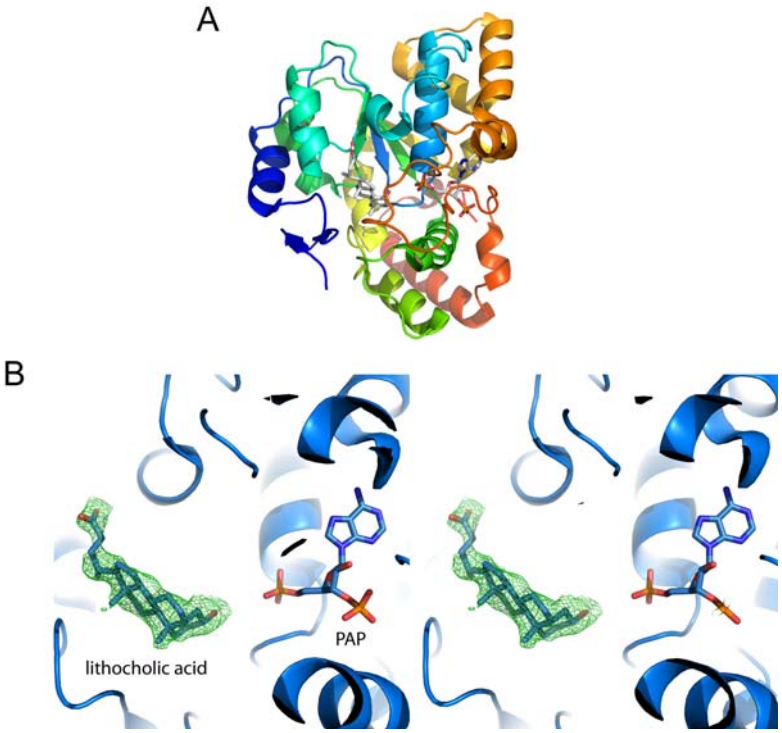


Figure 3.7 **SULT2A1 in complex with PAP and lithocholic acid.** A, Ribbon representation of SULT2A1 complex; B, stereo diagram of PAP and lithocholic acid. The F_o-F_c omit density map of lithocholic acid is shown in green, and contoured at 2σ.

3.3.5 Comparison of SULT2A1 ternary and binary structures

The SULT2A1·PAP·lithocholic acid ternary complex (blue) is compared to the binary structure of SULT2A1·PAP (light blue) in Figure 3.8A. The overall structure features seem to align quite well as expected. The only difference is the labeled loop which clashes into the substrate binding site when there is no compound bound. In the previous systemic structural characterization, it was noted that this loop becomes more ordered as a ternary complex is formed (Figure 2.2).

In Figure 3.8B, the ternary complex structure (blue) is compared to a binary structure of SULT2A1·ADT (yellow). Without the cofactor product PAP present, androsterone penetrates deeper in the pocket. In the presence of PAP, the backbone of lithocholic acid is almost parallel to that of androsterone except for the 3β – hydroxy attached 6 membered ring. This ring is oriented almost vertical to the plane of the other rings and as a result, the 3β – hydroxy end of the molecule is farther away from the center of the enzyme than androsterone. Dehydroandrosterone-complexed

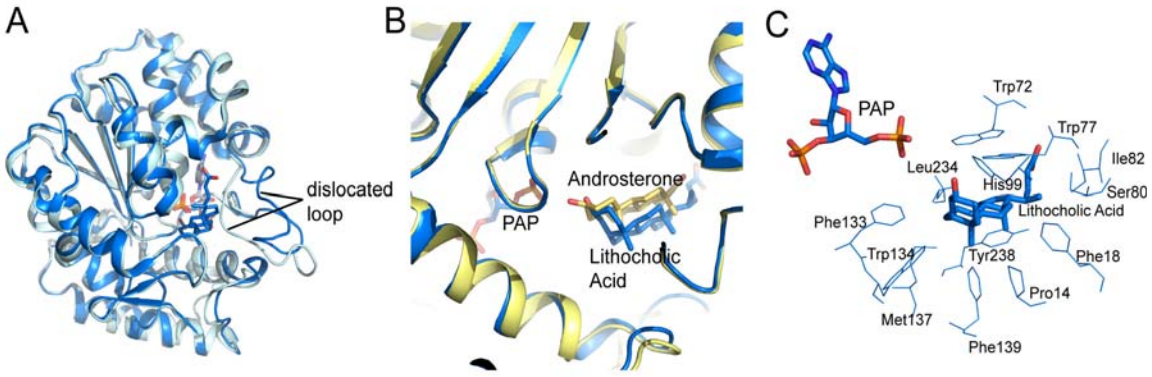


Figure 3.8 **Comparison of SULT2A1 ternary and binary complex structures.** *A*, SULT2A1 ternary complex structure (blue) superimposed onto the binary structure SULT2A1·PAP (PDB: 1EFH) shown in light blue. A substrate binding loop has variable conformations in these two structures. *B*, SULT2A1 ternary complex structure (blue) superimposed onto SULT2A1·androsterone (ADT) (PDB: 1OV4) shown in yellow. Lithocholic acid has a bent conformation compare to the relatively planar ADT. *C*, Active site residues within 4Å of lithocholic acid.

SULT2A1 structures (PDB: 1J99, 2QP3 and 2QP4) are all similar to the androsterone complexed structure in which the steroid molecule assumes an almost flat conformation. The ternary complex structure presented here may be a more realistic model of what the active side would look like during the sulfonation reaction. In addition, the cofactor PAPS has an extra sulfate group compared to PAP, the deep penetration of complexed molecules in enzyme-substrate binary complexes is unlikely due to predicted clashes with the cofactor.

Consistent with the binary structures of SULT2A1, lithocholic acid is bound in a mainly hydrophobic pocket (Figure 3.8C). The hydrophobicity of the SULT2A1 substrate pocket poses restraints on the molecules that can enter this site and appears to be the structural based determinant for enzyme specificity.

3.4 Conclusions

Based on the information provided in the previous family-wide chemical profiling of the human SULTs family, two human SULT ternary complex crystal structures – SULT1B1·PAP·resveratrol

and SULT2A1·PAP·lithocholic acid were solved to understand the observed substrate preferences. The resveratrol complexed SULT1B1 structure suggested the hydrogen-bonding coordination of the 5-OH group is the structural determinant for the observed substrate preference towards resveratrol by the SULT1 family. This structure also provides insight of how SULTs coordinate compounds with more than one hydroxy group. As the first ternary complex structure for SULT2A1, the lithocholic acid complexed SULT2A1 structure provides a more realistic and complete picture of the catalytic site and confirms that SULT2A1's specificity for steroids-related compounds lies in the high hydrophobicity of the substrate binding pocket.

Chapter 4

Histone Peptide Binding and Activity Profiles for Nuclear Associated SIRTs

4.1 Introduction

Sirtuins (SIRTs) belong to an ancient family of NAD^+-dependent protein deacetylases with targets ranging from transcriptional regulators to metabolic enzymes. As with the human cytosolic sulfotransferase family (SULTs), discovery of potent small molecules probes is critical to understand sirtuin biology and to further evaluate sirtuins as pharmacological targets. In comparison with the human SULTs family, human SIRTs are more complex due to their heterogeneous subcellular compartmentalization and the fact they share relatively low sequence identity (< 40%), yet converged to catalyze the same enzymatic reactions (Michishita, Park et al. 2005). In addition, SIRTs catalyze reactions on other protein substrates whereas SULTs exert their functions on small molecules. Despite the differences between the two protein families, the principles behind the experimental strategy for understanding the molecular mechanism of specificity and promiscuity used for the SULTs family can be applied to SIRTs.

4.1.1 Systematic biochemical/biophysical characterization of the nuclear sirtuins

Sirtuins family members have conserved catalytic mechanism (as discussed in Chapter 1), but little is known about the substrate preferences of the human sirtuins. No study to date has systematically compared human sirtuins across a panel of substrates and determined their preferences toward

acetylated substrates. Unlike the cytosolic sulfotransferases family, sirtuin members can be found in the cytosol, nucleus, and in mitochondria (Michishita, Park et al. 2005), and therefore have different tendencies to encounter certain protein substrates to exert biological functions depending on their cellular locations. For the purpose of my study, I focused on sirtuins that are expressed in the nucleus, namely SIRT1, 2, 6, and 7; these may be responsible for linking the cellular energy state (NAD^+) to epigenetic control. By focusing only on the sirtuins that localize to the nucleus, it also facilitated the compilation of a substrate library - acetylated histone proteins. Several acetylated histone marks, including H3K9Ac and H4K16Ac are well studied sirtuin physiological substrates (Vaquero, Scher et al. 2006; Taylor, Maxwell et al. 2008). In addition, since all possible acetylation sites on the histones are known, a comprehensive acetyl histone peptide library can be constructed to probe the nuclear sirtuins. Hits from the binding assay were further analyzed using secondary enzymatic assays. Validated peptide substrates can be used to support further structural biology studies of the nuclear associated sirtuin enzymes.

4.1.2 Initial binding assay – SPOT blot

The aim of the study was to have a general idea of the nuclear sirtuin substrate preference profile. Although it is desirable to study the activities of nuclear associated sirtuins in the context of native chromatin or purified nucleosomes, it is difficult to directly introduce systematic modification in the context of purified histones or nucleosomes enzymatically (difficulties involving generate a specific uniformly modified histone mark in an *in vitro* system) and chemically (chemical-ligation techniques are currently limited to introduce modifications at extreme termini) (He, Bauman et al. 2003; Shogren-Knaak, Fry et al. 2003; Shogren-Knaak 2007). For many binding, structural, and even specificity studies, it is often appropriate to use short synthetic histone peptide as histone protein surrogates since most histone marks are located on the natively disordered N-terminal histone tails and these tails do not assume secondary or tertiary structure in the absence of a binding partner.

There are several methods that can be used to assess the binding of protein and potential

synthetic peptide substrates. One strategy utilizes a protein-array approach that involves probing pre-selected immobilized glutathione-S-transferase fusion proteins with fluorescently tagged methylated histone peptides (Kim, Daniel et al. 2006). An alternative strategy is to use biotinylated histone peptide to 'pull down' histone binding proteins from a cell lysate (Wysocka 2006). For the purpose of identifying the substrate preferences of the four purified human nuclear sirtuins, I needed a method that can incorporate a library of histone acetylation sequences and provides a direct readout of binding profiles for each protein. I therefore selected the synthetic peptide arrays (SPOT blot analysis) (Nady, Min et al. 2008) for primary acetylated peptide substrate binding screening. The efficiency and quality of this method using standard L-amino acids for synthesis has been validated by mass spectrometry (Frank 2002; Hilpert, Winkler et al. 2007). The optimal peptide length for synthesis was between 6 and 18 amino acids but this greatly depended on the sequence (Toepert, Knaute et al. 2003). The incorporation of a bulky modification (i.e. the acetyl-Lys(Ac) required in my experiment) was evaluated as quite efficient as a preliminary screening tool (Nady, Min et al. 2008).

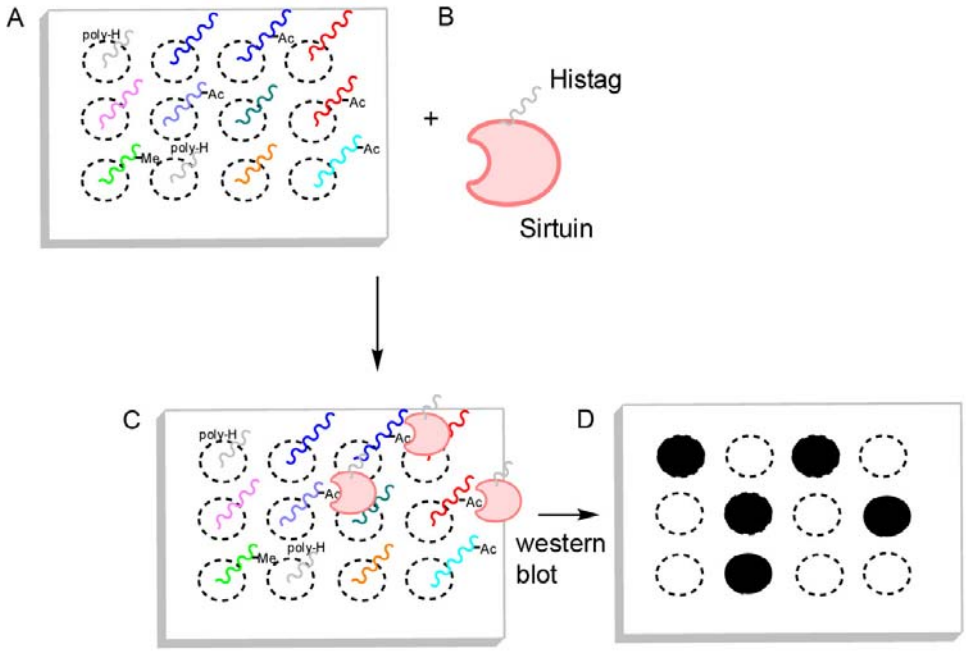


Figure 4.1 **Schematic of the sirtuin histone peptide array membrane SPOT blot method.** *A*, the array of histone peptides with predetermined sequences and modifications is synthesized on the membrane. *B*, the recombinant sirtuin protein of interest with poly-histidine tag. *C*, the membrane is incubated with the protein and *D*, visualization of hits via western blot techniques and an antibody against the His-tag. Positive interactions appear as a spot on the blot. Poly-His peptides are used as positive control.

The principle of this screen technique is illustrated in Figure 4.1. Acetylated histone peptides are synthesized directly on a cellulose membrane with a polyethylglycol linker. The membrane is then extensively blocked with a dilute protein solution followed by incubation with a His-tagged nuclear sirtuins. The presence of bound sirtuin to a specific acetylated histone mark is detected via Western blot analysis. The method has been reported to be very sensitive and can detect even weak peptide–antibody interactions with dissociation constants as high as in the range of 100μM to 1mM (Reineke, Sabat et al. 1996; Kramer, Reineke et al. 1999). The strength of the SPOT peptide assay is its unbiased, comprehensive and systematic approach to evaluate a given protein for binding to histone tails. As with any screening method, hits from SPOT blot method required careful examination. Due to the lack of quantitative assessment of the amount of peptides synthesized on each spot, the intensity of each spot after Western blot does not directly correlate with how strong the protein binds to the peptides. Because of this semi-quantitative nature of the SPOT method, identified hits need to be further evaluated by a secondary biochemical assay to eliminate false positives. The false negative rate for this method was reported to be very low according to the results from over 50 membrane screens for more than 15 different proteins with either literature data or follow-up studies (Nady, Min et al. 2008).

4.1.3 Secondary assay - sirtuin continuous enzyme-coupled assay

There are many available end-point sirtuin enzymatic assays that require either modified or artificial substrates. One class of sirtuin assays uses radioactive substrates to monitor the transfer of a [^{3}H]-labeled acetyl group on a peptide or protein substrate (Borra and Denu 2004) or the release of [^{14}C]-nicotinamide from [^{14}C]NAD^{+} (McDonagh, Hixon et al. 2005). Other sirtuin assays utilize fluorescent substrates, measuring fluorescence-resonance-energy-transfer (FRET) (Marcotte, Richardson et al. 2004) or fluorescence polarization (Milne, Lambert et al. 2007). The most representative assay in this class is the commercial Fluor-de-Lys assay (Biomol) (Wegener, Hildmann et al. 2003). It employs 7-amino-4-methylcoumarin (AMC), which is quenched through conjugation to the C-terminal end of a short p53 acetyl-lysine containing peptide. Once deacetylated,

the newly free ε-amino group of the lysine becomes a trypsin substrate. Trypsin cleavage releases the fluorophore resulting in an increase in fluorescence. This assay has been implemented in a high-throughput manner and identified several putative SIRT1 inhibitors and activators (Howitz, Bitterman et al. 2003; Napper, Hixon et al. 2005). However, SIRT1 activating compounds such as resveratrol that were identified using this assay do not correlate with activity measured using peptides or full-length proteins without the fluorescent tag (Borra, Smith et al. 2005; Kaeberlein, McDonagh et al. 2005). Activation with these compounds required the AMC portion of the Fluor de Lys substrate. Other sirtuin assays include those based on mobility shift assay (Liu, Gerber et al. 2008), bioluminescence (Liu, Gerber et al. 2008), capillary electrophoresis (Fan, Ludewig et al. 2008; Fan, Ludewig et al. 2009) and enzyme-linked immunosorbent assays (ELISA) (Schlicker, Gertz et al. 2008).

While all these methods have been proven to be useful in studying sirtuin enzymatic

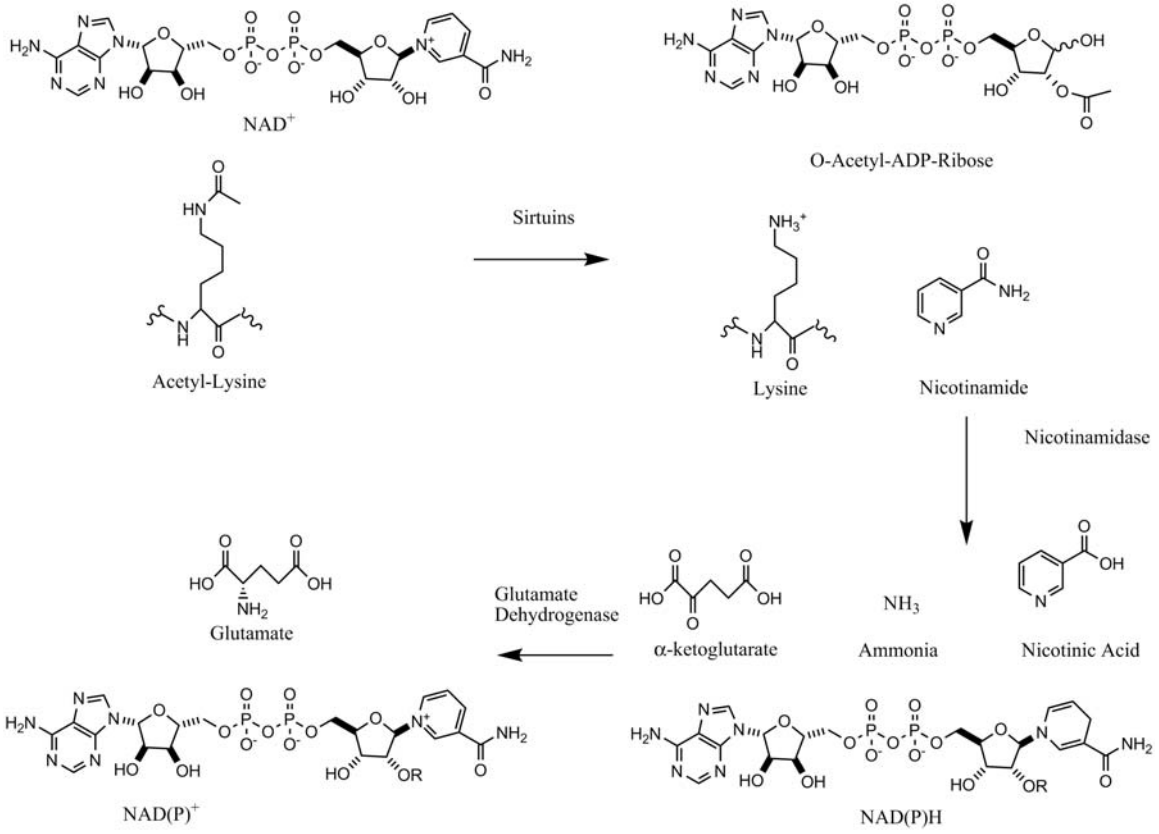


Figure 4.2 **General scheme of sirtuin continuous enzyme-coupled assay.** This system utilizes nicotinamide, which is released in the sirtuin catalyzed deacetylation reaction. Nicotinamide converted to nicotinic acid and ammonia by nicotinamidase. The ammonia is transferred to α-ketoglutarate via glutamate dehydrogenase, yielding glutamate and the oxidation of NAD(P)H to $NAD(P)^+$, which is measured spectrophotometrically at 340 nm.

activities, all of them rely on unnatural substrates to detect the enzymatic reaction and are endpoint assays. Smith et al. reported a continuous spectrophotometric assay (Smith, Hallows et al. 2009) that measures nicotinamide formation in the sirtuin catalyzed deacetylation reaction through an enzyme-coupled system with nicotinamidase and glutamate dehydrogenase and thereby eliminating the need for a labeled substrate. In this assay format, nicotinamidase hydrolyses nicotinamide to nicotinic acid and ammonia. Glutamate dehydrogenase then converts ammonia, α-ketoglutarate, and NAD(P)H to glutamate and $NAD(P)^+$. NAD(P)H oxidation/consumption is measured spectrophotometrically at 340 nm (Figure 4.2). Results obtained with the coupled assay were shown to be comparable to those obtained with a radioactive, charcoal-binding assay (Smith, Hallows et al. 2009).

I used the sirtuin enzyme-coupled assay as a secondary assay to confirm the acetylated histone binding peptide hits from SPOT blot assay. Since only unmodified acetylated histone peptides are needed to test the activity, this method eliminates the need to synthesize various labeled hits once the SPOT blot assay results are available. In addition, there is little chance for the native histone peptides to interfere with nicotinamidase or glutamate dehydrogenase activities and thus less susceptible to false positives or negatives.

4.2 Methods

4.2.1 Cloning, expression and purification of human nuclear associated SIRTs

Human SIRT2 (1-389), SIRT6 (3-318) SIRT7 (1-367) were amplified by PCR and sub-cloned into the pET28a-LIC vector downstream of the poly-histidine coding region. Human SIRT1 (156-664) was cloned into pET28a-LIC-CHis vector, which contains a non-cleavable C-terminal poly-histidine tag.

The recombinant sirtuins were expressed in BL21(DE3)-R3-pRARE2 cells in Terrific Broth (TB) in the presence of 50 μg/mL of kanamycin at 37 °C to an OD_{600} of 0.8. Cells were then

induced by isopropyl-1-thio-D-galactopyranoside (IPTG), with a final concentration 0.5 mM, and incubated overnight at 15°C. Cells were harvested by centrifugation at 7,000 rpm. The cell pellets were frozen in liquid nitrogen and stored at -80°C.

Cell pastes were thawed and resuspended in lysis buffer (1X PBS, pH 7.5, 0.5 M NaCl, 5% glycerol) with protease inhibitor (0.1mM phenylmethyl sulfonyl fluoride, PMSF). Cell lysis was accomplished by sonication (Virtis408912, Virsonic) on ice: the sonication protocol was 10 sec pulse at 80% maximal frequency (8.0), 10 second rest, for 5 minutes total sonication time. The cell lysates were clarified by centrifugation using a Beckman JLA-16.250 rotor at 15,500 rpms for 45 minutes at 4°C.

For SIRT1 and SIRT2, the clarified lysates were loaded onto 5 mL HiTrap Chelating columns (Amersham Biosciences), charged with Ni^{2+}. The columns were washed with 10 column volumes of 20 mM Tris, pH 8.0, containing 250 mM NaCl and 5% glycerol, and the protein eluted with elution buffer (20 mM Tris buffer, pH 8.0, 250 mM NaCl, 5% glycerol and 250 mM imidazole). Proteins were then loaded onto Superdex200 column (2.6x60 cm) (Amersham Biosciences), equilibrated with 20 mM Tris buffer, pH 8.0, and 150 mM NaCl, at flow rate 4 mL/min. The proteins were further purified to homogeneity by ion-exchange chromatography on Source 30Q column (1.0x10) (Amersham Biosciences), equilibrated with buffer 20 mM Tris buffer, pH 8.0, and eluted with a linear gradient of NaCl up to 500 mM concentration (30 colume volumes). Purification yield was 10 mg of SIRT1, and 8 mg of SIRT2 per 1L of cell culture.

For SIRT6 and SIRT7, their clarified lysates were loaded onto 5 mL HiTrap chelating columns (Amersham Biosciences), charged with Ni^{2+}. The columns were washed with 10 CV of 20 mM Hepes, pH 7.0, containing 250 mM NaCl and 5% glycerol, and the proteins were eluted with elution buffer (20 mM Hepes, pH 7.0, containing 250 mM NaCl, 5% glycerol and 250 mM imidazole). The protein fractions were then loaded onto a Superdex200 column (2.6x60 cm) (Amersham Biosciences), equilibrated with 20 mM Pipes buffer, pH 6.5, and 150 mM NaCl, at flow rate 4 mL/min. The proteins were further purified to homogeneity by ion-exchange chromatography on Source 30S column (1.0x10 cm) (Amersham Biosciences), equilibrated with buffer 20 mM Mes

buffer, pH 6.5, and eluted with a linear gradient of NaCl up to 500 mM concentration (30 column volumes). The yield for SIRT6 was 3mg /L cell culture, and SIRT7 5mg/L. All proteins were concentrated to 20-30 mg/mL and stored at -80 °C until they were used.

4.2.2 Human histone SPOT Blot – peptide array synthesis

Histone peptides were synthesized on the cellulose membranes (Intavis) using an Intavis MultiPep SPOT peptide arrayer using standard (Intavis) and modified (acetyl-lysine, Bachem) L-amino acid precursors. The synthesized peptides were 8–14 amino acids in length (full list of sequences in Table S4.1) and contained either a single acetyl-lysine residue or an unmodified counterpart. Synthesis was carried out in automated manner using standard Fmoc chemistry, building a peptide from C to N terminus as previously published procedures (Jenuwein and Allis 2001; Kouzarides 2007). These synthetic peptides were immobilized on the membrane by covalent linkage between the first C-terminal residue of the peptide and ethylglycol on the modified cellulose membrane. To ensure complete reaction, a double coupling process was used for the first seven amino acids, and triple coupling for each subsequent amino acid. Once the peptide synthesis was completed, the side-chain protection groups were cleaved using the manufacturer's recommended protocol and stained with Fast Green FCF dye (Sigma) for an easy and quick view to ensure uniform peptide synthesis in each spot.

4.2.3 Human histone SPOT blot – nuclear SIRTs binding

The membranes were first washed with PBS buffer with 0.1% Tween 20, pH 7.4 (PBS/T). Then the membranes were blocked with 5% skim milk (w/v) in PBS/T overnight at 4°C to minimize the non-specific binding of the protein to the membrane. Before the binding experiment, membranes were washed again with PBS/T followed by a single wash with PBS. To start the binding experiment, the membranes were immersed in 1-2 μM purified recombinant His-tagged SIRT1, 2, 6 and 7 in PBS buffer. The protein solutions and their membranes were sealed in tight containers and allowed

to interact overnight at 4°C.

The detection of the binding followed a general Western blot procedure. The membranes were washed with PBS/T and blocked with 10% skim milk (w/v) in PBS/T for one hour. After extensive washing with PBS/T, the membranes were immersed in a solution with the primary monoclonal anti-His antibody (Qiagen) with dilution 1:1000, and 10% skim milk (w/v) in PBS/T for one hour at room temperature. After washing with PBS/T, the membranes were transferred into solutions containing antirabbit HRP fragment secondary antibody (Amersham Biosciences) with 1:5000 dilutions, and 5% skim milk (w/v) in PBS/T for another hour. After three PBS/T washes (5 min each), the western blots were developed using ECL kit (Perkin Elmer) with a 1-minute reaction time and 10-30 seconds film exposure time.

4.2.4 Sirtuins continuous assay

The sequences of the acetylated peptides used in the assay are listed in Table 4.1. All of the peptides were purchased from the Tufts Institute peptide synthesis facility. MBP-PncA (maltose binding protein fused to nicotinamidase from *Salmonella enterica*) (Garrity, Gardner et al. 2007) was expressed and purified from Escherichia coli as described previously. The MBP-PncA plasmid was a generous gift from John Denu (UW-Madison). Glutamate dehydrogenase from bovine liver and glutamate dehydrogenase from Proteus were purchased from Sigma. Enzyme concentrations were determined by measuring the OD_{280} (NanoDrop). Enzyme aliquots were stored at -20 °C until use. All other chemicals used were of the highest purity commercially available and were purchased from Sigma (St. Louis, MO), and Aldrich (Milwaukee, WI).

Nuclear sirtuin activities were continuously monitored using a Biotek Synergy 2 Plate Reader (BioTek, Winooski, VT). A typical assay mixture contained 0.1 mM of acetylated histone peptide, 0.2 mM NAD^+, 0.2 mM NAD(P)H, 1 mM DTT, 3.33 mM α-ketoglutarate, 1 μM MBP-PncA (nicotinamidase), ~3 units of glutamate dehydrogenase from bovine liver (one unit is defined by the manufacturer to reduce 1.0 μmol of α-ketoglutarate to glutamate per min), and 2 μM of the sirtuin proteins in 20 mM Na/K phosphate buffer at pH 7.0.

Table 4.1 **Sequences of peptides used in the sirtuin continuous assay.**

peptide name	Sequence
H3K9Ac	QTARKacSTGG
H3K14Ac	KQTARKSTGGKacAPRKQ
H3K18Ac	GGKAPRKacQLATKAARKSAPATG
H3K23Ac	GGKAPRKQLATKacAARKSAPATG
H3K36Ac	PATGGVKacKPHRYRP
H3K56Ac	EIRRYQKacSTELLIR
H3K115Ac	HAKacRVTIQKKD
H4K5Ac	SGRGKacGGKGLGKGGAKR
H4K12Ac	SGRGKGGKGLGKacGGAKR
H4K16Ac	YKGGAKacRHRKVLRDNIQGIT
H4K79Ac	HAKRKacTVTSLD

The reactions were carried out in a final volume of 50 μL per well in a flat-bottom clear 384-well plate. All assay components except the sirtuin proteins were pre-incubated at 25 °C for 5 min or until absorbance at the 340 nm stabilized, and the reaction was initiated by addition of the protein. The measurements were taken continuously for 5 min with 10 seconds intervals by monitoring NAD(P)H consumption at 340 nm. The background rates of reactions lacking the proteins, resulting from the spontaneous formation of nicotinamide or ammonia, were subtracted from the initial velocities of the reactions. The initial velocities were directly calculated in the units of mOD/min using Gen5 data analysis software (BioTek, Winooski, VT).

4.3 Results and Discussion

4.3.1 Purification of nuclear associated human SIRTs

Recombinant human sirtuins have been purified by various groups around the world and some of them are commercially available. However, these enzymes are generally obtained in small scale for biochemical studies and have relatively low purity. The nuclear associated sirtuins that I purified

– SIRT1, 2, 6 and 7 - were obtained with relatively high yield and purity (> 98% purity on gel).

Although only human SIRT2 (1-389) construct is the full length protein, SIRT1 (156-664), SIRT6 (3-318) and SIRT7 (1-367) constructs all contain the sirtuin domain with extensive N and C terminal flanking sequences. Similar length constructs for SIRT1 and SIRT6 have been reported to function similarly to the wild type full length proteins (Milne, Lambert et al. 2007; Tennen, Berber et al. 2010). Such data is not available for the SIRT7 construct. The truncations at the extreme N and C termini for these proteins not only increased the yield, but also made them much more stable during the purification procedure. Except for SIRT1, which contains a non-cleavable C terminal His-tag, the N-terminal His-tags of SIRT2, 6, and 7 can be cleaved by thrombin. These tags were not cleaved for the binding experiments since the recognition of binding is depended on the antibody recognition of His-tagged sirtuins on the membrane. To make binding and activity data more comparable, these His-tagged proteins were used for the continuous enzymatic assays as well.

4.3.2 Histone peptide binding preferences for nuclear sirtuins

To have a general and systematic view of the substrate preference and selectivity for each human nuclear sirtuins, I synthesized a histone peptide array containing all possible single acetyl-lysine modification of human histones – H2A, H2B, H3, H4 and H1. This peptide array design was based on the acetyl-lysine subset of peptides in a comprehensive histone modification array (Nady, Min et al. 2008). Figure 4.3 displays the developed SPOT membranes for the nuclear sirtuins. A labeled grid was superimposed onto each membrane for easy visualization of the peptide identity in each spot. Poly-histidine peptides and methylated histone marks – H3K4me, me2, me3, and H3K9me were included in the peptide array as positive and negative controls, respectively. These controls are highlighted in red. The peptide array membrane is organized into sections according to each histone. The number on the top-left corner in each cell corresponds to an acetylated lysine containing short histone peptide while no labeling indicates a non-acetylated peptide. The grid location and the peptide sequence of H3 and H4 histone binding results are summarized in Table

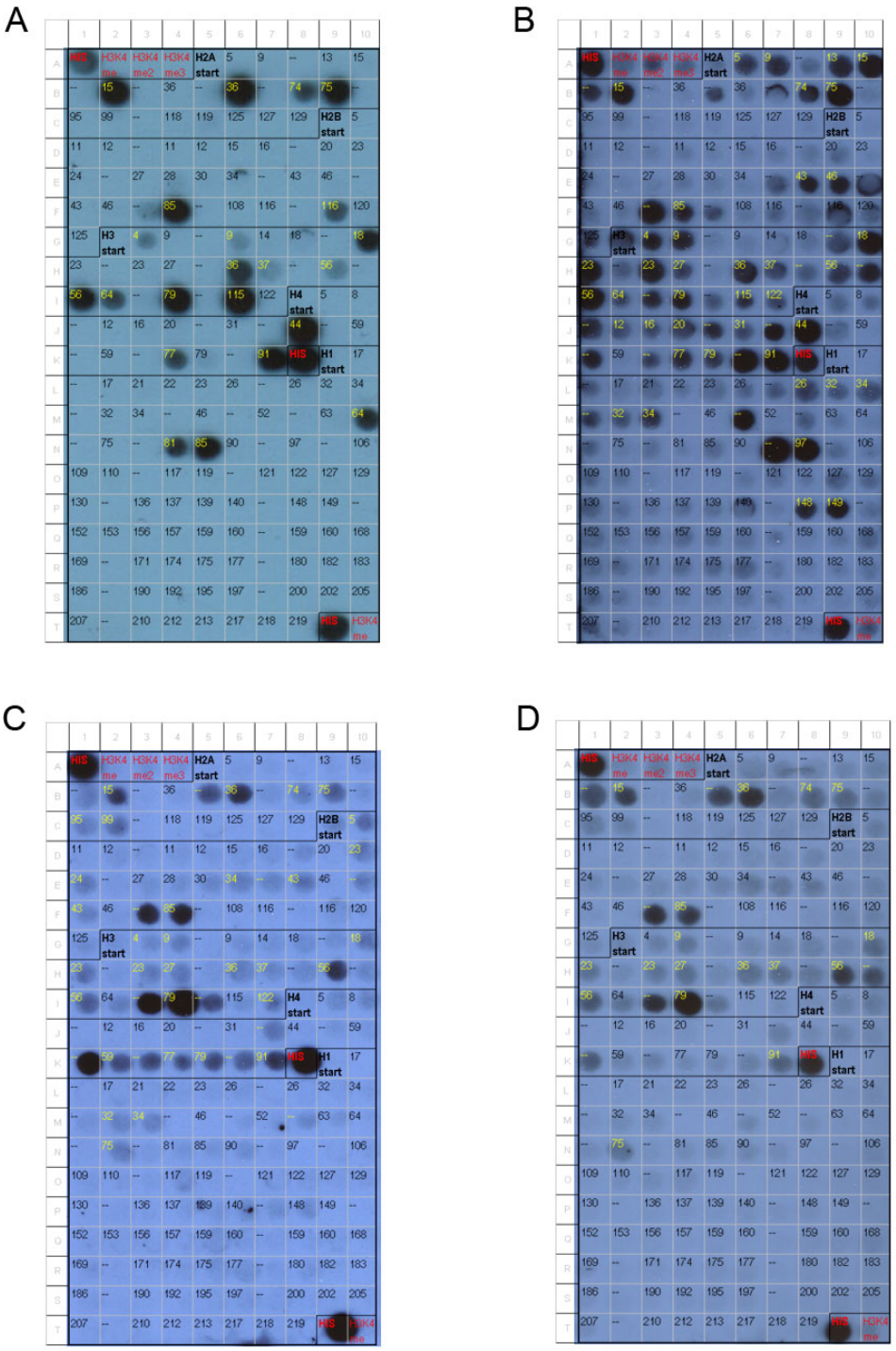


Figure 4.3 **Developed SPOT blots for nuclear associated sirtuin enzymes.** Binding profiles for A, SIRT1; B, SIRT2; C, SIRT6; D, SIRT7. Positive controls (poly-His) and negative controls (methylated peptides) are labeled in red. Highlighted in yellow are the peptides that are considered to be positive hits.

4.2. A full list of peptides sequences in the grid is provided in Table S4.1.

As shown in Figure 4.3, SIRT1 has a unique attribute in its binding profile as compared to the other nuclear sirtuins because its positive hit spots only appeared in grids of certain histone peptides with acetylated lysines while none is observed for any of the non-acetylated sequences. Within histone H2A, the hit spots for SIRT1 were located in the middle of the protein (Lys 15, 36, 74 and 75 acetylated sequences). For histone H2B, the two SIRT1 hits corresponded to the C-terminal acetyl-lysine residue (Lys85 and Lys116) containing sequences. Many of the H3 acetyl-lysine mark containing sequences in the middle and C-terminal regions of the H4 also appeared as SIRT1 hits. The association of SIRT1 with several linker histone H1 acetyl-lysine marks was also observed.

Table 4.2 **List of SPOT blot hits of acetylated H3 and H4 histone tails by SIRT1, 2, 6, and 7.**

SIRT1		SIRT2		SIRT6		SIRT7	
H3	location	H3	location	H3	location	H3	location
K4 - ac	G3	K4 - ac	G3	K9 - ac	G4	K9 - ac	G4
K9 - ac	G6	K9 - ac	G4	K9 - ac	G6	K18 - ac	G10
K18 - ac	G10	K18 - ac	G10	K23 - ac	H1	K23 - ac	H1
K36 - ac	H6	K23 - ac	H1	K27 - ac	H4	K23 - ac	H3
K37 - ac	H7	K23 - ac	H3	K56 - ac	H9	K27 - ac	H4
K56 - ac	H9	K27 - ac	H4	K56 - ac	I1	K36 - ac	H6
K56 - ac	I1	K36 - ac	H6	K79 - ac	I4	K37 - ac	H7
K64 - ac	I2	K37 - ac	H7	K122 - ac	I7	K56 - ac	H9
K79 - ac	I4	K56 - ac	H9			K56 - ac	I1
K115 - ac	I6	K56 - ac	I1			K79 - ac	I4
		K64 - ac	I2				
		K79 - ac	I4				
		K115 - ac	I6				
		K122 - ac	I7				
H4	location	H4	location	H4	location	H4	location
K44 - ac	J8	K12 - ac	J2	K59 - ac	K2	K91 - ac	K7
K77 - ac	K4	K16 - ac	J3	K77 - ac	K4		
K91 - ac	K7	K20 - ac	J4	K79 - ac	K5		
		K31 - ac	J6	K91 - ac	K7		
		K44 - ac	J8				
		K59 - ac	K2				
		K77 - ac	K4				
		K79 - ac	K5				
		K91 - ac	K7				

Although SIRT1 is largely a nuclear protein, its nucleo-cytoplasmic shuttling property

(North and Verdin 2007) allows it to modify cytoplasmic proteins in addition to the wide range of nuclear substrates including transcription factors, histones, and DNA repair proteins (Taylor, Maxwell et al. 2008). It is reasonable to speculate that SIRT1 may be able to bind protein substrates non-specifically and this leads to the large number observed in its substrate pool. In this binding assay, I observed that SIRT1 exhibited differential binding toward peptides of varying lengths containing the same acetyl lysine mark (Figure 4.3A). It recognized the grid G6 H3K9Ac (ARKAcSTGGKAPRKQL), but not the sequence in G4 grid (ARTKQTARKAcSTGG). Another example was the two H3K18Ac sequences (G8 grid no binding: ARKSTGGKAPRKAcQL and G10 grid binding: APRKAcQLATKAA). They were differentially recognized as well. For both cases, SIRT1 prefers peptides with longer C-terminal flanking sequences. The differential binding suggested SIRT1 has the ability to recognize differences in the acetyl-marker flanking sequences. Taken together with the above mentioned observation that SIRT1 only binds to sequences containing acetylated lysines, the SIRT1 SPOT membrane blot results suggested both the lysine-acetyl histone mark and its flanking sequence are key factors towards SIRT1 substrate specificity.

The positive binding hits for SIRT2 (Figure 4.3B) were located at the N-terminal half of the sequences of H2A and C-terminal half H2B core histones. Almost all of the peptides derived from the H3 and H4 histones appear to be recognized by SIRT2 including the known substrate H4K16Ac (Vaquero, Scher et al. 2006). Acetyl-lysine containing sequences from linker histone H1, such as sequences surrounding lysines 26 to 34, 97, and 148 and 149 also exhibited binding. All poly-His positive controls showed binding as expected for the four sirtuins - histone peptide binding (Figure 4.3). However, SIRT2 (Figure 4.3B) appeared to have a much higher background level and exhibited binding even towards methylated negative controls suggesting perhaps the SIRT2 binding is non-specific and/or that SIRT2 cannot distinguish different histone marks and its binding preference is solely based upon protein susbtrate amino acid sequences. Another possibility is the requirement of a SIRT2 interacting partner protein, such as the homeobox transcription factor HOXA10 (Bae, Swanson et al. 2004) or HDAC6 (North, Marshall et al. 2003; Nahhas, Dryden et al. 2007), to manifest the true substrate preference of SIRT2.

SIRT6 (Figure 4.3C) and SIRT7 (Figure 4.3D) had quite similar binding profiles and did not appear to differentiate acetylated or non-acetylated peptides. For both proteins, many hit sequences were located in the middle region of H2A. SIRT6 bound to peptide sequences containing acetyl-lysine 5, 23, 34, 43 or 85 while SIRT7 only had one hit corresponded to the lysine 85 containing sequences of H2B. In the core histone H3 region, both SIRT6 and SIRT7 appeared to associate with various sequences, with most binding to the non-acetylated and acetylated H3K79 region. Due to the semi-quantitative nature of the SPOT blot method, the intensity of a binding spot does not imply the strength of the peptide and protein binding. However, high intensity did suggest there was a relatively large amount of sirtuin binding. One hypothesis is that the H3K79 peptide (Table S4.1) is merely "sticky" due to its hydrophobicity. Indeed, acetylated H3K79 appears to be a hit for all the sirtuins tested. Within histone H4, SIRT6 appeared to preferentially interact with the C-terminal sequences while SIRT7 interacted with a subset of these C-terminal marks, including non-acetylated H4K59 and acetylated H4K91. In the linker histone H1 region, SIRT6 seemed to associate with acetylated H1K32, 34, and 75, whereas SIRT7 only associated with H1K75Ac.

It should be pointed out that the SPOT blot results presented here were prone to false hits due to the lack of quality control in peptide synthesis on the membrane. While it would be inadequate to make the conclusion of which peptide sequence the tested sirtuin bind or did not bind based solely on the SPOT blot analysis, the sirtuin binding profiles might suggest these proteins have some degree of substrate specificity. I tested this idea by subjecting individual hits to a secondary enzymatic screens.

4.3.3 Deacetylase activity towards histone peptides

To evaluate the SPOT blot binding hits, eleven purified synthetic histone peptides were used as substrates in the sirtuin continuous enzyme-coupled assays The upper detection limit of the SPOT blot screen - 100 μM of acetylated histone peptide, was chosen for all sirtuins in the presence of excess NAD^+. The sirtuin concentration used was 2 μM, which makes it comparable to the SPOT blot experiment. Figure 4.4 reports the slope of the change of absorbance curve over time (minutes

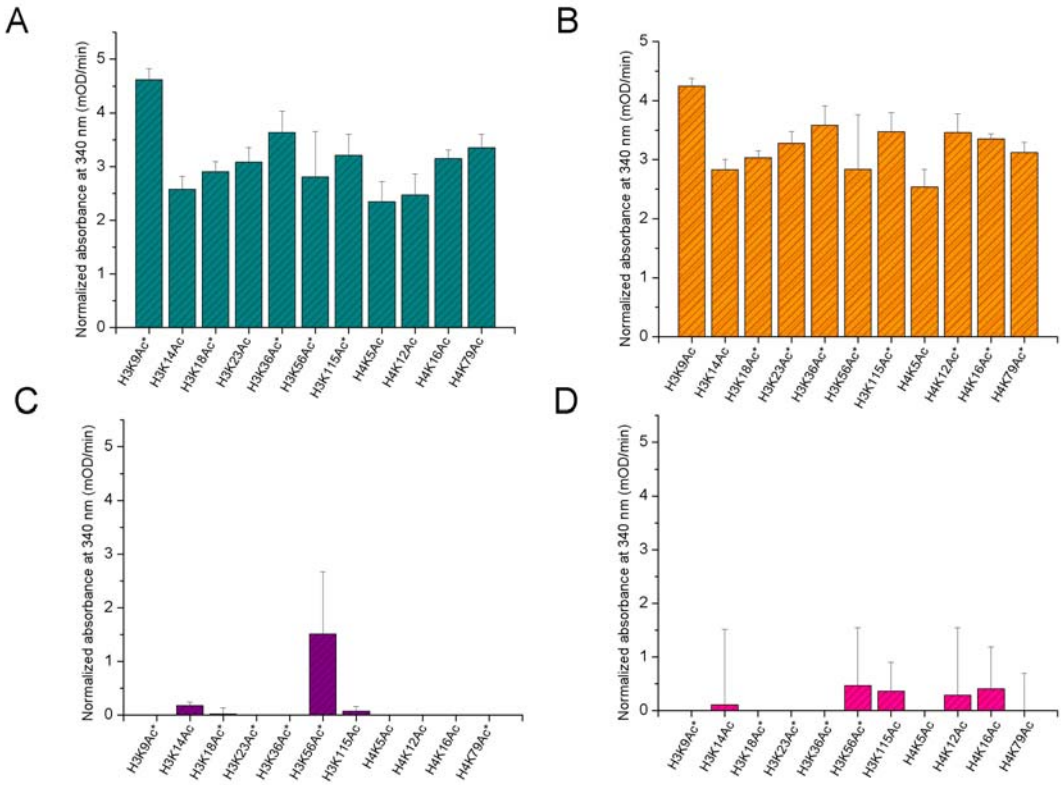


Figure 4.4 **Initial rates of sirtuin catalyzed reactions against eleven histone peptides using a continuous enzyme-couple assay.** *A*, SIRT1; *B*, SIRT2; *C*, SIRT6; *D*, SIRT7. Background rates have been subtracted from each reaction. Histone peptides labeled with a "*" indicates this peptide has been identified as a positive hit in the SPOT blot binding assay.

5 to 10 after the initiation of the experiment, 30 data points total with 10 sec interval) at 340 nm for each peptide tested.

In Figure 4.4A, SIRT1 showed some activity towards all eleven peptide tested whether they were hits or not in the first round SPOT blot assay. The same activity profile was observed for SIRT2 (Figure 4.3B). This agrees with previous studies of SIRT2, which suggested it does not appear to display a preference among all of the histone peptides tested (Borra and Denu 2004). In the deacetylase assay, SIRT1 and SIRT2 had a mild preference towards H3K9Ac (Figure 4.4A and B). This might be due to the fact that the H3K9Ac peptide used was the shortest of the eleven peptides (Table 4.1) having 4 amino acids flanking the lysine 9. SIRT6 did not deacetylate peptides tested except H3K56Ac (Figure 4.4C). The H3K56Ac histone mark was identified previously as a substrate for human SIRT6 (Michishita, McCord et al. 2009; Yang, Zwaans et al. 2009). SIRT7 did not have any deacetylase activity (Figure 4.4D), and this may be the reason that no substrates

had been reported for human SIRT7 so far. Recombinant mouse Sirt7 protein has been reported to deacetylate a commerically available p53 peptide included in the Sirt1 Fluorometric Drug Discovery kit (BIOMOL, AK-555) (Vakhrusheva, Smolka et al. 2008).

4.3.4 Comparison of binding and activity profiles

The activity profiles for each nuclear sirtuin for the selected eleven peptides from the SPOT blot data did not correspond well to their binding profiles. SIRT1 and SIRT2 belong to the Class I of the sirtuin family due to their sequence similarity. The mitochondrial SIRT3 also belongs to this class. Multiple substrates have been reported for all members of sirtuins in this class (Taylor, Maxwell et al. 2008). At the protein and peptide concentration used, SIRT1 and SIRT2 had activities on all peptides tested regardless of whether the sirtuin protein showed interaction toward these peptides in the binding assay. Preferences towards certain peptides were also observed and they appear to arise in part from the sequences flanking the acetyl-lysine mark. SIRT6 and SIRT7 constitute Class IV of the sirtuin family. Despite the rich amount of sequences that SIRT6 and SIRT7 were suggested to bind in the SPOT blot analysis, the activity of SIRT6 appeared to be very specific for H3K56Ac; and no activity for any peptides was observed for SIRT7. The discrepanies between the binding and activity data immediately suggests that protein/peptide binding is not the sole determinant of sirtuin activity.

One factor that might contribute to the discrepancies was the absence of cofactor NAD^+ in the binding assay. The rationale behind not including the NAD^+ was based on the previously reported isothermal titration calorimetry (ITC) results of many sirtuin family members (Borra, Langer et al. 2004; Milne, Lambert et al. 2007; Jin, Wei et al. 2009), suggesting peptide substrates bind prior to the binding of the cofactor NAD^+. In addition, the artificial His-tag and the N and C termini deletions for SIRT1, 6, and 7 might also affect the binding and activity profiles differently and lead to the observed discrepancies.

Another factor that might be responsible for the observed differences was the substrate peptide concentration used in the enzymatic assay. A single 100 μM peptide concentration was

used to assess the activity for all four nuclear sirtuins. Looking at a specific example: H3K9Ac and H3K56Ac are two physiological substrates for SIRT6 (Michishita, McCord et al. 2008; Michishita, McCord et al. 2009; Yang, Zwaans et al. 2009) and were both shown as positive hits in SIRT6 binding profile; I found SIRT6 had low activity towards H3K56Ac but not for H3K9Ac. This might suggest that the K_m for H3K9Ac is higher than the concentration used in my experiment (100 μM). In addition, it is possible the Class IV sirtuins (i.e. SIRT6 and SIRT7) are less active enzymes when compared to the Class I sirtuins (i.e. SIRT1 and SIRT2) when functioning alone and therefore require much higher substrate concentration to observe activity.

4.4 Conclusions

The nuclear associated sirtuins – SIRT1, 2, 6, and 7 – were subjected to an acetylated histone peptide array binding analysis. SIRT1 selectively bound many acetylated histone marks and distinguished acetylated and non-acetylated histone peptides. SIRT2, 6, and 7 did have preference for acetylated histone marks in general. The veracity of the binding data must be questioned because a secondary enzymatic screen suggested both SIRT1 and SIRT2 deacetylated all peptides tested regardless of whether they were binding hits in the primary screen or not. SIRT6 and SIRT7 had similar binding patterns, but almost all positive hits in the binding assays were not deacetylated in the enzymatic assay probably due to the fact they have inherently lower activity. These results raised the question of what governs the substrate preferences of the sirtuins enzymes and the molecular basis for the minimal activities observed for SIRT6 and SIRT7. As a special case study, the structural and biochemical properties of SIRT6, using its first identified physiological substrate - H3K9Ac, is further investigated in the next Chapter.

Chapter 5

Structure and Biochemical Functions of SIRT6

This Chapter is adapted from:

Patricia W. Pan, Jessica L. Feldman, Mark K. Devries, Aiping Dong, Aled M. Edwards, and John M. Denu 2011 Structure and Biochemical Functions of SIRT6. THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 16, pp. 14575–14581, April 22, 2011

My contributions towards this chapter:

I purified and crystallized SIRT6 proteins, solved SIRT6 complex structure with ADPr and *N*AADPr, performed structural analysis, carried out ITC experiments for SIRT1, 2, 3, 5, and 6, and contributed to data analysis and the manuscript. JLF and MKD performed the biochemical assays and ITCs for wild type and H131A mutant SIRT6 protein, and contributed to data analysis and the manuscript. AD helped with X-ray data collection and refinement. AME contributed to overall project design and the manuscript. JMD contributed to overall project design, and lead the preparation of the manuscript.

5.1 Introduction

Sirtuins comprise an ancient and diverse family of nicotinamide adenine dinucleotide (NAD^+)-dependent protein deacetylases that are evolutionarily conserved from bacteria to eukaryotes (Michan and Sinclair 2007). Unlike other classes of histone deacetylases (HDACs), which utilize an active site Zn^{2+} and involve direct attack of a water molecule on acetylated lysines, sirtuins transfer the acetyl group from the lysine side chain of a protein or peptide substrate to the co-factor NAD^+, generating nicotinamide, 2'-*O*-acetyl- ADP-ribose (*O*AADPr) (Sauve, Celic et al. 2001; Jackson and Denu 2002), and a deacetylated substrate. This unique requirement of NAD^+ suggests that sirtuins might act as sensors of the cellular metabolic state (Haigis and Guarente 2006), relaying changes in cellular metabolism to reverse acetylation-mediated pathways, which include transcription, cell cycle progression, genome maintenance, apoptosis, and organism longevity. The founding member of the sirtuin family, yeast *Sir2* (silent information regulator 2), has emerged as an important regulator in extending the life span of *Saccharomyces cerevisiae* (Kaeberlein, McVey et al. 1999). In other organisms, such as *Caenorhabditis elegans* and *Drosophila melanogaster*, sirtuin-activated pathways promote longevity (Tissenbaum and Guarente 2001; Rogina and Helfand 2004). It was recently demonstrated that mammalian SIRT3 deacetylates isocitrate dehydrogenase 2, reduces oxidative damage, and prevents age-related hearing loss in response to caloric restriction (Someya, Yu et al.). This result provides a direct mechanistic link between mammalian sirtuins and aging.

There are seven members in the human sirtuin family, SIRT1 to 7; these exhibit diversity and complexity in their cellular localization patterns and targets (Michishita, Park et al. 2005). SIRT1, being the putative ortholog to yeast *Sir2*, has been the most studied. A recent study reported that *Sirt6*-deficient mice have a striking degenerative phenotype leading to shortened life span (Mostoslavsky, Chua et al. 2006). At the cellular level, SIRT6 deficiency leads to marked metabolic and genomic instability caused by hypoglycemia and potential defects in base excision repair or DNA double strand break repair (Mostoslavsky, Chua et al. 2006). However, the molecular mechanism of the protective role of SIRT6 in preventing these processes is not fully understood.

The early *in vitro* analysis suggested that SIRT6 was not an active deacetylase and instead could undergo intramolecular mono-ADP-ribosylation utilizing NAD^+ as a substrate (Liszt, Ford et al. 2005). More recently, however, it was reported that SIRT6 does harbor NAD^+-dependent deacetylase activity toward the H3K9Ac (Michishita, McCord et al. 2008) and H3K56Ac (Michishita, McCord et al. 2009; Yang, Zwaans et al. 2009) histone marks as well as toward the double strand break resection protein, CtIP (C-terminal binding protein (CtBP)-interacting protein) (Kaidi, Weinert et al. 2010). Hundreds of genes are differentially expressed in *Sirt6* $^{-/-}$ mouse cells when compared with wild-type controls (Kawahara, Michishita et al. 2009), and the modulation of NF-κB-dependent pathways (Kawahara, Michishita et al. 2009) and HIF1α-dependent pathways (Zhong, D'Urso et al. 2010) was reported to be achieved through H3K9Ac deacetylation by SIRT6. Despite these recent advances in elucidating the biological functions of SIRT6, the lack of detailed biochemical and structural studies has hindered our mechanistic understanding of this class IV sirtuin. Three human sirtuins (SIRT2 (Finnin, Donigian et al. 2001), SIRT3 (Jin, Wei et al. 2009), and SIRT5 (Schuetz, Min et al. 2007)) have crystal structures available. Along with structures from other sirtuin homologs (Min, Landry et al. 2001; Avalos, Celic et al. 2002; Chang, Kim et al. 2002; Zhao, Chai et al. 2003; Zhao, Chai et al. 2003; Avalos, Boeke et al. 2004; Zhao, Chai et al. 2004; Zhao, Harshaw et al. 2004; Avalos, Bever et al. 2005; Cosgrove, Bever et al. 2006; Sanders, Zhao et al. 2007; Hawse, Hoff et al. 2008), these results have provided important insights into the mechanism for activity, inhibition, and substrate specificity of individual sirtuins. Crystal structures of the class IV sirtuins, which include mammalian SIRT6 and SIRT7, have not been solved.

Here we explore the biochemical function and structural details of human SIRT6 protein. We provide the first quantitative assessment of SIRT6 deacetylation and give direct evidence for *O*AADPr as one of the end products of a very inefficient SIRT6-catalyzed reaction.We also report the first sets of crystal structures of human SIRT6: a complex structure of SIRT6 with ADP-ribose (ADPr) and another complex with a synthetic *O*AADPr analog, 2'-*N*-acetyl-ADP-ribose (*N*AADPr) (Comstock and Denu 2007). Comparison of SIRT6 with other sirtuin structures reveals several unique features that provide a rationale for the low deacetylase activity and the ability

to bind NAD^+ with high affinity in the absence of acetylated substrate. Furthermore, isothermal titration binding and tryptophan fluorescence assays suggest that NAD^+ and ADPr may induce different structural changes upon binding. These novel features provide insight into the class IV sirtuins, which have poorly defined molecular functions.

5.2 Methods

5.2.1 Chemicals and reagents

NAD^+, NADH, ADP-ribose, dithiothreitol (DTT), *tris*(2-carboxyethyl)phosphine (TCEP), glutamate dehydrogenase from bovine liver, and α-ketoglutarate were purchased from Sigma. Tris-HCl, Tris base, sodium phosphate monobasic, sodium phosphate dibasic, and sodium chloride were purchased through Fisher. [^{3}H]Acetic anhydride (50–100 mCi/mmol) was purchased from American Radiolabeled Chemicals, Inc. Site directed mutagenesis kits from Stratagene were utilized to introduce a point mutation at His131 of SIRT6. 2'-*N*AADPr was synthesized as described previously (Comstock and Denu 2007).

5.2.2 Expression and purification of recombinant WT SIRT6, H131Y SIRT6, and Hst2 for deacetylation assays and isothermal titration calorimetry (ITC)

pQE-80L (Qiagen, Valencia, CA) His-tagged WT or H131Y SIRT6 were transformed into the competent *Escherichia coli* strain, BL21(DE3). Overexpression was initiated by growing cells to an OD_{600} of 0.6–0.8 at 37 °C. To induce expression, 0.5 mM isopropyl-1-thio-D-galactopyranoside was added, and cells were grown at room temperature for 18 h. Cells were harvested by centrifugation at 5,000 rpm for 10 min and stored at -20 °C. For purification, cells were resuspended in 50mM sodium phosphate, pH 7.2, 250mM NaCl, 5 mM imidazole, and 1 mM β-mercaptoethanol, lysed by sonication, and purified by nickel resin affinity chromatography. WT and H131Y SIRT6 were further purified via a HiTrap SP-Sepharose Fast Flow column (GE Healthcare) using a linear gradient

from 50 to 750mM NaCl in 50mM sodium phosphate, pH 7.2, and 1 mM β-mercaptoethanol. Fractions containing purified WT or H131Y SIRT6 were pooled, concentrated, and dialyzed into 50 mM Tris, pH 8.0 (4 °C), 150 mM NaCl, 100 μM TCEP, and 5% (w/v) glycerol. Hst2 was expressed and purified as described previously (Borra, Langer et al. 2004). Protein concentrations were determined by the Bradford reagent assay.

5.2.3 Synthesis and analysis of the H3K9 peptide

A peptide corresponding to residues 5–13 of histone H3 (QTARKSTGG) was synthesized by the University of Wisconsin-Madison Biotechnology Center and purified over a preparative C18 HPLC column. The chromatographic purity of the peptide was determined to be ≥95%, and mass spectrometric analysis on a Bruker REFLEX II: MALDI-TOF instrument confirmed the identification of the peptide.

5.2.4 Generation of a ^{3}H-labeled H3K9Ac peptide

To produce a ^{3}H-labeled H3K9Ac peptide, 4 mM H3K9 peptide was reacted with 20 mM tritium-labeled acetic anhydride in 50 mM HEPES, pH 8.0, and incubated at 4 °C for 4 h. The reaction was quenched by the addition of 100 mM Tris-HCl, pH 7.4 (final concentration). During synthesis, the N terminus was left unblocked, and therefore acetic anhydride reacted with the ε-amine of the lysine residue as well as the N-terminal amine to produce a peptide that was acetylated at both the lysine residue and the N-terminus. The presence of the doubly acetylated peptide was confirmed by mass spectrometric analysis on a Bruker REFLEX II: MALDI-TOF of a peptide acetylated with unlabeled acetic anhydride. The molar concentration of [^{3}H]H3K9Ac (3.7 mM) was determined by the amount of tritium in 2–3 μl of the resuspended peptide and calculated as described previously (Borra and Denu 2004).

5.2.5 Charcoal-binding assay determination of SIRT6 activity

The charcoal-binding assay was performed as described previously (Borra and Denu 2004).

Reactions were performed in 400 µl solution containing 600 µM NAD^+, 300 µM [^{3}H]H3K9Ac, 1 mM DTT, and 4 µM WT SIRT6 or H131Y SIRT6 in 50mM Tris, pH 7.5, at room temperature. Time points (0.5, 1, 1.5, 2, 2.5, and 3 h) were taken by adding 60 µl of the reaction mixture to 60 µl of activated charcoal, pH 9.5 (1:3 (w/v) of charcoal (Sigma, C3345) and 2:3 (v/v) of 2 M glycine, pH 9.5).

5.2.6 HPLC assay to monitor OAADPr formation

Reactions were performed in 40 µl containing 2 mM NAD^+, 300 µM [^{3}H]H3K9Ac, 1 mM DTT, and 4 µM WT SIRT6, Hst2, or no enzyme in 50 mM Tris, pH 7.5, at room temperature. The reaction was quenched after 3 h by the addition of 1% (v/v) trifluoroacetic acid (TFA) to 20 µl of the reaction mixture. The reaction mixture was separated on an analytical C18 HPLC column by a gradient separation at 0.5 ml/min: 0–8% solvent B (0.02% TFA in acetonitrile) over 2.5 min and 8–100% solvent B over 3 min, followed by 100% solvent B for 10 min. Fractions (1 min) were collected, and radioactivity was quantified by scintillation counting 150 µl of each fraction and plotted versus the fractions. The peak corresponding to [^{3}H]*O*AADPr was confirmed by alignment with purified *O*AADPr.

5.2.7 Sirtuin enzyme-coupled assay

SIRT6 activity was measured continuously using the method previously reported (Smith, Hallows et al. 2009) with minor variations. The assay solution contained 2 µM maltose - binding protein fused-PncA (nicotinamidase), 0.2 mM NADH, 0.6 mM NAD^+, 3.3 mM α-ketoglutarate, 1 mM DTT, 2 units of glutamate dehydrogenase from bovine liver, 300 µM H3K9Ac peptide, and 0 or 8 µM SIRT6 enzyme in 20 mM phosphate buffer (pH 7.5). Continuous readings were taken every 20 s for a total duration of 1 h and 30 min. The rate of nicotinamide formation was determined by the use of the coupled enzymes nicotinamidase and glutamate dehydrogenase. Nicotinamidase converts nicotinamide, formed by the cleavage of NAD^+, into nicotinic acid and ammonia. The ammonia is then utilized by glutamate dehydrogenase to convert α-ketoglutarate to glutamate, consuming

NADH in the process. The oxidation/consumption of NADH was monitored continuously at 340 nm, and subsequently the rate of nicotinamide formation could be determined.

5.2.8 Protein cloning, expression, and purification for crystallization

Human SIRT6(3–318) was amplified by PCR and subcloned into the pET28a-LIC vector downstream of the polyhistidine coding region. The recombinant protein was overexpressed in *E. coli* BL21(DE3)-R3-pRARE2 cells in Terrific Broth (TB) in the presence of 50 μg/ml kanamycin at 37 °C to an OD_{600} of 0.8. Protein expression was induced by the addition of isopropyl-1-thio-D-galactopyranoside to a final concentration 0.5 mM, and the cells were then incubated overnight at 15 °C. Cells were harvested by centrifugation. The cell pellets were frozen in liquid nitrogen and stored at -80 °C. Upon purification, the cell paste was thawed and resuspended in lysis buffer (1× PBS, pH 7.5, 0.5 M NaCl, 5% glycerol) with protease inhibitor (0.1 mM PMSF). Cell lysis was accomplished by sonication (Virtis408912, Virsonic) on ice; the sonication protocol was a 10 s pulse at 80% maximal frequency (8.0) and a 10 s rest for a 5-min total sonication time. The cell lysate was clarified by centrifugation using a Beckman JLA-16.250 rotor at 15,500 rpm for 45 min at 4 °C. The clarified lysate was loaded onto a 5 ml HiTrap chelating column (Amersham Biosciences), charged with Ni^{2+}. The column was washed with 10 column volumes of 20 mM Hepes, pH 7.0, containing 250 mM NaCl and 5% glycerol, and the protein was eluted with elution buffer (20mM Hepes buffer, pH 7.0, 250mM NaCl, 5% glycerol, and 250mM imidazole). The protein was loaded onto a Superdex 200 column (2.6 × 60 cm) (Amersham Biosciences), equilibrated with 20 mM Pipes buffer, pH 6.5, and 150mM NaCl, at a flow rate of 4 ml/min. Thrombin was added to combined fractions containing SIRT6 at 4 °C overnight to cleave off the N-terminal poly-His tag. The protein was further purified to apparent homogeneity by ion exchange chromatography on a Source 30S column (1.0 × 10 cm) (Amersham Biosciences), equilibrated with 20mM Mes buffer, pH 6.5, and eluted with a linear gradient of NaCl up to 500mM concentration (30 column volumes).

5.2.9 SIRT6 protein crystallization

The H3K9Ac peptide (TKQTARK$_{Ac}$STGGKAPY; Tufts University Core Facility, Peptide Synthesis Service) was dissolved in a buffer containing 20 mM Mes, pH 6.5, and 150 mM NaCl and then adjusted to pH 7.0 with a final concentration of 100 mM. Subtilisin A (Sigma) was dissolved in 10 mM NaOAC and 5 mM $Ca(OAC)_2$ to a final concentration of 1 mg/ml. The protein solution contained 10 mg/ml purified SIRT6 mixed with 10 mM NAD^+/ADPr (Sigma)/*N*AADPr and 3 mM H3K9Ac and 0.01 mg/ml subtilisin A (Dong, Xu et al. 2007) in 20 mM Mes buffer, pH 6.5. SIRT6·ADPr and SIRT6·*N*AADPr crystals were obtained using hanging drop vapor diffusion at 20 °C by mixing 2 µl of the protein mix with 2 µl of the reservoir solution containing 1.8 –1.9 M $(NH_4)_2SO_4$, 2% PEG 400, BisTris, pH 5.6–6.2. All crystals were soaked in the corresponding mother liquor supplemented with 15% ethylene glycol as cryoprotectant before freezing in liquid nitrogen.

5.2.10 Data collection and structure determination

X-ray diffraction data were collected at 100 K at beamline 19ID of the Advanced Photon Source at Argonne National Laboratory and a Rigaku FR-E home source. Data were processed using the HKL-2000 software suite (Minor 1997). The SIRT6 ADP-ribose structure (PDB code 3K35) was solved by molecular replacement using the program Phaser (McCoy, Grosse-Kunstleve et al. 2005). A mixed model of Sir2Tm (PDB code 1YC5) generated by SCWRL and Jackal modeling method on the FFAS03 search server (Jaroszewski, Rychlewski et al. 2005) was used as template. Refinement was carried out with Refmac (Murshudov, Vagin et al. 1997). SIRT6·ADPr (PDB code 3PKI) and SIRT6·*N*AADPr (PDB code 3PKJ) were solved by molecular replacement using the initial SIRT6 structure (PDB code 3K35; ADPr and all other small molecules were removed) as a search model. Subsequent refinements were performed using Refmac and Buster-TNT (Blanc, Roversi et al. 2004), respectively. The graphics program COOT (Emsley and Cowtan 2004) was used for model building and visualization. Data statistics are reported in Table 5.1.

5.2.11 Isothermal titration calorimetry studies of WT and H131Y SIRT6 with NAD$^+$ metabolites

ITC measurements were taken using a VP-ITC microcalorimeter from MicroCal, LLC (Northampton, MA). Binding experiments were all performed in 50mM Tris, pH 7.5 (25 °C), 150mM NaCl, 100 μM TCEP, and 5% (w/v) glycerol. Purified WT and H131Y SIRT6 were dialyzed against the previously described buffer, and ligands were suspended in the dialysis buffer. Experimental data were fitted to a one-site binding model using Origin scientific plotting software. Enthalpy (ΔH), binding constant (K_b), and number of binding sites (n) were flexible parameters, whereas the free energy (ΔG) and entropy (ΔS) were calculated according to the following equation.

$$\Delta G = \Delta H - T\Delta S = -RT\ln K_b \quad \text{(Eq. 1)}$$

For each experiment, 37 automatic injections (1–8 μl) were titrated into the cell (initial cell volume 1.42 ml) while being stirred at 300 rpm. NAD$^+$ (1.40 mM) was titrated into 27 μM WT SIRT6. NAD$^+$ (1.63 mM) was titrated into 50 μM Hst2. NADH (450 μM) was titrated into 25 μM WT SIRT6. ADPr (300 μM) was titrated into 28 μM WT SIRT6. 2'-*N*AADPr (2.96 μM) was titrated into 38 μM WT SIRT6. NAD$^+$ (767 μM) was titrated into 33 μM H131Y SIRT6, and 1.45 mM ADPr was titrated into 37 μM H131Y SIRT6. All titrations were also performed in the absence of enzyme to account for the heat caused by ligand dilution.

5.2.12 Intrinsic tryptophan fluorescence emission studies of WT and H131Y SIRT6 with NAD$^+$ metabolites

Fluorescence emission was monitored on a FluoroMax®-4 spectrofluorometer from HORIBA Jobin Yvon. Samples were excited at 295 nm (slit width 2 mm), and emission was monitored from 320 to 390nm (slit width 3 mm). Experiments under native conditions were performed in 50 mM Tris, pH 7.5, at room temperature. Increasing concentrations of NAD$^+$ from 0 to 600 μM or ADPr from 0 to 400 μM were added to 1 μM WT SIRT6, and the emission spectrum was monitored. Likewise, increasing concentrations of NAD$^+$ (0 – 200 μM) or ADPr (0 – 1.2 mM) were

added to 1 μM H131Y SIRT6. The peak maximum was determined for each concentration. To account for decreases in fluorescence from nonspecific quenching due to increasing concentration of ligand, the same experiments were carried out in 7 M urea. The peak maxima were obtained, and the correction factor was determined by (F_{obs}/F_0) where F_{obs} is the maximum intensity at each concentration, and F_0 is the maximum intensity at 0 μM ligand. The intensities from experiments under native conditions were divided by their correction factors, and the fraction of substrate bound (F_{SB}) was determined by Equation 2,

$$F_{SB} = \frac{F_{obs} - F_0}{F_s - F_0} \quad \text{(Eq. 2)}$$

where F_s represents the maximum intensity under saturating ligand, and F_{obs} and F_0 are the same as stated previously. Eq. 2 is a normalized substrate bound form to free form ratio adapted for the fluorescence emission study. The F_{SB} values were plotted *versus* ligand concentration, and the data were fitted to a one-site binding equation to determine the K_d,

$$F_{SB} = \left(\frac{K_d + P_c + L_c - \sqrt{(K_d + P_c + L_c)^2 - 4L_cP_c}}{2P_c} \right) \times C \quad \text{(Eq. 3)}$$

where K_d is the dissociation constant, P_c is the protein concentration (1 μM), L_c is the varying ligand concentration, and C is the y axis maximum. This quadratic equation assumes equilibrium binding with ligand depletion and was originally used to estimate concentration of a drug that will localize *in vivo* (Eckelman, Reba et al. 1979).

5.3 Results

5.3.1 SIRT6 deacetylase activity

It was initially reported that SIRT6 was not an active deacetylase but rather was capable of performing intramolecular auto-ADP-ribosylation using NAD^+ as a co-substrate (Liszt, Ford et al.

2005). More recently, Michishita et al. (Michishita, McCord et al. 2008) used mass spectrometry and Western blot analysis to support SIRT6 as a histone H3K9 deacetylase. However, SIRT6 did not deacetylate H3K9 with high efficiency; noncatalytic amounts of SIRT6 (~100 pmol) were needed to observe deacetylation of ~65 pmol of acetylated histone by Western blot analysis, which detects loss of immunoreactivity from an anti-acetyl lysine antibody. In mass spectrometry assays, ~100 pmol of SIRT6 was needed to convert ~10% of 360 pmol of H3K9Ac peptide to the deacetylated form (Michishita, McCord et al. 2008). It is still unknown whether SIRT6 can form 2'-*O*-acetyl-ADP-ribose as the product of the deacetylase reaction. Thus, the molecular function(s) of SIRT6 remains unclear. A detailed quantitative assessment of the SIRT6 reaction is essential toward our understanding of its cellular functions.

To monitor the deacetylase activity of purified recombinant SIRT6, we used an *in vitro* charcoal-binding assay (Borra and Denu 2004) that measures acetate released from deacetylation of a tritium-labeled H3K9Ac peptide. [^{3}H]Acetate, formed through the hydrolysis of [^{3}H]*O*AADPr, was measured in the presence of 300 μM [^{3}H]H3K9Ac, 600 μM NAD^+ and 4 μM SIRT6. After subtracting the control reaction without SIRT6, reactions containing SIRT6 exhibited a linear increase in acetate formation over time (Fig. 5.1A). The rate of deacetylation was 3.2 ± 0.2 μM/h, with a specific activity of 0.80 ± 0.05 h^{-1} (0.00022 ± 0.00001 s^{-1}). This low level of activity is in general agreement with the previous qualitative assays (Michishita, McCord et al. 2008) and indicates that SIRT6 is ~1000 times less active than the well characterized yeast Hst2 enzyme, which deacetylates the same H3K9Ac peptide with a k_{cat} of 0.32 ± 0.08 s^{-1} (Borra, Langer et al. 2004).

Such low activity raised the concern that the deacetylation activity could be due to a contaminating *E. coli* sirtuin co-purifying with SIRT6. To ensure that the activity measured was dependent on SIRT6, the invariant catalytic base histidine residue (His131) (Sauve, Celic et al. 2001; Smith and Denu 2006) was mutated to a tyrosine (H131Y). If the low level of deacetylase activity emanated from SIRT6, the H131Y mutant purified in exactly the same manner (Figure S5.1) should yield no [^{3}H]H3K9 deacetylation. H131Y SIRT6 deacetylase activity was monitored by the charcoal binding assay. No detectable activity was observed above a control reaction containing

no SIRT6 (Fig. 5.1A), thereby providing further evidence that the low level of deacetylase activity observed in the wide type (WT) reactions stemmed from catalytically active SIRT6.

To provide direct evidence that the acetate formed through the charcoal-binding assay was due to deacetylation by SIRT6 and subsequent hydrolysis of *O*AADPr, an HPLC assay was utilized to monitor the formation of [^{3}H]*O*AADPr. Reactions containing 300 μM [^{3}H]H3K9Ac, 2 mM NAD^+, and either 4 μM SIRT6, 4 μM Hst2 or no enzyme were incubated for 3 h and quenched by the addition of 1% (v/v) TFA. As a positive control for *O*AADPr formation, deacetylation of the

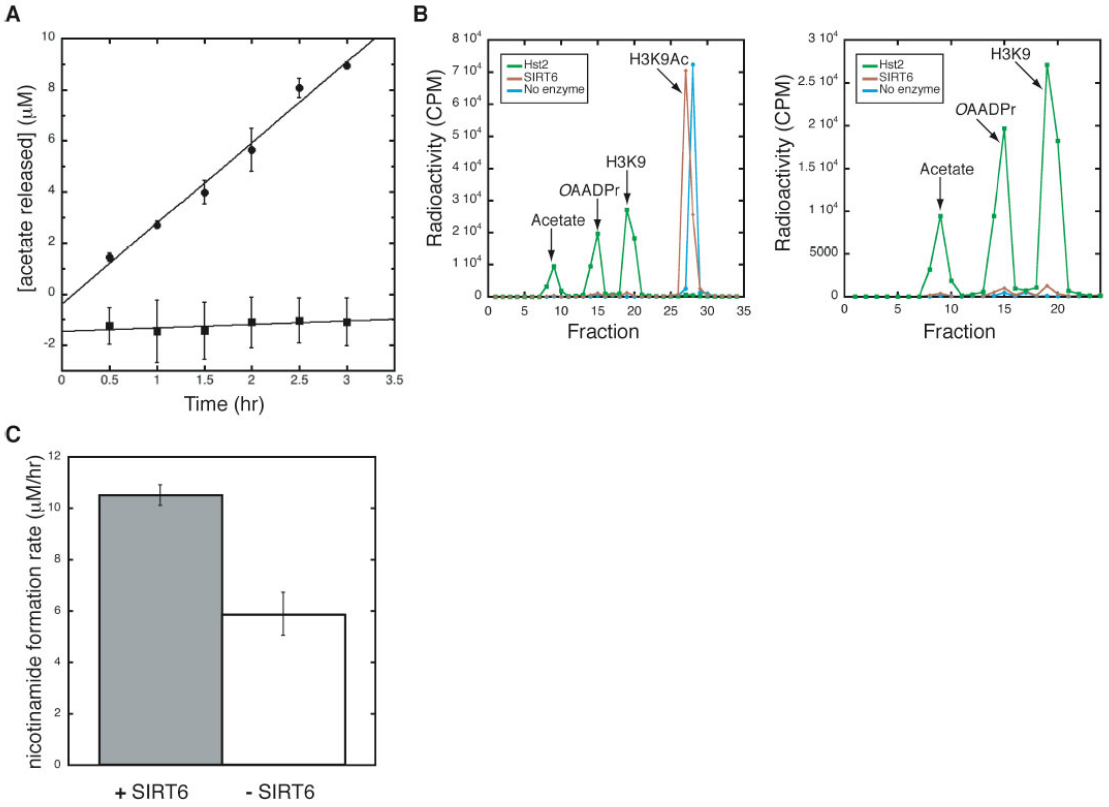


Figure 5.1. **SIRT6 deacetylation assays.** *A*, charcoal-binding assay measuring the production of OAADPr. All assays were carried out in the presence of 4 μM SIRT6 WT (●) or H131Y (■), 300 μM [^{3}H]H3K9Ac peptide, and 600 μM NAD^+. Shown is the average of three experiments for WT and H131Y. The rate of *O*AADPr production was 3.2 ± 0.2 μM/h based on linear regression analysis with an R^2 of 0.98612. *B*, HPLC separation of ^{3}H-labeled products monitoring deacetylation of a [^{3}H]H3K9Ac peptide. Reactions were carried out in the presence of 300 μM[^{3}H]H3K9Ac peptide, 2mM NAD^+, and either 4 μM SIRT6, 4 μM Hst2 or no enzyme. Shown are all counts for the Hst2 deacetylation reaction (*green*), SIRT6 (*brown*), and a no enzyme control (*blue*) as well as a zoomed in view displaying counts for [^{3}H]acetate, [^{3}H]*O*AADPr, and H3K9 ([^{3}H]acetylated N terminus). *C*, continuous assay monitoring the release of nicotinamide over time. Shown is the average of three trials with assays carried out in the presence of 8 μM (*gray*) or 0 μM (*white*) SIRT6, 300 μM H3K9Ac, and 0.6mM NAD^+. The rate of nicotinamide release in the presence of SIRT6 was 10.5 ± 0.4 μM/h and the rate in its absence was 5.9 ± 0.8 μM/h. *Error bars*, S.D.

H3K9 peptide by Hst2 (Borra, Langer et al. 2004) was monitored. Hst2 deacetylated the H3K9Ac peptide after 3 h, as demonstrated by the loss of a peak corresponding to [^{3}H]H3K9Ac peptide and generation of peaks corresponding to [^{3}H]*O*AADPr and [^{3}H]acetate, which was formed through the hydrolysis of [^{3}H]*O*AADPr (Figure 5.1B).

Compared with Hst2, SIRT6 (4 μM) was able to weakly deacetylate the [^{3}H]H3K9Ac peptide (Figure 5.1B), producing 13 μM *O*AADPr during the 3-h incubation. Thus, SIRT6 deacetylated H3K9Ac at a rate of 4.3 μM/h and had a specific activity of 1.1 h^{-1} (0.0003 s^{-1}), which is consistent with the results from the charcoal-binding assay (Figure 5.1A). Along with the charcoal binding assay, the results provide the first evidence that SIRT6 can perform multiple turnovers and can generate *O*AADPr through NAD^+-dependent protein deacetylation.

To provide an independent method for determining SIRT6 activity toward acetylated peptides, a continuous assay was utilized (Smith, Hallows et al. 2009). The continuous assay measures nicotinamide formation, which is the first product released from either the NAD^+-dependent deacetylation or an ADP-ribosylation reaction. Therefore, the assay provides a good indication of the coupling between nicotinamide cleavage and *O*AADPr formation. H3K9Ac peptide (300 μM) was incubated with 0.6 mM NAD^+ and either 0 or 8 μM SIRT6, and a linear increase of nicotinamide was observed over 1.5 h. The rate of nicotinamide formation was greater in the presence of SIRT6 compared with spontaneous NAD^+ hydrolysis (10.5 ± 0.4 μM/h versus 5.9 ± 0.8 μM/h), indicating that the SIRT6-dependent rate of NAD^+ cleavage is 4.6 ± 0.9 μM/h with a specific activity of 0.6 ± 0.1 h^{-1} (Figure 5.1C). The three different assays (Figures 5.1A–C) are in excellent agreement and confirm the extremely low level of deacetylase activity displayed by SIRT6. Importantly, the fact that the rate of nicotinamide cleavage is similar to that for *O*AADPr formation indicates that NAD^+ cleavage is tightly coupled to deacetylation and argues against a significant contribution from protein ADP-ribosylation that does not require deacetylation.

5.3.2 SIRT6 protein construct for crystallization

It is unclear why SIRT6 displays such low deacetylase activity (~1000-fold lower) compared with other sirtuins. To investigate the molecular basis for this low activity, the crystal structures

of human SIRT6 bound to ADPr and 2'-*N*AADPr were solved. Full-length human SIRT6 has 355 amino acids and consists of a putative catalytic sirtuin core with N- and C-terminal flanking extensions. In our attempts to obtain SIRT6 protein crystals, we generated several minimal truncations of SIRT6 as well as full-length protein. Among the various constructs examined, the SIRT6 (3–318) construct was expressed and purified from *E. coli* in the 10–20 milligram per liter range, compared with lower yields obtained with the full-length construct. Therefore, SIRT6 (3–318) was selected for crystallization trials. This construct has both the N-terminal extension and the core catalytic domain, but lacks the C-terminal extension. It was previously reported that a SIRT6 construct lacking the C-terminal extension is enzymatically active but is impaired in proper nuclear localization (Tennen, Berber et al. 2010).

5.3.3 In situ proteolysis and the two ADPr-bound SIRT6 structures

We initially attempted to co-crystallize SIRT6 with H3K9Ac and the cofactor NAD^+, but no crystals were observed. We then attempted to promote crystallization using *in situ* proteolysis (Dong, Xu et al. 2007). After adding trace amounts of the protease, subtilisin A, into the protein mix, protein crystals appeared within 24 h after setting up the crystallization experiment. The crystals were refined to achieve X-ray diffraction quality. In the solved structure (PDB code 3K35), we did not observe the nicotinamide moiety of NAD^+. The nicotinamide ribose ring adopted a C3'-*endo* conformation and remained as the β anomer. We repeated the crystallization experiment, substituting ADPr for NAD^+ in the protein mix, and obtained crystals within 24 h. Although the space group changed from $P2_1$ to P1, the number of molecules in the asymmetric unit remained the same, and monomer structural features were nearly identical. The only difference was that the ADPr now became the α anomer. Because both the α and β anomers of the ribose ring interconvert in aqueous solution, we do not believe that there is any significant biological differences between the two SIRT6·ADPr structures. However, at similar resolution (~2 Å), the model (PDB code 3PKI) was built on a more complete data set with higher redundancy and had lower R_{work} and R_{free} values (Table 5.1). In describing the structure of SIRT6·ADPr, we will be referring to the model/data set in which ADPr was used for co-crystallization.

Table 5.1 **Data collection and refinement statistics.**

PDB	3K35	3PKI	3PKJ
Data Collection	SIRT6+ADP-ribose	SIRT6+ADP-ribose	SIRT6+*N*-acetyl-ADP-ribose
Beamline	RIGAKU FR-E	APS 19ID	RIGAKU FR-E
Wavelength (Å)	1.54	0.97945	1.54
Space group	$P2_1$	P1	$P2_1$
Cell dimensions			
a , *b* , *c* (Å)	89.4, 136.3,89.3	77.4, 90.2, 90.8	89.0, 135.8, 89.1
α, β, γ (°)	90,0, 119.9, 90.0	118.1, 91.4, 115.8	90.0,120.0,90.0
Resolution (Å)*	20.0-2.00(2.07-2.00)	50.0-2.00(2.07-2.03)	44.5 - 2.10(2.17-2.12)
R_{sym}*	0.086(0.702)	0.120(0.965)	0.120(0.930)
I/(σI)*	18.96(2.25)	15.50(1.64)	15.90(2.02)
Completeness (%)*	97.1(94.3)	98.1(97.0)	96.2 (97.5)
Redundancy*	3.7(3.6)	4.5(4.1)	5.4 (5.4)
Refinement			
Resolution (Å)	19.95-2.00	29.9-2.04	20.5 – 2.12
No. reflections	119269	115679	99730
R_{work}†/R_{free}‡	0.202/0.267	0.182/0.213	0.237/0.270
Mean *B* -factors (Å ^{2})	30.84	33.39	37.82
R.m.s. deviations			
Bond lengths (Å)	0.013	0.010	0.010
Bond angles (°)	1.506	0.94	0.940
Ramachandran plot			
Favored regions (%)	99.0	99.03	99.09
Additionally allowed (%)	1.0	0.97	0.91
Disallowed regions (%)	0	0	0

* Values in parentheses are for the highest-resolution shell.
† R_{work} is defined as $\Sigma||F_{obs}| - |F_{calc}|| / \Sigma|F_{obs}|$, where F_{obs} and F_{calc} are observed and calculated structure-factor amplitudes, respectively.
‡ R_{free} is the R factor for the test set (1-10 % of the data).

5.3.4 Overall Structure of SIRT6 in Complex with ADP-ribose

To gain insight into the structural basis for SIRT6 enzymatic function, we initially solved a 2.0 Å crystal structure of the human SIRT6 protein in complex with ADPr. The asymmetric unit consists of six molecules, and each monomer contains one zinc atom, one molecule of ADPr, and several

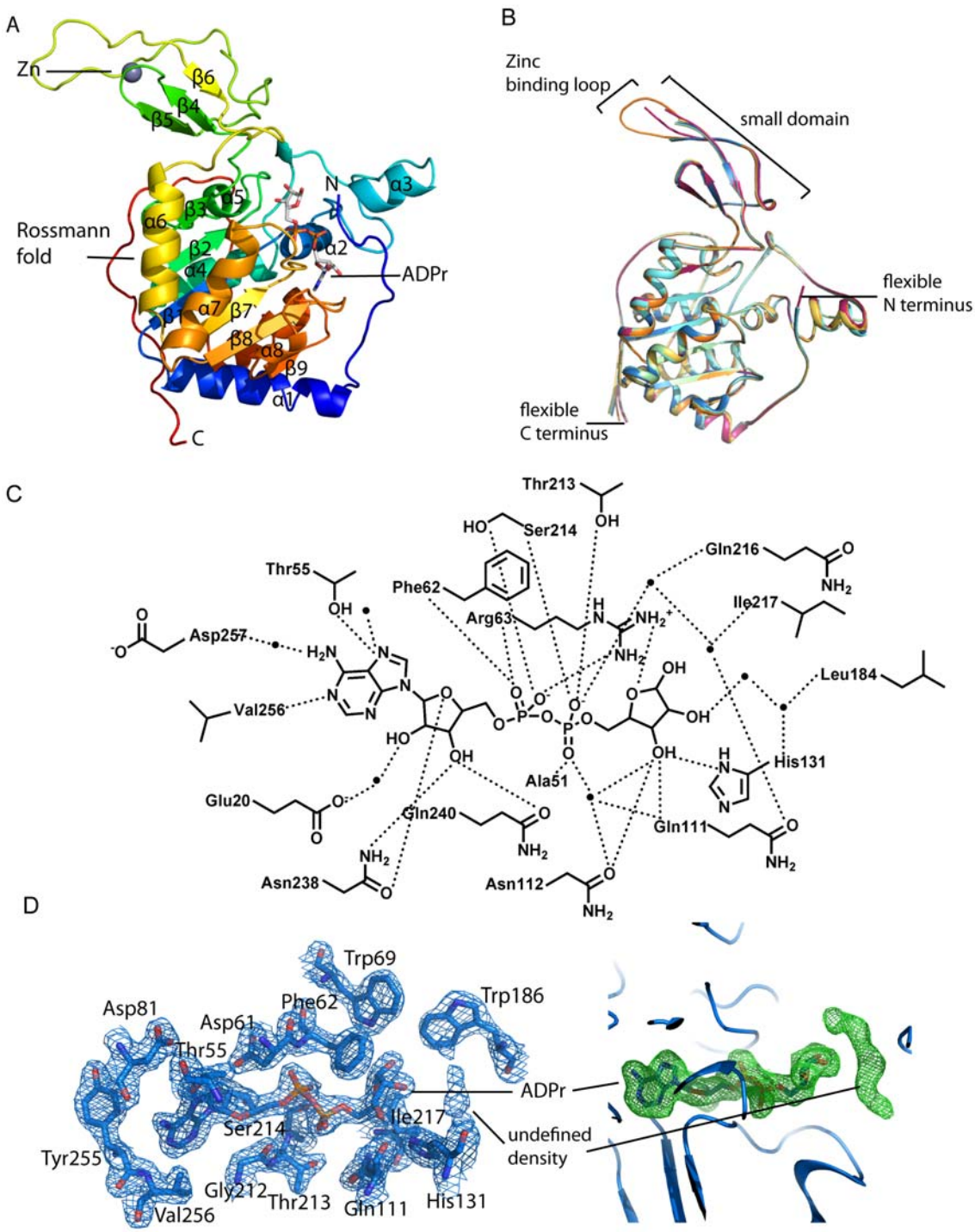


Figure 5.2. **Structure of human SIRT6 in complex with ADP-ribose.** *A*, overall structural features of SIRT6 monomer. *B*, superimposition of the six molecules in the asymmetric unit. *Red*, chain A; *green*, chain B; *dark blue*, chain C; *orange*, chain D; *cyan*, chain E; *yellow*, chain F. *C*, schematic illustration of the hydrogen bonding network surrounding ADPr; hydrogen bonds are indicated as dashed lines, and water molecules are shown as spheres. *D*, left, SIRT6·ADPr $2F_o - F_c$ electron density map (*blue mesh*, 1.5σ) of the residues within 4 Å of ADPr. *Right*, $F_o - F_c$ omit electron density map (*green mesh*, 2σ) of the ADPr molecule; the putative peptide binding site contains an unidentified electron density.

sulfate molecules, which were present in the crystallization buffer. Residues 11–295 of SIRT6 (3–318) were visible in the crystal structure. It is possible that the N-terminal (positions 3–10) and C-terminal (positions 295–318) residues were cleaved by trace amounts of the subtilisin that was added to promote crystallization (Dong, Xu et al. 2007), or alternatively these regions are disordered. Both the N- and C-terminal residues visible in the crystal structure form long unstructured coils, providing support that the unseen residues of the SIRT6 constructs are highly flexible and disordered.

Consistent with other solved sirtuin structures, SIRT6 contains two globular domains composed of eight α-helices and nine β-strands: a large Rossmann fold for NAD^+ binding (residues 25–128 and 191–266) and a smaller domain, which contains a zinc-binding motif (residues 129–190) (Figure 5.2A). The large Rossmann fold domain is formed by a six-stranded (β1, β2, β3, β7, β8, and β9) parallel β-sheet sandwiched between two helices (α6 and α7) on one side and four helices (α1, α4, α5, and α8) on the other side. The small domain is formed by two extending loops (connecting β3 and α6) from the large domain and consists of a three-stranded antiparallel β-sheet (β4, β5, and β6). Interestingly, although most sirtuins contain a Cys-*X*-*X*-Cys-$X_{15–20}$-Cys-*X*-*X*-Cys sequence motif for Zn^{2+}-binding, SIRT6 contains a 10-residue insertion between the second set of cysteines, resulting in an extended long loop (Figure S5.2). The six molecules in the asymmetric unit appear to be almost identical (Figure 5.2B), with variations in the extended loop of the zinc-binding motif as well as in the N and C termini. The extended loop is highly flexible, and only one of the six molecules in the asymmetric unit has an ordered extended loop.

The NAD^+ binding pocket of sirtuins has been divided into three regions: sites A, B, and C (Min, Landry et al. 2001). Similar to all other solved sirtuin structures, the adenosine moiety of ADPr is bound in the A site of SIRT6, and the nicotinamide ribose moiety is bound in the B site. Many of the residues involved in ADPr binding are conserved between SIRT6 and the other sirtuins (Figure S5.2). The ADPr interactions in the SIRT6 active site are shown in Figure 5.2C. For all six molecules in the asymmetric unit, the ribose moiety of ADPr assumes the C2' *exo* conformation in its α anomer form (Figure 5.2D).

Although acetylated H3K9Ac peptide was included during co-crystallization, we did not

observe the peptide in the solved SIRT6 structure. However, almost perpendicular to the ADPr electron density, we observed elongated density (Figure 5.2D). Given that this location is where peptide binding is predicted (Jin, Wei et al. 2009), we attempted to model a lysine side chain or a PEG molecule into this density. However, we were unable to unambiguously fit the density, and therefore we left this region unassigned.

5.3.5 Structural comparison of SIRT6·ADP-ribose with SIRT6·2'-N-Acetyl-ADP-ribose

As we have demonstrated in Figure 5.1, SIRT6 forms 2'-*O*AADPr as one product of the reaction, and therefore to gain insight into the interactions between 2'-*O*AADPr and SIRT6, we co-crystallized

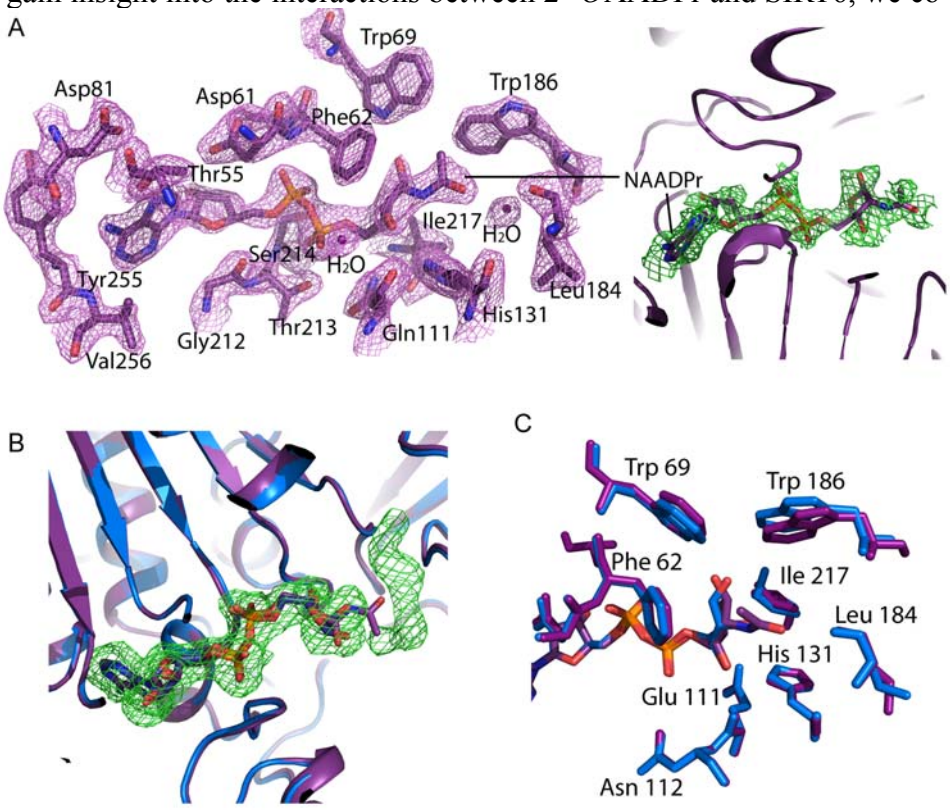


Figure 5.3. **Comparison of SIRT6·ADP-ribose and SIRT6·2'-*N*-acetyl-ADP ribose structures.** *A*, *left*, SIRT6·*N*AADPr $2F_o$ - F_c electron density map (*purple mesh*, 1.5σ) of the residues within 4 Å of *N*AADPr. *Right*, F_o - F_c omit electron density map (*green mesh*, 1σ) of the *N*AADPr molecule. *B*, superimposition of SIRT6·ADPr (*blue*) onto SIRT6·NAADPr (*purple*); the unknown density (F_o - F_c omit electron density map contoured at 2σ, *green mesh*) observed in the SIRT6·ADPr structure is not present in the SIRT6·NAADPr structure, and the *N*-acetyl group occupies part of the unidentified density space. *C*, residues immediately surrounding the *N*-acetyl group of *N*AADPr (*purple*) and their corresponding residues (*blue*) in the SIRT6·ADPr structure.

SIRT6 with 2'-*N*AADPr (Figure 5.3A), a non-hydrolyzable analog of *O*AADPr in which the *O*-acetyl moiety is substituted with an N-acetyl group (Comstock and Denu 2007). The overall structural features are identical between the SIRT6·ADPr and SIRT6·*N*AADPr crystal structures (Figure 5.3B). This is not surprising because *N*AADPr only contains an extra N-acetyl group compared with ADPr. There are no significant changes in the orientation of residues surrounding the NAD^+ binding site (Figure 5.3C). It is interesting to note that the unknown density observed in the SIRT6·ADPr structure was not present in the SIRT6·*N*AADPr structure (Figure 5.3B). The *N*-acetyl group protrudes into the space that is otherwise occupied by the unassigned chemical entity in the SIRT6·ADPr structure.

5.3.6 Structural comparison with other solved human sirtuins (SIRT2, SIRT3, and SIRT5)

Sirtuins have been phylogenetically divided into five subclasses: I–IV and U (Frye 1999; Frye 2000; Tanner, Landry et al. 2000). Among the seven human sirtuins, SIRT1 to 3 belong to class I; SIRT4 and 5 belong to class II and III, respectively; and SIRT6 and 7 are both class IV sirtuins. The three previously characterized human sirtuins provide structural insight into classes I and III. The SIRT6 structure constitutes the first solved sirtuin structure in class IV. Structural features revealed by SIRT6 provide additional information on how the different evolutionary lines developed in the sirtuin family.

Although SIRT6 shares the overall domain architecture with SIRT2 (Finnin, Donigian et al. 2001), SIRT3 (Jin, Wei et al. 2009), and SIRT5 (Schuetz, Min et al. 2007), there are several differences on the surface of the protein. In all previously solved sirtuin structures, a conserved "cofactor binding loop" (Figure S5.2) (Sanders, Jackson et al. 2010) is involved in NAD^+ binding; this loop adopts several different conformations depending on the ligand(s) bound in the active site. A small region of the cofactor binding loop is disordered in the ADPr-bound SIRT5 structure, but in the suramin·SIRT5 complex structure, this loop becomes ordered and even forms β strands and a small helix, which directly interact with suramin, an anti-parasitic drug that also binds sirtuins with

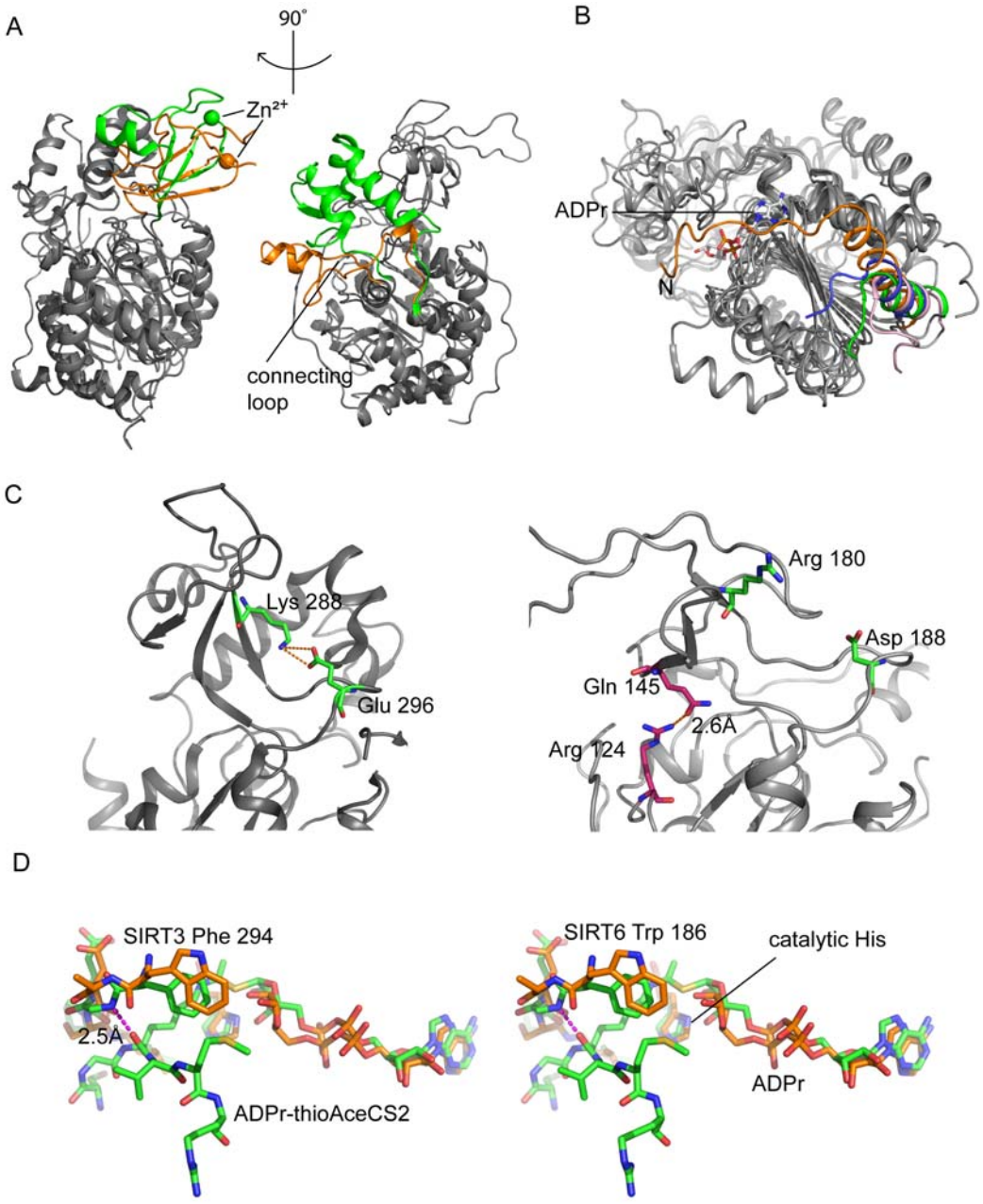


Figure 5.4. **SIRT6·ADPr structure compared with other solved human SIRTs.** *A*, SIRT6 structure (PDB code 3PKI, *orange*) superimposed onto SIRT3 (PDB code 3GLR, *green*). The left panel highlights differences in the zinc-binding domain, and the right panel shows a missing helix bundle. *B*, comparison of the N-terminal loop of SIRT6 (3PKI, *orange*), SIRT2 (1J8F, *pink*), SIRT3 (3GLR, *green*), and SIRT5 (2B4Y, *blue*); SIRT6 contains a long loop covering the NAD^+ binding site. *C*, SIRT6 lacks the conserved salt bridge. *Left*, SIRT3 (3GLR) structure with conventional salt bridge in the sirtuin family. *Right*, the *green pair* (Arg180 and Asp188) is where the bridge would be predicted to form on the SIRT6 structure according to other sirtuin crystal structures. *Pink pair* (Arg124 and Gln145), the actual hydrogen bonding pair found in SIRT6 structure. *D*, stereo diagram of flipped FGE*X*L (WEDSL) loop of SIRT6 (PDB code 3PKI, *orange*) complexed with ADPr (*orange*) compared with SIRT3 (PDB code 3GLT, *green*) with an ADPr·thioAceCS2 (*green*) complex.

micromolar affinity (Schuetz, Min et al. 2007). The changes in the cofactor-binding loop are also observed in the set of SIRT3 structures (Jin, Wei et al. 2009). SIRT6 lacks the cofactor-binding loop, which is replaced by a single helix (α3) containing several NAD^+ binding residues (Figure 5.4A). Unlike the cofactor-binding loop in the ADPr-bound SIRT5 structure and the SIRT3 apo structure, helix α3 appears to be ordered in the SIRT6·ADPr and SIRT6·*N*AADPr structures. This is an indication that the SIRT6 substrate-binding pocket may be less flexible and does not significantly vary its conformation when different ligands are bound. In addition, helix α1 in SIRT6 is much longer, and the N-terminal unstructured coil folds back toward and may structurally stabilize the NAD^+ binding site (Figure 5.4B), although there is no direct interaction between the random coil and ADPr.

In contrast to class I, II, and III human sirtuins, class IV sirtuins contain a deletion in the sequence immediately following the α3 helix (Figure S5.2). The result is that SIRT6 does not have a helix bundle in its small domain (Figure 5.4A). The helix bundle is replaced by a short loop, which interacts with the loop between α2 and α3 and contacts a small region on the zinc-binding module. The lack of a helix bundle to form extensive interactions with the β-sheets in the zinc-binding module provides one explanation for why SIRT6 has a splayed small domain. This unique structural feature is likely to be adopted by SIRT7, another class IV sirtuin that also contains a deletion in this region (Figure S5.2) but that has not yet been structurally and functionally characterized. This unique feature may be responsible for the observed lower activity (~1000 times) of SIRT6. To date, there are no reports of detectable deacetylase activity for human recombinant SIRT7.

In addition to the missing helical bundle, another reason for the splayed small domain in SIRT6 is the loss of a conserved salt bridge, which has been reported to contribute to the positioning of the zinc-binding motif with respect to the Rossmann fold domain and substrate-binding site (Min, Landry et al. 2001; Schuetz, Min et al. 2007). Based on precedent from other sirtuins, a salt bridge should be formed between Arg180 and Asp188 (Figure 5.4C) in SIRT6. However, this interaction is not observed. The side chain of Arg180 is directed toward the surface of the protein rather than toward the Asp188 residue. Instead, Arg124 on the loop connecting α5 and β3 and Gln145 on β5 form a hydrogen bond to tether the zinc-binding motif and the large domain together

(Figure 5.4C). The loss of the previously described salt bridge removes the constraint on the two loops, which would otherwise result in a more closed conformation. The new Arg124–Gln145 hydrogen bond in SIRT6 pulls the two domains together on the opposite side, and therefore, the small domain is not directly above the Rossmann fold but rather in an open conformation tilted in a 45° angle.

It is generally observed that peptide substrates interact with sirtuins via nonspecific backbone interactions (Jin, Wei et al. 2009). There is a conserved FGE*X*L loop that forms hydrogen bonds with the substrate peptide (Avalos, Celic et al. 2002). However, in the case of SIRT6, the sequence is WEDSL (Figure S5.2). The amide nitrogen of Trp186 appears to be flipped, whereas its analog in other sirtuins is involved in a hydrogen bond with a carbonyl group from the peptide substrate backbone. Upon superimposition of the SIRT6 structure with the SIRT3-acetylated peptide structure (PDB code 3GLT) (Jin, Wei et al. 2009), we observed that the backbone amide between the tryptophan and glutamate side chains is further away and cannot form a hydrogen bond with the carbonyl in the amide backbone of the lysine (Figure 5.4D). This poses an issue for the peptide-protein anti-parallel β-sheet formation. It is possible that a β-sheet configuration might still form, but a kink near the active site would adversely affect the interactions between the peptide substrate and the protein. Therefore, a less stable substrate-protein interaction might be the cause of the poor deacetylase activity for SIRT6 and the reason why the H3K9Ac peptide does not exhibit binding saturation by ITC (Figure S5.3).

Sirtuins catalyze deacetylation through a sequential mechanism in which the substrate peptide binds first to form an ordered NAD^+-binding pocket for the subsequent co-substrate binding. SIRT1, SIRT2, and SIRT3 were all found to require a substrate/ligand binding prior to NAD^+ binding (Borra, Langer et al. 2004; Milne, Lambert et al. 2007; Jin, Wei et al. 2009). The unique SIRT6 structural features suggest that SIRT6 utilizes a different binding mechanism, which does not require peptide substrate binding prior to NAD^+ binding. These features include an ordered helix regardless of substrate binding, a large open channel in the substrate and NAD^+ binding site due to the missing helix bundle, and a NAD^+ site potentially reinforced by the N-terminal loop. To test this hypothesis, we performed a series of equilibrium binding experiments with SIRT6.

5.3.7 NAD⁺ metabolite binding

Isothermal titration calorimetry was used to monitor the binding of NAD^+ in the absence of an acetyl lysine substrate. Consistent with previous kinetic results, Hst2, SIRT1 to 3, and SIRT5 did not bind NAD^+ in the absence of the acetyl lysine substrate (Figures 5.5A and S5.4). Strikingly, NAD^+ bound to SIRT6 with a K_d of 27 ± 1 μM (Figure 5.5B), providing the first physical evidence that SIRT6, unlike other characterized sirtuins, can bind NAD^+ efficiently in the absence of an acetylated substrate. We detected no binding when NADH was titrated into SIRT6 (Figure S5.5A), suggesting that SIRT6 cannot accommodate the puckered conformation of the nicotinamide ring of NADH. Interestingly, ADPr bound to SIRT6 with a K_d of 4.7 ± 0.5 μM (Figures S5.5B and Figure 5.5B), suggesting that the binding pocket of SIRT6 distinguishes structural differences

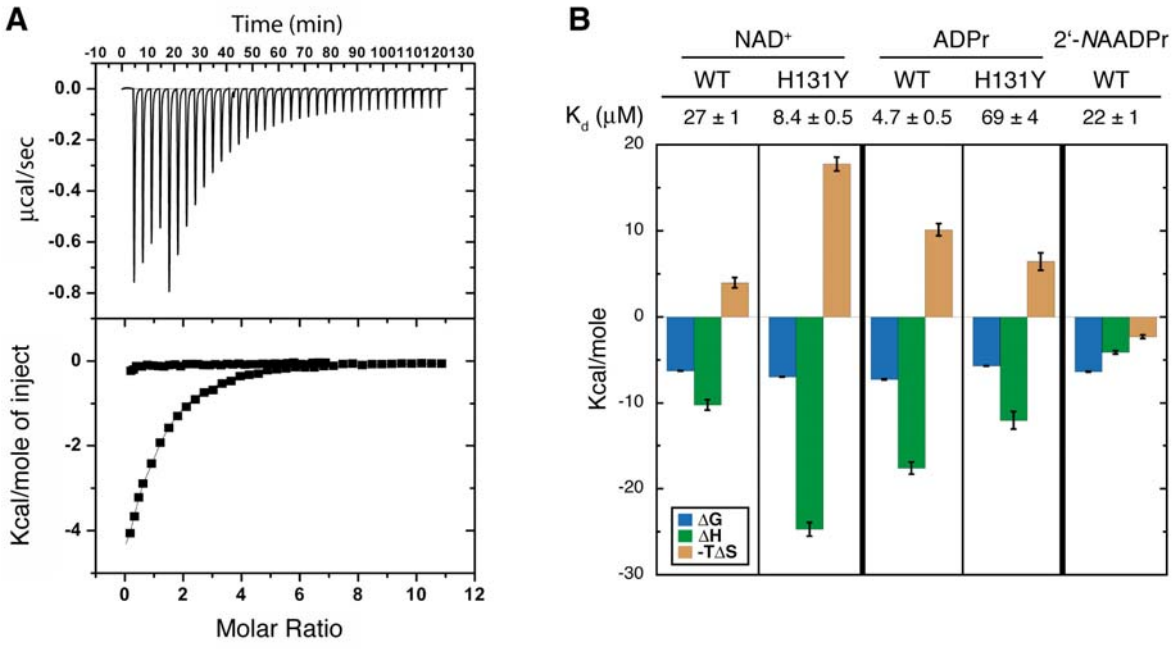


Figure 5.5. **Isothermal titration calorimetry studies.** *A*, representative ITC trace. The top graph shows data obtained for 37 automatic injections (1–8 μl) of 1.4 mM NAD^+ titrated into 27 μM SIRT6. The bottom graph represents integrated curves of the experimentally generated heats. The data were fitted to a one-site binding curve (s*olid line*). NAD^+ (1.40 or 1.63mM) was titrated into 27 μM WT SIRT6 (■) or 50 μM Hst2 (●), respectively. *B*, dissociation constants and column graph displaying thermodynamic parameters for WT and H131Y SIRT6 binding to NAD^+, ADPr, and 2'-*N*AADPr. NAD^+ (1.40 mM or 767 μM) was titrated into 27 μM or 33 μM WT or H131Y SIRT6, respectively. ADPr (390 μM or 1.45mM) was titrated into 28 or 37 μM WT or H131Y SIRT6, respectively. 2'-*N*AADPr (2.96mM) was titrated into 38 μM WT SIRT6. The change in Gibbs free energy (ΔG) is shown in *black*. Change in enthalpy (ΔH) is shown is *gray*, and $-T\Delta S$ is displayed in *light gray*. *Error bars*, S.D.

between NAD^+ and ADPr. The construct (residues 3–318) used for SIRT6 crystallization was able to bind NAD^+ and ADPr with similar affinities compared with the full-length SIRT6 construct (Figure S5.6), suggesting that there is negligible variability between the full-length construct used for kinetic assays and the crystallized constructs. Consistent with the crystal structure, the non-hydrolyzable product analog 2'-*N*AADPr bound to SIRT6 with a K_d of 22 ± 2 μM (Figures S5.5C and Figure 5.5B). Interestingly, unlike NAD^+ and ADPr, 2'-*N*AADPr bound with a negative $-T\Delta S$ (Figures 5.5A and B), suggesting that there is more disorder when 2'-*N*AADPr binds. This result is consistent with the crystal structure, which showed a high degree of variability and poor electron density for the 2'-*N*-acetyl-ribose moiety of 2'-*N*AADPr.

The invariant catalytic base histidine residue of Hst2 is important for both NAD^+ binding and activation of the 2'-hydroxyl of the α-1'-*O*-alkylamidate intermediate during the deacetylation reaction (Smith and Denu 2006). To investigate the role of SIRT6 His131 in NAD^+ and ADPr binding, ITC was used to monitor the effects of the H131Y substitution on the ability of SIRT6 to bind NAD^+ metabolites. The substitution to tyrosine led to a dramatic effect in the ability of SIRT6 to bind ADPr. The K_d of 69 ± 4 μM was ~15 times higher than the K_d for WT SIRT6 (Figures S5.5D and Figure 5.5B). The enthalpy increased 6 ± 1 kcal/mol, consistent with the loss of hydrogen bonds to the 2'- and 3'-hydroxyl groups of the ribose, observed in the crystal structure of SIRT6 bound to ADPr (Figure 5.2C). Interestingly, the H131Y mutant bound NAD^+ with a K_d of 8.4 ± 0.5 μM, or ~3 times lower than the K_d for WT SIRT6 (Figures S5.5E and Figure 5.5B). This is a 47-fold change in specificity for NAD^+ over ADPr from WT SIRT6 to H131Y SIRT6. The enthalpy of binding NAD^+ decreased 15 ± 1 kcal/mol, whereas $-T\Delta S$ became less favorable as it increased 14 ± 1 kcal/mol from WT to H131Y (Figure 5.5B). These changes probably reflect increased hydrogen bonding and van der Waal's contacts between the nicotinamide moiety of NAD^+ and H131Y of SIRT6, causing the ΔH to become more negative, and subsequent ordering of both NAD^+ and SIRT6, leading in turn to the increase in the entropy noted in the $-T\Delta S$ value.

As an independent method for analyzing NAD^+ metabolite binding, we monitored changes in tryptophan fluorescence as a function of ligand binding. SIRT6 contains two tryptophan residues near the active site (Trp69 and Trp186, Figure 5.3C). We observed a decrease in the tryptophan

fluorescence signal of SIRT6 with increasing concentrations of NAD^+ (Figure S5.7A), suggesting a change in the local environment surrounding the tryptophan residues upon NAD^+ binding. The fluorescence intensities were adjusted for nonspecific quenching of tryptophan residues located outside of the active site and the data were fitted to a one-site binding curve (Equation 3) (Eckelman, Reba et al. 1979). The K_d of 24 ± 7 μM is in excellent agreement with the K_d determined by ITC (K_d of 27 μM). However, at increasing concentrations of ADPr, no change in fluorescence was observed (Figure S5.7B), despite the fact that by ITC, ADPr binds to SIRT6 with a K_d of 4.7 μM. The results suggest that there are different structural rearrangements in the active site, depending on the NAD^+ metabolite.

In agreement with the ITC results, a decrease in the fluorescence of H131Y SIRT6 was observed at increasing concentrations of NAD^+, yielding a K_d of 3.5 ± 0.6 μM for NAD^+ (Figure S5.7C). Interestingly, an *increase* in fluorescence was observed with H131Y SIRT6 upon binding ADPr (Figure S5.7D), suggesting that the environment surrounding the tryptophan residues is more non-polar when ADPr binds. The K_d of 51 ± 8 μM is in agreement with the previous ITC results (K_d of 69 μM; Figure 5.5B). These results further highlight the conclusion that NAD^+ and ADPr induce different structural perturbations on SIRT6 upon binding.

5.4 Discussion

Several physiological and cellular investigations have suggested that SIRT6 plays a role in genome maintenance and metabolic regulation (Mostoslavsky, Chua et al. 2006; Kanfi, Shalman et al. 2008; Michishita, McCord et al. 2008; Kawahara, Michishita et al. 2009; McCord, Michishita et al. 2009; Kaidi, Weinert et al. 2010; Kim, Xiao et al. 2010; Xiao, Kim et al. 2010; Zhong, D'Urso et al. 2010). Although these studies have provided insight into the biological role of SIRT6, the molecular functions of SIRT6 remain unclear. Here we have reported the first quantitative assessment of SIRT6 activity, providing direct evidence that *O*AADPr is formed as a product of an extremely inefficient deacetylase reaction. Unlike other sirtuins, SIRT6 displays tight binding

of NAD^+ in the absence of an acetylated substrate. Furthermore, tryptophan fluorescence results suggest that NAD^+ and ADPr induce different structural arrangements upon binding.We have also solved the first crystal structures of SIRT6, a member of the class IV sirtuins. The crystal structures provide insight into the biochemical functions presented in this study and reveal several significant differences between SIRT6 and the other solved sirtuin structures.

SIRT6 was reported to act as an H3K9 deacetylase (Michishita, McCord et al. 2008). Here, we utilized three quantitative assays to establish that SIRT6 forms *O*AADPr, but at a specific activity of ~0.0002 s^{-1} toward H3K9Ac, which is ~1,000 times slower than other highly active sirtuins (Borra and Denu 2004). Furthermore, the rate of nicotinamide cleavage is similar to the rate of *O*AADPr formation, suggesting that SIRT6 functions as a NAD^+-dependent deacetylase rather than as an ADP-ribosyltransferase. It is unlikely that the low deacetylase activity is caused by inhibition of the undefined molecule observed in the acetyl lysine-binding site of the SIRT6·ADPr structure. The SIRT6·*N*AADPr structure shows no evidence of this molecule, suggesting that it was either readily displaced by *N*AADPr or was not present during crystallization. It is important to note that the acetyl group of *N*AADPr occupied the same space as that of the undefined molecule in the SIRT6·ADPr structure. If this molecule were responsible for the low activity of SIRT6, we would expect to observe major differences in binding affinity among *N*AADPr, NAD^+, and ADPr. The fact that *N*AADPr and NAD^+ displayed nearly identical binding affinity suggests that this molecule, if present, did not impede *N*AADPr binding and therefore would not significantly contribute to inhibiting SIRT6 activity. In addition, the results are consistent with the previous analyses, which demonstrated that SIRT6 is not an efficient H3K9 deacetylase (Michishita, McCord et al. 2008). The fact that SIRT6 has been implicated in important processes, including several metabolic pathways (Kanfi, Peshti et al. 2008; Kim, Xiao et al. 2010; Xiao, Kim et al. 2010; Zhong, D'Urso et al. 2010), double strand break repair (McCord, Michishita et al. 2009; Kaidi, Weinert et al. 2010), base excision repair (Mostoslavsky, Chua et al. 2006), and telomere maintenance (Michishita, McCord et al. 2008; Michishita, McCord et al. 2009), raises the question of why a seemingly important sirtuin would display such low deacetylase activity.

The newly solved crystal structures of SIRT6 in complex with ADPr and *N*AADPr provide

some insight. SIRT6 shares an overall similar architecture with the previously solved sirtuin structures and contains the conserved catalytic residues. However, the SIRT6 structure revealed many unique features. SIRT6 exists in a more open conformation, containing a zinc-binding motif in the small domain, which is splayed from the larger Rossmann fold domain. Several structural features contribute to the displacement of the small domain relative to its position on other sirtuins, including the lack of a conserved salt bridge between the zinc-binding motif and the substrate-binding loop. This conserved interaction was previously demonstrated to be important for maintaining the proper orientation of the small domain with respect to the large domain and is important for forming the acetylated substrate-binding pocket (Min, Landry et al. 2001; Schuetz, Min et al. 2007). In SIRT6, however, a new hydrogen bond connects the zinc-binding motif to the lower portion of the Rossmann fold, stabilizing SIRT6 in an open conformation. Furthermore, the class IV sirtuins, which include SIRT6, lack a conserved helix bundle region that forms contacts between the small domain and the Rossmann fold domain. The consequence of this feature contributes to the open conformation observed in the SIRT6 crystal structure. The open conformation might disrupt the structural integrity of the substrate-binding pocket, providing one explanation for why H3K9Ac does not exhibit binding site saturation, which in turn leads to poor deacetylase efficiency.

Instead of containing a highly conserved cofactor binding loop (Sanders, Jackson et al. 2010) that aids in NAD^+ binding, SIRT6 contains a single helix that forms interactions with ADPr and *N*AADPr. In several sirtuin structures, the orientation of the loop and the ordering of the residues depend on the substrate(s) bound in the active site (Avalos, Boeke et al. 2004; Zhao, Harshaw et al. 2004). The single helix in SIRT6, however, appears to be ordered in both structures, indicating that the substrate binding pocket is less flexible and does not widely vary its conformation. This led us to hypothesize that, unlike all other studied sirtuins, SIRT6 would be able to bind NAD^+ in the absence of an acetylated substrate.

We found that SIRT6 can bind NAD^+ with a relatively high affinity (K_d = 27 μM) in the absence of an acetylated substrate. Furthermore, we determined that SIRT6 can bind both ADPr (4.7 μM) and *N*AADPr (22 μM) with relatively similar affinities compared with NAD^+. However,

the fact that a change in tryptophan fluorescence was observed when the concentration of NAD^+ was increased and not when ADPr was increased suggests that there may be differing structural changes near the tryptophan residues when the ligands bind; this provides evidence that SIRT6 is able to distinguish between the ligands that are bound in the active site. To investigate NAD^+ binding more closely, we mutated the catalytic histidine residue (His131), which was previously reported to be important not only for catalysis but also for NAD^+ binding (Smith and Denu 2006). We found that H131Y SIRT6 can still bind NAD^+ (8.4 μM) but has a decreased ability to bind ADPr (69 μM). In addition, although a decrease in the tryptophan fluorescence signal was seen with increasing NAD^+, an increase in fluorescence was seen with increasing ADPr. Together, these results further support the conclusion drawn for WT SIRT6, that there is an inherent structural difference near the tryptophan residues in SIRT6 between NAD^+ and ADPr binding. Therefore, this structural difference may play an important role in controlling SIRT6 activity.

The fact the NAD^+ binds to SIRT6 in the absence of an acetyl lysine substrate is a unique feature of this sirtuin. Also, the low basal activity and the unique structural features suggest that regulatory mechanisms may be needed to activate SIRT6 in cells. The open structure of SIRT6 might provide a surface for regulatory protein binding and stimulation of activity. Similar mechanisms have been demonstrated for other histone-modifying proteins, including the yeast acetyltransferase RTT109, which requires the histone chaperones Asf1 (Tsubota, Berndsen et al. 2007) and Vps75 (Berndsen, Tsubota et al. 2008) for efficient activity. SIRT6 interacts with several transcription factors, including NF-κB (Kawahara, Michishita et al. 2009) and Hif1α (Zhong, D'Urso et al. 2010); these interactions might stabilize an active conformation of SIRT6 and subsequently stimulate histone deacetylation at the site of action. It is also possible that SIRT6 is indeed a robust enzyme, but we and others have not supplied the appropriate protein substrate. Alternatively, the deacetylase activity of SIRT6 might not be regulated by activation, but instead this intrinsically low deacetylase activity might suggest that SIRT6 has a different molecular function.

Interestingly, SIRT6 has been implicated in many of the same pathways as SIRT1, including the targeting of H3K9 for deacetylation (Vaquero, Scher et al. 2004). SIRT1 is a highly active deacetylase, begging the question of why purified SIRT6 would also deacetylate H3K9 but very

inefficiently. The ability of SIRT6 to bind NAD^+ in the absence of an acetylated substrate is unique, and the possibility that SIRT6 may adopt different structural arrangements depending on which ligand is bound suggests that an alternative hypothesis for SIRT6 molecular function is plausible: that SIRT6 has evolved away from a functioning NAD^+-dependent deacetylase and toward a function as a NAD^+ metabolite sensor. There are many examples of enzymes that have evolved away from their catalytic role and toward a regulatory function that takes advantage of the ability to bind ligand. These include signaling proteins that belong to the large family of protein kinases (Anamika, Abhinandan et al. 2009) and phosphatases (Pils and Schultz 2004).

SIRT6 controls flux into glycolysis by inhibiting Hif1α-dependent transcription of several glycolytic genes to promote mitochondrial respiration (Zhong, D'Urso et al. 2010). In addition, SIRT6 has also been reported to negatively regulate PPARγ-mediated transcription to prevent triglyceride formation (Kanfi, Peshti et al. 2010). These are two pathways controlled by cellular NAD^+ levels and furthermore, they are also pathways in which SIRT1 has been reported to play an important role (Picard, Kurtev et al. 2004; Rodgers, Lerin et al. 2005). SIRT6 is found at the promoters of several metabolic genes involved in these pathways, and its interaction with transcription factors suggests that NAD^+ levels might control SIRT6 function. Presumably, under high levels of NAD^+, SIRT6 could bind to transcription factors at promoters of metabolic genes preventing triglyceride formation and flux into glycolysis. Alternatively, below K_d levels of NAD^+, the interaction would be destabilized, allowing transcription factors to access the targeted genes. Therefore, SIRT6 might ensure metabolic homeostasis and promote more efficient energy usage through its ability to sense the levels of NAD^+ (and related metabolites) and regulate protein-protein interactions with transcription factors. Although the concentration of NAD^+ in mammalian nucleus has been estimated to be in the range of 90 μM (Zhang, Piston et al. 2002), which is about 2 times above the K_d value determined for SIRT6, the true physiological dynamic range of NAD^+ concentration still remains to be determined. In addition, the SIRT6 K_d value for NAD^+ is determined *in vitro*, which does not taken into account protein partners that might affect the ability of SIRT6 to bind to NAD^+. Future *in vivo* studies will be needed to elucide these speculations and further explore the possibility that SIRT6 can function as a NAD^+ cellular sensor.

Chapter 6

Conclusions and Future Perspectives

6.1 Principal findings and contributions

A major challenge in chemical biology and drug discovery is to understand the structural basis for protein specificity and promiscuity of homologous enzymes within the same family. The main objective of this thesis is to provide structural insights that relate protein local sequences to their observed binding and activity profiles through the study of two human protein families – SULTs and SIRTs.

6.1.1 Major observations and findings of the human SULT family

As a Phase II detoxification enzyme family, hSULTs are known to exhibit broad and overlapping substrate specificities. The knowledge of the structural basis for their observed promiscuities is particularly desirable in order to predict the fate of xenobiotics, hormones and drug candidates in humans.

In Chapter 2, I presented a family-wide chemical profile of the hSULT enzyme family through binding and activity assays followed by a structural profile which enabled us to compare structural features that determine how enzyme interacts with certain type of molecules. While observing the known substrate preferences for well-studied SULTs, we have identified a number of novel compounds that bind to the less well-characterized SULT1C3 and SULT4A1. We also identified three additional broad-spectrum hSULT inhibitors: PLP, AMP-PNP, and quercetin. Through the structural profiling, I revealed a structural role for the cofactor, PAP(S) for priming of

the flexible substrate binding loops, which are responsible for the wide repertoire of compounds that can be accommodated in the substrate binding pocket. I also provided evidence for a novel SULTs inhibition mechanism by solving a ternary complex SULT1C2·PAP·PCP structure. PCP and possibly other polychlorinated phenolic compounds can bind hSULTs in a catalytically competent conformation. However, they remain unreactive due to their weak acidities.

In Chapter 3, I presented two detailed structural case studies: SULT1B1·PAP·resveratrol and SULT2A1·PAP·lithocholic acid ternary complex structures, based on the observed substrate preferences provided in the previous family-wide chemical profiling of the human SULTs family. The hydrogen-bonding coordination of the 5-OH group is the structural determinant for the observed substrate preference towards resveratrol (and possibly other polyphenols contain similar OH moieties) by the SULT1 family. The lithocholic acid complexed SULT2A1 structure captures a more realistic and complete picture of the catalytic site with both cofactor product PAP and a substrate bind at their corresponding binding sties at the same time. This structure also confirms that the specificity of SULT2A1 for lithocholic acid and steroids-related compounds lies in its high hydrophobicity in the substrate binding pocket.

6.1.2 Major observations and findings of the nuclear associated sirtuins

The varied localization and substrates of the human sirtuins place them at the center of important cellular pathways including histone remodeling, gene silencing, regulation of transcription factors, cell cycle regulation, and lifespan extension. There has been much interest in identifying and characterizing small molecules that modify the ability of sirtuins to deacetylate substrate proteins. Understanding the molecular details of how each sirtuin interacts with their physiological substrates will aid in developing small molecule regulators which in turn holds promise for drug discovery efforts in multiple therapeutic areas.

In chapter 4, I presented a systematic human acetylated histone peptide array binding profile for the nuclear associated sirtuins – SIRT1, 2, 6, and 7. SIRT1 can selectively bind to many acetylated histone marks while SIRT2, 6, and 7 can bind to both acetylated and non-acetylated

histone sequences suggesting the specificity of SIRT1 lies not only in the sequence but also the presence of the acetyl mark. On the other hand, the bindings of SIRT2, 6 and 7 are predominantly relying on the peptide sequence. The secondary enzymatic screen suggests both SIRT1 and SIRT2 deacetylate all peptides tested regardless of whether they were primary screen binding hits. The class IV sirtuins, namely SIRT6 and SIRT7 have similar binding patterns, but almost all positive hits in the binding assays were not deacetylated in the enzymatic assay.

The structural and molecular basis governing observed substrate preferences for Class I and Class IV sirtuins were discussed in Chapter 5 where we specifically discuss SIRT6. Using three different assays, I presented biochemical and kinetic evidence that SIRT6-dependent histone deacetylation produces *O*AADPr but at a rate ~1,000 times slower than other highly active sirtuins. I solved the first set of crystal structures of this class IV Sirtuin in complex with ADPr and the non-hydrolyzable analog of *O*AADPr – *N*AADPr to understand the molecular basis for such low deacetylase activity. The structures revealed human SIRT6 has a splayed zinc-binding domain, lacks a helix bundle and the conserved, highly flexible, NAD^+-binding loop. Using ITC, we also demonstrated SIRT6, unlike all other studied sirtuins, is capable of binding NAD^+ (K_d = 27μM) in the absence of an acetylated substrate.

6.2 Future perspectives

I have shown that the hSULT family promiscuity derives from the considerable flexibility or plasticity of the hSULT binding sites. Although a full understanding of specificity will require multiple three dimensional structures for each hSULT in complex with substrates and inhibitors, as well as knowledge of the full spectrum of small molecules that bind in both productive and nonproductive conformations, the work I have presented in this thesis has progressed our understanding of the human SULTs substrate promiscuity and selectivity. Along with many other reported biological characterizations of SULTs, we now have a framework for the design of selective specific small molecule modulators to regulate the sulfontation pathway. Some of the interactions of substrate/

inhibitor and the active site residues may be easily optimized to improve probe potency. Some may require more complicated derivatization or even re-engineering. In addition, we can perform both virtual screening and chemical database mining in search for available analogs based on our chemical profiling to act as more specific and potent modulators. If these molecules have low toxicity, high potency and specificity, they can be rapidly adapted to perturb a relevant sulfonation biological function in cells and eventually to advance our understanding in the *in vivo* sulfonation process in human.

The high specificity and low activity of SIRT6 (possibly SIRT7 as well) observed in our *in vitro* biochemical experiments certainly raises the question if this is indeed the case *in vivo*. Several SIRT6 interacting partners have been reported including NF-κB (Kawahara, Michishita et al. 2009), GCIP (Ma, Stafford et al. 2007), CtIP (Kaidi, Weinert et al. 2010). The ligand profiling could be reperformed in the presence of these proposed partners. If the deaceylation activity is increased, it would be interesting to obtain sructures of the protein-protein complex to understand what interactions cause the activation of SIRT6. Although SIRT7 is the closest related sirtuin member to SIRT6 based on sequence, its deacetylase activity has very limited characterization. The binding and activity assays presented in my thesis could be expanded to include more potential physiological substrates which might be able to give some insights in the biological functions of SIRT7.

Many sirtuin activators and inhibitors have been identified in chemical screens. However, our structural understanding of how these molecules exert their function on the sirtuins is lagging. So far, only SIRT3 has crystal structures solved with its physiological substrates (Jin, Wei et al. 2009) and only SIRT5 has an X-ray structure with inhibitor bound (Schuetz, Min et al. 2007). The *N*AADPr complexed SIRT6 structure presented in Chapter 5 suggested a possible mechanism of inhibition, which can might be exploited when designing inhibitors. To further our knowledge in this area, complex structures with activators and inhibitors for each sirtuin family members are therefore needed.

6.3 Concluding Remarks

Though my investigation of the two human protein families, SULTs (a family with high sequence identity members), and SIRTs (a family with low sequence identity members), I have shed some light on the structural basis for their observed ligand specificities. The ligand profiling approach utilizes simple, medium throughput binding and activity screens and can be used to profile properties of purified enzymes. While each screen method used has its pros and cons, when used in combination, they are useful tools to enable identification of novel substrates and to assist in the analysis of enzyme specificity, and can be easily adapted to become high throughput screens for large scale ligand libraries.

I have demonstrated that the relationship between sequence/structure and function within the SULT family is remarkably complex and differences in activity can reflect just a few amino acid changes at critical locations within the active site of the protein. For the sirtuins family, I have presented some structural evidences that the preference and specificity probably lies in the different structural features rather than amino acid level. The data presented here form a basis for further detailed biochemical and structural studies for both protein families. By studying members of the entire protein family simultaneously, it is hoped the characterization results will lead to the discovery of new protein functions and open doors for engineering specific modulators that would ultimately be used for therapeutic intervention.

Appendix

Supporting Figures and Tables

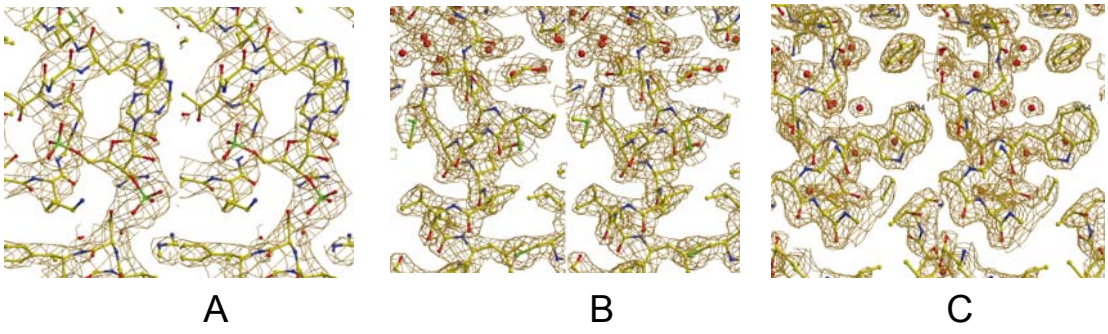


Figure S2.1. **Stereo images of representative model-phased electron density.** *A*, coordinates of SULT1C3 (PDB ID 2H8K); *B*, SULT4A1 (PDB ID 1ZD1); and *C*, SULT1C2 (PDB ID 2AD1).

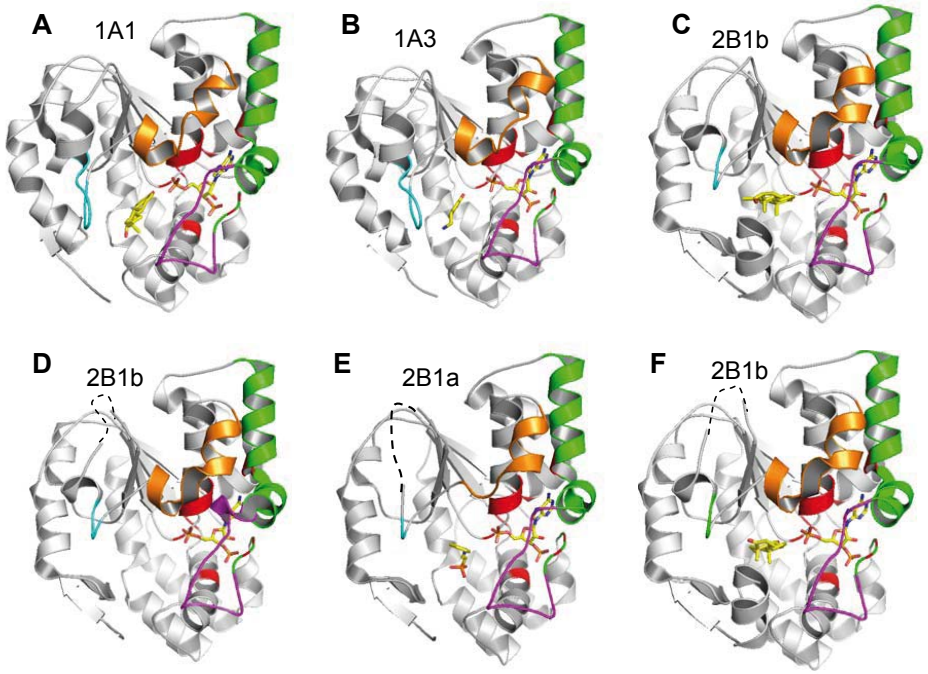


Figure S2.2**. Additional structures of human SULTs from the PDB.** *A*, SULT1A1 with PAP and estradiol (2D06); *B*, SULT1A3 with PAP and dopamine (2A3R); *C*, SULT2B1b with PAP and pregenolone (1Q20); *D*, SULT2B1b with PAP (1Q1Z); *E*, SULT2B1a with PAP and 2-[N-cyclohexylamino] ethane sulfonic acid (1Q1Q); and *F*, SULT2B1b with PAP and DHEA (1Q22). For more details see Figure 2.2 legend in thesis.

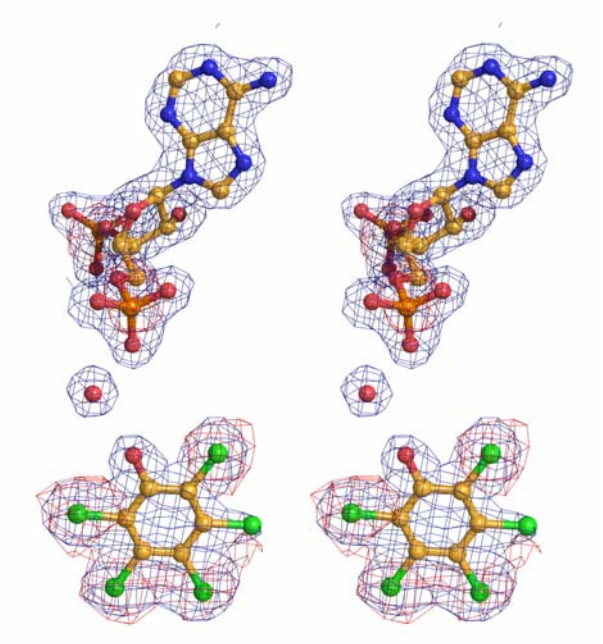


Figure S2.3. **Stereo image of model-phased $2F_o$- F_c (1.8 Å resolution, 1.4σ, blue) and anomalous difference Fourier (2.5Å, 3σ, red) maps for the SULT1C2·PAP·PCP complex (PDB ID 2GWH).**

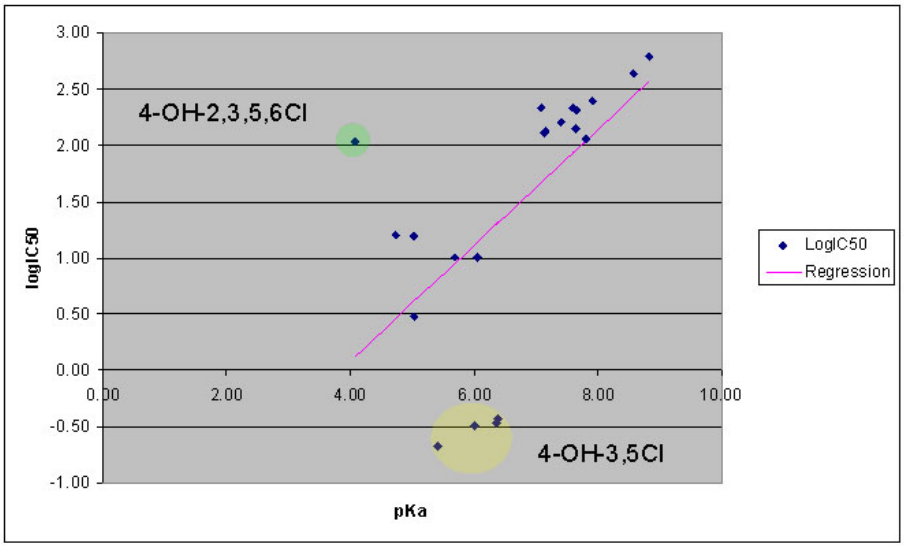


Figure S2.4. **Correlation between computed pK_a values and published IC_{50} values of hydroxylated polychlorinated biphenyls on hSULT1E1.** Compounds **2-22** of Kester et. al (Kester, Bulduk et al. 2000) comprise a series of 4-OH substituted polychlorinated biphenyls and are plotted above. All data points were included in the regression analysis. The "outliers" from the regression line correspond to **4-OH-2,3,5,6,**2',4',5'-heptachlorobiphenyl (highlighted in green), and **4-OH-3,5**,3',4'-tetrachlorobiphenyl, **4-OH-3,5**,3',5'-tetrachlorobiphenyl, **4-OH-3,5,**2',3',4'-pentachlorobiphenyl, **4-OH-3,5,**3',4',5'-pentachlorobiphenyl (highlighted in yellow). The latter series (yellow) are particularly potent inhibitory compounds and also have a linear relationship between LogIC_{50} vs computed pK_a within this structural series.

Table S2.1. **Data collection and refinement statistics.** Reflections were indexed, integrated and scaled with the HKL2000 suite (Otwinowski and Minor 1997). Structures were solved by molecular replacement using MOLREP (Vagin and Teplyakov 1997) and PHASER (McCoy, Grosse-Kunstleve et al. 2005). Except for SULT1C3, an initial model was build automatically by ARP/wARP (Perrakis, Morris et al. 1999). Restrained coordinate and temperature factor refinement in CNS (Brunger, Adams et al. 1998) and REFMAC5 (Murshudow, Vagin et al. 1997) and manual rebuilding in O (Jones, Zou et al. 1991) or COOT (Emsley and Cowtan 2004) were iterated for several cycles.

	SULT1C2	SULT1C3-PAP	SULT4A1	SULT1C2-PAP-PCP
PDB code	2AD1	2H8K	1ZD1	2GWH
Data collection				
X-ray source	Rigaku FR-E	APS 17ID	Rigaku FR-E	Rigaku FR-E
Wavelength (Å)	1.5418	1.0000	1.5418	1.5418
Space group	$P3_221$	$C222_1$	C2	I422
Cell dimensions				
a, *b*, *c* (Å)	58.82, 58.82, 189.29	87.16, 108.72, 136.92	98.94, 74.20, 85.83	135.22, 135.22, 158.02
α, β, γ (°)	90, 90, 120	90, 90, 90	90, 105.52, 90	90, 90, 90
Resolution (Å)*	50.00-2.00 (2.07-2.00)	100-3.00 (3.11-3.00)	50.00-2.24 (2.28-2.24)	30.00-1.80 (1.85-1.80)
R_{sym} or R_{merge}	0.059 (0.767)	0.149 (0.651)	0.087 (0.600)	0.111 (0.000)
$I / \sigma I$	14.6 (2.9)	10.5 (1.1)	9.2 (2.6)	26.4 (1.9)
Completeness (%)	93.7 (72.5)	98.4 (87.9)	99.1 (97.6)	99.9 (99.5)
Redundancy	10.7 (9.5)	5.7 (4.0)	3.8 (3.6)	14.0 (10.8)
Molecular replacement				
Search model PDB code	1ZHE (Dombrovski, Dong et al. 2006)	2AD1	1CJM (Bidwell, McManus et al. 1999)	2AD1
Refinement				
Resolution (Å)	50.00-2.00	30.00-3.20	50.00-2.24	30.00-1.80
No. reflections	23766	10487	27265	65377
R_{work} / R_{free}	0.218/0.262	0.292/0.343	0.186/0.262	0.191/0.219
No. atoms	2204	3616	4326	5131
Protein	2051	3562	4098	4726
Ligand/ion	N/A	54	6	90
Water	153	N/A	222	315
B-factors (Å^2)	42.75	65.88	37.44	25.24
Protein	42.49	65.97	37.42	25.01
Ligand/ion	N/A	60.13	58.34	26.30
Water	46.16	N/A	37.18	28.60
R.m.s deviations				
Bond lengths (Å)	0.021	0.016	0.025	0.016
Bond angles (°)	1.7	1.4	2.0	1.3
Ramachandran plot				
% residues(2)				
Favored	90.6	82.8	91.7	91.3
Additional allowed	8.9	15.8	6.2	8.7
Generously allowed	0.0	1.4	1.8	0.0
Disallowed	0.5	0.0	0.2	0.0

*Highest resolution shell is shown in parenthesis.

Table S2.2. **List of 90 compounds in SULT compound library.**

Compound
mefenamic acid
salbutamol
R-(-)-apomorphine hydrochloride hemihydrate
salicylic acid
quercetin dihydrate
resveratrol
acetaminophen (paracetamol)
isoprenaline hydrochloride
vanillin
tyramine
p-cresol
4-aminophenol
(±)-epinephrine
pentachlorophenol
2,6-dinitrotoluene
2,4-dinitrotoluene
α-zearalenol
2-hydroxyestradiol
4-ethylphenol
lithocholic acid
(±)-chlorpheniramine maleate salt
4-nitrophenol
clomiphene citrate salt
adenosine-3',5'-bisphosphate
N-(iodoacetyl-aminoethyl)-5-N'-naphthylamin-1-sulfonic acid (1,5-IAEDANS)
N-bromosuccinimide
p-chloromercuribenzoate
2,6-dichloro-4-nitrophenol
3-hydroxyindole (1H-indol-3-ol dihydrogen phosphate ester, disodium salt)
dopamine (hydrochloride)
minoxidil
(+)-(S)-1,2,3,4-tetrahydro-1-naphthol
1-naphthol
2,3-butanedione
2-ethylphenol
2-*n*-propylphenol
2-*sec*-butylphenol
4-*n*-amylphenol (4-pentylphenol)
4-*n*-heptylphenol
3,5-dibromo-4-hydroxy-benzoic acid (6,8-dichloro-4-oxo-4H-chromen-3-ylmethylene)-hydrazide
3,5-dibromo-4-hydroxy-benzoic acid (6-chloro-4-oxo-4H-chromen-3-ylmethylene)-hydrazide
N-acetylglucosamine 6-sulfate sodium salt, mixed anomers
sucrose
D-(+) glucose
2-aminofluorene
17α-ethynylestradiol

dexamethasone
estrone
thienylbutyl isothiocyanate
thienyldecyl isothiocyanate
thienyldodecyl isothiocyanate
γ-aminobutyric acid - GABA
acetylcholine chloride
(-)-norepinephrine
histamine dihydrochloride
cytidine 5'-diphosphocholine
pyridoxal 5'-phosphate
thiamine monophosphate
niacin - vitamin B3
argininosuccinic acid disodium salt
DL-glyceraldehyde
L-ornithine monohydrochloride
L-citrulline
coenzyme B12
ADP - Na
CDP - Na
GDP - Na
UDP - Na
ATP - diNa
CTP - Na
GTP - Na
ITP - triNa
UTP - Na
AMP-PNP (adenosine 5′-(β,γ-imido)triphosphate)
GMP-PNP (5'-guanosyl-imido-triphosphate)
guanosine
T2 (3,5-diiodo-L-thyronine)
T3 (3,3′,5-triiodo-L-thyronine sodium salt)
T4 (3,3′,5,5'-tetraiodo-L-thyronine) (L-thyroxine)
dehydroisoandrosterone 3-sulfate sodium salt
2-naphthol
dATP
dGTP
dCTP
dTTP
Rp-cAMPS triethylammonium salt (Rp-Adenosine 3′,5′-cyclic monophosphorothioate triethylammonium salt hydrate)
D-(+)-glucosamine hydrochloride
D-glucosamine 6-sulfate
H-7 (1-(5-isoquinolinesulfonyl)-2-methylpiperazine)
H-7 dihydrochloride (1-(5-isoquinolinesulfonyl)-2-methylpiperazine dihydrochloride)
control 100 mM HEPES 150 mM NaCl pH 7.5 1% DMSO
control 100 mM HEPES 150 mM NaCl pH 7.5

Table S4.1 **List of peptides on the SPOT blot membrane.** All possible single acetyl-lysine modifications of human histones H2A, H2B, H3, H4 and H1 are included in this membrane design. Poly-histidines serves as positive control since the anti-body used is anti-His and should always give a positive signal for this peptide. Four popular histone marks (H3K4me, me2, me3 and H3K9me) are included as negative controls. This peptide array design is based on the acetyl-lysine subset of peptides that were used in the original comprehensive histone modification (Nady, Min et al. 2008).

Positive/Negative Controls			
Histone Region	**Modifications**	**Array location**	**Sequence (N to C)**
1 - 12	polyHis	A1	H H H H H H H H H H H H
1 - 13	H3 K4 - me	A2	A R T Kme Q T A R K S T G G
1 - 13	H3 K4 - me2	A3	A R T Kme2 Q T A R K S T G G
1 - 13	H3 K4 - me3	A4	A R T Kme3 Q T A R K S T G G
Histone - H2A			
1 - 12	none	A5	S G R G K Q G G K A R A
1 - 12	K5 - ac	A6	S G R G Kac Q G G K A R A
1 - 12	K9 - ac	A7	S G R G K Q G G Kac A R A
10 - 21	none	A8	A R A K A K T R S S R A
10 - 21	K13 - ac	A9	A R A Kac A K T R S S R A
10 - 21	K15 - ac	A10	A R A K A Kac T R S S R A
14 - 27	none	B1	A K T R S S R A G L Q F P V
14 - 27	K15 - ac	B2	A Kac T R S S R A G L Q F P V
25 - 38	none	B3	F P V G R V H R L L R K G N
25 - 38	K36 - ac	B4	F P V G R V H R L L R Kac G N
33 - 45	none	B5	L L R K G N Y S E R V G A
33 - 45	K36 - ac	B6	L L R Kac G N Y S E R V G A
70 - 83	none	B7	A R D N K K T R I I P R H L
70 - 83	K74 - ac	B8	A R D N Kac K T R I I P R H L
70 - 83	K75 - ac	B9	A R D N K Kac T R I I P R H L
93 - 103	none	B10	L N K L L G K V T I A
93 - 103	K95 - ac	C1	L N Kac L L G K V T I A
93 - 103	K99 - ac	C2	L N K L L G Kac V T I A
116 - 129	none	C3	L P K K T E S H H K A K G K
116 - 129	K118 - ac	C4	L P Kac K T E S H H K A K G K
116 - 129	K119 - ac	C5	L P K Kac T E S H H K A K G K
116 - 129	K125 - ac	C6	L P K K T E S H H Kac A K G K
116 - 129	K127 - ac	C7	L P K K T E S H H K A Kac G K
116 - 129	K129 - ac	C8	L P K K T E S H H K A K G Kac
Histone - H2B			
1 - 13	none	C9	P E P A K S A P A P K K G
1 - 13	K5 - ac	C10	P E P A Kac S A P A P K K G
1 - 13	K11 - ac	D1	P E P A K S A P A P Kac K G
1 - 13	K12 - ac	D2	P E P A K S A P A P K Kac G
7 - 18	none	D3	A P A P K K G S K K A V
7 - 18	K11 - ac	D4	A P A P Kac K G S K K A V
7 - 18	K12 - ac	D5	A P A P K Kac G S K K A V
7 - 18	K15 - ac	D6	A P A P K K G S Kac K A V

7 - 18	K16 - ac	D7	A P A P K K G S K Kac A V
17 - 26	none	D8	A V T K A Q K K D G
17 - 26	K20 - ac	D9	A V T Kac A Q K K D G
17 - 26	K23 - ac	D10	A V T K A Q Kac K D G
17 - 26	K24 - ac	E1	A V T K A Q K Kac D G
25 - 39	none	E2	D G K K R K R S R K E S Y S V
25 - 39	K27 - ac	E3	D G Kac K R K R S R K E S Y S V
25 - 39	K28 - ac	E4	D G K Kac R K R S R K E S Y S V
25 - 39	K30 - ac	E5	D G K K R Kac R S R K E S Y S V
25 - 39	K34 - ac	E6	D G K K R K R S R Kac E S Y S V
35 - 47	none	E7	E S Y S V Y V Y K V L K Q
35 - 47	K43 - ac	E8	E S Y S V Y V Y Kac V L K Q
35 - 47	K46 - ac	E9	E S Y S V Y V Y K V L Kac Q
41 - 53	none	E10	V Y K V L K Q V H P D T G
41 - 53	K43 - ac	F1	V Y Kac V L K Q V H P D T G
41 - 53	K46 - ac	F2	V Y K V L Kac Q V H P D T G
81 - 93	none	F3	A H Y N K R S T I T S R E
81 - 93	K85 - ac	F4	A H Y N Kac R S T I T S R E
106 - 118	none	F5	L A K H A V S E G T K A V
106 - 118	K108 - ac	F6	L A Kac H A V S E G T K A V
106 - 118	K116 - ac	F7	L A K H A V S E G T Kac A V
114 - 125	none	F8	G T K A V T K Y T S S K
114 - 125	K116 - ac	F9	G T Kac A V T K Y T S S K
114 - 125	K120 - ac	F10	G T K A V T Kac Y T S S K
114 - 125	K125 - ac	G1	G T K A V T K Y T S S Kac

Histone - H3

peptide identity	Modifications		Sequence (N - C)
1 - 13	none	G2	A R T K Q T A R K S T G G
1 - 13	K4 - ac	G3	A R T Kac Q T A R K S T G G
1 - 13	K9 - ac	G4	A R T K Q T A R Kac S T G G
7 - 20	none	G5	A R K S T G G K A P R K Q L
7 - 20	K9 - ac	G6	A R Kac S T G G K A P R K Q L
7 - 20	K14 - ac	G7	A R K S T G G Kac A P R K Q L
7 - 20	K18 - ac	G8	A R K S T G G K A P R Kac Q L
15 - 25	none	G9	A P R K Q L A T K A A
15 - 25	K18 - ac	G10	A P R Kac Q L A T K A A
15 - 25	K23 - ac	H1	A P R K Q L A T Kac A A
21 - 33	none	H2	A T K A A R K S A P A T G
21 - 33	K23 - ac	H3	A T Kac A A R K S A P A T G
21 - 33	K27 - ac	H4	A T K A A R Kac S A P A T G
31 - 44	none	H5	A T G G V K K P H R Y R P G
31 - 44	K36 - ac	H6	A T G G V Kac K P H R Y R P G
31 - 44	K37 - ac	H7	A T G G V K Kac P H R Y R P G
44 - 57	none	H8	G T V A L R E I R R Y Q K S
44 - 57	K56 - ac	H9	G T V A L R E I R R Y Q Kac S
54 - 67	none	H10	Y Q K S T E L L I R K L P F
54 - 67	K56 - ac	I1	Y Q Kac S T E L L I R K L P F
54 - 67	K64 - ac	I2	Y Q K S T E L L I R Kac L P F
75 - 88	none	I3	A Q D F K T D L R F Q S S A

75 - 88	K79 - ac	I4	A Q D F Kac T D L R F Q S S A
114 - 126	none	I5	A K R V T I M P K D I Q L
114 - 126	K115 - ac	I6	A Kac R V T I M P K D I Q L
114 - 126	K122 - ac	I7	A K R V T I M P Kac D I Q L

Histone - H4

peptide identity	Modifications		Sequence (N - C)
1 - 11	none	I8	S G R G K G G K G L G
1 - 11	K5 - ac	I9	S G R G Kac G G K G L G
1 - 11	K8 - ac	I10	S G R G K G G Kac G L G
10 - 22	none	J1	L G K G G A K R H R K V L
10 - 22	K12 - ac	J2	L G Kac G G A K R H R K V L
10 - 22	K16 - ac	J3	L G K G G A Kac R H R K V L
10 - 22	K20 - ac	J4	L G K G G A K R H R Kac V L
21 - 33	none	J5	V L R D N I Q G I T K P A
21 - 33	K31 - ac	J6	V L R D N I Q G I T Kac P A
41 - 50	none	J7	G G V K R I S G L I
41 - 50	K44 - ac	J8	G G V Kac R I S G L I
51 - 60	none	J9	Y E E T R G V L K V
51 - 60	K59 - ac	J10	Y E E T R G V L Kac V
56 - 66	none	K1	G V L K V F L E N V I
56 - 66	K59 - ac	K2	G V L Kac V F L E N V I
76 - 86	none	K3	A K R K T V T A M D V
76 - 86	K77 - ac	K4	A Kac R K T V T A M D V
76 - 86	K79 - ac	K5	A K R Kac T V T A M D V
87 - 97	none	K6	V Y A L K R Q G R T L
87 - 97	K91 - ac	K7	V Y A L Kac R Q G R T L

1 - 12	polyHis	K8	H H H H H H H H H H H H

Histone - H1.4

peptide identity	Modifications		Sequence
10 - 20	none	K9	A A P A P A E K T P V
10 - 20	K17 - ac	K10	A A P A P A E Kac T P V
15 - 28	none	L1	A E K T P V K K K A R K S A
15 - 28	K17 - ac	L2	A E Kac T P V K K K A R K S A
15 - 28	K21 - ac	L3	A E K T P V Kac K K A R K S A
15 - 28	K22 - ac	L4	A E K T P V K Kac K A R K S A
15 - 28	K23 - ac	L5	A E K T P V K K Kac A R K S A
15 - 28	K26 - ac	L6	A E K T P V K K K A R Kac S A
24 - 35	none	L7	A R K S A G A A K R K A
24 - 35	K26 - ac	L8	A R Kac S A G A A K R K A
24 - 35	K32 - ac	L9	A R K S A G A A Kac R K A
24 - 35	K34 - ac	L10	A R K S A G A A K R Kac A
30 - 40	none	M1	A A K R K A S G P P V
30 - 40	K32 - ac	M2	A A Kac R K A S G P P V
30 - 40	K34 - ac	M3	A A K R Kac A S G P P V
40 - 49	none	M4	V S E L I T K A V A

40 - 49	K46 - ac	M5	V S E L I T Kac A V A
49 - 60	none	M6	A A S K E R S G V S L A
49 - 60	K52 - ac	M7	A A S Kac E R S G V S L A
61 - 70	none	M8	A L K K A L A A A G
61 - 70	K63 - ac	M9	A L Kac K A L A A A G
61 - 70	K64 - ac	M10	A L K Kac A L A A A G
71 - 80	none	N1	Y D V E K N N S R I
71 - 80	K75 - ac	N2	Y D V E Kac N N S R I
80 - 91	none	N3	I K L G L K S L V S K G
80 - 91	K81 - ac	N4	I Kac L G L K S L V S K G
80 - 91	K85 - ac	N5	I K L G L Kac S L V S K G
80 - 91	K90 - ac	N6	I K L G L K S L V S Kac G
91 - 100	none	N7	G T L V Q T K G T G
91 - 100	K97 - ac	N8	G T L V Q T Kac G T G
101 - 111	none	N9	A S G S F K L N K K A
101 - 111	K106 - ac	N10	A S G S F Kac L N K K A
101 - 111	K109 - ac	O1	A S G S F K L N Kac K A
101 - 111	K110 - ac	O2	A S G S F K L N K Kac A
111 - 120	none	O3	A A S G E A K P K A
111 - 120	K117 - ac	O4	A A S G E A Kac P K A
111 - 120	K119 - ac	O5	A A S G E A K P Kac A
120 - 132	none	O6	A K K A G A A K A K K P A
120 - 132	K121 - ac	O7	A Kac K A G A A K A K K P A
120 - 132	K122 - ac	O8	A K Kac A G A A K A K K P A
120 - 132	K127 - ac	O9	A K K A G A A Kac A K K P A
120 - 132	K129 - ac	O10	A K K A G A A K A Kac K P A
120 - 132	K130 - ac	P1	A K K A G A A K A K Kac P A
132 - 144	none	P2	A G A A K K P K K A T G A
132 - 144	K136 - ac	P3	A G A A Kac K P K K A T G A
132 - 144	K137 - ac	P4	A G A A K Kac P K K A T G A
132 - 144	K139 - ac	P5	A G A A K K P Kac K A T G A
132 - 144	K140 - ac	P6	A G A A K K P K Kac A T G A
141 - 151	none	P7	A T G A A T P K K S A
141 - 151	K148 - ac	P8	A T G A A T P Kac K S A
141 - 151	K149 - ac	P9	A T G A A T P K Kac S A
151 - 162	none	P10	A K K T P K K A K K P A
151 - 162	K152 - ac	Q1	A Kac K T P K K A K K P A
151 - 162	K153 - ac	Q2	A K Kac T P K K A K K P A
151 - 162	K156 - ac	Q3	A K K T P Kac K A K K P A
151 - 162	K157 - ac	Q4	A K K T P K Kac A K K P A
151 - 162	K159 - ac	Q5	A K K T P K K A Kac K P A
151 - 162	K160 - ac	Q6	A K K T P K K A K Kac P A
158 - 170	none	Q7	A K K P A A A A G A K K A
158 - 170	K159 - ac	Q8	A Kac K P A A A A G A K K A
158 - 170	K160 - ac	Q9	A K Kac P A A A A G A K K A
158 - 170	K168 - ac	Q10	A K K P A A A A G A Kac K A
158 - 170	K169 - ac	R1	A K K P A A A A G A K Kac A
170 - 179	none	R2	A K S P K K A K A A
170 - 179	K171 - ac	R3	A Kac S P K K A K A A
170 - 179	K174 - ac	R4	A K S P Kac K A K A A

170 - 179	K175 - ac	R5	A K S P K Kac A K A A
170 - 179	K177 - ac	R6	A K S P K K A Kac A A
178 - 189	none	R7	A A K P K K A P K S P A
178 - 189	K180 - ac	R8	A A Kac P K K A P K S P A
178 - 189	K182 - ac	R9	A A K P Kac K A P K S P A
178 - 189	K183 - ac	R10	A A K P K Kac A P K S P A
178 - 189	K186 - ac	S1	A A K P K K A P Kac S P A
189 - 199	none	S2	A K A K A V K P K A A
189 - 199	K190 - ac	S3	A Kac A K A V K P K A A
189 - 199	K192 - ac	S4	A K A Kac A V K P K A A
189 - 199	K195 - ac	S5	A K A K A V Kac P K A A
189 - 199	K197 - ac	S6	A K A K A V K P Kac A A
198 - 209	none	S7	A A K P K T A K P K A A
198 - 209	K200 - ac	S8	A A Kac P K T A K P K A A
198 - 209	K202 - ac	S9	A A K P Kac T A K P K A A
198 - 209	K205 - ac	S10	A A K P K T A Kac P K A A
198 - 209	K207 - ac	T1	A A K P K T A K P Kac A A
208 - 219	none	T2	A A K P K K A A A K K K
208 - 219	K210 - ac	T3	A A Kac P K K A A A K K K
208 - 219	K212 - ac	T4	A A K P Kac K A A A K K K
208 - 219	K213 - ac	T5	A A K P K Kac A A A K K K
208 - 219	K217 - ac	T6	A A K P K K A A A Kac K K
208 - 219	K218 - ac	T7	A A K P K K A A A K Kac K
208 - 219	K219 - ac	T8	A A K P K K A A A K K Kac

Positive/Negative Controls

peptide identity	Modifications		Sequence
1 - 12	polyHis	T9	H H H H H H H H H H H H
1 - 13	H3 K9 - me3	T10	A R T K Q T A R Kme3 S T G G

Figure S5.1. **SDS-PAGE analysis of protein purification.** Loaded 2 μg of WT or H131Y SIRT6.

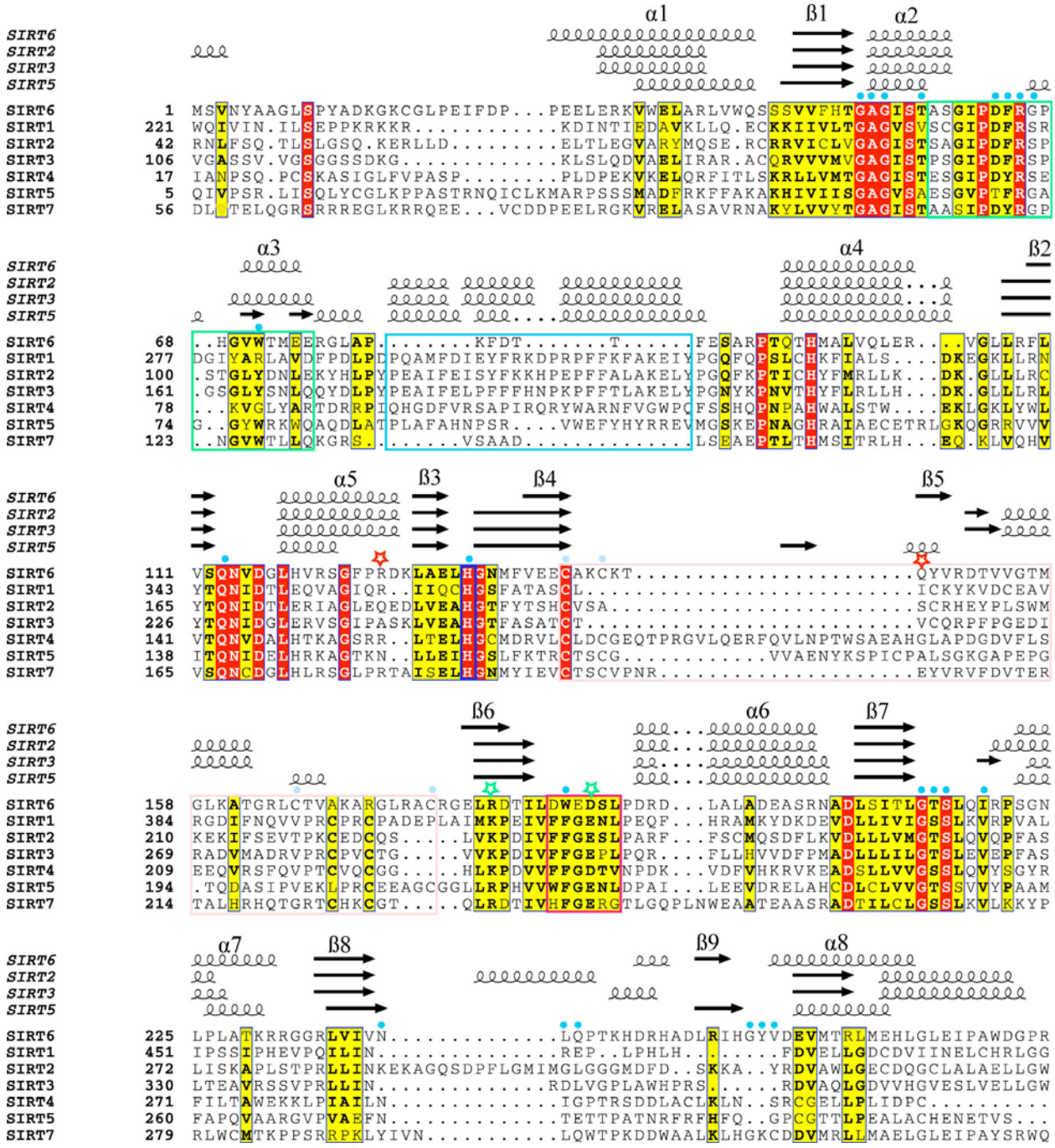


Figure S5.2. **Sequence alignment of the sirtuin domains of human SIRT1-7.** Cysteine containing zinc-binding sequences are highlighted in a *pink box* with *light blue dots* indicating zinc-binding cysteines in SIRT6; Conserved catalytic histidine is in the *blue box*; SIRT6 and SIRT7 missing helix region is highlighted in a *sky blue box*; *green stars* indicating conventional salt bridge while *red stars* corresponding to the hydrogen-bonding pairs in SIRT6; the conserved FGExL loop is labeled in *a pink box*; ADPr interacting residues are denoted with *sky blue dots*; the cofactor binding loop is indicated in the *green box*.

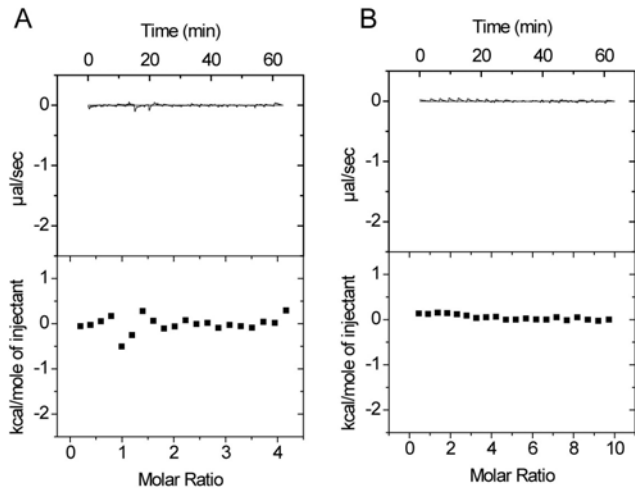


Figure S5.3. **Peptide substrate ITC with crystallized SIRT6 construct (3-318)**. *A*, 10 mM H3K9Ac peptide titrated to 1mM SIRT6; *B*, 10 mM H3K9Ac peptide titrated to 1mM SIRT6 pre-equilibrated with saturating amount of NAD^+.

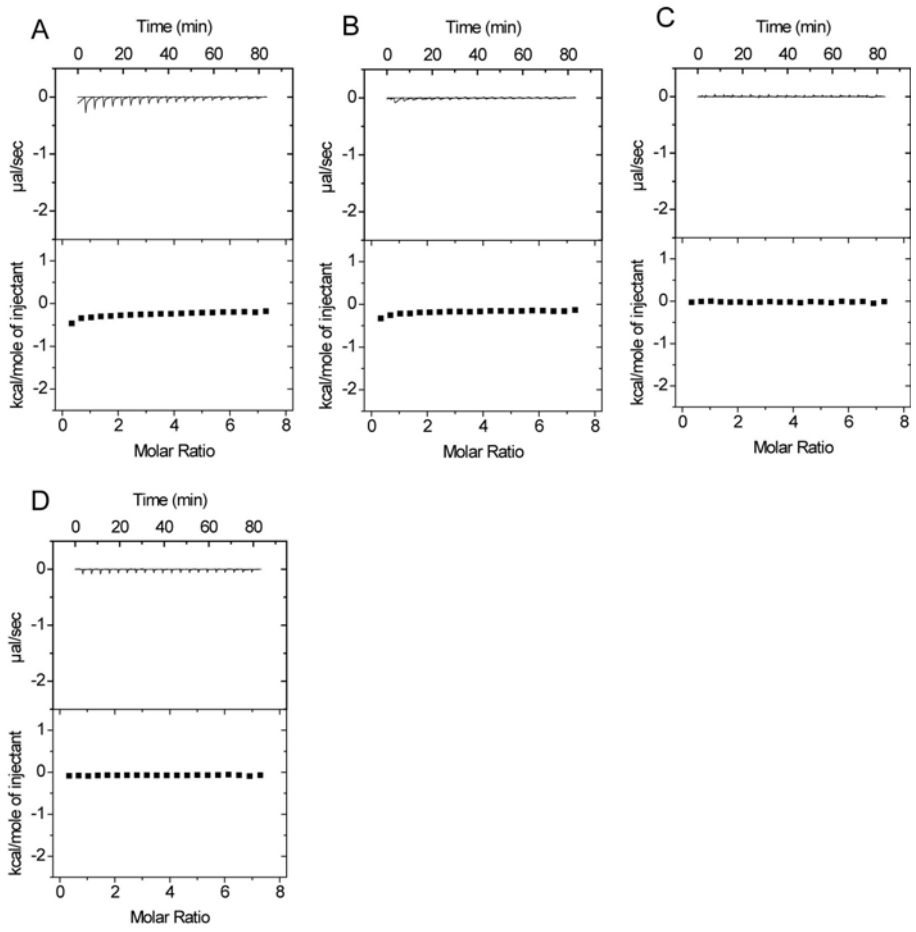


Figure S5.4. **Isothermal titration calorimetry binding assay of NAD^+ and SIRT1, 2, 3, 5 binding.** 5 mM NAD^+ titrated 50-100 µM sirtuin proteins. *A*, SIRT1. *B*, SIRT2. *C*, SIRT3. *D*, SIRT5.

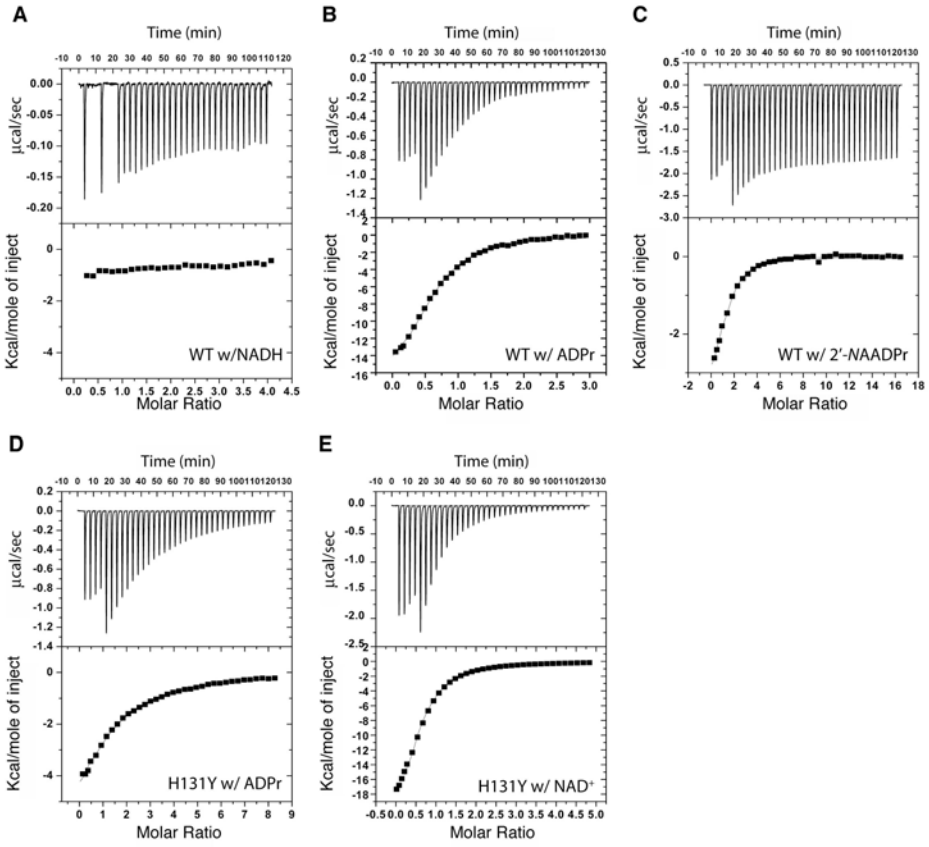


Figure S5.5. **Isothermal titration calorimetry curves for WT and H131Y SIRT6 with various ligands.** Top panel shows data obtained for 37 automatic injections (1-8 µL) and bottom panel represents integrated curves of the experimentally generated heats. *A*, 450 µM NADH was titrated into 25 µM SIRT6. *B*, 390 µM ADPr was titrated into 28 µM WT SIRT6. *C*, 2.96 mM 2'-*N*AADPr was titrated into 38 µM WT SIRT6. *D*, 1.45 mM ADPr was titrated into 37 µM H131Y SIRT6. *E*, 767 µM NAD^+ was titrated into 33 µM H131Y SIRT6.

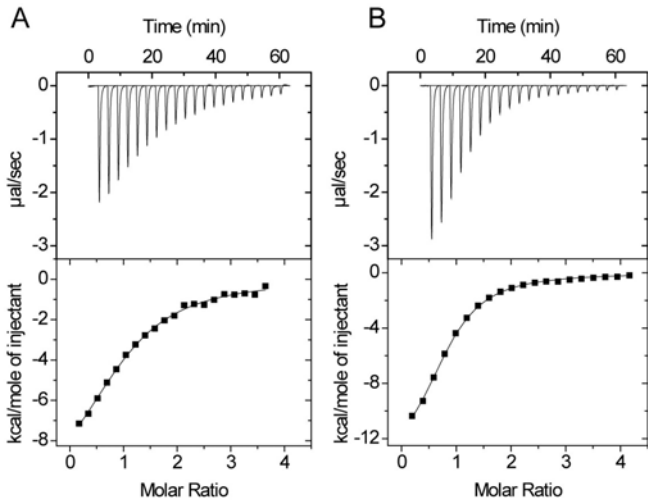


Figure S5.6. **ITC study of crystallized SIRT6 construct (3-318) with NAD^+ and ADPr.** *A*, 1 mM NAD^+ titrated to 50 µM SIRT6. *B*, 1 mM ADPr titrated to 50 µM SIRT6.

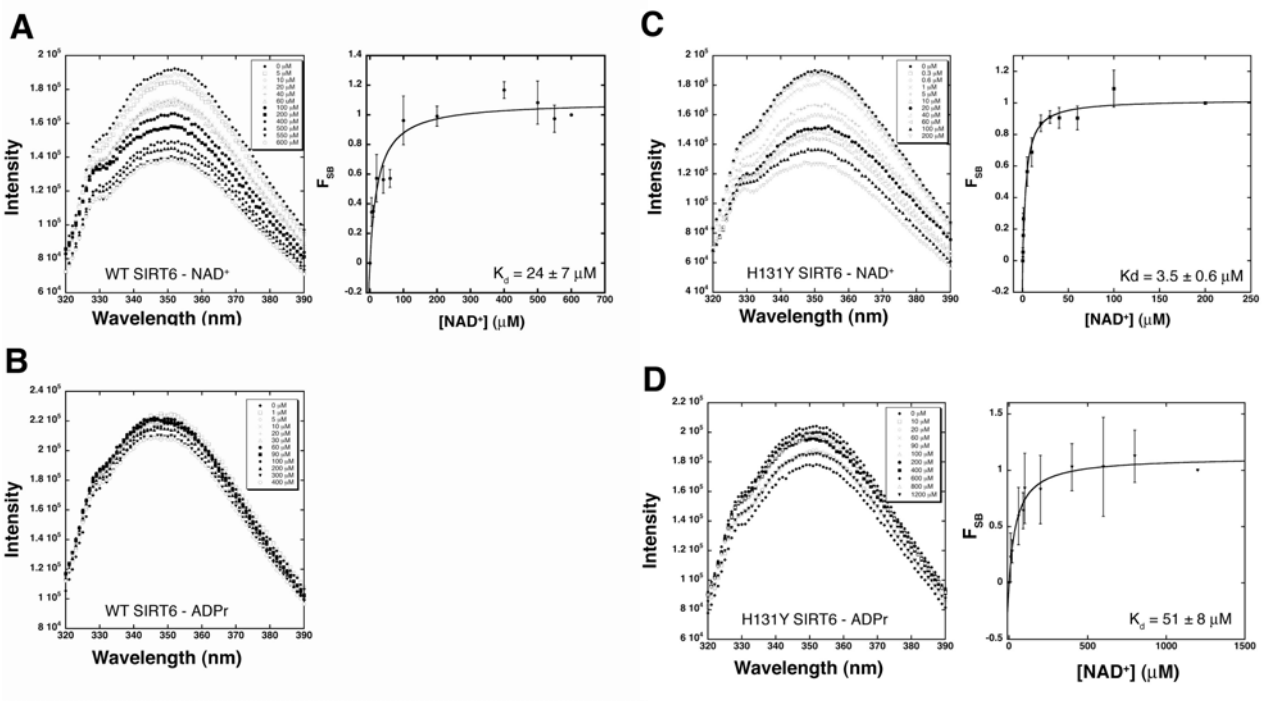


Figure S5.7. **Tryptophan fluorescence emission spectrum and saturation binding curves.** Emission spectra are not adjusted for non-specific tryptophan quenching due to increasing ligand concentrations. Saturation binding curves were repeated in triplicate and fitted to a one-site binding equation. *A,* increasing concentrations of NAD^+ (0 – 600 μM) were added to 1 μM WT SIRT6. *B*, increasing concentrations of ADPr (0 – 400 μM) were added to 1 μM ADPr. *C*, increasing concentrations of NAD^+ (0 – 200 μM) were added to 1 μM H131Y SIRT6. *D*, increasing concentrations of ADPr (0 – 1.2 mM) were added to 1 μM H131Y SIRT6.

References

Adjei, A. A. and R. M. Weinshilboum (2002). "Catecholestrogen sulfation: possible role in carcinogenesis." Biochem Biophys Res Commun **292**(2): 402-408.

Allali-Hassani, A., T. L. Campbell, et al. (2004). "Probing the active site of YjeE: a vital Escherichia coli protein of unknown function." Biochem J **384**(Pt 3): 577-584.

Allali-Hassani, A., P. W. Pan, et al. (2007). "Structural and chemical profiling of the human cytosolic sulfotransferases." PLoS Biol **5**(5): e97.

Anamika, K., K. R. Abhinandan, et al. (2009). "Classification of nonenzymatic homologues of protein kinases." Comp Funct Genomics: 365637.

Araya, Z. and K. Wikvall (1999). "6alpha-hydroxylation of taurochenodeoxycholic acid and lithocholic acid by CYP3A4 in human liver microsomes." Biochim Biophys Acta **1438**(1): 47-54.

Arun, K. S., T. S. Huang, et al. (1987). "Least-Squares Fitting Of 2 3-D Point Sets." Ieee Transactions On Pattern Analysis And Machine Intelligence **9**(5): 699-700.

Asaba, T., T. Suzuki, et al. (2009). "Inhibition of human sirtuins by in situ generation of an acetylated lysine-ADP-ribose conjugate." J Am Chem Soc **131**(20): 6989-6996.

Avalos, J. L., K. M. Bever, et al. (2005). "Mechanism of sirtuin inhibition by nicotinamide: altering the NAD(+) cosubstrate specificity of a Sir2 enzyme." Mol Cell **17**(6): 855-868.

Avalos, J. L., J. D. Boeke, et al. (2004). "Structural basis for the mechanism and regulation of Sir2 enzymes." Mol Cell **13**(5): 639-648.

Avalos, J. L., I. Celic, et al. (2002). "Structure of a Sir2 enzyme bound to an acetylated p53 peptide." Mol Cell **10**(3): 523-535.

Bae, N. S., M. J. Swanson, et al. (2004). "Human histone deacetylase SIRT2 interacts with the homeobox transcription factor HOXA10." J Biochem **135**(6): 695-700.

Bamforth, K. J., K. Dalgliesh, et al. (1992). "Inhibition of human liver steroid sulfotransferase activities by drugs: a novel mechanism of drug toxicity?" Eur J Pharmacol **228**(1): 15-21.

Bamforth, K. J., A. L. Jones, et al. (1993). "Common food additives are potent inhibitors of human liver 17 alpha-ethinyloestradiol and dopamine sulphotransferases." Biochem Pharmacol **46**(10): 1713-1720.

Barnett, A. C., S. Tsvetanov, et al. (2004). "Active site mutations and substrate inhibition in human sulfotransferase 1A1 and 1A3." J Biol Chem **279**(18): 18799-18805.

Bateman, A., L. Coin, et al. (2004). "The Pfam protein families database." Nucleic Acids Res **32**(Database issue): D138-141.

Baur, J. A., K. J. Pearson, et al. (2006). "Resveratrol improves health and survival of mice on a high-calorie diet." Nature **444**(7117): 337-342.

Baur, J. A. and D. A. Sinclair (2006). "Therapeutic potential of resveratrol: the in vivo evidence." Nat Rev Drug Discov **5**(6): 493-506.

Bedalov, A., T. Gatbonton, et al. (2001). "Identification of a small molecule inhibitor of Sir2p." Proc Natl Acad Sci U S A **98**(26): 15113-15118.

Belenky, P., K. L. Bogan, et al. (2007). "NAD^+ metabolism in health and disease." Trends Biochem Sci **32**(1): 12-19.

Bereshchenko, O. R., W. Gu, et al. (2002). "Acetylation inactivates the transcriptional repressor BCL6." Nat Genet **32**(4): 606-613.

Berndsen, C. E., T. Tsubota, et al. (2008). "Molecular functions of the histone acetyltransferase chaperone complex Rtt109-Vps75." Nat Struct Mol Biol **15**(9): 948-956.

Bertelli, A. A., L. Giovannini, et al. (1995). "Antiplatelet activity of synthetic and natural resveratrol in red wine." Int J Tissue React **17**(1): 1-3.

Bidwell, L. M., M. E. McManus, et al. (1999). "Crystal structure of human catecholamine sulfotransferase." J Mol Biol **293**(3): 521-530.

Blanc, E., P. Roversi, et al. (2004). "Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT." Acta Crystallogr D Biol Crystallogr **60**: 2210-2221.

Blanchard, R. L., R. R. Freimuth, et al. (2004). "A proposed nomenclature system for the cytosolic sulfotransferase (SULT) superfamily." Pharmacogenetics **14**(3): 199-211.

Bordone, L., D. Cohen, et al. (2007). "SIRT1 transgenic mice show phenotypes resembling calorie restriction." Aging Cell **6**(6): 759-767.

Borra, M. T. and J. M. Denu (2004). "Quantitative assays for characterization of the Sir2 family of NAD(+)-dependent deacetylases." Methods Enzymol **376**: 171-187.

Borra, M. T., M. R. Langer, et al. (2004). "Substrate specificity and kinetic mechanism of the Sir2 family of NAD^+-dependent histone/protein deacetylases." Biochemistry **43**(30): 9877-9887.

Borra, M. T., B. C. Smith, et al. (2005). "Mechanism of human SIRT1 activation by resveratrol." J Biol Chem **280**(17): 17187-17195.

Bradbury, C. A., F. L. Khanim, et al. (2005). "Histone deacetylases in acute myeloid leukaemia show a distinctive pattern of expression that changes selectively in response to deacetylase inhibitors." Leukemia **19**(10): 1751-1759.

Brix, L. A., A. C. Barnett, et al. (1999). "Analysis of the substrate specificity of human sulfotransferases SULT1A1 and SULT1A3: site-directed mutagenesis and kinetic studies." Biochemistry **38**(32): 10474-10479.

Bron, C. and J. Kerbosch (1973). "Algorithm 457: finding all cliques of an undirected graph." Communications of the ACM **16**(9): 575-577.

Brunger, A. T., P. D. Adams, et al. (1998). "Crystallography & MNR system: A new software suite for macromolecular structure determination." Acta Crystallographica Section D - Biological Crystallography **54** pt 5: 905-921.

Buirchell, B. J. and R. Hahnel (1975). "Metabolism of estradiol-17beta in human endometrium during the menstrual cycle." J Steroid Biochem **6**(11-12): 1489-1494.

Carey, M. C., S. F. Wu, et al. (1979). "Solution properties of sulfated monohydroxy bile salts. Relative insolubility of the disodium salt of glycolithocholate sulfate." Biochim Biophys Acta **575**(1): 16-26.

Chang, H. J., R. Shi, et al. (2004). "Identifying androsterone (ADT) as a cognate substrate for human dehydroepiandrosterone sulfotransferase (DHEA-ST) important for steroid homeostasis: structure of the enzyme-ADT complex." J Biol Chem **279**(4): 2689-2696.

Chang, J. H., H. C. Kim, et al. (2002). "Structural basis for the NAD-dependent deacetylase mechanism of Sir2." J Biol Chem **277**(37): 34489-34498.

Chapman, E., M. D. Best, et al. (2004). "Sulfotransferases: structure, mechanism, biological activity, inhibition, and synthetic utility." Angew Chem Int Ed Engl **43**(27): 3526-3548.

Chen, G., E. Banoglu, et al. (1996). "Influence of substrate structure on the catalytic efficiency of hydroxysteroid sulfotransferase STa in the sulfation of alcohols." Chem Res Toxicol **9**(1): 67-74.

Comer, K. A., J. L. Falany, et al. (1993). "Cloning and expression of human liver dehydroepiandrosterone sulphotransferase." Biochem J **289** (Pt 1): 233-240.

Comstock, L. R. and J. M. Denu (2007). "Synthesis and biochemical evaluation of O-acetyl-ADP-ribose and N-acetyl analogs." Org Biomol Chem **5**(19): 3087-3091.

Constant, J. (1997). "Alcohol, ischemic heart disease, and the French paradox." Coron Artery Dis **8**(10): 645-649.

Cosgrove, M. S., K. Bever, et al. (2006). "The structural basis of sirtuin substrate affinity." Biochemistry **45**(24): 7511-7521.

Coughtrie, M. W. (1998). "Catecholamine sulfation in health and disease." Adv Pharmacol **42**: 339-342.

Coughtrie, M. W. and L. E. Johnston (2001). "Interactions between dietary chemicals and human sulfotransferases-molecular mechanisms and clinical significance." Drug Metab Dispos **29**(4 Pt 2): 522-528.

Coughtrie, M. W., S. Sharp, et al. (1998). "Biology and function of the reversible sulfation pathway catalysed by human sulfotransferases and sulfatases." Chem Biol Interact **109**(1-3): 3-27.

Dajani, R., A. Cleasby, et al. (1999). "X-ray crystal structure of human dopamine sulfotransferase, SULT1A3. Molecular modeling and quantitative structure-activity relationship analysis demonstrate a molecular basis for sulfotransferase substrate specificity." J Biol Chem **274**(53): 37862-37868.

De Santi, C., A. Pietrabissa, et al. (2000). "Sulphation of resveratrol, a natural compound present in wine, and its inhibition by natural flavonoids." Xenobiotica **30**(9): 857-866.

De Santi, C., A. Pietrabissa, et al. (2000). "Sulphation of resveratrol, a natural product present in grapes and wine, in the human liver and duodenum." Xenobiotica **30**(6): 609-617.

Denu, J. M. (2005). "The Sir 2 family of protein deacetylases." Curr Opin Chem Biol **9**(5): 431-440.

Deo, A. K. and S. M. Bandiera (2008). "Biotransformation of lithocholic acid by rat hepatic microsomes: metabolite analysis by liquid chromatography/mass spectrometry." Drug Metab Dispos **36**(2): 442-451.

Dionne, S., B. Tuchweber, et al. (1994). "Phase I and phase II metabolism of lithocholic acid in hepatic acinar zone 3 necrosis. Evaluation in rats by combined radiochromatography and gas-liquid chromatography-mass spectrometry."

Biochem Pharmacol **48**(6): 1187-1197.

Dombrovski, L., A. Dong, et al. (2006). "Crystal structures of human sulfotransferases SULT1B1 and SULT1C1 complexed with the cofactor product adenosine-3'- 5'-diphosphate (PAP)." Proteins **64**(4): 1091-1094.

Dong, A., X. Xu, et al. (2007). "*In situ* proteolysis for protein crystallization and structure determination." Nat Methods **4**(12): 1019-1021.

Dooley, T. P. (1998). "Cloning of the human phenol sulfotransferase gene family: three genes implicated in the metabolism of catecholamines, thyroid hormones and drugs." Chem Biol Interact **109**(1-3): 29-41.

Eckelman, W. C., R. C. Reba, et al. (1979). "Receptor-binding radiotracers: a class of potential radiopharmaceuticals." J Nucl Med **20**(4): 350-357.

Eddy, S. R. (1998). "Profile hidden markov models." Bioinformatics **14**: 755-766.

Eisenhofer, G., M. W. Coughtrie, et al. (1999). "Dopamine sulphate: an enigma resolved." Clin Exp Pharmacol Physiol Suppl **26**: S41-53.

Emsley, P. and K. Cowtan (2004). "Coot: Model-building tools for molecular graphics." Acta Crystallographica Section D - Biological Crystallography **60**: 2126-2132.

Eryilmaz, J., P. Pan, et al. (2009). "Structural studies of a four-MBT repeat protein MBTD1." PLoS One **4**(10): e7274.

Everitt, B. (2001). Cluster Analysis of Subjects, Hierarchical Methods. Encyclopaedia of Biostatistics. P. Armitage and T. Colton, John Wiley & Sons, inc.

Fabian, M. A., W. H. Biggs, 3rd, et al. (2005). "A small molecule-kinase interaction map for clinical kinase inhibitors." Nat Biotechnol **23**(3): 329-336.

Falany, C. N. (1997). "Enzymology of human cytosolic sulfotransferases." Faseb J **11**(4): 206-216.

Falany, C. N., K. A. Comer, et al. (1995). "Human dehydroepiandrosterone sulfotransferase. Purification, molecular cloning, and characterization." Ann N Y Acad Sci **774**: 59-72.

Falany, C. N., V. Krasnykh, et al. (1995). "Bacterial expression and characterization of a cDNA for human liver estrogen sulfotransferase." J Steroid Biochem Mol Biol **52**(6): 529-539.

Falany, C. N., J. Wheeler, et al. (1994). "Steroid sulfation by expressed human cytosolic sulfotransferases." J Steroid Biochem Mol Biol **48**(4): 369-375.

Falany, C. N., X. Xie, et al. (2000). "Molecular cloning and expression of novel sulphotransferase-like cDNAs from human and rat brain." Biochem J **346** Pt 3: 857-864.

Falany, J. L., N. Macrina, et al. (2002). "Regulation of MCF-7 breast cancer cell growth by beta-estradiol sulfation." Breast Cancer Res Treat **74**(2): 167-176.

Fan, Y., R. Ludewig, et al. (2008). "Development of a capillary electrophoresis-based assay of sirtuin enzymes." Electrophoresis **29**(18): 3717-3723.

Fan, Y., R. Ludewig, et al. (2009). "9-Fluorenylmethoxycarbonyl-labeled peptides as substrates in a capillary electrophoresis-based assay for sirtuin enzymes." Anal Biochem **387**(2): 243-248.

Fatkins, D. G. and W. Zheng (2008). "Substituting N-thioacetyl-lysine for N-acetyl-lysine in Peptide Substrates as a General Approach to Inhibiting Human NAD-dependent Protein Deacetylases." Int J Mol Sci **9**(1): 1-11.

Finnin, M. S., J. R. Donigian, et al. (2001). "Structure of the histone deacetylase SIRT2." Nat Struct Biol **8**(7): 621-625.

Firestein, R., G. Blander, et al. (2008). "The SIRT1 deacetylase suppresses intestinal tumorigenesis and colon cancer growth." PLoS One **3**(4): e2020.

Ford, E., R. Voit, et al. (2006). "Mammalian Sir2 homolog SIRT7 is an activator of RNA polymerase I transcription." Genes Dev **20**(9): 1075-1080.

Frank, R. (2002). "The SPOT-synthesis technique. Synthetic peptide arrays on membrane supports--principles and applications." J Immunol Methods **267**(1): 13-26.

Freimuth, R. R., M. Wiepert, et al. (2004). "Human cytosolic sulfotransferase database mining: identification of seven novel genes and pseudogenes." Pharmacogenomics J **4**(1): 54-65.

Frye, R. A. (1999). "Characterization of five human cDNAs with homology to the yeast SIR2 gene: Sir2-like proteins (sirtuins) metabolize NAD and may have protein ADP-ribosyltransferase activity." Biochem Biophys Res Commun **260**(1): 273-279.

Frye, R. A. (2000). "Phylogenetic classification of prokaryotic and eukaryotic Sir2-like proteins." Biochem Biophys Res Commun **273**(2): 793-798.

Fuda, H., Y. C. Lee, et al. (2002). "Mutational analysis of human hydroxysteroid sulfotransferase SULT2B1 isoforms reveals that exon 1B of the SULT2B1 gene produces cholesterol sulfotransferase, whereas exon 1A yields pregnenolone

sulfotransferase." J Biol Chem **277**(39): 36161-36166.

Funk, C., C. Ponelle, et al. (2001). "Cholestatic potential of troglitazone as a possible factor contributing to troglitazone-induced hepatotoxicity: in vivo and in vitro interaction at the canalicular bile salt export pump (Bsep) in the rat." Mol Pharmacol **59**(3): 627-635.

Gamage, N., A. Barnett, et al. (2006). "Human sulfotransferases and their role in chemical metabolism." Toxicol Sci **90**(1): 5-22.

Gamage, N. U., R. G. Duggleby, et al. (2003). "Structure of a human carcinogen-converting enzyme, SULT1A1. Structural and kinetic implications of substrate inhibition." J Biol Chem **278**(9): 7655-7662.

Gamage, N. U., S. Tsvetanov, et al. (2005). "The structure of human SULT1A1 crystallized with estradiol. An insight into active site plasticity and substrate inhibition with multi-ring substrates." J Biol Chem **280**(50): 41482-41486.

Garrity, J., J. G. Gardner, et al. (2007). "N-lysine propionylation controls the activity of propionyl-CoA synthetase." J Biol Chem **282**(41): 30239-30245.

Glatt, H. (1997). "Sulfation and sulfotransferases 4: bioactivation of mutagens via sulfation." Faseb J **11**(5): 314-321.

Glatt, H. (2000). "Sulfotransferases in the bioactivation of xenobiotics." Chem Biol Interact **129**(1-2): 141-170.

Glatt, H., C. E. Engelke, et al. (2000). "Sulfotransferases: genetics and role in toxicology." Toxicol Lett **112-113**: 341-348.

Glatt, H. and W. Meinl (2005). "Sulfotransferases and acetyltransferases in mutagenicity testing: technical aspects." Methods Enzymol **400**: 230-249.

Grubisha, O., B. C. Smith, et al. (2005). "Small molecule regulation of Sir2 protein deacetylases." FEBS J **272**(18): 4607-4616.

Guarente, L. and F. Picard (2005). "Calorie restriction--the SIR2 connection." Cell **120**(4): 473-482.

Guo, Y., N. Nady, et al. (2009). "Methylation-state-specific recognition of histones by the MBT repeat protein L3MBTL2." Nucleic Acids Res **37**(7): 2204-2210.

Haigis, M. C. and L. P. Guarente (2006). "Mammalian sirtuins--emerging roles in physiology, aging, and calorie restriction." Genes Dev **20**(21): 2913-2921.

Haigis, M. C., R. Mostoslavsky, et al. (2006). "SIRT4 inhibits glutamate dehydrogenase and opposes the effects of calorie restriction in pancreatic beta cells." Cell **126**(5): 941-954.

Hawse, W. F., K. G. Hoff, et al. (2008). "Structural insights into intermediate steps in the Sir2 deacetylation reaction." Structure **16**(9): 1368-1377.

He, S., D. Bauman, et al. (2003). "Facile synthesis of site-specifically acetylated and methylated histone proteins: reagents for evaluation of the histone code hypothesis." Proc Natl Acad Sci U S A **100**(21): 12033-12038.

Heinemann, U., K. Bussow, et al. (2003). "Facilities and methods for the high-throughput crystal structural analysis of human proteins." Acc Chem Res **36**(3): 157-163.

Heltweg, B., T. Gatbonton, et al. (2006). "Antitumor activity of a small-molecule inhibitor of human silent information regulator 2 enzymes." Cancer Res **66**(8): 4368-4377.

Hempel, N., A. Barnet, et al. (2005). Human SULT1A SULTs. Human Cytosolic Sulfotransferases. G. M. Pacifici and M. W. Coughtrie. Boca Raton, FL, CRC Press: 179-230.

Her, C., G. P. Kaur, et al. (1997). "Human sulfotransferase SULT1C1: cDNA cloning, tissue-specific expression, and chromosomal localization." Genomics **41**(3): 467-470.

Her, C., R. Raftogianis, et al. (1996). "Human phenol sulfotransferase STP2 gene: molecular cloning, structural characterization, and chromosomal localization." Genomics **33**(3): 409-420.

Hildebrandt, M. A., O. E. Salavaggione, et al. (2004). "Human SULT1A3 pharmacogenetics: gene duplication and functional genomic studies." Biochem Biophys Res Commun **321**(4): 870-878.

Hilpert, K., D. F. Winkler, et al. (2007). "Peptide arrays on cellulose support: SPOT synthesis, a time and cost efficient method for synthesis of large numbers of peptides in a parallel and addressable fashion." Nat Protoc **2**(6): 1333-1349.

Hirao, M., J. Posakony, et al. (2003). "Identification of selective inhibitors of NAD^+-dependent deacetylases using phenotypic screens in yeast." J Biol Chem **278**(52): 52773-52782.

Hodgson, J. (2001). "ADMET--turning chemicals into drugs." Nat Biotechnol **19**(8): 722-726.

Hoff, R. H., P. G. Czyryca, et al. (2006). "Transition state of the sulfuryl transfer reaction of estrogen sulfotransferase." J Biol Chem **281**(41): 30645-30649.

Hofmann, A. F. (2004). "Detoxification of lithocholic acid, a toxic bile acid: relevance to drug hepatotoxicity." Drug Metab Rev **36**(3-4): 703-722.

Horvath, L. G., S. M. Henshall, et al. (2005). "Lower levels of nuclear beta-catenin predict for a poorer prognosis in localized prostate cancer." Int J Cancer **113**(3): 415-422.

Horwitz, J. P., R. S. Misra, et al. (1978). "IV. Synthesis and assay of analogs of adenosine 3',5'-diphosphate as inhibitors of bovine adrenal estrogen sulfotranferase." Biochim Biophys Acta **525**(2): 364-372.

Howitz, K. T., K. J. Bitterman, et al. (2003). "Small molecule activators of sirtuins extend Saccharomyces cerevisiae lifespan." Nature **425**(6954): 191-196.

Iyer, V. K., W. B. Butler, et al. (1983). "Some adenine and adenosine methylene-bridged estrogens." J Med Chem **26**(2): 162-166.

Jackson, M. D. and J. M. Denu (2002). "Structural identification of 2'- and 3'-O-acetyl-ADP-ribose as novel metabolites derived from the Sir2 family of beta -NAD^+-dependent histone/protein deacetylases." J Biol Chem **277**(21): 18535-18544.

Jang, M., L. Cai, et al. (1997). "Cancer chemopreventive activity of resveratrol, a natural product derived from grapes." Science **275**(5297): 218-220.

Jaroszewski, L., L. Rychlewski, et al. (2005). "FFAS03: a server for profile--profile sequence alignments." Nucleic Acids Res **33**(Web Server issue): W284-288.

Javitt, N. B. (1966). "Cholestasis in rats induced by taurolithocholate." Nature **210**(5042): 1262-1263.

Jenuwein, T. and C. D. Allis (2001). "Translating the histone code." Science **293**(5532): 1074-1080.

Jin, L., W. Wei, et al. (2009). "Crystal structures of human SIRT3 displaying substrate-induced conformational changes." J Biol Chem **284**(36): 24394-24405.

Jones, T. A., J. Y. Zou, et al. (1991). "Improved methods for building protein models in the electron-density maps and the location of errors in these models." Acta Crystallographica Section A **47**: 110-119.

Kaeberlein, M., T. McDonagh, et al. (2005). "Substrate-specific activation of sirtuins by resveratrol." J Biol Chem **280**(17): 17038-17045.

Kaeberlein, M., M. McVey, et al. (1999). "The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms." Genes Dev **13**(19): 2570-2580.

Kaidi, A., B. T. Weinert, et al. (2010). "Human SIRT6 promotes DNA end resection through CtIP deacetylation." Science **329**(5997): 1348-1353.

Kakuta, Y., L. C. Pedersen, et al. (1998). "Mouse steroid sulfotransferases: substrate specificity and preliminary X-ray crystallographic analysis." Biochem Pharmacol **55**(3): 313-317.

Kamio, K., K. Honke, et al. (1995). "Pyridoxal 5'-phosphate binds to a lysine residue in the adenosine 3'-phosphate 5'-phosphosulfate recognition site of glycolipid sulfotransferase from human renal cancer cells." Glycoconj J **12**(6): 762-766.

Kanfi, Y., V. Peshti, et al. (2010). "SIRT6 protects against pathological damage caused by diet-induced obesity." Aging Cell **9**(2): 162-173.

Kanfi, Y., V. Peshti, et al. (2008). "Regulation of SIRT1 protein levels by nutrient availability." FEBS Lett **582**(16): 2417-2423.

Kanfi, Y., R. Shalman, et al. (2008). "Regulation of SIRT6 protein levels by nutrient availability." FEBS Lett **582**(5): 543-548.

Kawahara, T. L., E. Michishita, et al. (2009). "SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span." Cell **136**(1): 62-74.

Kester, M. H., S. Bulduk, et al. (2000). "Potent inhibition of estrogen sulfotransferase by hydroxylated PCB metabolites: a novel pathway explaining the estrogenic activity of PCBs." Endocrinology **141**(5): 1897-1900.

Kester, M. H., S. Bulduk, et al. (2002). "Potent inhibition of estrogen sulfotransferase by hydroxylated metabolites of polyhalogenated aromatic hydrocarbons reveals alternative mechanism for estrogenic activity of endocrine disrupters." J Clin Endocrinol Metab **87**(3): 1142-1150.

Kim, D., M. D. Nguyen, et al. (2007). "SIRT1 deacetylase protects against neurodegeneration in models for Alzheimer's disease and amyotrophic lateral sclerosis." EMBO J **26**(13): 3169-3179.

Kim, H. S., C. Xiao, et al. (2010). "Hepatic-Specific Disruption of SIRT6 in Mice Results in Fatty Liver Formation Due to Enhanced Glycolysis and Triglyceride Synthesis." Cell Metab **12**(3): 224-236.

Kim, J., J. Daniel, et al. (2006). "Tudor, MBT and chromo domains gauge the degree of lysine methylation." EMBO Rep **7**(4): 397-403.

Kitada, H., M. Miyata, et al. (2003). "Protective role of hydroxysteroid sulfotransferase in lithocholic acid-induced liver toxicity." J Biol Chem **278**(20): 17838-17844.

Kopp, P. (1998). "Resveratrol, a phytoestrogen found in red wine. A possible explanation for the conundrum of the 'French paradox'?" Eur J Endocrinol **138**(6): 619-620.

Kouzarides, T. (2007). "Chromatin modifications and their function." Cell **128**(4): 693-705.

Kozoni, V., G. Tsioulias, et al. (2000). "The effect of lithocholic acid on proliferation and apoptosis during the early stages of colon carcinogenesis: differential effect on apoptosis in the presence of a colon carcinogen." Carcinogenesis **21**(5): 999-1005.

Kramer, A., U. Reineke, et al. (1999). "Spot synthesis: observations and optimizations." J Pept Res **54**(4): 319-327.

Kuipers, F., H. Heslinga, et al. (1986). "Intestinal absorption of lithocholic acid sulfates in the rat: inhibitory effects of calcium." Am J Physiol **251**(2 Pt 1): G189-194.

Kundu, J. K. and Y. J. Surh (2008). "Cancer chemopreventive and therapeutic potential of resveratrol: mechanistic perspectives." Cancer Lett **269**(2): 243-261.

Kuzmichev, A., R. Margueron, et al. (2005). "Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation." Proc Natl Acad Sci U S A **102**(6): 1859-1864.

Lain, S., J. J. Hollick, et al. (2008). "Discovery, in vivo activity, and mechanism of action of a small-molecule p53 activator." Cancer Cell **13**(5): 454-463.

Lapatto, R., T. Blundell, et al. (1989). "X-ray analysis of HIV-1 proteinase at 2.7 A resolution confirms structural homology among retroviral enzymes." Nature **342**(6247): 299-302.

Laskowski, R. A. (1995). "Surfnet - a Program for Visualizing Molecular-Surfaces, Cavities, and Intermolecular Interactions." Journal of Molecular Graphics **13**(5): 323-330.

Lee, K. A., H. Fuda, et al. (2003). "Crystal structure of human cholesterol sulfotransferase (SULT2B1b) in the presence of pregnenolone and 3'-phosphoadenosine 5'-phosphate. Rationale for specificity differences between prototypical SULT2A1 and the SULT2BG1 isoforms." J Biol Chem **278**(45): 44593-44599.

Leonard, S. S., C. Xia, et al. (2003). "Resveratrol scavenges reactive oxygen species and effects radical-induced cellular responses." Biochem Biophys Res Commun **309**(4): 1017-1026.

Lesley, S. A., P. Kuhn, et al. (2002). "Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline." Proc Natl Acad Sci U S A **99**(18): 11664-11669.

Li, X., C. A. Corsa, et al. (2010). "MOF and H4 K16 acetylation play important roles in DNA damage repair by modulating recruitment of DNA damage repair protein Mdc1." Mol Cell Biol **30**(22): 5335-5347.

Likhitwitayawuid, K., K. Sawasdee, et al. (2002). "Flavonoids and stilbenoids with COX-1 and COX-2 inhibitory activity from Dracaena loureiri." Planta Med **68**(9): 841-3.

Liszt, G., E. Ford, et al. (2005). "Mouse Sir2 homolog SIRT6 is a nuclear ADP-ribosyltransferase." J Biol Chem **280**(22): 21313-21320.

Liu, M. C., M. Suiko, et al. (2000). "Mutational analysis of the substrate binding/catalytic domains of human M form and P form phenol sulfotransferases." J Biol Chem **275**(18): 13460-13464.

Liu, Y., R. Gerber, et al. (2008). "High-throughput assays for sirtuin enzymes: a microfluidic mobility shift assay and a bioluminescence assay." Anal Biochem **378**(1): 53-59.

Lombard, D. B., F. W. Alt, et al. (2007). "Mammalian Sir2 homolog SIRT3 regulates global mitochondrial lysine acetylation." Mol Cell Biol **27**(24): 8807-8814.

Longo, V. D. and B. K. Kennedy (2006). "Sirtuins in aging and age-related disease." Cell **126**(2): 257-268.

Lu, J. H., H. T. Li, et al. (2005). "Crystal structure of human sulfotransferase SULT1A3 in complex with dopamine and 3'-phosphoadenosine 5'-phosphate." Biochem Biophys Res Commun **335**(2): 417-423.

Lu, L. Y., Y. C. Hsieh, et al. (2008). "Identification and characterization of two amino acids critical for the substrate inhibition of human dehydroepiandrosterone sulfotransferase (SULT2A1)." Mol Pharmacol **73**(3): 660-668.

Luo, J. and D. C. Altieri (2008). "SIRTing through breast cancer is just a survivin' game." Mol Cell **32**(2): 159-160.

Ma, W., L. J. Stafford, et al. (2007). "GCIP/CCNDBP1, a helix-loop-helix protein, suppresses tumorigenesis." J Cell Biochem **100**(6): 1376-1386.

Makishima, M., T. T. Lu, et al. (2002). "Vitamin D receptor as an intestinal bile acid sensor." Science **296**(5571): 1313-1316.

Marcotte, P. A., P. L. Richardson, et al. (2004). "Fluorescence assay of SIRT protein deacetylases using an acetylated peptide substrate and a secondary trypsin reaction." Anal Biochem **332**(1): 90-99.

Matsui, M., M. Takahashi, et al. (1995). "Structure-activity relationships of alkylamines that inhibit rat liver hydroxysteroid sulfotransferase activities in vitro." Biochem Pharmacol **49**(5): 739-741.

McCord, R. A., E. Michishita, et al. (2009). "SIRT6 stabilizes DNA-dependent protein kinase at chromatin for DNA

double-strand break repair." Aging (Albany NY) **1**(1): 109-121.

McCoy, A. J., R. W. Grosse-Kunstleve, et al. (2005). "Likelihood-enhanced fast translation functions." Acta Crystallographica Section D - Biological Crystallography **61**: 458-464.

McDonagh, T., J. Hixon, et al. (2005). "Microplate filtration assay for nicotinamide release from NAD using a boronic acid resin." Methods **36**(4): 346-350.

Michan, S. and D. Sinclair (2007). "Sirtuins in mammals: insights into their biological function." Biochem J **404**(1): 1-13.

Michishita, E., R. A. McCord, et al. (2008). "SIRT6 is a histone H3 lysine 9 deacetylase that modulates telomeric chromatin." Nature **452**(7186): 492-496.

Michishita, E., R. A. McCord, et al. (2009). "Cell cycle-dependent deacetylation of telomeric histone H3 lysine K56 by human SIRT6." Cell Cycle **8**(16): 2664-2666.

Michishita, E., J. Y. Park, et al. (2005). "Evolutionarily conserved and nonconserved cellular localizations and functions of human SIRT proteins." Mol Biol Cell **16**(10): 4623-4635.

Miksits, M., K. Wlcek, et al. (2009). "Antitumor activity of resveratrol and its sulfated metabolites against human breast cancer cells." Planta Med **75**(11): 1227-1230.

Miller, M., J. Schneider, et al. (1989). "Structure of complex of synthetic HIV-1 protease with a substrate-based inhibitor at 2.3 A resolution." Science **246**(4934): 1149-1152.

Milne, J. C., P. D. Lambert, et al. (2007). "Small molecule activators of SIRT1 as therapeutics for the treatment of type 2 diabetes." Nature **450**(7170): 712-716.

Min, J., J. Landry, et al. (2001). "Crystal structure of a SIR2 homolog-NAD complex." Cell **105**(2): 269-279.

Minor, Z. O. a. W. (1997). "Processing of X-ray Diffraction Data Collected in Oscillation Mode." Methods in Enzymology **276**(Macromolecular Crystallography, part A): 307-326.

Miyai, K., V. M. Price, et al. (1971). "Bile acid metabolism in mammals. Ultrastructural studies on the intrahepatic cholestasis induced by lithocholic and chenodeoxycholic acids in the rat." Lab Invest **24**(4): 292-302.

Mostoslavsky, R., K. F. Chua, et al. (2006). "Genomic instability and aging-like phenotype in the absence of mammalian SIRT6." Cell **124**(2): 315-329.

Mulder, G. J. and E. Scholtens (1977). "Phenol sulphotransferase and uridine diphosphate glucuronyltransferase from rat liver in vivo and vitro. 2,6-Dichloro-4-nitrophenol as selective inhibitor of sulphation."Biochem J **165**(3): 553-559.

Murias, M., N. Handler, et al. (2004). "Resveratrol analogues as selective cyclooxygenase-2 inhibitors: synthesis and structure-activity relationship." Bioorg Med Chem **12**(21): 5571-5578.

Murias, M., W. Jager, et al. (2005). "Antioxidant, prooxidant and cytotoxic activity of hydroxylated resveratrol analogues: structure-activity relationship." Biochem Pharmacol **69**(6): 903-912.

Murias, M., M. Miksits, et al. (2008). "Metabolism of resveratrol in breast cancer cell lines: impact of sulfotransferase 1A1 expression on cell growth inhibition." Cancer Lett **261**(2): 172-182.

Murshudow, G. N., A. A. Vagin, et al. (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallographica Section D - Biological Crystallography **53**: 240-255.

Nady, N., J. Min, et al. (2008). "A SPOT on the chromatin landscape? Histone peptide arrays as a tool for epigenetic research." Trends Biochem Sci **33**(7): 305-313.

Nagata, K. and Y. Yamazoe (2000). "Pharmacogenetics of sulfotransferase." Annu Rev Pharmacol Toxicol **40**: 159-176.

Nahhas, F., S. C. Dryden, et al. (2007). "Mutations in SIRT2 deacetylase which regulate enzymatic activity but not its interaction with HDAC6 and tubulin." Mol Cell Biochem **303**(1-2): 221-30.

Najmanovich, R. J., A. Allali-Hassani, et al. (2007). "Analysis of binding site similarity, small-molecule similarity and experimental binding profiles in the human cytosolic sulfotransferase family." Bioinformatics **23**(2): e104-109.

Nakagawa, T., D. J. Lomb, et al. (2009). "SIRT5 Deacetylates carbamoyl phosphate synthetase 1 and regulates the urea cycle." Cell **137**(3): 560-570.

Napper, A. D., J. Hixon, et al. (2005). "Discovery of indoles as potent and selective inhibitors of the deacetylase SIRT1." J Med Chem **48**(25): 8045-8054.

Narisawa, T., N. E. Magadia, et al. (1974). "Promoting effect of bile acids on colon carcinogenesis after intrarectal instillation of N-methyl-N'-nitro-N-nitrosoguanidine in rats." J Natl Cancer Inst **53**(4): 1093-1097.

North, B. J., B. L. Marshall, et al. (2003). "The human Sir2 ortholog, SIRT2, is an NAD^+-dependent tubulin deacetylase." Mol Cell **11**(2): 437-444.

North, B. J. and E. Verdin (2007). "Interphase nucleo-cytoplasmic shuttling and localization of SIRT2 during mitosis."

PLoS One **2**(8): e784.

Oelberg, D. G., M. V. Chari, et al. (1984). "Lithocholate glucuronide is a cholestatic agent." J Clin Invest **73**(6): 1507-1514.

Ota, H., E. Tokunaga, et al. (2006). "Sirt1 inhibitor, Sirtinol, induces senescence-like growth arrest with attenuated Ras-MAPK signaling in human cancer cells." Oncogene **25**(2): 176-185.

Otwinowski, Z. and W. Minor (1997). "Processing of X-ray diffraction data collected in oscillation mode." Macromolecular Crysallography, Pt a **276**: 307-326.

Outeiro, T. F., E. Kontopoulos, et al. (2007). "Sirtuin 2 inhibitors rescue alpha-synuclein-mediated toxicity in models of Parkinson's disease." Science **317**(5837): 516-519.

Ozawa, S., K. Nagata, et al. (1995). "Primary structures and properties of two related forms of aryl sulfotransferases in human liver." Pharmacogenetics **5** Spec No: S135-140.

Pacholec, M., J. E. Bleasdale, et al. (2010). "SRT1720, SRT2183, SRT1460, and resveratrol are not direct activators of SIRT1." J Biol Chem **285**(11): 8340-51.

Pacifici, G. M. (2004). "Inhibition of human liver and duodenum sulfotransferases by drugs and dietary chemicals: a review of the literature." Int J Clin Pharmacol Ther **42**(9): 488-495.

Pan, P. W., J. L. Feldman, et al. (2011). "Structure and Biochemical Functions of SIRT6." J Biol Chem **286**(16): 14575-14587.

Pedersen, L. C., E. Petrotchenko, et al. (2002). "Crystal structure of the human estrogen sulfotransferase-PAPS complex: evidence for catalytic role of Ser137 in the sulfuryl transfer reaction." J Biol Chem **277**(20): 17928-17932.

Pedersen, L. C., E. V. Petrotchenko, et al. (2000). "Crystal structure of SULT2A3, human hydroxysteroid sulfotransferase." FEBS Lett **475**(1): 61-64.

Perrakis, A., R. Morris, et al. (1999). "Automated protein model building combined with iterative structure refinement." Nature Structural Biology **6**(5): 458-463.

Petrotchenko, E. V., M. E. Doerflein, et al. (1999). "Substrate gating confers steroid specificity to estrogen sulfotransferase." J Biol Chem **274**(42): 30019-30022.

Picard, F., M. Kurtev, et al. (2004). "Sirt1 promotes fat mobilization in white adipocytes by repressing PPAR-gamma." Nature **429**(6993): 771-776.

Pils, B. and J. Schultz (2004). "Evolution of the multifunctional protein tyrosine phosphatase family." Mol Biol Evol **21**(4): 625-631.

Porcu, M. and A. Chiarugi (2005). "The emerging therapeutic potential of sirtuin-interacting drugs: from cell death to lifespan extension." Trends Pharmacol Sci **26**(2): 94-103.

Qian, Y. M., X. J. Sun, et al. (2001). "Targeted disruption of the mouse estrogen sulfotransferase gene reveals a role of estrogen metabolism in intracrine and paracrine estrogen regulation." Endocrinology **142**(12): 5342-5350.

Raftogianis, R. B., T. C. Wood, et al. (1999). "Human phenol sulfotransferases SULT1A2 and SULT1A1: genetic polymorphisms, allozyme properties, and human liver genotype-phenotype correlations." Biochem Pharmacol **58**(4): 605-616.

Rao, S. I. and M. W. Duffel (1991). "Inhibition of aryl sulfotransferase by carboxylic acids." Drug Metab Dispos **19**(2): 543-545.

Rehse, P. H., M. Zhou, et al. (2002). "Crystal structure of human dehydroepiandrosterone sulphotransferase in complex with substrate." Biochem J **364**(Pt 1): 165-171.

Reineke, U., R. Sabat, et al. (1996). "Mapping protein-protein contact sites using cellulose-bound peptide scans." Mol Divers **1**(3): 141-148.

Richard, K., R. Hume, et al. (2001). "Sulfation of thyroid hormone and dopamine during human development: ontogeny of phenol sulfotransferases and arylsulfatase in liver, lung, and brain." J Clin Endocrinol Metab **86**(6): 2734-2742.

Riches, Z., E. L. Stanley, et al. (2009). "Quantitative evaluation of the expression and activity of five major sulfotransferases (SULTs) in human tissues: the SULT "pie"." Drug Metab Dispos **37**(11): 2255-2261.

Rodgers, J. T., C. Lerin, et al. (2005). "Nutrient control of glucose homeostasis through a complex of PGC-1alpha and SIRT1." Nature **434**(7029): 113-118.

Rogina, B. and S. L. Helfand (2004). "Sir2 mediates longevity in the fly through a pathway related to calorie restriction." Proc Natl Acad Sci U S A **101**(45): 15998-16003.

Rubin, G. L., A. J. Harrold, et al. (1999). "Regulation of sulphotransferase expression in the endometrium during the menstrual cycle, by oral contraceptives and during early pregnancy." Mol Hum Reprod **5**(11): 995-1002.

Rupp, B. (2003). "High-throughput crystallography at an affordable cost: the TB Structural Genomics Consortium

Crystallization Facility." Acc Chem Res **36**(3): 173-181.

Safe, S. H. (1994). "Polychlorinated biphenyls (PCBs): environmental impact, biochemical and toxic responses, and implications for risk assessment." Crit Rev Toxicol **24**(2): 87-149.

Sakakibara, Y., M. Suiko, et al. (2002). "Highly conserved mouse and human brain sulfotransferases: molecular cloning, expression, and functional characterization." Gene **285**(1-2): 39-47.

Sakakibara, Y., Y. Takami, et al. (1998). "Localization and functional analysis of the substrate specificity/catalytic domains of human M-form and P-form phenol sulfotransferases." J Biol Chem **273**(11): 6242-6247.

Sakakibara, Y., K. Yanagisawa, et al. (1998). "Molecular cloning, expression, and characterization of novel human SULT1C sulfotransferases that catalyze the sulfonation of N-hydroxy-2-acetylaminofluorene." J Biol Chem **273**(51): 33929-33935.

Sanders, B. D., B. Jackson, et al. (2010). "Structural basis for sirtuin function: what we know and what we don't." Biochim Biophys Acta **1804**(8): 1604-1616.

Sanders, B. D., K. Zhao, et al. (2007). "Structural basis for nicotinamide inhibition and base exchange in Sir2 enzymes." Mol Cell **25**(3): 463-472.

Sauve, A. A., I. Celic, et al. (2001). "Chemistry of gene silencing: the mechanism of NAD+-dependent deacetylation reactions." Biochemistry **40**(51): 15456-15463.

Schlicker, C., M. Gertz, et al. (2008). "Substrates and regulation mechanisms for the human mitochondrial sirtuins Sirt3 and Sirt5." J Mol Biol **382**(3): 790-801.

Schrag, M. L., D. Cui, et al. (2004). "Sulfotransferase 1E1 is a low km isoform mediating the 3-O-sulfation of ethinyl estradiol." Drug Metab Dispos **32**(11): 1299-1303.

Schuetz, A., J. Min, et al. (2007). "Structural basis of inhibition of the human NAD^+-dependent deacetylase SIRT5 by suramin." Structure **15**(3): 377-389.

Schwer, B., B. Schumacher, et al. (2010). "Neural sirtuin 6 (Sirt6) ablation attenuates somatic growth and causes obesity." Proc Natl Acad Sci U S A **107**(50): 21790-21794.

Senisterra, G. A., E. Markin, et al. (2006). "Screening for ligands using a generic and high-throughput light-scattering-based assay." J Biomol Screen **11**(8): 940-948.

Service, R. F. (2002). "Structural genomics. Tapping DNA for structures produces a trickle." Science **298**(5595): 948-950.

Shevtsov, S., E. V. Petrotchenko, et al. (2003). "Crystallographic analysis of a hydroxylated polychlorinated biphenyl (OH-PCB) bound to the catalytic estrogen binding site of human estrogen sulfotransferase." Environ Health Perspect **111**(7): 884-888.

Shogren-Knaak, M. A. (2007). "Mimicking methylated histones." ACS Chem Biol **2**(4): 225-227.

Shogren-Knaak, M. A., C. J. Fry, et al. (2003). "A native peptide ligation strategy for deciphering nucleosomal histone modifications." J Biol Chem **278**(18): 15744-15748.

Smith, B. C. and J. M. Denu (2006). "Sir2 protein deacetylases: evidence for chemical intermediates and functions of a conserved histidine." Biochemistry **45**(1): 272-282.

Smith, B. C., W. C. Hallows, et al. (2009). "A continuous microplate assay for sirtuins and nicotinamide-producing enzymes." Anal Biochem **394**(1): 101-109.

Sobolev, V., A. Sorokine, et al. (1999). "Automated analysis of interatomic contacts in proteins." Bioinformatics **15**(4): 327-332.

Someya, S., W. Yu, et al. "Sirt3 mediates reduction of oxidative damage and prevention of age-related hearing loss under caloric restriction." Cell **143**(5): 802-812.

Stiehl, A. (1974). "Sulfation of bile salts: a new metabolic pathway." Digestion **11**(5-6): 406-413.

Stiehl, A., D. L. Earnest, et al. (1975). "Sulfation and renal excretion of bile salts in patients with cirrhosis of the liver." Gastroenterology **68**(3): 534-544.

Takikawa, H., J. Tomita, et al. (1991). "Cytotoxic effect and uptake mechanism by isolated rat hepatocytes of lithocholate and its glucuronide and sulfate." Biochim Biophys Acta **1091**(2): 173-178.

Tanner, K. G., J. Landry, et al. (2000). "Silent information regulator 2 family of NAD- dependent histone/protein deacetylases generates a unique product, 1-O-acetyl-ADP-ribose." Proc Natl Acad Sci U S A **97**(26): 14178-14182.

Tanno, M., J. Sakamoto, et al. (2007). "Nucleocytoplasmic shuttling of the NAD^+-dependent histone deacetylase SIRT1." J Biol Chem **282**(9): 6823-6832.

Taylor, D. M., M. M. Maxwell, et al. (2008). "Biological and potential therapeutic roles of sirtuin deacetylases." Cell Mol Life Sci **65**(24): 4000-4018.

Tennen, R. I., E. Berber, et al. (2010). "Functional dissection of SIRT6: identification of domains that regulate histone deacetylase activity and chromatin localization." Mech Ageing Dev **131**(3): 185-192.

Thompson, J. D., D. G. Higgins, et al. (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Res **22**(22): 4673-4680.

Tissenbaum, H. A. and L. Guarente (2001). "Increased dosage of a sir-2 gene extends lifespan in Caenorhabditis elegans." Nature **410**(6825): 227-230.

Todd, A. E., C. A. Orengo, et al. (2001). "Evolution of function in protein superfamilies, from a structural perspective." Journal of Molecular Biology **307**(4): 1113-1143.

Toepert, F., T. Knaute, et al. (2003). "Combining SPOT synthesis and native peptide ligation to create large arrays of WW protein domains." Angew Chem Int Ed Engl **42**(10): 1136-1140.

Tsubota, T., C. E. Berndsen, et al. (2007). "Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes." Mol Cell **25**(5): 703-712.

Urpi-Sarda, M., O. Jauregui, et al. (2005). "Uptake of diet resveratrol into the human low-density lipoprotein. Identification and quantification of resveratrol metabolites by liquid chromatography coupled with tandem mass spectrometry." Anal Chem **77**(10): 3149-3155.

Vagin, A. and A. Teplyakov (1997). "MOLREP: An automated program for molecular replacement." Journal of Applied Crystallography **30**: 1022-1025.

Vakhrusheva, O., C. Smolka, et al. (2008). "Sirt7 increases stress resistance of cardiomyocytes and prevents apoptosis and inflammatory cardiomyopathy in mice." Circ Res **102**(6): 703-710.

Vaquero, A., M. Scher, et al. (2004). "Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin." Mol Cell **16**(1): 93-105.

Vaquero, A., M. B. Scher, et al. (2006). "SirT2 is a histone deacetylase with preference for histone H4 Lys 16 during mitosis." Genes Dev **20**(10): 1256-1261.

Varghese, J. N. (1999). "Development of neuraminidase inhibitors as anti-influenza virus drugs." Drug Dev. Res. **46**(3-4): 176-196.

Vaziri, H., S. K. Dessain, et al. (2001). "hSIR2(SIRT1) functions as an NAD-dependent p53 deacetylase." Cell **107**(2): 149-159.

Vedadi, M., F. H. Niesen, et al. (2006). "Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination." Proc Natl Acad Sci U S A **103**(43): 15835-15840.

Veronese, M. E., W. Burgess, et al. (1994). "Functional characterization of two human sulphotransferase cDNAs that encode monoamine- and phenol-sulphating forms of phenol sulphotransferase: substrate kinetics, thermal-stability and inhibitor-sensitivity studies." Biochem J **302** (Pt 2): 497-502.

Visser, T. J. (1996). "Pathways of thyroid hormone metabolism." Acta Med Austriaca **23**(1-2): 10-16.

Walle, T., E. A. Eaton, et al. (1995). "Quercetin, a potent and specific inhibitor of the human P-form phenosulfotransferase." Biochem Pharmacol **50**(5): 731-734.

Walle, T., F. Hsieh, et al. (2004). "High absorption but very low bioavailability of oral resveratrol in humans." Drug Metab Dispos **32**(12): 1377-1382.

Wang, J., J. L. Falany, et al. (1998). "Expression and characterization of a novel thyroid hormone-sulfating form of cytosolic sulfotransferase from human liver." Mol Pharmacol **53**(2): 274-282.

Wang, R. H., K. Sengupta, et al. (2008). "Impaired DNA damage response, genome instability, and tumorigenesis in SIRT1 mutant mice." Cancer Cell **14**(4): 312-323.

Wegener, D., C. Hildmann, et al. (2003). "Improved fluorogenic histone deacetylase assay for high-throughput-screening applications." Anal Biochem **321**(2): 202-208.

Weinshilboum, R. M., D. M. Otterness, et al. (1997). "Sulfation and sulfotransferases 1: Sulfotransferase molecular biology: cDNAs and genes." Faseb J **11**(1): 3-14.

Wenzel, E. and V. Somoza (2005). "Metabolism and bioavailability of trans-resveratrol." Mol Nutr Food Res **49**(5): 472-481.

Wood, E. R., A. T. Truesdale, et al. (2004). "A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells." Cancer Res **64**(18): 6652-6659.

Wu, J. M., Z. R. Wang, et al. (2001). "Mechanism of cardioprotection by resveratrol, a phenolic antioxidant present in red wine (Review)." Int J Mol Med **8**(1): 3-17.

Wysocka, J. (2006). "Identifying novel proteins recognizing histone modifications using peptide pull-down assay."

Methods **40**(4): 339-343.

Xiao, C., H. S. Kim, et al. (2010). "SIRT6 deficiency results in severe hypoglycemia by enhancing both basal and insulin-stimulated glucose uptake in mice." J Biol Chem **285**(47): 36776-36784.

Xu, C., C. Bian, et al. (2010). "Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2)." Proc Natl Acad Sci U S A **107**(45): 19266-19271.

Yamazoe, Y., K. Nagata, et al. (1999). "Sulfotransferase catalyzing sulfation of heterocyclic amines." Cancer Lett **143**(2): 103-107.

Yang, B., B. M. Zwaans, et al. (2009). "The sirtuin SIRT6 deacetylates H3 K56Ac in vivo to promote genomic stability." Cell Cycle **8**(16): 2662-2663.

Yeung, F., J. E. Hoberg, et al. (2004). "Modulation of NF-kappaB-dependent transcription and cell survival by the SIRT1 deacetylase." EMBO J **23**(12): 2369-2380.

Zhang, Q., D. W. Piston, et al. (2002). "Regulation of corepressor function by nuclear NADH." Science **295**(5561): 1895-1897.

Zhang, Y., Q. Au, et al. (2009). "Identification of a small molecule SIRT2 inhibitor with selective tumor cytotoxicity." Biochem Biophys Res Commun **386**(4): 729-733.

Zhao, K., X. Chai, et al. (2003). "Structure and autoregulation of the yeast Hst2 homolog of Sir2." Nat Struct Biol **10**(10): 864-871.

Zhao, K., X. Chai, et al. (2003). "Structure of the yeast Hst2 protein deacetylase in ternary complex with 2'-O-acetyl ADP ribose and histone peptide." Structure **11**(11): 1403-1411.

Zhao, K., X. Chai, et al. (2004). "Structure and substrate binding properties of cobB, a Sir2 homolog protein deacetylase from Escherichia coli." J Mol Biol **337**(3): 731-741.

Zhao, K., R. Harshaw, et al. (2004). "Structural basis for nicotinamide cleavage and ADP-ribose transfer by NAD(+)-dependent Sir2 histone/protein deacetylases." Proc Natl Acad Sci U S A **101**(23): 8563-8568.

Zhong, L., A. D'Urso, et al. (2010). "The histone deacetylase Sirt6 regulates glucose homeostasis via Hif1alpha." Cell **140**(2): 280-293.

Zhu, X., M. E. Veronese, et al. (1996). "cDNA cloning and expression of a new form of human aryl sulfotransferase." Int J Biochem Cell Biol **28**(5): 565-571.