

Structural & functional studies of L-PGDS and SMPDL3A enzymes in lipid signaling family

Lim, Sing Mei

2015

Lim, S. M. (2016). Structural & functional studies of L-PGDS and SMPDL3A enzymes in lipid signaling family. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/66456>

<https://doi.org/10.32657/10356/66456>



**NANYANG
TECHNOLOGICAL
UNIVERSITY**

**STRUCTURAL & FUNCTIONAL STUDIES OF
L-PGDS AND SMPDL3A,
ENZYMES IN LIPID SIGNALING FAMILY**

LIM SING MEI

SCHOOL OF BIOLOGICAL SCIENCES

2015

From the Department of Medical Biochemistry and Biophysics
Karolinska Institutet, Sweden, 2015

STRUCTURAL & FUNCTIONAL STUDIES OF L-PGDS AND SMPDL3A, ENZYMES IN LIPID SIGNALING FAMILY

Sing Mei LIM



**Karolinska
Institutet**

Nanyang Technological University-Karolinska Institutet Joint Ph.D

Stockholm 2015

All previously published papers were reproduced with permission from the publisher.

Published by Karolinska Institutet.

Printed by E-print AB

© Sing Mei Lim, 2015

ISBN 978-91-7676-170-0

Structural and functional studies of L-PGDS & SMPDL3a, enzymes in lipid signaling family

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Sing Mei LIM

Principal Supervisor:

Pär Nordlund
Karolinska Institutet
Department of Medical Biochemistry and
Biophysics
Division of Biophysics

Co-supervisor:

Konstantin Pervushin
Nanyang Technological University
School of Biological Sciences
Division Structural Biology and Biochemistry

Opponent:

Andrea Mattevi
University of Pavia
Department of Biology and Biotechnology

Examination Board:

Sven Pettersson
Nanyang Technological University
Lee Kong Cien School of Medicine
&
Karolinska Institutet
Department of Microbiology, Tumor and Cell
biology

Günter Schneider
Karolinska Institutet
Department of Medical Biochemistry and
Biophysics
Division of Molecular Structural Biology

Mats Sandgren
Department of Chemistry and Biotechnology
Swedish University of Agricultural Sciences

To my beloved

*Would it not be strange if a universe without purpose accidentally
created humans who are so obsessed with purpose?*

Sir John Templeton

ABSTRACT

Enzymes are indispensable in maintaining the biological system. They metabolize complex molecules to supply nutrients, to produce energy, to regulate transcription of gene expression, and to control the concentration of effective signaling molecules in a cell, thus maintaining the homeostasis of biological system. This thesis summarizes the study of the structure and function of two enzymes in lipid signaling family using integrative application of X-ray crystallography, solution NMR spectroscopy, light scattering, ITC and thermal shift assay.

Lipocalin prostaglandin D synthase (L-PGDS) is a tissue specific prostaglandin D₂ producing enzyme with a lipocalin fold. Apart from its enzymatic role, it is known to act as a lipophilic ligand carrier. Crystal structure of human L-PGDS and substrate analog altogether with NMR spectroscopy experiments revealed binding sites for substrate catalysis and entry. NMR titration experiments with membrane mimetic showed that L-PGDS has intrinsic membrane binding affinity depending on the ligand bound. These results allowed a model of substrate catalysis and product egression to be proposed, hence, converging the enzymatic and transporter role that has been reported in literature previously. Since prostaglandin D₂ is a pivotal inflammatory signaling molecule, molecular understanding of L-PGDS is important to facilitate future regulation of the prostaglandin isomerase. The dynamics of substrate-product exchange may guide future design of this lipophilic carrier as vehicle for drug delivery.

The second enzyme, human acid sphingomyelinase like 3a (SMPDL3a), belongs to a metallophosphodiesterase family and shares close sequence identity with human acid sphingomyelinase (aSMase). SMPDL3a's structure is reported for the first time revealing its binuclear catalytic core site bound with Zn metal. Even though it was presumed to be part of the lipid hydrolase family, enzymatic assays showed that it metabolizes nucleotides and modified nucleotides like CDP-choline, CDP-ethanolamine and ADP-ribose. Subsequently, CDP-choline soaked crystal revealed 5' cytidine monophosphate (CMP) ligand bound in the catalytic site due to spontaneous catalysis. Its α -phosphate forms key interactions with histidine residues in the binuclear center. Based on this CMP-enzyme structure, general catalytic mechanism of aSMase family can be proposed. Besides, SMPDL3a also serves as a template for aSMase catalytic domain homology modeling. Further study on enzymes in the acid sphingomyelinase family can now be guided by the newly available structural information.

LIST OF SCIENTIFIC PAPERS

- I. **Sing Mei Lim**, Dan Chen, Hsiangling Teo, Annette Roos, Anna Elisabet Jansson, Tomas Nyman, Lionel Trésaugues, Konstantin Pervushin, and Pär Nordlund. Structural and dynamic insights into substrate binding and catalysis of human lipocalin prostaglandin D synthase. *The Journal of Lipid Research*. June 2013; 54(6):1630-1643.

- II. **Sing Mei Lim**, Kit Yeung, Lionel Trésaugues, Hsiangling Teo, and Pär Nordlund. The structure and mechanism of human sphingomyelin phosphodiesterase like 3A, an acid sphingomyelinase homolog - with novel nucleotide hydrolase activity. *FEBS Journal*. March 2016. 283(2016):1107-1123.

Paper not included in this thesis:

Sing Mei Lim, Kit Yeung, Hsiangling Teo and Pär Nordlund. Structure and mutational analysis of human 15LOX-2 apoenzyme provides insights into substrate binding. *Manuscript*.

TABLE OF CONTENTS

1	Enzymes – Catalysts of the chemistry of life	3
1.1	Enzyme structure and function	5
1.2	The line between specificity and promiscuity	6
1.3	Enzymes in metabolic pathways.....	7
2	Aim of thesis	9
3	From Protein to Structure and Function	11
3.1	Recombinant Protein expression.....	11
3.1.1	Bacterial expression systems	11
3.1.2	Insect cells system.....	12
3.1.3	Protein purification	13
3.2	Thermal stability assay	14
3.2.1	Thermal aggregation method.....	14
3.2.2	Dot-blot method	15
3.3	Isothermal titration calorimetry (ITC)	16
3.4	Multi-angle Light Scattering (MALS)	17
3.5	Macromolecular X-ray crystallography.....	18
3.5.1	Principles of X-ray diffraction.....	18
3.5.2	Protein crystallization	19
3.5.3	Data collection	20
3.5.4	Solving the phase problem.....	21
3.6	Solution NMR spectroscopy	23
3.6.1	1D experiment : ^1H	24
3.6.2	2D experiment: Heteronuclear Single Quantum Coherence (HSQC)	24
3.6.3	3D experiments: HNCA, CBCA(CO)NH, ^{15}N NOESY-HSQC, HNCO	27
4	Paper I: The study of Lipocalin prostglandin D synthase (L-PGDS).....	29
4.1	L-PGDS in eicosanoids signalling	29
4.2	Results summary	30
5	Paper II: Unraveling a novel phosphodiesterase in structure & function	32
5.1	Sphingomyelinase phosphodiesterase like 3A (SMPDL3a)	32
5.1.1	ASMase in sphingolipid hydrolysis.....	32
5.1.2	Calcineurin-like-phosphodiesterase family	32
5.2	Results summary	33
6	Acknowledgements	36
7	References	38
8	Appendices.....	42

LIST OF ABBREVIATIONS

L-PGDS	Lipocalin Prostaglandin D synthase
ITC	Isothermal titration calorimetry
MALS	Multiangle Light scattering
HSQC	Heteronuclear Single Quantum coherence
NOESY	Nuclear Overhauser Effect spectroscopy
Sf9	Spodoptera Frugiperda
SMPDL3a	Sphingomyelin phosphodiesterase like 3a
ASMase	Acid sphingomyelinase
IMAC	Immobilization affinity chromatography
SEC	Size exclusion column
TEV	Tobacco Etch Virus
PGH ₂	Prostaglandin H ₂
MME	Monomethyl ether
SAD	Single-wavelength Anomalous Dispersion
MAD	Multiple-wavelength Anomalous Dispersion
SIR	Single isomorphous replacement
MIR	Multiple isomorphous replacement
RF	Radio frequency
FID	Free induction decay
CSP	Chemical shifts perturbation
DPC	Dodecylphosphocholine
c.m.c	Critical micelle concentration
AA	Arachidonic acid
COX	Cyclooxygenase
LOX	Lipoxygenase
H-PGDS	Hematopoietic prostaglandin D ₂ synthase
SA	Substrate analog
PA	Product analog
NPD	Niemann Pick Disease
PAP	Purple acid phosphatase
ATP	Adenosine triphosphate
CDP	Cytidine diphosphate
CMP	Cytidine monophosphate
cAMP	Cyclic adenosine monophosphate
CLP	Calcineurin like phosphodiesterase

1 ENZYMES – CATALYSTS OF THE CHEMISTRY OF LIFE

In the mid-18th century, scientists began to observe that saliva is able to breakdown starch [1], and almost a few decades later, Payen and Persoz discovered and isolated the starch-decomposing compound which they called “diastase”, now commonly known as amylase [2]. Twenty years after the release of Payen’s discovery, Louis Pasteur proposed that a “vital force” in yeast promotes fermentation of sugar to alcohol in 1858 [3] and Wilhem Kühne used the word “enzyme” to describe the process of fermentation. It was subsequently used to refer to the protein that catalyzes the biochemistry reaction. In fact, the term “catalyst” was coined by Swedish chemist, Jöns Jakob Berzelius in 1835 based on work from other scientists whom he collected and observed. An excerpt of his published statement:

“It is then shown that several simple and compound bodies, soluble and insoluble, have the property of exercising on other bodies and action very different from chemical affinity. The body effecting the changes does not take part in the reaction and remains unaltered through the reaction.” – [4]

This sets the fundamental foundation for the concept of catalysis. Now, we understand that a catalyst acts to speed up a chemical reaction by lowering its activation energy while remaining unchanged at the end of the reaction (Figure 1). This energy barrier is important to prevent chemical reactions from occurring spontaneously. However, if unaided, most chemical reactions would occur very slowly due to the high activation energy required. Therefore the presence of a catalyst is vital to increase the speed of reactions and allows only certain reaction to take place at a time. These controlled-chemical processes in a biological system are normally governed by enzymes.

At the end of 19th century, many scientists began to draw great interest in the study of enzymes and this have contributed significantly in shaping the current understanding of enzymology and cellular processes where they are involved. Buchner was the first to describe enzymatic reaction outside of living cells and he was awarded the Nobel prize in Chemistry for his discovery of cell-free fermentation [5]. This enabled the application of cell free enzyme systems in food and wine processing during the 20th century.

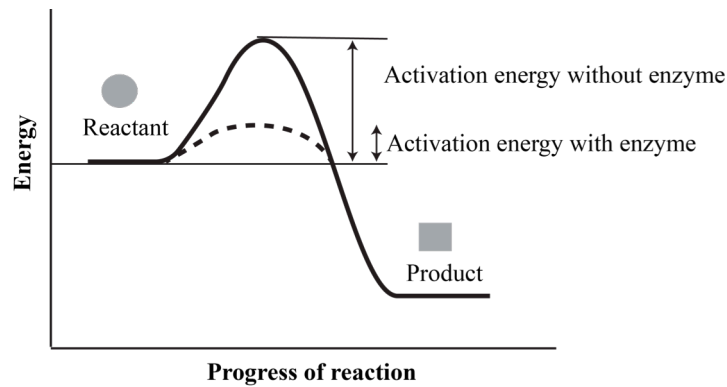
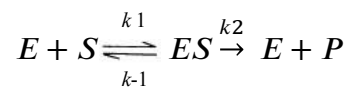


Figure 1: Energy barrier of chemical reactions (also termed activation energy) is lowered in the presence of enzyme.

Around 1890s, scientists began to study the reaction rate of enzymatic induced biochemical reaction when they realize that enzymatic reaction can be observed outside of the living cells. Many deduced that enzyme performs its reaction by forming complex with its substrate and a chemist, Emil Fischer, proposed a lock and key model to explain the enzyme-substrate interaction [6]. Subsequently L. Michaelis and M. Menten proposed the basic idea of enzymatic reaction in the following reaction scheme.



For a single transition state, the reaction would result in forming the first intermediate (ES) before proceeding to yield the product (P). The rate of constant k_1 described the binding of substrate to enzyme whereas the k_2 described the catalytic rate of reaction turning from substrate to product. The process can also be reversed with rate constant k_{-1} . When a reaction reaches steady state, the ES intermediate concentration is assumed to be relatively constant where the rate of ES formation is close to or same as the rate of ES dissociation. This steady state approximation allows derivation of the Michaelis-Menten equation [7].

$$V_0 = \frac{v_{max}[S]}{K_M + [S]}$$

The Michaelis-Menten equation explains kinetics of enzymatic reaction. From the equation, the reaction rate dependence on substrate and enzyme concentration can be explained. For example, a low K_M value indicates a strong binding between substrate and enzyme in which small amount of substrate concentration result in 50% of maximum reaction velocity. Enzyme kinetics enables scientists to make important observations such as measuring the change of reaction rate with respect to change of pH and temperature. This pH and temperature dependence is attributed to the protein nature of enzyme but the identity of

enzyme was still unknown at that time. It was not until 1926 when J. B. Sumner showed that the enzyme urease is a protein and it can be isolated and crystallized. Despite confronted with much skepticism, his findings were soon confirmed by J. H. Northrop and W.M. Stanley [8-11].

It is now well established that enzymes are versatile catalysts that can have exquisite selectivity. They facilitate most chemical reactions in biological system including processes controlling gene expression, signal transduction, energy production, molecular motors and cellular homeostasis. The enzyme's function can be assisted by catalytic cofactors such as prosthetic group, metal ions and co-enzyme, or further regulated by posttranslational modifications, interacting proteins or allosteric modulation. Most enzymes are proteins but many proteins are not enzymes. Proteins have three-dimensional structures comprising a chain of amino acids connected by peptide bonds. Apart from proteins, RNA has also being identified with enzymatic ability [12]. Nonetheless, this thesis will focus only on protein enzymes.

1.1 ENZYME STRUCTURE AND FUNCTION

Sumner's urease extraction and crystallization work did not just provided a breakthrough in the understanding of enzymes but it also set the stage for much later work in the field of structural biology. From 1912 to 1930s, there were many small chemical structures being determined using X-ray crystallography. The physical method was introduced into the biological world when Crowfoot, D. first demonstrated that protein crystals showed distinct X-ray diffraction patterns like chemical crystals [13]. Based on fiber diffraction method, J. Watson, F. Crick and R. Franklin first uncovered the structure of DNA nucleic acid [14]. It was an extraordinary breakthrough and an important step towards the formulation of the central dogma of molecular biology that explains the flow of information from DNA to mRNA and subsequently translated to proteins. The excitement of viewing DNA structure spurred scientist like M. Perutz and J.C. Kendrew to work towards the first protein structure – myoglobin in just five years later [15]. It was soon followed by the first enzyme structure, the hen egg white lysozyme, in 1965 [16, 17].

Structural elucidations of enzymes have contributed significantly in establishing the principles of enzyme function. The emergence of serine proteases structures including trypsin, pepsin, chymotrypsin and papain gave early insights into the structural basis for catalysis and substrate specificity of these enzymes [18, 19]. Following the structural information explosion, especially in the last 20 years, the field now provides a comprehensive

view of many important biological systems on the structural level, including enzyme functionalities. Since many drug targets are enzymes, these structural information, and accompanying strategies to rapidly determine protein structure, now serve as a valuable source to guide rational drug design.

1.2 THE LINE BETWEEN SPECIFICITY AND PROMISCUITY

Enzyme has an active site where the substrate is bound and chemical reaction takes place. Fischer's lock and key model in 1894 proposed that the substrate has to be perfectly complimentary to the active site for productive interactions [6]. In this model, the specificity of enzyme is governed by the fact that geometrically incompatible substrate would not be the right 'key' to activate the 'lock' (enzyme) (Figure 2A). In 1958 D.E. Koshland suggested a modification of the prevailing model to the 'induced-fit model' based on his study on hexokinase [20]. In the 'induced-fit model', both the enzyme and substrate were deemed to be moderately flexible and the encounter of enzyme-substrate in the correct orientation would induce conformational change in the active site to promote catalysis (Figure 2B). Scientists gradually disfavor the 'lock-and-key' model when more and more enzyme structures with and without their substrate(s) have been unveiled. These atomic resolution structures showed that enzymes in their substrate-bound and unbound states could adopt different conformations.

The enzyme substrate interactions are often quite specific in order to tightly control the *in vivo* chemical transformations. As a result, enzymes are generally classified by the types of chemical transformation they catalyze which usually include a defined substrate. However, there are also many enzymes which have a relatively broad substrate specificity such as pyruvate decarboxylase [21], pepsin [22] and chymotrypsin [23]. The phenomenon where enzyme catalyzes substrate other than the substrate it physiologically specialized is termed substrate promiscuity [24, 25]. Substrate promiscuity can be achieved by for example structural plasticity at the active site where it could adopt alternate conformations [26, 27].

The molecular level of specificity can also be regulated by gene expression, compartmentalization and allosteric modification to impart specificity of action at the system level. For example, many enzymes are not constitutively expressed or always in the active form, instead of protein level, localization and activity of enzyme can be controlled by cellular or external stimulation. Even if a promiscuous enzyme was expressed constitutively, its cellular localization, activation by effectors and availability of substrates set boundaries for

its enzymatic action. Nonetheless, most enzymes are still highly specific and efficient in their role.

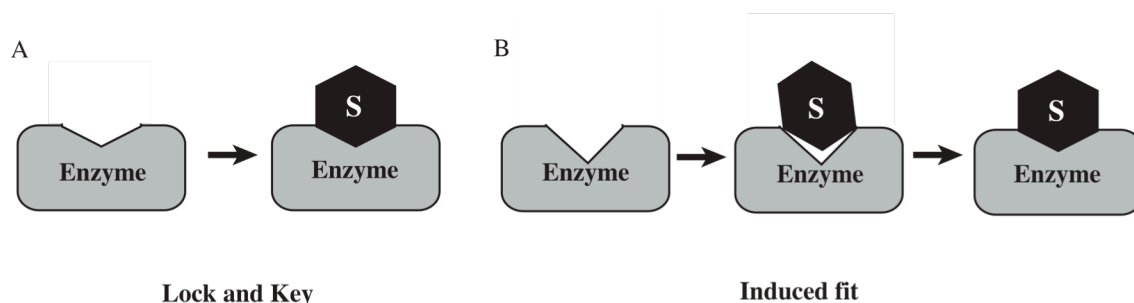


Figure 2. Cartoon depiction of Fischer's lock and key model in (A) versus Koshland's Induced-fit model in (B)

1.3 ENZYMES IN METABOLIC PATHWAYS

Enzymes, given the above-discussed nature and modus operandi, are indispensable and paramount in all types pathways and processes in biological systems. A series of enzymes can be localized to catalyze stepwise biochemical reactions that are more or less interdependent of one another. The arachidonic acid metabolism pathway is one such example. The initial substrate phospholipid is converted to an arachidonic acid (AA) as the first committed step in the pathway. The long hydrocarbon chain of AA can be further oxygenated via three specific enzymes into three distinct signaling pathways. Therefore metabolism of arachidonic acids generates a class of lipid signaling molecules known as eicosanoids. Eicosanoids signaling is further elaborated in Chapter 4 in the study of enzyme lipocalin prostaglandin D synthase (L-PGDS).

The second enzyme in our study is a homolog of the lipid modifying enzyme, acid sphingomyelinase (aSMase), that is involved in ceramide metabolism. Ceramide belongs to a class of lipids known as sphingolipids, which carries a sphingoid backbone and it can be produced via three pathways; de novo synthesis from palmitate [28] and serine, salvage pathway from recycled sphingosines [29] or breakdown of sphingomyelin [30]. Enzyme aSMase produces ceramide by breaking the phosphodiester bond of sphingomyelin. However, our studies and other literature reports found that the conserved catalytic domain of the acid sphingomyelinase like 3a (SMPDL3a) enzyme selects nucleotide substrate instead of sphingomyelin [31]. This is a classical example of enzymes with conserved structural domains but with modified substrate pocket for substrate specificity. From an evolutionary point of view, this is efficient way to develop specialization. Nonetheless, the exact

physiological substrate and biological function of SMPDL3a is unfortunately still unclear. The work in Paper II helps to establish a structural and biochemical framework for future study of this protein and the definition of its biological role.

2 AIM OF THESIS

A total of 110,789 protein, DNA and RNA structures were deposited in the Protein Data Bank (www.rcsb.org) as of August 2015. Among those, 98720 were solved with X-ray crystallography, 11077 by nuclear magnetic resonance (NMR) spectroscopy, and 809 by cryo-electron microscopy (cryo-EM) [32]. This exponential growth of structures is fueled by innovative techniques at different stages of the structure elucidation process and high-throughput automation. X-ray crystallography celebrated its 100 years anniversary in 2014 and is still the dominating method in structure biology. Most importantly, together with X-ray crystallography, other structural methods such as NMR, high-resolution cryo-EM and small-angle X-ray scattering (SAXS) can be used in parallel to address a specific structural problem. This idea of integrative structural biology is slowly gaining ground as well as further integrating biophysical, biochemical analysis, cellular studies, bioinformatics and computational biology to give a more holistic view of protein's function and mechanism.

The aim of this thesis is to gain structural and functional insights into two important lipid enzyme families in human cells by integrating several techniques including X-ray crystallography and solution NMR spectroscopy as well as biophysical binding studies and biochemical activity assays.

Paper I

Human lipocalin Prostaglandin D synthase (L-PGDS) has previously been in the spotlight for structural study using X-ray crystallography, NMR spectroscopy and SAXS individually. Despite several mouse L-PGDS structures being reported, these structures usually harbor a catalytic residue mutation with an empty substrate binding pocket and unclear the structural basis of the enzyme and transporter activities. Therefore the objective of this project was to uncover the molecular mechanism of substrate binding, catalysis and product release of wild type L-PGDS utilizing a combination of X-ray crystallography and NMR.

Paper II

Sphingomyelin phosphodiesterase like 3a (SMPDL3a) has been proposed to share similar function with its close homolog acid sphingomyelinase (aSMase), based on the conservation of its metallophosphodiesterase domain. Apart from that, little is known about this protein except its regulation by liver X receptor, cAMP activation and cholesterol loading on macrophages as well as absence of sphingomyelinase activity. This project aims to shed light

on the structure, function and mechanism of SMPDL3a, with possible implications for an understanding of the aSMase family.

3 FROM PROTEIN TO STRUCTURE AND FUNCTION

3.1 RECOMBINANT PROTEIN EXPRESSION

In order to obtain large amount of highly pure and homogenous proteins for structural studies, recombinant protein over-expression is most often an essential requirement. Proteins can be over-expressed in either a cell-based system or a cell free system but currently cell-based expression is still most commonly used. Two types of cell-based systems were used in this work and they will be discussed further in the following sections.

3.1.1 Bacterial expression systems

The bacterial system is the oldest and most widely used cell-based system for recombinant protein expression due to its efficiency, simplicity and economical attributes. Prokaryotic cells have simpler transcription-translation machinery and a fast growth rate. It is genetically easy to manipulate and quick to obtain large cell mass containing proteins of interest. The target gene is cloned into bacteria expression vectors and subsequently the plasmid is transformed into bacteria cells for amplification. Under antibiotic selection pressure specific to the recombinant plasmid, only bacteria cells with the target gene is multiplied and subsequently induced to express the heterologous gene. The bacteria system can be used to also express various types of isotope labeled proteins for NMR experiments.

3.1.1.1 *L-PGDS recombinant protein expression*

In Paper I, the human lipocalin prostaglandin D synthase cDNA was cloned into pNIC-CH3 vector containing a C-terminal non-cleavable hexa-His tag [33]. *Escherichia coli* (E.coli) of BL21 (DE3) Rosetta T1R strain was transformed with the vector containing the gene of interest, then cultured in Terrific-broth (TB) media. The production of L-PGDS was induced using isopropyl-1-thio- β -D-galactopyranoside (IPTG) and harvested after 16-hours growth at 37°C.

Isotope-labeled proteins for NMR studies were expressed similarly in M9 minimal media supplemented with ^{15}N -labelled ammonium chloride (HN_4Cl) for single labeled protein. Doubled labeled protein was expressed with both ^{15}N - HN_4Cl and ^{13}C -labeled glucose added. The M9 media was also supplemented with ^{15}N - HN_4Cl and ^{13}C -labelled leucine and alanine to label the respective residue type in a non-deuterated environment.

3.1.2 Insect cells system

Another increasingly popular cell-based system is the binary baculovirus-insect cells system. Baculovirus is referred to a family of virus with large, circular double stranded DNA that has a rather specific host range. It is known to infect primarily arthropods like insects [34]. Years of development have led to a more optimized Baculovirus genome for recombinant expression of foreign protein in insect cells. A commercially available Bac-to-Bac method based on Luckow et al's paper is adopted in our production of human acid sphingomyelinase like protein (SMPDL3a) in Paper II [35].

This system employs a baculovirus shuttle vector known as bacmid. The gene of interest is first cloned into a vector and transformed into *DH10Bac* competent cells that carry a bacmid with transposition insertion site (attTn7) and a transposition helper plasmid. This allows the transposition of gene of interest from the vector to the bacmid which disrupts the *lacZa* reading frame, thus making the cells unable to produce β -galactosidase. These cells form only white colonies on blue-gal containing agar (Figure 3A) [35]. The recombinant bacmids are then isolated and validated with PCR (Figure 3B) before transfection.

Common insect cell strains selected for Baculovirus propagation are Sf9 or Sf21 cells derived from the pupal ovarian tissue of *Spodoptera frugiperda* worm and High-Five™ cells from cabbage looper [34]. They can be grown either as suspension cells or adherent cells at 27°C. Transfection of recombinant bacmid by transfecting agents like cellfectin induced the insect cells to produce baculovirus that express the heterologous gene. After 72 hours of infection, the insect cells will lysed, releasing viruses into the media, which is known as P0 virus stock. The P0 virus stock is used for subsequent infection and propagation to P1 and P2 viruses for large-scale (volume) protein production.

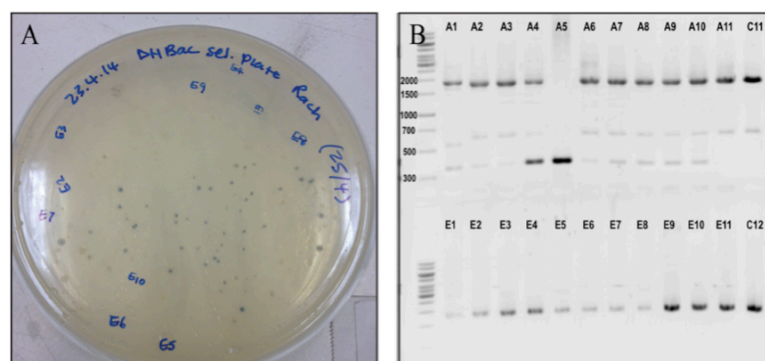


Figure 3: (A) Blue white colony screening after DH10Bac transposition. (B) PCR validation of white colonies extracted plasmids to verify the size of transposed gene of interest.

3.1.2.1 *SMPDL3a* protein expression

The gene that encodes SMPDL3a was cloned into pFastBac-Sec vector. This vector includes a secretory signal peptide in order to express target gene as a secreted protein with an N-terminal cleavable His tag for affinity purification. Bacmid carrying SMPDL3a gene is transfected in Sf9 cells for protein expression. Since SMPDL3a is produced as a secreted protein, the supernatant was collected for subsequent protein purification.

The baculovirus insect cell system is especially beneficial for SMPDL3a protein as it requires post-translational modification like N-linked glycosylation. However insect cells are incapable of complex glycosylation as they are able to form only simple oligo-mannose sugar chain [36]. This may affect the biological function of the enzyme if complex glycans are essential for its *in vivo* activity.

3.1.3 Protein purification

After protein over-expression, proper purification is necessary and critical for structural studies to ensure that the protein is soluble (except for solid-state NMR), pure, homogenous, and in good amount for structural characterization. This is usually achieved by chromatography where the protein is separated based on selective binding, isoelectric point, hydrophobicity or size.

3.1.3.1 *Affinity tag purification*

Affinity chromatography is a common and effective method to separate target protein from protein mixture based on selective interaction with a ligand. The ligand that is specific towards target protein is coated on a stationary phase and will bind primarily to the target protein when a protein mixture (mobile phase) is passed through. Proteins expressed with the hexa-histidine fusion tag have high affinity towards nickel beads and this strategy is often used for initial purification of recombinant proteins. A nickel-NTA beads slurry packed in a column can act as the stationary phase to selectively bind the His-tagged protein. Proteins that do not bind specifically to the Ni-NTA will flow through the column upon washing. His-tagged protein can be recovered by increasing the imidazole concentration or by altering the pH of the washing buffer (mobile phase). Imidazole can compete for interactions with nickel NTA while changing the pH alters the protonation state of imidazole nitrogen on histidine ($pK_a = 6.0$) and disrupts the histidine-nickel interaction.

Both L-PGDS and SMPDL3a were expressed with His-tag fusions but purified in a slightly different manner. As L-PGDS protein is expressed in bacteria cells, it is necessary to lyse the

bacteria cell wall by sonication and/or other disruption methods to extract the protein before chromatography. Meanwhile SMPDL3a was secreted into the media, hence no cell lysis was required. The media containing target protein was collected for purification. However, since the histidine-nickel interaction is sensitive to pH, media was first titrated with Tris pH 8 buffer to adjust the slightly acidic pH to a neutral pH. After affinity purification, SMPDL3a was proteolysed by Tobacco Etch Virus (TEV) to remove the His-tag. In some cases, host proteins that are rich in histidines can be isolated together with the target protein. Hence, a second step of chromatography is required to ensure the purity.

3.1.3.2 Size exclusion chromatography (SEC)

This method separates molecules in solution by size (molecular weight). In this case the stationary phase consists of inert porous matrix that retards smaller proteins while allowing larger proteins to flow through, separating the mixture in an isocratic manner. Both L-PGDS and SMPDL3a were subjected to size exclusion chromatography. An interesting observation was that L-PGDS protein elution fractions appeared yellow but the color reduced with increasing elution time. The yellow fraction and colorless fractions were separated for crystallization and NMR experiments initially. No crystals were formed from the yellow fraction and it gave poor NMR spectra (detailed discussion in *Section 3.6.2*) hence only the colorless fractions of L-PGDS were pooled and used for all subsequent crystallization and NMR measurements.

3.2 THERMAL STABILITY ASSAY

3.2.1 Thermal aggregation method

All proteins are stable within a specific pH and temperature range where its structural integrity is preserved to allow it to perform its function. Thermal induced denaturation of proteins has been a vital tool to access protein conformation stability. Binding of ligands such as co-factors, effectors, products or inhibitors often increase protein stability, but sometimes decrease stability (Figure 4). Ligands that induced significant stabilization have been shown to increase the probability of forming protein-

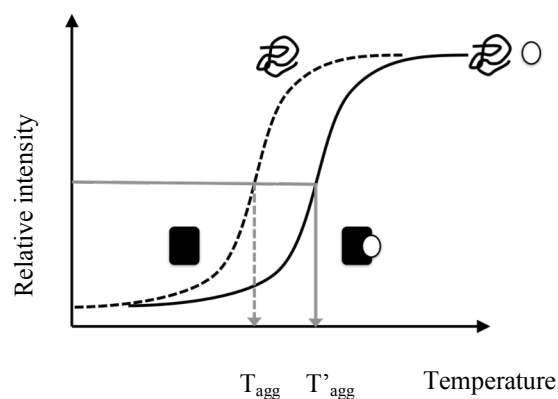


Figure 4: Protein stabilization curve in the presence and absence of ligands.

ligand complexes in crystal structures [37, 38]. Measurement of protein thermal stability can be done in a various ways [37, 39]. Differential scanning fluorometry that adopts SYPRO orange dye to bind to hydrophobic core of protein upon unfolding is not suitable for proteins with hydrophobic patches like L-PGDS. Therefore, in Paper I, a Harbinger Technology, Stargazer-384™ instrument is employed, it uses the concept that when heat denatured proteins are unfold they often aggregate and these aggregates can be detected by light scattering. This approach constitutes a convenient and high-throughput method to screen for ligands and compounds suitable for protein stabilization and ligand binding. The stargazer method is used to investigate whether different compounds were ligands for purified L-PGDS and as a guide for strategizing complexes generation for crystallographic studies. Substrate analog (SA U44069) was shown to give strongest stabilization and thus was used for the co-crystallization experiments that eventually gave the enzyme-SA complex structure.

3.2.2 Dot-blot method

The stargazer method requires relatively high concentrations of proteins and it has a maximum temperature detection of about 80°C. Modified detection method based on similar principle that has been discussed in *Section 3.2.1* can be used to determine thermal stability of protein that are less soluble or with aggregation temperature above 80°C. Instead of monitoring the aggregated proteins, the remaining soluble proteins are detected in a dot-blot after a centrifugation or filtration step. In practice, the protein sample is heated in a PCR iCycler followed by high-speed centrifugation to pellet aggregated protein. Aliquots of supernatant from each temperature point can be taken for SDS PAGE analysis (Figure 5). In order to facilitate large sample size analysis, aliquots are dotted on nitrocellulous membrane instead and subsequently probed with a protein-specific antibody, or in the case of His-tagged SMPDL3a, a His-probe antibody (Figure 5). Relative intensity of coomassie stained band or Western blot signal is measured by ImageJ/ImageLab software. Normalized relative intensity is plotted against each temperature point and fitted with non-linear regression in Graphpad. This was useful for the analysis of SMPDL3a where its native aggregation temperature is approximately 80°C and further stabilization by ligands and metals was not possible to be measured by Stargazer instrument at these high temperatures.

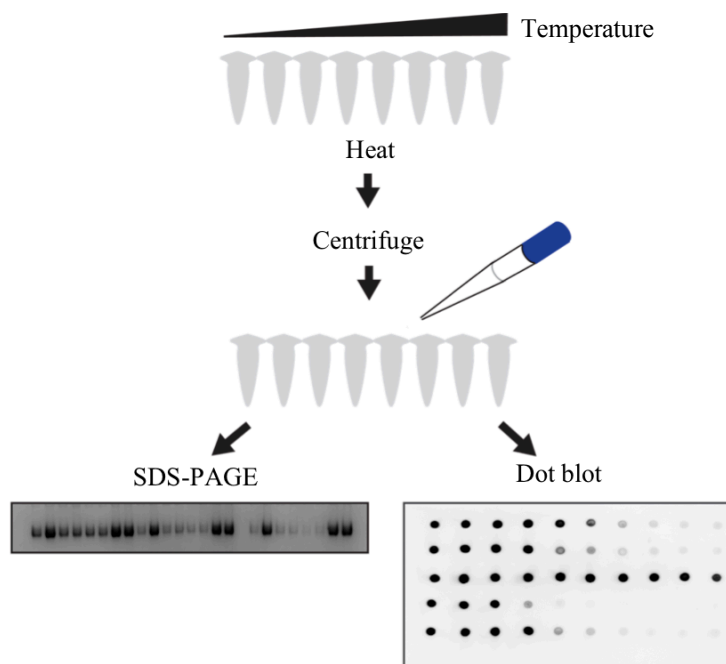


Figure 5: Cartoon description of thermal stability assay for protein SMPDL3a.

3.3 ISOTHERMAL TITRATION CALORIMETRY (ITC)

Isothermal titration calorimetry (ITC) can be utilized to characterize the thermodynamics of enzyme-ligand binding interactions. Since chemical reactions and binding is accompanied by change in enthalpy, measuring the heat released (exothermic) or gained (endothermic) during enzymatic reaction can be correlated to the rate of the reaction [40]. In the case where non-substrate ligand is supplied to the system, ITC can measure the affinity and thermodynamic parameters for binding instead. The experiment setup of MicroCal iTC200 consists of a reference cell for water, a sample cell for protein sample and a syringe for titrant. A series of incremental injections of the titrant are made within specific time intervals throughout the course of the titration experiment. For each injection, the change of enthalpy is measured and plotted against the concentration of ligands introduced into the protein system. Binding affinity and occupancy of ligand(s) can subsequently be determined. Due to the high reactivity of L-PGDS substrate, PGH₂, it is challenging to measure the binding affinity of enzyme-substrate. Therefore the more stable product, PGD₂ was used as a titrant in the study.

3.4 MULTI-ANGLE LIGHT SCATTERING (MALS)

Multi angle light scattering uses the principle of Rayleigh scattering to assess protein homogeneity, molar mass and oligomerization state. Polarizable macromolecules in solution are able to scatter light. The intensity of scattered light is proportional to the concentration of macromolecules and it is enhanced when oligomer is present because oligomer contributes to scattering in coherent phase. Wyatt Technology MALS system is coupled with a size exclusion column (Superdex 200 5/150 GL) and consists of three detectors: absorbance detector (UV), a static light scattering detector and a refractive index detector. MALS was used to determine to oligomeric state of SMPDL3a and the experiment revealed a monodisperse peak revealing that SMPDL3a exists as a monomer in solution (Figure 6).

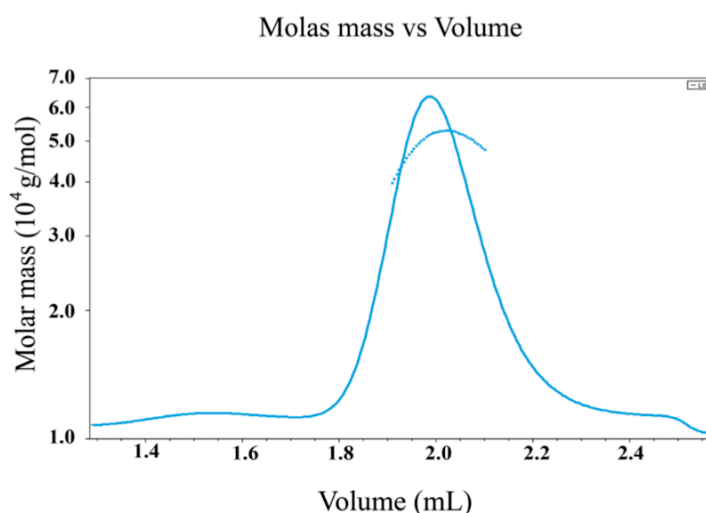


Figure 6: MALS profile for SMPDL3a protein in 20 mM HEPES pH7.5, 300 mM NaCl and 10% glycerol. Sample is monomeric and homogeneous eluting at mass 50.6 g/mol.

3.5 MACROMOLECULAR X-RAY CRYSTALLOGRAPHY

As discussed above, X-ray crystallography is the most commonly used method for elucidating molecular structure of nucleic acid and proteins. The foundation of the method is based on the scattering property of light whereby diffraction of radiation with wavelength in atomic range can resolve atomic structures of macromolecules, if X-ray radiation of sufficiently low wavelength is used, typically in the order of 0.1 nm. In 1913, W.L. Bragg and W.H. Bragg proposed Bragg's Law to describe the scattering properties of X-rays when they encounter atoms organized in an ordered lattice (crystal). This phenomenon termed X-ray diffraction can be used to determine the atomic structure of the molecule(s) in the ordered crystal.

3.5.1 Principles of X-ray diffraction

Bragg's Law states that when two in-phase X-rays incident beams are scattered by atoms that are separated by constant distance d in a crystal lattice, the path differences between the two waves is $2d\sin\theta$ [41] (Figure 7). This constructive interference of two scattering waves reflected by the consecutive lattice plane will produce an intense peak when measured with an X-ray detector.

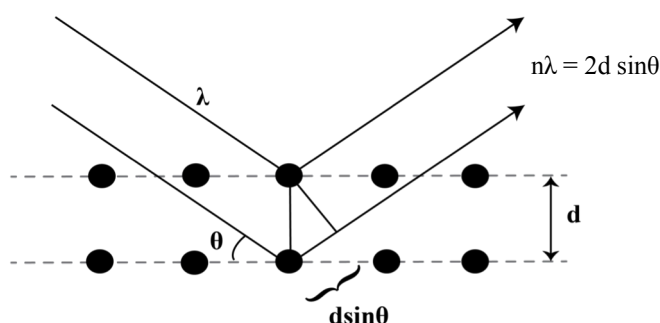


Figure 7: Depiction of Bragg's Law.

Bragg's Law provides the location of diffraction peaks corresponding to the atom arrangement but measurements of the diffraction peaks' intensities are required to calculate the electron density distribution in order to construct the protein structure. The following Fourier transform describes how this electron density distribution relates to the strength of the diffraction peak.

$$F(hkl) = \iiint_{x,y,z=0}^1 \rho(x,y,z) e^{2\pi i (hx+ky+lz)} dx dy dz$$

$F(hkl)$ is referred as structure factor and the reverse Fourier transform of the above equation allows for the calculation of the electron density distribution in a unit cell,

$$\rho(x, y, z) = \frac{1}{V} \sum_{h,k,l} F(hkl) e^{2\pi i (hx+ky+lz)}$$

and,

$$F(hkl) = |F(hkl)| e^{i\alpha}$$

The structure factor amplitude $|F(hkl)|$ can be measured directly from the intensities of diffraction peaks but the phase α has to be obtained indirectly. The later is done either by additional experiments where site specific distortions are generated in the structure leading to small changes in diffraction intensities, or by using a homologues structure to generate initial phase information. When the electron density map is of sufficient quality, the atomic modeling of the molecule(s) in the unit cell can be constructed and refined. A unit cell can house one or more molecule(s) depending on how the crystal is packed and it is the smallest repeating unit in a crystal.

3.5.2 Protein crystallization

A key challenge in macromolecule crystallography is to crystallize the subject of study, which may be proteins, nucleic acids or protein-ligand complexes [42]. In order to drive macromolecules to form crystals, a common method used is vapor diffusion. In this method, a small drop of highly concentrated soluble and pure protein solution is mixed with a precipitant in a specific ratio, placed next to a volume of precipitant reservoir in a sealed environment. The difference of vapor pressure between the droplet and precipitant reservoir will gradually drive the droplet concentration to achieve equilibrium in the closed environment. If the condition is optimal, this allows the protein solution to enter supersaturation phase and subsequently undergoes nucleation that will lead to crystal formation [42].

There are many variables that could affect crystal formation [43]. Firstly, the quality of protein is very important. Proteins that are soluble, homogenous and stable (above crystallization temperature) have a higher propensity to form crystals. Other factors to be considered include temperature, pH, precipitant choice, precipitant concentration, precipitant-protein ratio, presence of additive or ligands starting protein concentration, etc. The choice of precipitant and buffer composition is usually made by trial and error. Nowadays, commercially available crystallization screens are used as an initial setup of crystallization trials. Once the initial crystallization conditions are identified, subsequent optimization by altering the variables described above can further improve the quality of the crystals.

3.5.2.1 Crystallization of L-PGDS

L-PGDS was crystallized in the presence 2 mM substrate analog 9,11-epoxymethano PGH₂ (SA-U44069) but no crystals were formed in the absence of SA-U44069 (Figure 8A). In order to study the structural differences of L-PGDS with and without SA-U44069, an apoenzyme structure was needed. Microcrystals of L-PGDS and SA-U44069 complex (SAC)

were streaked on protein-precipitant drop without substrate analog. This is known as seeding and commonly used in protein crystallization to provide nucleation sites for molecules to assemble under the right condition. Crystal grown from the seed usually adopts the characteristics of the template crystals.

3.5.2.2 Crystallization of SMPDL3a

SMPDL3a, with 21.6 mg/ml concentration, took 21 days to form crystals at 24°C. Initial crystals were small and diffracted to 3-3.5 Å resolution (Figure 8B). The optimizations of crystallization condition include changing the additive concentration and temperature of incubation helped to reduce the period of crystal formation down to 7 days and improved the diffraction to 2.3 Å resolution (Figure 8C). This was sufficient for heavy atom based phasing experiments that were used to solve the structure.

3.5.3 Data collection

Once a crystal is obtained of a macromolecule, it is subjected to a focused monochromatic X-ray beam and if it diffracts, the diffraction data can be detected and collected. The basic setup used nowadays has the crystal mounted in a drop containing a cryo-protectant like glycerol, sitting in a cryo loop and frozen under a stream of liquid nitrogen gas. The maintenance of the crystal at cryo temperatures often dramatically increases crystal life time and data quality. An X-ray beam from a copper source is directed onto the crystal loop mounted on a goniometer that can be rotated in almost any orientation with respect to the beam. A detector then collects the diffraction spots from the rotating crystal (Figure 8D). The intensities of the spots, also known as reflections, can now be integrated to generate a diffraction dataset.

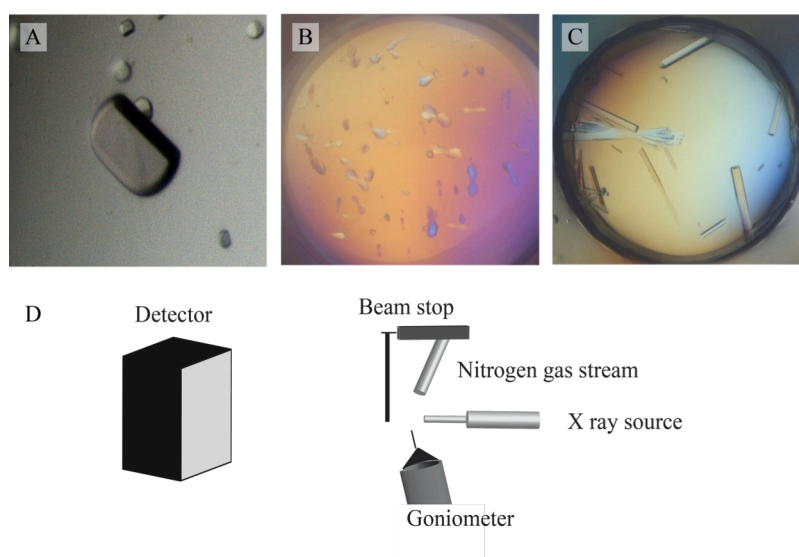


Figure 8. (A) L-PGDS crystal after optimization (B) SMPDL3a crystals without additive (C) SMPDL3a crystal with 0.1M TCEP as additive (D) Basic setup of a synchrotron beamline where the sample is mounted on a goniometer cooled with nitrogen gas stream. When subjected to an X-ray beam, the diffraction pattern is collected on the detector.

X rays can be generated by X-ray tubes, rotating anodes or a synchrotron radiation sources. The synchrotron radiation is where the most powerful and highest-quality X-ray source available. Complete datasets can be collected within minutes at high-end synchrotron stations. Besides, it can produce monochromatic X-ray beam at selected wavelength that is useful for experimental phasing (*Section 3.5.4.2*).

Dataset collections for L-PGDS and SMPDL3a have been carried out at the following synchrotron facilities; Australia Synchrotron (Melbourne, Australia), BESSY Helmholtz-Zentrum (Berlin, Germany), Diamond Light Source (Oxford, UK), European Synchrotron Radiation Facility (Grenoble, France), Soleil (Gif sur Yvette, France) and National Synchrotron Research Radiation Center (Hsinchu, Taiwan).

3.5.4 Solving the phase problem

In order to solve the macromolecule structure, both reflection intensities and the phase information are required. This is depicted in the Fourier summation that generates the electron density in which each reflection from the diffracting crystal is represented by an amplitude and a relative phase (α) (*Section 3.5.1*). The phase information unlike the intensities (amplitudes) cannot be obtained directly using current instruments [44], this lack of phase information is commonly known as the phase problem.

Nonetheless, several methods can be used to solve this problem. Phases can be determined experimentally by isomorphous replacement (single isomorphous replacement, SIR; multiple isomorphous replacement, MIR), anomalous scattering (single wavelength anomalous dispersion, SAD; multi-wavelength anomalous dispersion, MAD) or combination of both, SIRAS / MIRAS. On the other hand, if a related protein structure has been previously solved, molecular replacement can be employed. This section will discuss two specific methodologies in more detail.

3.5.4.1 Molecular Replacement (Paper I)

Molecular replacement (MR) is a method that uses the known phase information of a related structure to generate an initial phase estimate for crystals of a new protein. As a rule of thumb, it is good to have at least 25% of amino acid sequence identity between the previously solved structure - a “search model” and the unknown structure known as a “target molecule” for better chance of success [45]. Normally, the best search model shares more than 30% of sequence identity between the two. MR sequentially employs rotation and translation functions in finding the appropriate orientation and position of the search model in the asymmetric unit (Figure 9). If successful, approximate phases can be calculated from this model and when combined with the experimentally derived structure factor amplitudes of the target molecule’s crystals, an electron density map can be calculated for the target molecule. Subsequently, a model of the target molecule can be generated in this electron density, potential guided by the search model. Finally, this model is improved and refined in a cyclic process. Since mouse L-PGDS that shares 72.5% sequence identity with human L-PGDS had

been previously determined, it made a good search model to solve the structure of human L-PGDS using molecular replacement.

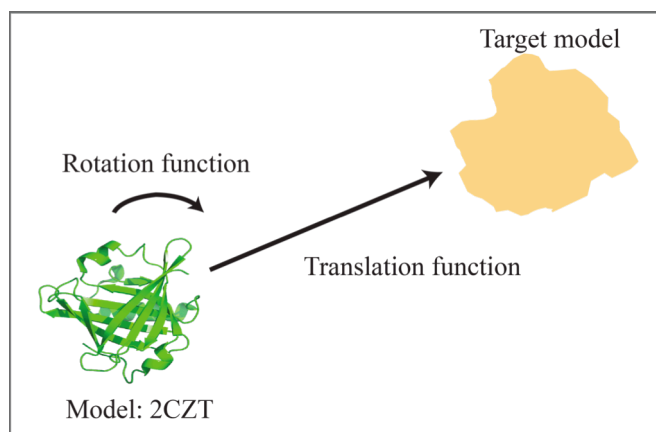


Figure 9. Illustration of rotation and translation functions applied onto mouse L-PGDS (PDB i.d.: 2CZT) as search model to solve the phase of target model, human L-PGDS.

3.5.4.2 *Single Anomalous Dispersion (SAD) (Paper II)*

Heavy atoms are rich in electrons, and contribute strong anomalous scattering when excited at a wavelength close to the atom's absorption edge. Radiation is scattered with an altered phase due to the anomalous scattering and this altered phase is dependent on the heavy atom position in the unit cell. The site(s) of heavy atoms in the unit cell have to be determined to allow the phases to be calculated [46]. This method requires a good incorporation of heavy atoms into the crystal sample and an X-ray source with a tunable wavelength. One way of adding heavy atoms directly to the protein during protein expression is by using selenium modified methionine amino acids. Alternatively, heavy atoms can be incorporated into the protein crystals through co-crystallization or soaking. Some common choices of heavy atoms include $\text{Pt}^{2+}/\text{Pt}^{4+}$, Hg^{2+} , Au^{3+} and Pb^{2+} , which have strong affinity towards amino acids like histidines, cysteines and methionine. They could be added as their salts but also in the form of metalo-organic compounds. Apart from these commonly used heavy atoms, some natural occurring metals such as manganese, copper, iron and zinc are also suitable anomalous scatterers.

SAD phasing requires data collection at the wavelength where absorption peak of the heavy atom is significant (f') and anomalous differences is strong. In MAD phasing, additional diffraction data at the inflection wavelength (f'') and remote wavelength is included to maximize the dispersive difference thereby increasing the quality of the phase information. Ideally, all anomalous dataset collection should come from a single crystal to minimize non-isomorphous difference. However, in practice radiation damage usually limits data quality.

Zn^{2+} is found to stabilize the SMPDL3a protein significantly in the thermal shift assay. As Zn^{2+} has a strong anomalous signal ($f'' = 3.9$ electrons at its K edge), it was a primary choice for experimental phasing of SMPDL3a. Twelve heavy atom sites were located in one asymmetric unit cell based on data collected at the Zn edge (9.661 keV). Phases determined

by autoSHARP version 2.60 [47] yielded an electron density map that allows poly-alanine chain building in Coot [48] and subsequently a complete model construction by ARP/wARP version 7.5 from the poly-alanine model (Figure 10) [49].

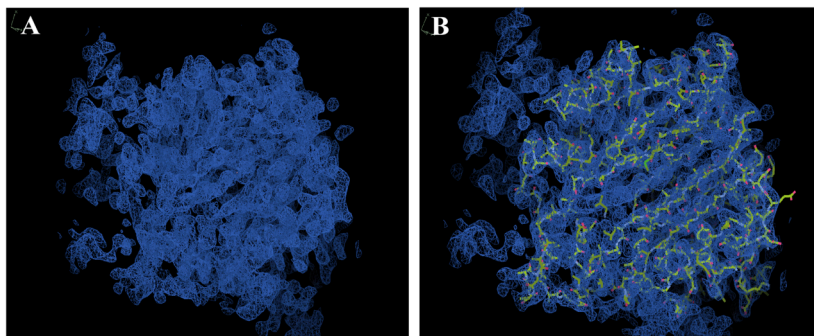


Figure 10. (A) Electron density map from SHARP. (B) Poly-alanine model built in Coot using the map as a guide.

3.6 SOLUTION NMR SPECTROSCOPY

NMR spectroscopy is a powerful and sensitive tool for structural study and is extremely advantageous to complement work of X-ray crystallography. Solution NMR study the protein structure in the solution, hence no crystal is needed. NMR can also provide both qualitative and quantitative measurements of protein-protein or protein-ligand interactions as well as to investigate internal protein dynamics.

Atoms are made of neutron, proton and electron each possessing the property of spin that comes in a magnitude of $\frac{1}{2}$. Only nuclei with a total non-zero spin are “visible” in NMR spectroscopy, such as ^1H , ^2H , ^{13}C , ^{15}N and ^{31}P . When these nuclei are placed in the presence of static magnetic field (B_0), the nuclei spins will act as magnetic moment and align in the magnetic field. Under thermodynamics equilibrium, the nuclei spins are populated in high and low energy levels based on a Boltzmann distribution. The energy required for the transition between two levels is represented as follow, where h is Planck constant and γ is the gyromagnetic ratio of nuclei.

$$E = h\nu B_0$$

When an energy that matches the energy difference (E) is applied in a form of radio frequency (RF), the nucleus absorbs the energy into a higher energy state and precesses opposing or perpendicular to B_0 depending on the type of RF irradiation. The precession generates electric current in the detection coil and since magnetization will decay in time to its equilibrium state, the signal is also known as free induction decay (FID). FID is subjected to Fourier transform to produce the NMR spectrum for further analysis [50].

The NMR spectrum comprises of NMR lines located at specific frequency known as chemical shifts, δ . The frequency detected is a “shift” of nucleus magnetization signal from

B_0 due to electron shielding of the molecule's the chemical environment relative to the standard nucleus magnetization. The standard used in L-PGDS experiments is 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS). Since δ is dependent on the chemical environment of atoms, it is able to reflect the chemical structure of molecules studied. Usually a macromolecules solution NMR experiment would begin with a one-dimensional (1D) experiment followed by two-dimensional (2D) and subsequently three-dimensional (3D) experiments for backbone assignment and structural calculation.

3.6.1 1D experiment : ^1H

Since ^1H is present in all biological samples, naturally and abundantly, no isotope labeling is required in preparing the sample. Observation of ^1H magnetization of the sample is known as 1D experiment. However the high abundance of proton signal from buffer and solvent in the sample dwarfs the protein peak. Most importantly the NMR lines corresponding to protein are highly overlapped and difficult to be interpreted. Therefore it is common to adopt multi-dimensional (2D and 3D) experiments involving ^1H , ^{15}N and ^{13}C isotope labeled protein to selectively increase the signal sensitivity and reduce signal overlap from the protein sample. Nonetheless, 1D-experiment is routinely used to enable a brief qualitative overview of the protein secondary and tertiary structure. 1D-NMR spectra of L-PGDS showed a good spread of side chains nitrogen bound protons (H^{N}) and backbone H^{N} chemical shifts from 5.5 ppm to 10 ppm indicating the protein is globular and consists of mainly β -sheets structure (Figure 11).

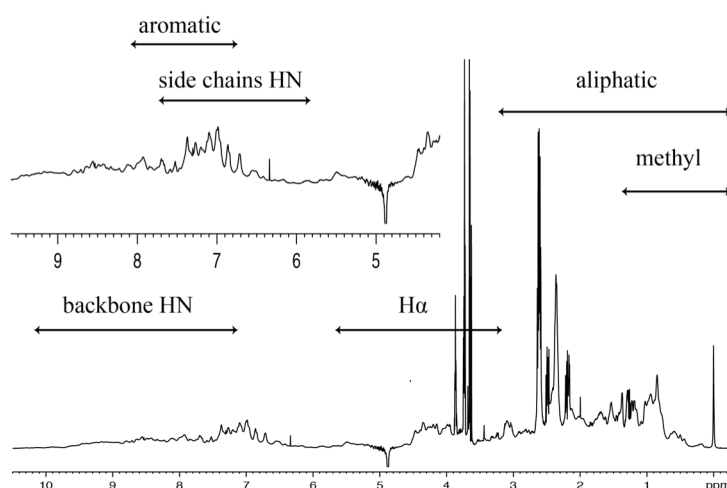


Figure 11: 1D-NMR spectra of 0.38 mM L-PGDS in 50 mM sodium phosphate pH 6.5, 20 mM NaCl, 2mM TCEP , 5% D_2O and 0.5mM DSS

3.6.2 2D experiment: Heteronuclear Single Quantum Coherence (HSQC)

Multidimensional experiments are key to circumvent the limitations of 1D-experiments. They resolve overlapping NMR lines and provide further information of connectivity between the same or different nuclei in the molecule. They serve as a basis for chemical shifts assignment and structural elucidation. Two-dimensional ^1H - ^{15}N heteronuclear single quantum coherence

(HSQC) experiment involves the transfer of magnetization from ^1H to its directly bonded ^{15}N nuclei and back to ^1H for detection (Figure 13B). ^1H - ^{15}N HSQC spectrum maps the chemical shifts correlation between ^1H (x-axis) and its directly bonded ^{15}N (y-axis) in distribution of cross peaks. Since proteins are made of amino acids linked with peptide bonds, the majority of these cross peaks are related to the protein amide backbone. In addition, signals from the side chain NH bonds of tryptophan, arginine, lysine, histidine, glutamine and asparagine are also “visible”. However, proline is “invisible” due to its cyclic backbone. Most importantly the positions of these cross peaks are unique to every protein like a fingerprint, because it is based on the amino acids’ chemical and structural environment. Therefore it serves as a reference map for chemical shifts assignments in combination with data from other multi-dimensional experiments. Other use of HSQC experiment involves protein quality inspection and protein-ligand interaction studies.

3.6.2.1 Conformational analysis

^{15}N labeled L-PGDS is prepared in buffer 20 mM HEPES pH 6.5, 150 mM NaCl, 2mM TCEP, 10% D_2O and 0.5 mM DSS for ^1H - ^{15}N HSQC measurement. As mentioned previously in Section 3.1.3.2, L-PGDS elution from SEC has a yellow fraction (Fraction 1) and a colorless fraction (Fraction 2). Even though both fractions appeared as single band in SDS PAGE analysis, their ^1H - ^{15}N HSQC NMR spectra are distinctively different (Figure 12). The cross peaks from Fraction 2 are better resolved and higher in quantity as compared to Fraction 1. It is likely that these two fractions are conformational heterogeneous. Therefore Fraction 2 was pooled for both NMR and crystallography studies.

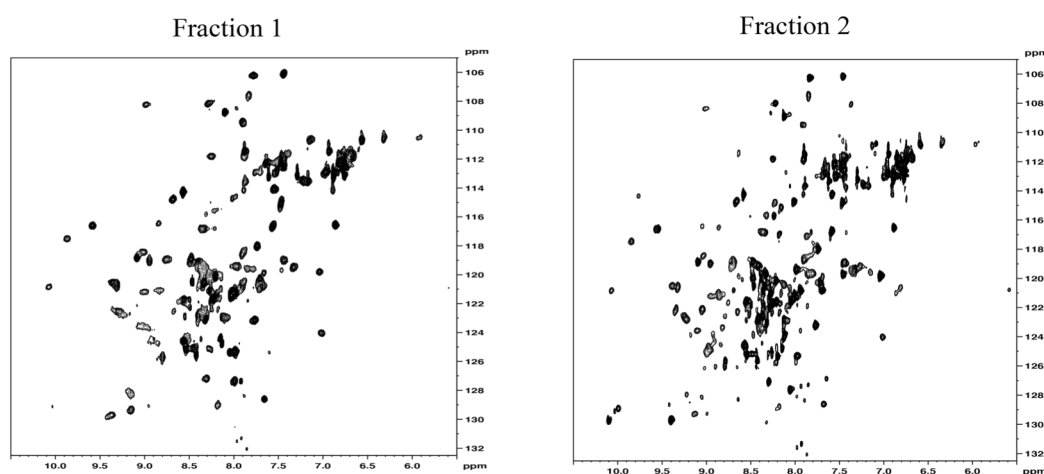


Figure 12: 2D-HSQC spectrum of L-PGDS Fraction 1 and Fraction 2

3.6.2.2 Ligand titration

In the event of ligand addition, the distribution of cross peaks will alter, as chemical shifts of the nuclei are sensitive to structural changes. This is known as ligand-induced chemical shifts perturbation (CSP). The ^1H - ^{15}N HSQC experiment is able to locate ligand-binding site(s) in the protein of interest if its chemical shifts has been assigned. ^1H - ^{15}N HSQC experiments of

^{15}N L-PGDS were recorded in the absence and presence of substrate analog (SA U44069) and product analog (PA 12415). Upon the addition of ligands, cross peaks of amino acids involved in binding and conformational changes can be detected (Figure 13A). Both analogs were titrated in concentrations of 2 mM, 3 mM and 4 mM. The CSP was measured for each cross peak and a threshold of 0.1 ppm was marked to distinguish between real binding and background perturbation [51]. It was also observed that HSQC spectrum of ^{15}N L-PGDS with 1 μL of substrate analog has significantly more cross peaks that are more resolved as compared to those without ligand (Appendix I & II). Therefore this spectrum is used as a reference sample for backbone assignment in order to assign as many residues as possible (Appendix III).

Furthermore, in order to test the hypothesis of possible interactions between L-PGDS and membrane during hydrophobic substrate binding and product release, titration of the analog bound ^{15}N L-PGDS complex with 3 mM dodecylphosphocholine (DPC) was carried out. DPC is a common detergent used as a membrane mimetics in NMR, especially in the study of interfacial membrane proteins [52]. It spontaneously forms micelle above its critical micelle concentration (c.m.c = 1.2 mM). Interaction of DPC micelles with ^{15}N L-PGDS was observed by the changes in CSP and alteration of intensity for all detectable cross peaks. Apo-enzyme titration with DPC micelles was used as a control experiment. In order to specifically locate the site of ligand binding and membrane mimetic interactions, sequential assignment of L-PGDS amide backbone was required. The backbone assignment was accomplished by several 3D experiments that will be discussed in the next section.

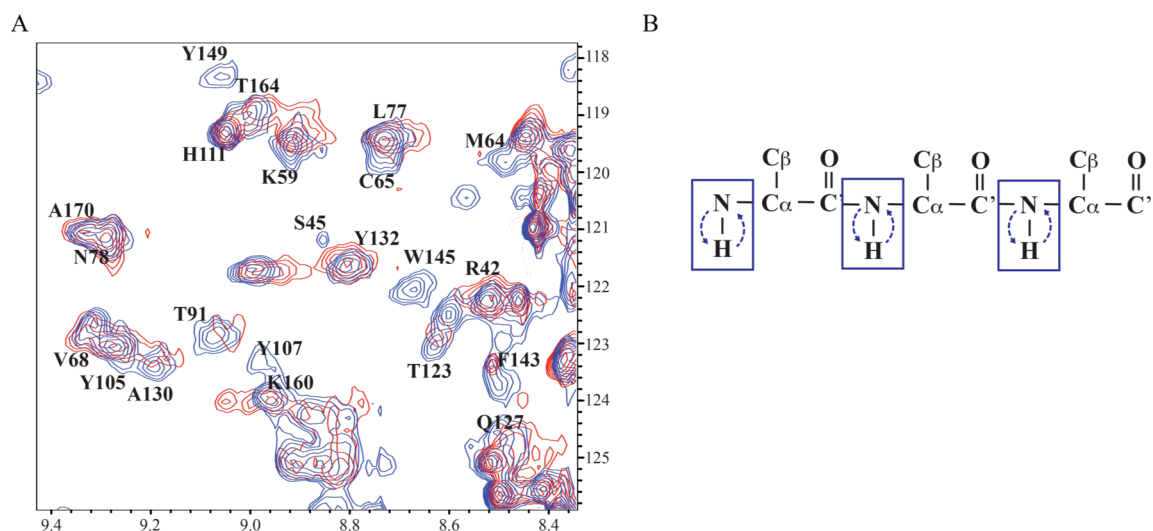


Figure 13: (A) ^{15}N – HSQC of L-PGDS apoenzyme in red and with 2 mM substrate analog U44069 in blue. Residues Y149, M64 and Y107 are among the chemical shifts that showed significant perturbation in the presence of substrate analog. (B) The transfer of magnetization from ^1H to its directly bonded ^{15}N nuclei and back to ^1H for detection in HSQC experiment.

3.6.3 3D experiments: HNCA, CBCA(CO)NH, ^{15}N NOESY-HSQC, HNCO

3.6.3.1 HNCA and CBCA(CO)NH

Both HNCA and CBCA(CO)NH experiments require ^{13}C and ^{15}N enriched protein sample. The HNCA experiment transfers magnetization via J-coupling from amide proton to amide nitrogen and then to its own alpha carbon, C^α_i and previous residue C^α_{i-1} subsequently back to amide proton for detection (Figure 14B). The magnetization transfer correlates to the amide cross peaks in ^1H - ^{15}N dimension with its intra and inter-residue $\text{C}\alpha$ in the ^{15}N - ^{13}C dimension. The C^α_i and C^α_{i-1} cross peaks are viewed in strips where the signal of C^α_i is usually stronger than C^α_{i-1} , chemical shifts of these $\text{C}\alpha$ would overlapped if they are adjacent residues (Figure 14A).

Meanwhile the CBCA(CO)NH experiment correlates the amide cross peak with its previous $\text{C}\alpha$ and $\text{C}\beta$ residue also via J-couple magnetization transferred (Figure 14B). Together with HNCA, they provide sequential information of the amide cross peaks. Since the chemical shifts of $\text{C}\alpha$ and $\text{C}\beta$ are distinguishable between amino acids due to their chemical structure, the identity of amino acids and its sequence can now be pieced together.

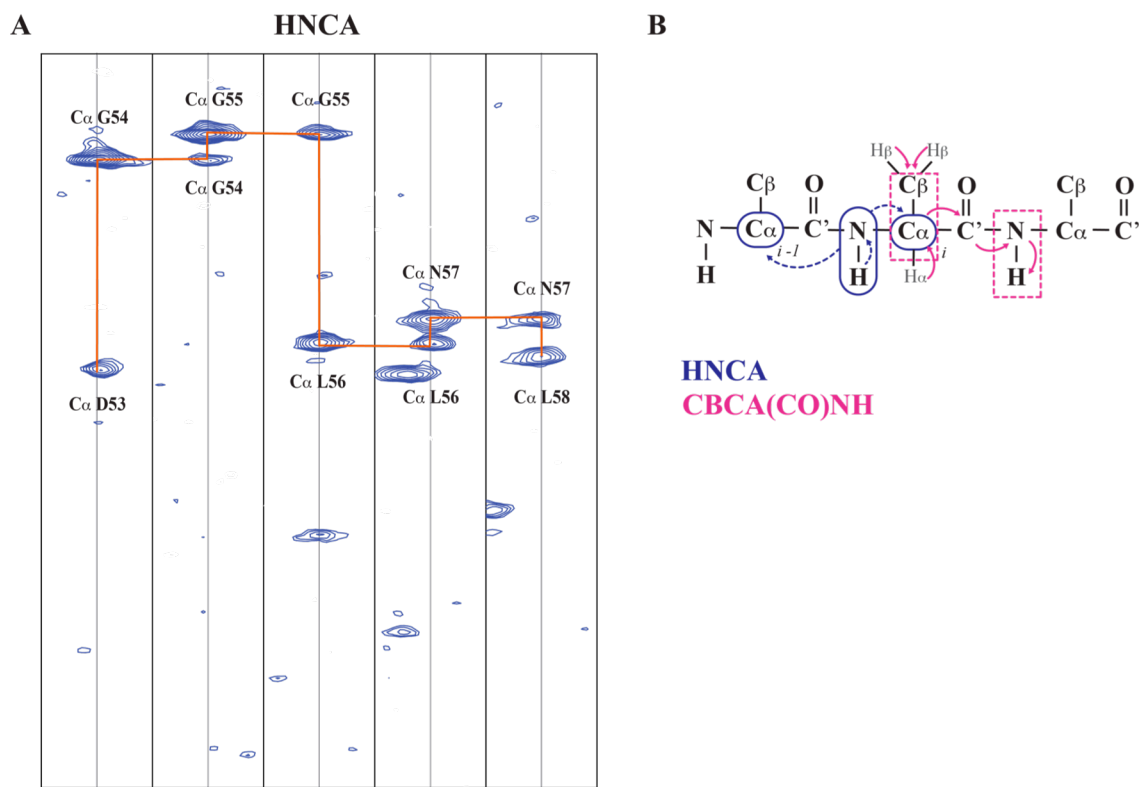


Figure 14: (A) An assigned and connected strip of HNCA in ^{15}N - ^{13}C dimension. (B) The magnetization transfer route for both HNCA and CBCA(CO)NH experiments.

3.6.3.2 ^{15}N (NOESY)- HSQC

Nuclear Overhauser effect (NOE) is a phenomenon observed when NMR signal intensity of a nucleus is enhanced due to the effect of resonance frequency saturation of another nucleus in close spatial proximity. This effect provides information of intermolecular distances with the intensity of NOE proportional to $1/r^6$ where r is the distance between two nuclei (usually protons). Therefore, a proton-proton distance within 5Å will give a NOESY signal. A NOESY-HSQC experiment allows magnetization exchange between all protons via NOE and then transferred back to the amide proton for detection. In this case it acquires knowledge of structural proximity in relation to the amide proton. Therefore, together with L-PGDS crystal structure, NOESY conformational dependent information can be used to validate backbone assignment of the protein.

3.6.3.3 *HNCO with specific residue labeling*

Due to the close $\text{C}\alpha$ $\text{C}\beta$ chemical shifts of residues like Leucine and Alanine, a labeling by residue type (LBRT) strategy was adopted to validate their assignments. This method required preparation of protein samples with ^{14}N ^{13}C -carbonyl labeled Leucine (or Alanine in a separate protein expression batch) in ^{15}N M9 minimal media. HNCO experiments transfers magnetization from the amide proton to the preceding carbonyl C (CO). The signals select the ^{15}N labeled residue that comes after ^{14}N ^{13}C -carbonyl labeled Leucine or Alanine in the sequence (Appendix IV & V). This information helps to resolve the ambiguity of assignment between these chemically alike residues.

4 PAPER I: THE STUDY OF LIPOCALIN PROSTAGLANDIN D SYNTHASE (L-PGDS)

4.1 L-PGDS IN EICOSANOIDS SIGNALLING

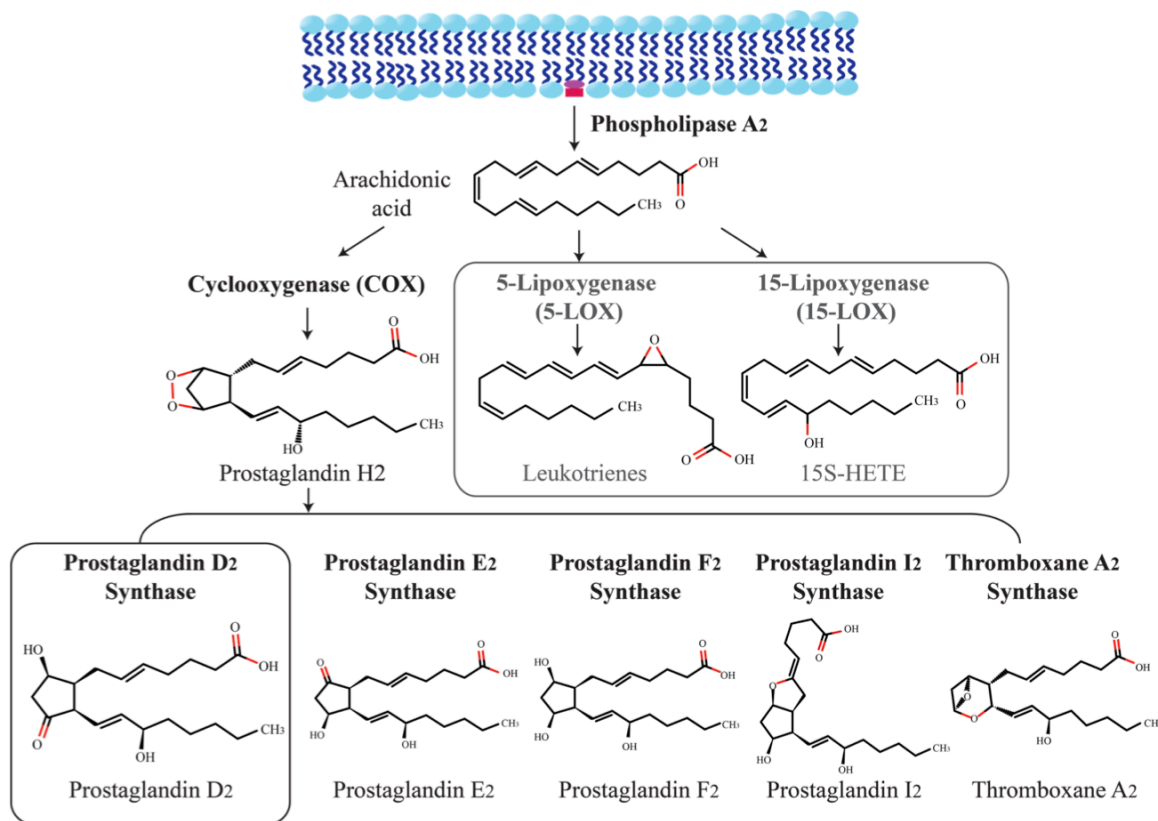


Figure 15: The arachidonic acid pathway.

Eicosanoids signaling is a major lipid-signaling pathway involved in acute inflammatory responses, nociception, platelet aggregation and immunoregulation [53]. External signal like stress triggers cytoplasmic phospholipase to activate membrane phosphatidylcholines lipid into arachidonic acid (AA) [54]. AA can be converted to prostaglandins by cyclooxygenase (COX), or turned into leukotrienes and lipoxins by the action of lipoxygenases (LOX). The product of COX, Prostaglandin H₂ is a common substrate for thromboxane synthase and tissue-specific prostaglandin isomerases like Prostaglandin D₂ synthase, Prostaglandin E₂ synthase, Prostaglandin F₂ synthase and Prostaglandin I₂ synthase (Figure 15). Prostaglandins function in autocrine and paracrine manner interacting with cognate prostaglandin receptors on plasma membrane. These ligands controlled G-protein coupled receptor then activates downstream signaling either by intracellular changes in cyclic AMP (cAMP) concentration or Ca²⁺ mobilization [55].

The work in Paper I focused on Lipocalin prostaglandin D₂ synthase (L-PGDS). Two types of PGD₂ synthase (PGDS) have been characterized based on its tissue distribution, hematopoietic PGDS (H-PGDS) and L-PGDS. H-PGDS is highly expressed in blood, mast cells and immune cells while lipocalin PGDS (L-PGDS) is mainly found in the central

nervous system, heart and reproductive organs [56]. They are distinct in overall structure, catalytic site and co-enzyme dependence despite catalyzing the same chemical reaction. H-PGDS catalysis is glutathione dependent but L-PGDS catalysis involves a cysteine residue. L-PGDS also moonlights as lipophilic ligand transporter, binding molecules like retinoic acid, bilirubin and biliverdin. The enzyme catalytic mechanism of L-PGDS has been previously speculated based on mouse apoenzyme structure but the substrate positioning and product delivery is unknown. Paper I addresses substrate binding and mechanistic aspects of L-PGDS substrate binding, catalysis and egression.

4.2 RESULTS SUMMARY

L-PGDS – substrate analog complex (SAC) crystal structures showed two sites of plausible ligand interaction but the density for substrate analog SA U44069 is ambiguous even at 1.8 Å. This uncertainty is confirmed by NMR titration studies with SA U44069 where the interaction sites coincide with what have been observed in the crystal structure. The hydrophobic moieties of the prostanoid substrate (PGH₂) has only three sites for hydrogen bonding inside the substrate-binding pocket; a cyclic endoperoxide group, a hydroxyl group on aliphatic ω chain and a carboxylate group on the acid tail. L-PGDS has a large hydrophobic beta barrel pocket relative to the size of the substrate. This increases the degree of freedom for PGH₂ and SA flexible hydrophobic tails to reside in the binding pocket. Furthermore, SA U44069 comprises of a non-hydrolysable bicyclicpentane ring that reduces it's the number of hydrophilic anchoring sites. It is possible that all these factors contribute to the poor ligand occupancy and discontinuous difference density (Fo-Fc) of SA U44069 in the crystal structure. Modeling of SA U44069 in the active sites identified three key areas for interactions. First, the bicyclopentane head group can be stabilized by Phe 83 to be positioned facing Cys 65 for catalysis. Secondly its carboxylate tail can be anchored by polar residue like Tyr 107 or His 111 and lastly the aliphatic hydrophobic chain can be inserted deep into the barrel lined with serine and leucine residues.

Co-crystal structures of L-PGDS also revealed important conformational changes of the Ω helix when compared to the apoenzyme structure. This observation was also supported by NMR titration results with both SA U44069 and product analog (PA 12415) whereby cross peaks of residues on Ω helix undergo CSP due to ligand-induced binding. Furthermore, CSP analysis also supported the “stop-plug” conformational changes proposed by Zhou et al where the residues at the bottom of the beta barrel undergo measurable structural changes in the event of ligand binding [57]. Similar patterns of ligand-induced CSP were observed when the L-PGDS-PA complex was titrated with the DPC micelles. In fact, upon DPC micelles titration, CSP of L-PGDS-PA complex were more pronounced than those of L-PGDS-SA complex. The peak intensity analysis for DPC-L-PGDS-PA spectrum showed that for some residues, the intensity enhancement of cross peaks during initial addition of PA to an apoenzyme were reversed or further attenuated when DPC micelles were added to the PA-enzyme complex. It is likely due to the reversal of PA binding upon DPC micelles titration. The contrasting CSP and cross peaks intensity alteration between DPC-L-PGDS-PA and

DPC-L-PGDS-SA experiments revealed different enzyme-ligand and enzyme-membrane interaction depending on the type of cargo bound.

Enzymes are known to have developed diverse strategies in handling substrate and product of seemingly opposing chemical properties [58]. This cargo-dependent interaction may likely reflect the *in-vivo* transport nature of soluble L-PGDS utilized for hydrophobic enzymatic product delivery. After catalysis, the enzyme may release its product PGD₂ into endoplasmic reticulum membrane with transient attachment facilitated by Trp 54 on Ω loop and Trp 112 on E-F loop, and potentially undergo conformational changes on Ω helix and E-F loop. PGD₂ may be delivered to plasma membrane via vesicular transport and released to the extracellular space by multi-drug resistant transporters [59]. This would allow PGD₂ to activate downstream signaling through cell surface receptors DP1 or CRTH2 (Figure 16) [60, 61]. All in all, an integrative structural biology strategy had provided a more holistic and resolved view of L-PGDS function, both as an enzyme and a transporter.

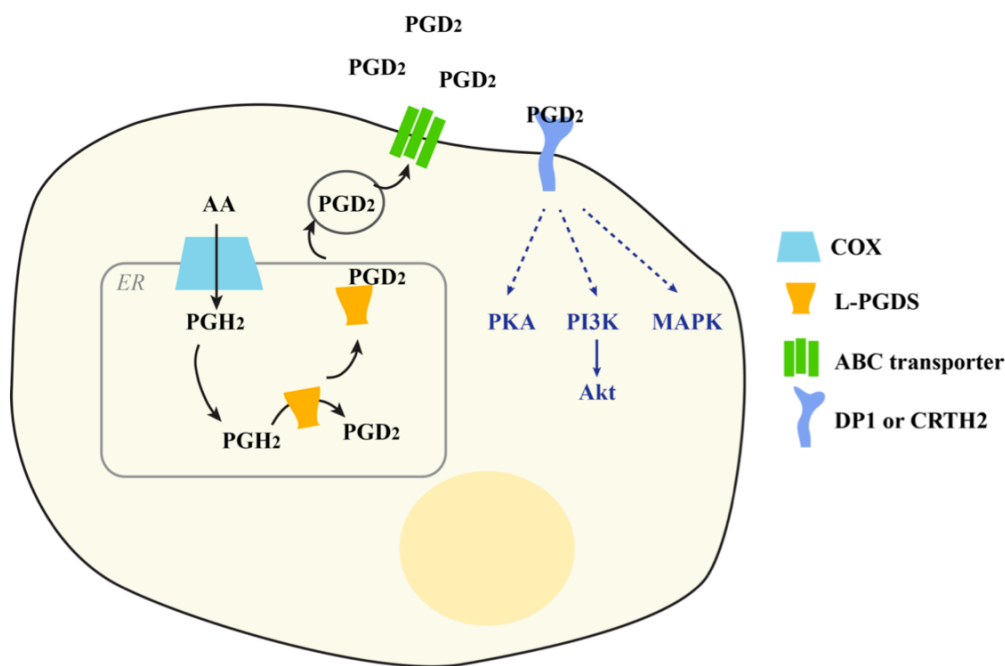


Figure 16. Proposed model of L-PGDS as an enzyme and transporter in the cell.

5 PAPER II: UNRAVELING A NOVEL PHOSPHODIESTERASE IN STRUCTURE & FUNCTION

5.1 SPHINGOMYELINASE PHOSPHODIESTERASE LIKE 3A (SMPDL3A)

Since the published of human genome in 2001, transcriptomic and proteomic studies have been accelerated by the wealth of genomic information [62, 63]. Collaborative effort in bridging the genomics proteomics gap has been ongoing for the past fourteen years. Presently, there are many proteins which their function and regulation is still unknown. Computational driven function prediction for uncharacterized protein have gone beyond sequence homology to include cell distribution, phylogenetic profiles, protein's domain analysis and gene position in chromosomes to provide a framework for experimentalist to plan their investigation [64]. Unknown enzymes can infer its function based on such predictions and are further grouped in respective family. Sphingomyelin phosphodiesterase like 3a (SMPDL3a) belongs to the aforementioned. With only a few published papers in the literature, it has been associated based on sequence homology and protein domain conservation to the function of the well-studied sphingomyelin phosphodiesterase 1 (SMPD1) or commonly known as acid sphingomyelinase (aSMase).

5.1.1 ASMase in sphingolipid hydrolysis

ASMase is an enzyme in the sphingolipid pathway that cleaves the phosphodiester bond on sphingomyelin to yield ceramide, an important signaling sphingolipid. Ceramide can activate cell apoptosis, senescence or cell-cycle arrest. It can be converted further to ceramide-1-phosphate, sphingosine or sphingosine-1-phosphate to regulate diverse cellular responses. ASMase together with other sphingolipid transforming enzymes are extremely important agents in governing the level of ceramide in the cells. Nonetheless, the structural basis for the function of aSMase is still unknown.

Dysfunctional aSMase causes lysosomal accumulation of sphingomyelin, resulting in a metabolic disorder known as Niemann Pick disease (NPD) type A or type B. Patients suffering from NPD-A and B are found to possess missense mutations on their aSMase gene leading to the loss of function phenotype [65]. In more severe cases, these mutations are detrimental to the neurophysiology of patients and cause death in their early childhood. Unfortunately, the molecular effects of the disease causing mutations is still poorly understood.

5.1.2 Calcineurin-like-phosphodiesterase family

aSMase and SMPDL3a share significant sequence homology (31%) and highly conserved binuclear metal center containing phosphodiesterase domain, leading to the presumption that SMPDL3a operates in a similar manner. The phosphodiesterase domain scaffolds a catalytic dimetal site that composed of primarily histidine, aspartate and glutamate residues for metal

coordination [66]. This domain was first identified in the Ca^{2+} dependent serine/threonine protein phosphatase - calcineurin, therefore also known as calcineurin like phosphodiesterase domain. Despite structural similarities, many enzymes with this conserved domain were found to vary in their substrate selection ranging from phosphorylated proteins, cyclic nucleotides, glycerophosphoesters, sphingolipids, nucleotide phosphates and its metabolites. Furthermore they also have diverse metal cofactor composition including Fe^{2+} , Mn^{2+} , Ni^{2+} , Mg^{2+} , Zn^{2+} , Ca^{2+} and combinations thereof [66]. These differences led to further sub-classification based on substrate and cofactor dependence, having groups like purple acid phosphatase (PAP), sphingomyelinase, nucleotidase, cyclic phosphodiesterase, alkaline phosphatase and many others (Figure 17). In 2014, Traini et al showed that despite sharing the same catalytic domain as predicted from sequence similarity, SMPDL3a has no sphingomyelinase activity [31]. In a very recent study, it is reported that SMPDL3a is the key enzyme in activating a prodrug that targets fructose-1,6-bisphosphosphate (FBPase) for Type II diabetes treatment [67]. The work in paper II reported the structure of SMPDL3a, as the first structure of this novel sub-class of the calcineurin-like phosphodiesterase. Based on the structural and biochemical information, a possible mechanism that distinguished SMPDL3a from aSMase was proposed.

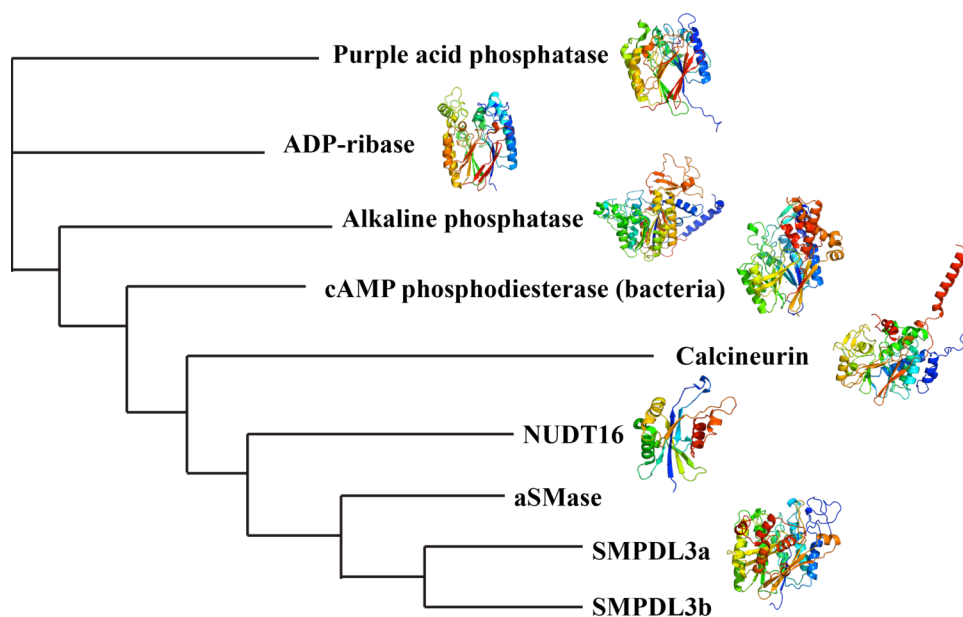


Figure 17. Multiple sequence alignment of SMPDL3a with other members of the calcineurin-like-phosphodiesterase family is represented in a cladogram generated by *Clustal Omega*.

5.2 RESULTS SUMMARY

The crystal structure of human SMPDL3a revealed the expected calcineurin like metallophosphodiesterase fold with two metals coordinated by six histidines, two aspartates and one asparagine. These residues are conserved among the aSMase and aSMase-like proteins, which are distinct as compared to other metallophosphodiesterase proteins. The

protein maintains a $\beta\alpha\beta\beta$ motif that comprises almost equal number of α helices and β strands with interspersed flexible loops. There are four N-acetylglucosamines observed in each protein. The functional role of these glycosylation sites is unknown although this glycosylation coat has been shown to protect the protein from thermal induced denaturation. The metals bound in native crystals were identified to be of Zn element based on a fluorescence scan. Additional Zn^{2+} (mM concentration) stabilized the protein significantly but inhibits its catalytic activity.

This study confirmed that the substrate of recombinant SMPDL3a *in vitro* is ATP but not sphingomyelin. Furthermore it is found to be able to hydrolyze substituted nucleotides such as CDP-choline, ADP-ribose and CDP-ethanolamine. These modified nucleotides are known substrates for another calcineurin like member, Mn-dependent ADP-ribose diphosphatase [68]. In order to better understand the substrate binding and catalysis mechanism, the modified nucleotide and various ATP analogs were soaked into SMPDL3a crystals to obtain protein-ligand complexes. In the best complex, CMP moieties, derived from CDP-choline, can be modeled into the ligand density of molecule A and C. Meanwhile only density that corresponds to phosphoribose was observed in molecule B (Figure 18). The CMP product is likely to be trapped inside the protein when the crystal was snap frozen after catalysis. Enzymatic assay and mass spectrometry analysis of the products confirmed that SMPDL3a hydrolysis of CDP-choline yields phosphocholine and CMP. The base moiety of CMP reveals specific interactions with the enzyme in molecule A. Meanwhile the ligands' α -phosphates interact with the active site's metals, His 114 and His 152 residues in all three molecules of the asymmetric unit (Figure 18). Residues His 114 and His 152 are likely to play a key role anchoring the substrate at the metal site for catalysis, they are also conserved in aSMase .

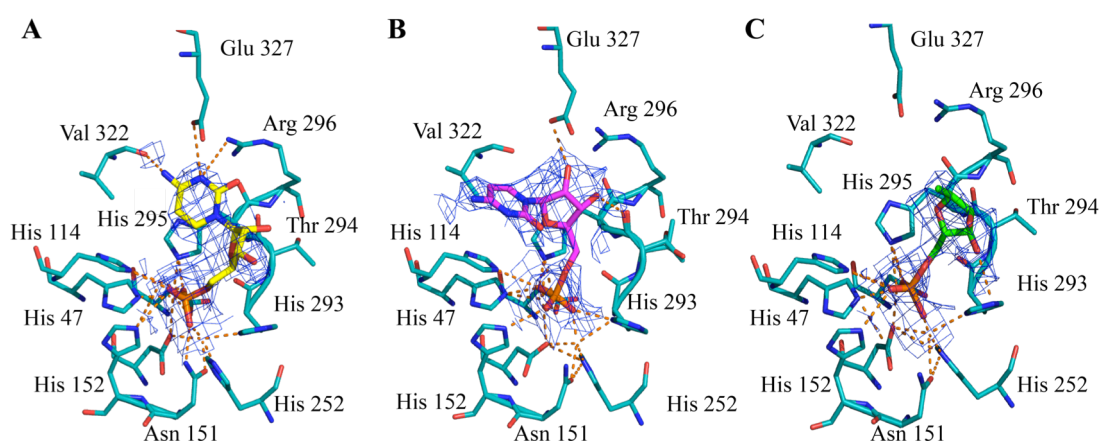


Figure 18. (A) Molecule A of the three subunits of CMP-bound SMPDL3a crystal structure showed interaction of CMP base with enzyme active site residues. (B) CMP in molecule C showed less electrostatic interaction of CMP base and enzyme. (C) Only electron density corresponding to the phosphoribose moiety is observed in molecule B.

The newly available structural data of SMPDL3a can also be used as template for aSMase homology modeling. Strong structural and sequence conservation at the binuclear catalytic

site supports similarities in the chemical mechanism between aSMase and SMPDL3a while large differences in the rest of the substrate-binding site suggest that they are likely to catalyze different substrates. The model also allowed us to map known NPD mutations onto the model of aSMase to shed light on possible molecular basis of the enzyme inactivation in cases whereby mutated proteins are expressed and shuttled to lysosome like the wild type aSMase [69].

Overall, the structural and biochemical data provide better understanding on the function and catalytic mechanism of SMPDL3a. It is still early to completely rule out the sphingomyelinase activity of SMPDL3a based on current *in vitro* studies. However, other activities, such as hydrolysis of nucleotides, appear more likely based on the present data. Further work is however required for conclusive determination of the *in-vivo* function of SMPDL3a.

6 ACKNOWLEDGEMENTS

My journey of scientific research would not be possible without a group wonderful people whom I am forever grateful of. They have filled this journey with many hues and flavors, sharing my ups and downs along the way.

First, I would like to thank my supervisor **Pär Nordlund** for giving me this opportunity to do a joint Ph.D in Karolinska Institutet and NTU, supporting me and grooming me to be an independent researcher. To my co-supervisor **Konstantin Pervushin** who is always full of creative ideas, thank you for trusting me in my work and always readily to discuss and share your NMR expertise with me.

I would like to thank **Hsiangling T** for being such a wonderful mentor in science and in life. The wonderful memories we shared in the lab and in conference will be always in my mind. Many thanks also to **Kit Y** for being such a helpful and hardworking colleague to work with as we crushed the walls of our challenging projects. To **Andreas L**, **Anna J** and **Sue-li D**, I am thankful for your advices and encouragement when I faced difficulties. And thank you **Kelly H**, for sharing my joy and distress throughout my Ph.D and ever so willing to teach, to encourage and to pray for me. I also appreciate **David G** for taking the time to teach me the tips and tricks of experimental phasing on my second project. Many thanks to the **PPP** old and new members, **Tobias C**, **Agnes H**, **Martina N**, **Lucy P**, **Dina D** and **Daniel S** for helping out with the initial screening of my target constructs in Singapore. I will also remember the good times during our lunch sessions in Biopolis with **Stephanie L**, **Dan C** and **Jeslin P**.

As for my colleagues in TROSY lab, **Baiyang**, **Leo W**, **Alistair I**, **Margaret P**, **Shubadra P**, **Rubing L**, **Edward T**, **Li Yan** and **Zhao Jing**, I have learnt so much from your inputs and experiences during our NMR study group and discussions. To **David L**, thank you for guiding me throughout my first few NMR experiments and imparting me with basic scientific skills that I am still holding on. Thank you, **Chandra Verma**, for your encouragement whenever we met up in Biopolis and on your visit to Stockholm.

During the last two and half years of my PhD in Stockholm, I would not be able to swiftly adjust and focus on my research work without the help of many. Thanks to **Lionel T** who is always willingly avails himself for meeting when I need scientific guidance, be it at the corridor or in the coffee room. I am also grateful to the responsible **Pelle**, who maintains the lab order diligently and ever so patient in dealing with all my administrative questions. To the bubbly and cheerful **Fatma G**, a great companion to have in the same office and always generously shares your laughter from KI to all synchrotrons. To **Rebecca D**, our gym trainer and pastry chef, thanks for getting me into the gym to work off some fats before eating your cakes, it's worth it! To **Rozbeh J** and **Michael R**, the men of our Ekeberg fishing team, thanks for all the hardwork and 'patience' in accommodating my 'honesty'. To **Christian L**, **Jens F**, **Daniel M**, **Esben Q**, **Catrine S** and **Michael D**, it's great to have you guys around

constantly injecting the lab with jokes, gossips, and craziness making ‘labbet’ such a relax place to work in. To the new members, **Sara, Annette, and Johan**, good luck to you all and may the new era of PN group be another exciting chapter. To our neighbor Divne group – **Christina D, Rosaria G, TC and Noor**, PN lab would not be complete without your pleasant company. Thanks to **Rosaria G**, my lab “mama”, lunch buddy and synchrotron partner for your encouragement and support. To ‘uncle’ **TC (Tien-Chye)**, whom I can speak Singlish with, your presence makes me feel not too far from Singapore. I am also grateful to the **PSF** members in KI, the reliable **Martin M** for taking care of the crystallization facilities, keeping a conducive environment for all crystallization projects to run smoothly; **Tomas** for assisting me in working with various biophysical machines; **Helena, Emma, Eleonora and Ida** for helping me in construct screening for various targets. To **Herwig, Aja, Torun, Filipa and Tobias K**, thank you for being warm and welcoming to me and ever so helpful when I needed to use some instruments from your group.

To my best friends since varsity years, **Kek Yee L, Eugene C, Kenshaun Y** for sharing my scientist’s woes no matter where I am geographically. To my dearest family-in-Christ, **Serene T, Yixian H, Susan Y, Sharmaine Z, Paul L, Sarah S, Kenny L, Emma N, Wilson W and James C**, thanks for sharing the major phase of my doctoral life in Singapore. To my “sisters” (bonded in Vietnam): **Alycia F, Xinyi Y and Rachel F**, you girls always warm my heart with all the Whatsapp messages and pictures though we are apart.

Thanks to **Lena & Björn D, Diane, Ellen, Shannon and Michelle** who warmly welcome me to Stockholm with Singapore food. I am also blessed to have known **Denise L, Grace C and Yan Han**, whom together we shared the many wonderful gatherings. To **Mariam Z**, the first person I got in touch in Stockholm and have helped me so much in settling down, you are truly an angel. To **Deborah & Pieter T, Polly W, Nienke K, Jessica P, Faezzah, Nurzian & Andreas, Sharon, Jingmei, Sam & Jinny, Xiaohui, Evelyn, Ying Chun, Saba, Madhan, Sandra, Ida H, Nina & Mirco, Linnea, Joseph & Cindy, Alvin T, Ming Wei and Yossa**, thank you for blessing me in one way or another in Stockholm.

To my dearest family **Mum, Dad and Belle** who have seen me through this journey, loving me and praying for me, thank you for never ceasing to encourage me and always lend a listening ear when I am in distress. Of course to **Jia Chi Y**, thank you for your love, encouragement and support that keep me going strong and inspire me to be the best I can be.

Lastly and most importantly, all glory and praise to **God** for His faithfulness, provision and grace throughout all these years. It would be impossible without Him!

7 REFERENCES

1. de Réaumur, R., *Observations sur la digestion des oiseaux*. Histoire de l'academie royale des sciences, 1752. **1752**(266).
2. Payen A, P.J.F., *Memoir on diastase, the principal products of its reactions, and their applications to the industrial arts*. . Annales de chimie et de physique, 1833. **2**(53): p. 19.
3. Pasteur, L., *Nouveaux faits concernant l'histoire de la fermentation alcoolique*. Annales de Chimie et de Physique, 1858. **3**(52): p. 14.
4. Berzelius, J.J., *Årsberättelsen om framsteg i fysik och kemi*. Royal Swedish Academy of Sciences, 1835.
5. Buchner, E., *Alcholic fermentation without yeast*. Chem. Ber, 1897. **30**: p. 8.
6. Fischer, E., *Einfluss der Configuration auf die Wirkung der Enzyme*. Berichte der Deutschen Chemischen Gesellschaft 1894. **27**(9): p. 2985.
7. Michaelis, L., et al., *The original Michaelis constant: translation of the 1913 Michaelis-Menten paper*. Biochemistry, 2011. **50**(39): p. 8264-9.
8. Sumner, J.B., *The Isolation and Crystallization of the Enzyme Urease*. The Journal of Biological Chemistry, 1926(69): p. 6.
9. Sumner, J.B., *The Digestion and Inactivation of Crystalline Urease by Pepsin and by Papain*. The Journal of Biological Chemistry, 1932(98): p. 9.
10. Northrop, J.H., *Crystalline Pepsin*. Science, 1929. **69**(1796): p. 580.
11. Stanley, W.M., *Isolation of a Crystalline Protein Possessing the Properties of Tobacco-Mosaic Virus*. Science, 1935. **81**(2113): p. 644-5.
12. Cech, T.R., *A model for the RNA-catalyzed replication of RNA*. Proceedings of the National Academy of Sciences of the United States of America, 1986. **83**(12): p. 4360-3.
13. Bernal, J.D., Crowfoot D. , *X-Ray Photographs of Crystalline Pepsin*. Nature, 1934(133): p. 2.
14. Watson, J.D. and F.H. Crick, *Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid*. Nature, 1953. **171**(4356): p. 737-8.
15. Kendrew, J.C., et al., *A three-dimensional model of the myoglobin molecule obtained by x-ray analysis*. Nature, 1958. **181**(4610): p. 662-6.
16. Blake, C.C., et al., *Structure of hen egg-white lysozyme. A three-dimensional Fourier synthesis at 2 Angstrom resolution*. Nature, 1965. **206**(4986): p. 757-61.
17. Johnson, L.N., Phillips, D.C., *Structure of Some Crystalline Lysozyme-Inhibitor Complexes determined by X-Ray Analysis and 6Å Resolution*. Nature, 1965(206): p. 2.
18. Matthews, B.W., et al., *Three-dimensional structure of tosyl-alpha-chymotrypsin*. Nature, 1967. **214**(5089): p. 652-6.
19. Kraut, J., *Serine proteases: structure and mechanism of catalysis*. Annual review of biochemistry, 1977. **46**: p. 331-58.

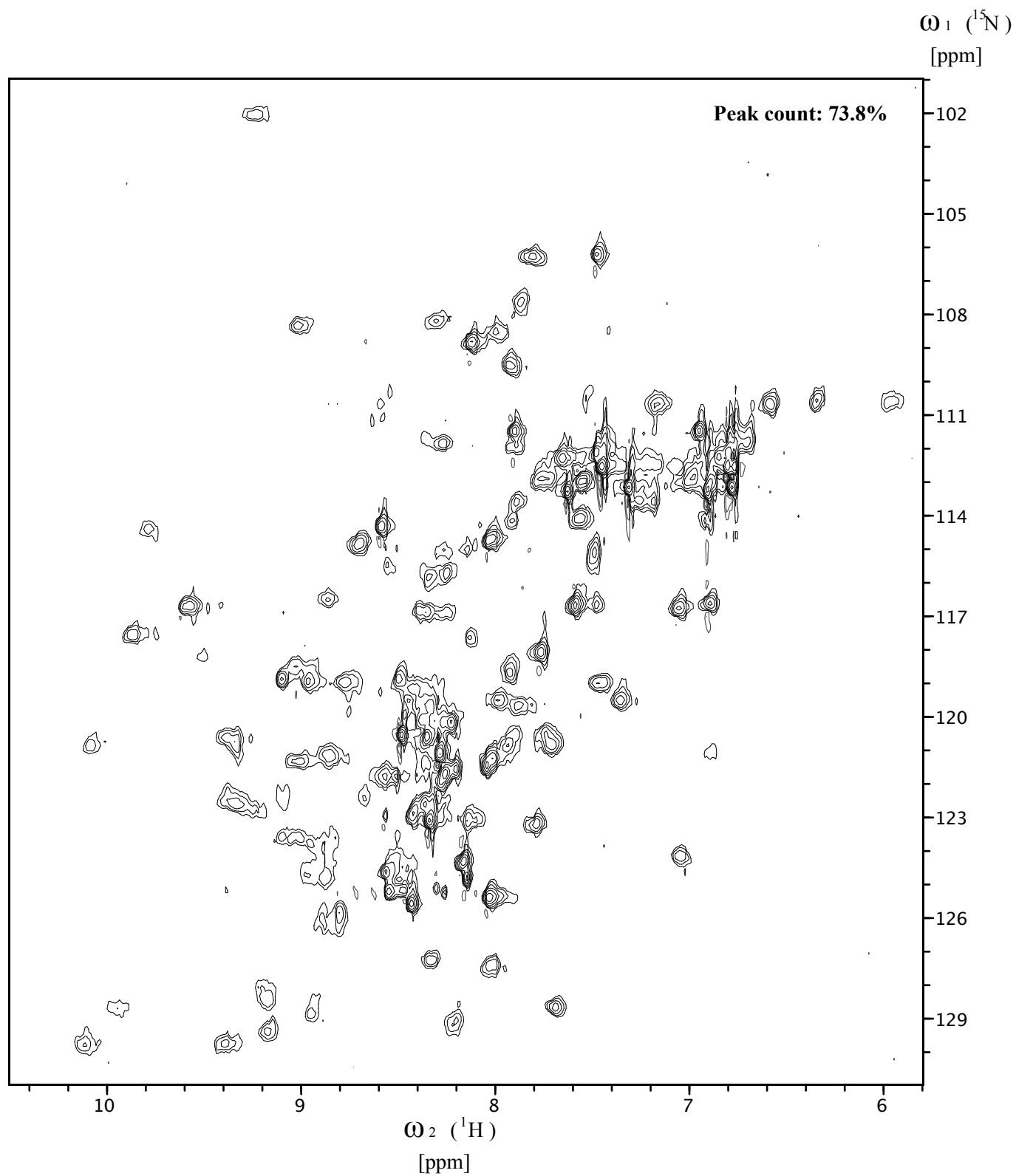
20. Koshland, D.E., *Application of a Theory of Enzyme Specificity to Protein Synthesis*. Proceedings of the National Academy of Sciences of the United States of America, 1958. **44**(2): p. 98-104.
21. Hult, K. and P. Berglund, *Enzyme promiscuity: mechanism and applications*. Trends in biotechnology, 2007. **25**(5): p. 231-8.
22. Reid, T.W. and D. Fahrney, *The pepsin-catalyzed hydrolysis of sulfite esters*. Journal of the American Chemical Society, 1967. **89**(15): p. 3941-3.
23. Nakagawa, Y. and M.L. Bender, *Modification of alpha-chymotrypsin by methyl p-nitrobenzenesulfonate*. Journal of the American Chemical Society, 1969. **91**(6): p. 1566-7.
24. O'Brien, P.J. and D. Herschlag, *Catalytic promiscuity and the evolution of new enzymatic activities*. Chemistry & biology, 1999. **6**(4): p. R91-R105.
25. Copley, S.D., *Enzymes with extra talents: moonlighting functions and catalytic promiscuity*. Current opinion in chemical biology, 2003. **7**(2): p. 265-72.
26. Gamage, N.U., et al., *The structure of human SULT1A1 crystallized with estradiol. An insight into active site plasticity and substrate inhibition with multi-ring substrates*. The Journal of biological chemistry, 2005. **280**(50): p. 41482-6.
27. Theodossis, A., et al., *The structural basis for substrate promiscuity in 2-keto-3-deoxygluconate aldolase from the Entner-Doudoroff pathway in Sulfolobus solfataricus*. The Journal of biological chemistry, 2004. **279**(42): p. 43886-92.
28. Guillas, I., et al., *Human homologues of LAG1 reconstitute Acyl-CoA-dependent ceramide synthesis in yeast*. The Journal of biological chemistry, 2003. **278**(39): p. 37083-91.
29. Kitatani, K., J. Idkowiak-Baldys, and Y.A. Hannun, *The sphingolipid salvage pathway in ceramide metabolism and signaling*. Cellular signalling, 2008. **20**(6): p. 1010-8.
30. Tomiuk, S., et al., *Cloned mammalian neutral sphingomyelinase: functions in sphingolipid signaling?* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(7): p. 3638-43.
31. Traini, M., et al., *Sphingomyelin phosphodiesterase acid-like 3A (SMPDL3A) is a novel nucleotide phosphodiesterase regulated by cholesterol in human macrophages*. The Journal of biological chemistry, 2014. **289**(47): p. 32895-913.
32. Berman, H.M., et al., *The Protein Data Bank*. Nucleic acids research, 2000. **28**(1): p. 235-42.
33. Savitsky, P., et al., *High-throughput production of human proteins for crystallization: the SGC experience*. Journal of structural biology, 2010. **172**(1): p. 3-13.
34. Jarvis, D.L., *Baculovirus-insect cell expression systems*. Methods in enzymology, 2009. **463**: p. 191-222.
35. Luckow, V.A., et al., *Efficient generation of infectious recombinant baculoviruses by site-specific transposon-mediated insertion of foreign genes into a baculovirus genome propagated in Escherichia coli*. Journal of virology, 1993. **67**(8): p. 4566-79.

36. Harrison, R.L. and D.L. Jarvis, *Protein N-glycosylation in the baculovirus-insect cell expression system and engineering of insect cells to produce "mammalianized" recombinant glycoproteins*. Advances in virus research, 2006. **68**: p. 159-91.
37. Ericsson, U.B., et al., *Thermofluor-based high-throughput stability optimization of proteins for structural studies*. Analytical biochemistry, 2006. **357**(2): p. 289-98.
38. Vedadi, M., et al., *Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination*. Proceedings of the National Academy of Sciences of the United States of America, 2006. **103**(43): p. 15835-40.
39. Benjwal, S., et al., *Monitoring protein aggregation during thermal unfolding in circular dichroism experiments*. Protein science : a publication of the Protein Society, 2006. **15**(3): p. 635-9.
40. Freyer, M.W. and E.A. Lewis, *Isothermal titration calorimetry: experimental design, data analysis, and probing macromolecule/ligand binding and kinetic interactions*. Methods in cell biology, 2008. **84**: p. 79-113.
41. Bragg, W.H. and W.L. Bragg, *The Reflection of X-rays by Crystals*. Proceedings of The Royal Society A, 1913. **88**(605): p. 10.
42. Bergfors, T., ed. *Protein Crystallization*. 2009, International University Line: San Diego.
43. McPherson, A., *Introduction to protein crystallization*. Methods, 2004. **34**(3): p. 254-65.
44. D. Sherwood, J.C., *Crystals, X-rays and Proteins*. 2011, Oxford: Oxford University Press.
45. Scapin, G., *Molecular replacement then and now*. Acta crystallographica. Section D, Biological crystallography, 2013. **69**(Pt 11): p. 2266-75.
46. Rhodes, G., *Crystallography made crystal clear: A guide for Users of Macromolecular Models*. 2000, San Diego: Academic Press.
47. Vonnrhein, C., et al., *Automated structure solution with autoSHARP*. Methods in molecular biology, 2007. **364**: p. 215-30.
48. Emsley, P., et al., *Features and development of Coot*. Acta crystallographica. Section D, Biological crystallography, 2010. **66**(Pt 4): p. 486-501.
49. Langer, G., et al., *Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7*. Nature protocols, 2008. **3**(7): p. 1171-9.
50. Wuthrich, K., *NMR of Protein and Nucleic Acids* 1986, USA: John Wiley & Sons Inc.
51. Williamson, M.P., *Using chemical shift perturbation to characterise ligand binding*. Progress in Nuclear Magnetic Resonance Spectroscopy, 2013(73): p. 16.
52. Warschawski, D.E., et al., *Choosing membrane mimetics for NMR structural studies of transmembrane proteins*. Biochimica et biophysica acta, 2011. **1808**(8): p. 1957-74.
53. Harizi, H., J.B. Corcuff, and N. Gualde, *Arachidonic-acid-derived eicosanoids: roles in biology and immunopathology*. Trends in molecular medicine, 2008. **14**(10): p. 461-9.

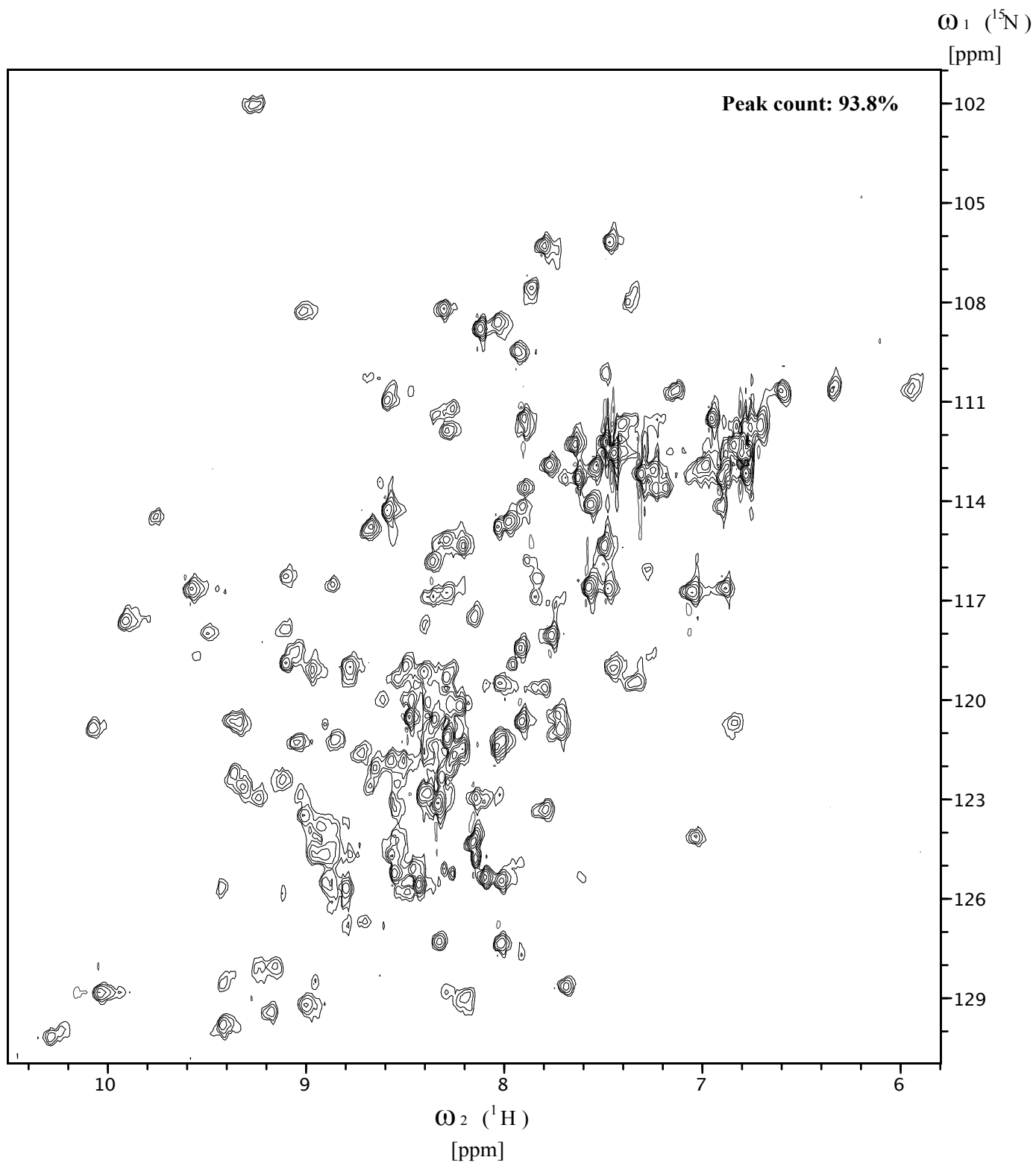
54. Wymann, M.P. and R. Schneiter, *Lipid signalling in disease*. Nature reviews. Molecular cell biology, 2008. **9**(2): p. 162-76.
55. Jabbour, H.N. and K.J. Sales, *Prostaglandin receptor signalling and function in human endometrial pathology*. Trends in endocrinology and metabolism: TEM, 2004. **15**(8): p. 398-404.
56. Urade, Y. and N. Eguchi, *Lipocalin-type and hematopoietic prostaglandin D synthases as a novel example of functional convergence*. Prostaglandins & other lipid mediators, 2002. **68-69**: p. 375-82.
57. Zhou, Y., et al., *Structure-function analysis of human l-prostaglandin D synthase bound with fatty acid molecules*. FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 2010. **24**(12): p. 4668-77.
58. Forneris, F. and A. Mattevi, *Enzymes without borders: mobilizing substrates, delivering products*. Science, 2008. **321**(5886): p. 213-6.
59. de Waart, D.R., et al., *Multidrug resistance associated protein 2 mediates transport of prostaglandin E2*. Liver international : official journal of the International Association for the Study of the Liver, 2006. **26**(3): p. 362-8.
60. Matsuoka, T., et al., *Prostaglandin D2 as a mediator of allergic asthma*. Science, 2000. **287**(5460): p. 2013-7.
61. Sawyer, N., et al., *Molecular pharmacology of the human prostaglandin D2 receptor, CRTH2*. British journal of pharmacology, 2002. **137**(8): p. 1163-72.
62. Venter, J.C., et al., *The sequence of the human genome*. Science, 2001. **291**(5507): p. 1304-51.
63. Tyers, M. and M. Mann, *From genomics to proteomics*. Nature, 2003. **422**(6928): p. 193-7.
64. Eisenberg, D., et al., *Protein function in the post-genomic era*. Nature, 2000. **405**(6788): p. 823-6.
65. Ferlinz, K., R. Hurwitz, and K. Sandhoff, *Molecular basis of acid sphingomyelinase deficiency in a patient with Niemann-Pick disease type A*. Biochemical and biophysical research communications, 1991. **179**(3): p. 1187-91.
66. Wilcox, D.E., *Binuclear Metallohydrolases*. Chemical reviews, 1996. **96**(7): p. 2435-2458.
67. Kubota, K., et al., *Identification of activating enzymes of a novel FBPase inhibitor prodrug, CS-917*. Pharmacology research & perspectives, 2015. **3**(3): p. e00138.
68. Canales, J., et al., *Mn²⁺-dependent ADP-ribose/CDP-alcohol pyrophosphatase: a novel metallophosphoesterase family preferentially expressed in rodent immune cells*. The Biochemical journal, 2008. **413**(1): p. 103-13.
69. Ferlinz, K., et al., *Molecular analysis of the acid sphingomyelinase deficiency in a family with an intermediate form of Niemann-Pick disease*. American journal of human genetics, 1995. **56**(6): p. 1343-9.

8 APPENDICES

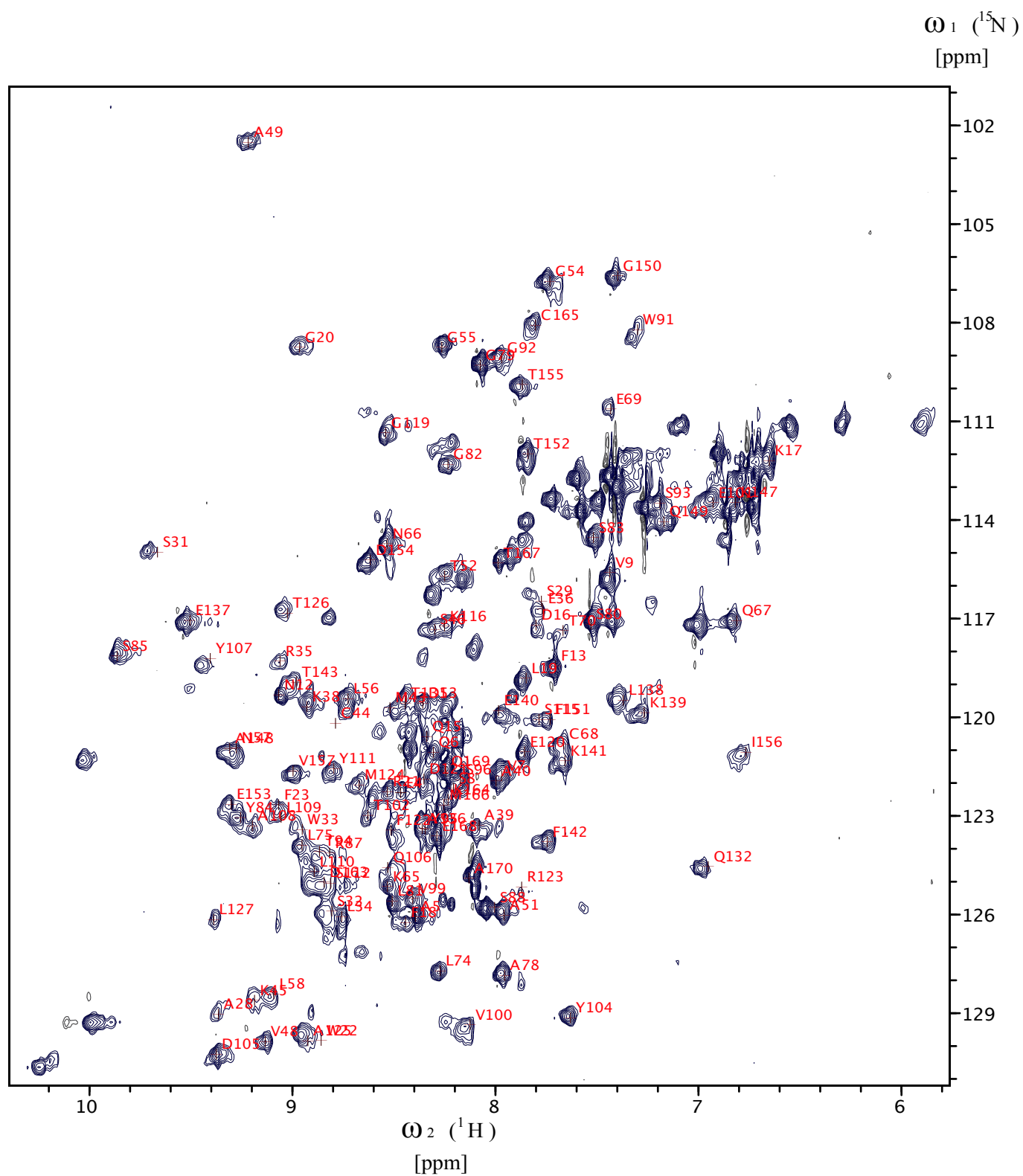
Appendix I : 2D HSQC of 0.39 mM L-PGDS in 20 mM HEPES pH 6.5, 150 mM NaCl, 2 mM MTCEP, 10% D₂O, 0.5 mM DSS



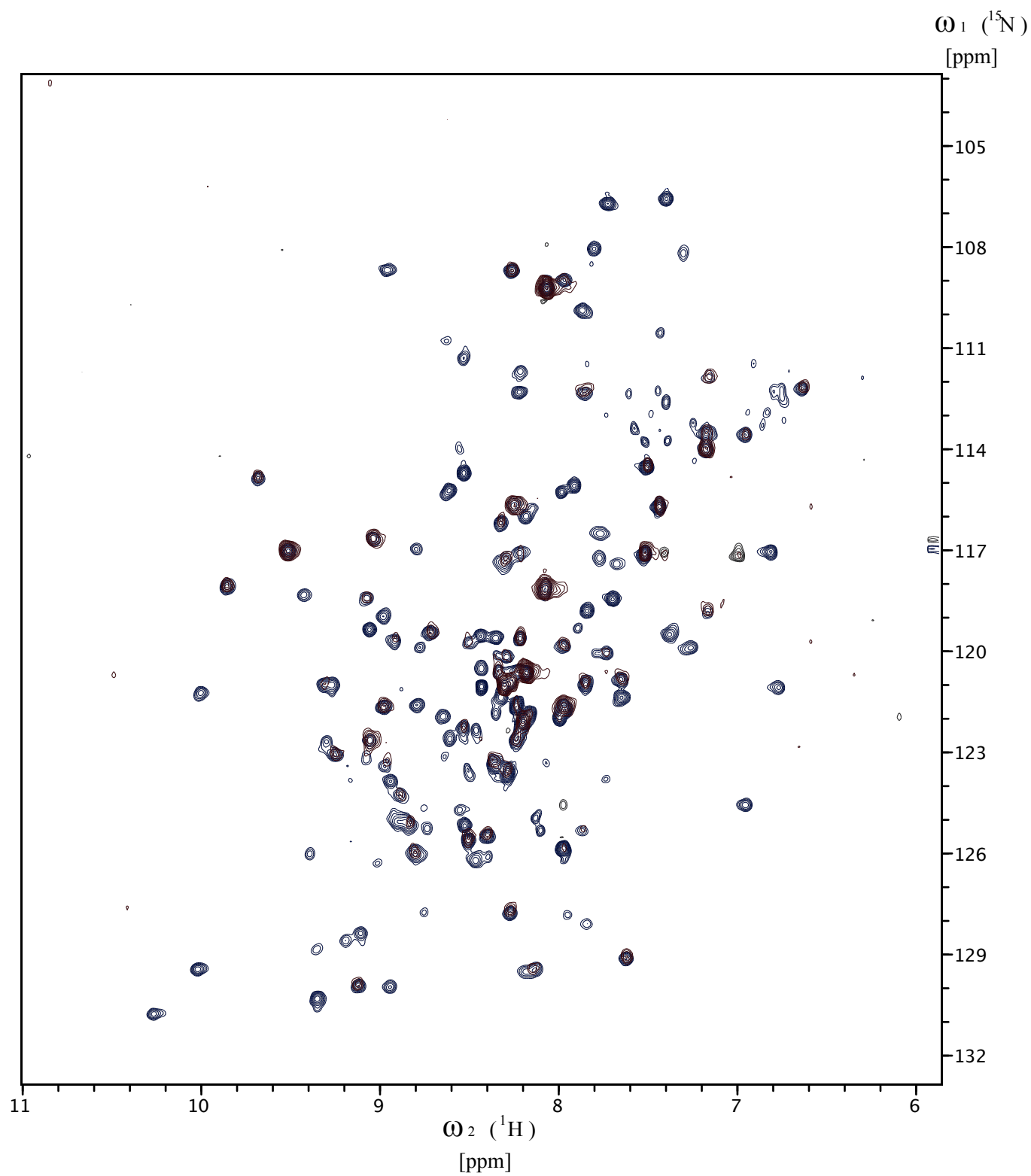
Appendix II : 2D HSQC of 0.39 mM L-PGDS in 20 mM HEPES pH6.5, 150 mM NaCl, 2 mM TCEP
10% D₂O, 0.5 mM DSS and 2.6 mM SA U44069



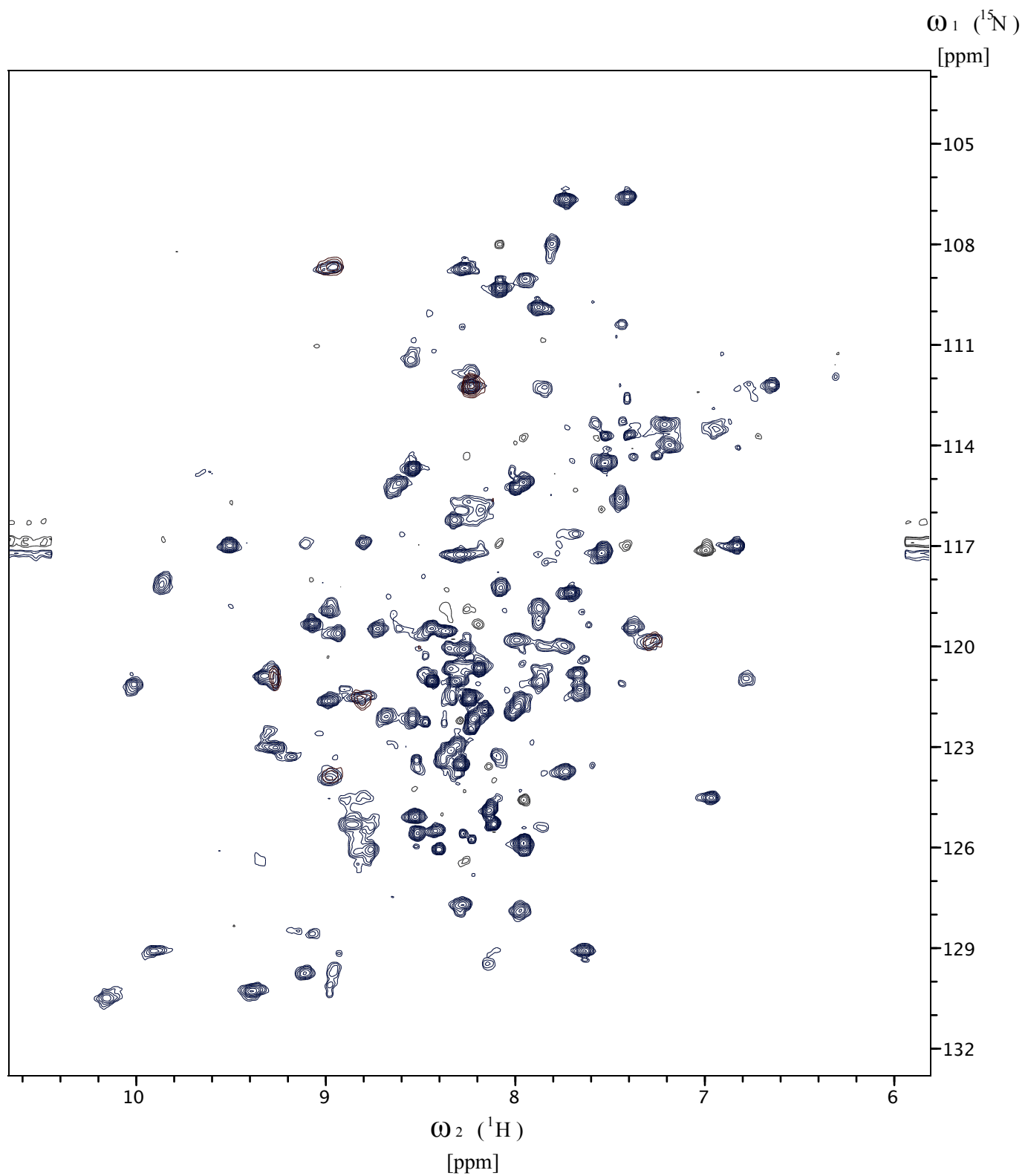
Appendix III : Backbone assignment of L-PGDS protein



Appendix IV : Overlay of ^{13}C Alanine labeled sample HSQC spectrum (blue) with HNCO spectrum (red).



Appendix V : Overlay of ^{13}C Leucine labeled HSQC spectrum (blue) with HNCO spectrum (red).



Paper I

Structural and dynamic insights into substrate binding and catalysis of human lipocalin prostaglandin D synthase^[S]

Sing Mei Lim,^{*,†} Dan Chen,^{*} Hsiangling Teo,^{*} Annette Roos,^{1,†} Anna Elisabet Jansson,^{*} Tomas Nyman,[†] Lionel Trésaugues,[†] Konstantin Pervushin,^{2,*} and Pär Nordlund^{2,*†}

Division of Structural Biology and Biochemistry,^{*} Nanyang Technological University, Singapore; and Department of Medical Biochemistry and Biophysics,[†] Karolinska Institutet, Stockholm, Sweden

Abstract Lipocalin prostaglandin D synthase (L-PGDS) regulates synthesis of an important inflammatory and signaling mediator, prostaglandin D₂ (PGD₂). Here, we used structural, biophysical, and biochemical approaches to address the mechanistic aspects of substrate entry, catalysis, and product exit of this enzyme. Structure of human L-PGDS was solved in a complex with a substrate analog (SA) and in ligand-free form. Its catalytic Cys 65 thiol group was found in two different conformations, each making a distinct hydrogen bond network to neighboring residues. These help in elucidating the mechanism of the cysteine nucleophile activation. Electron density for ligand observed in the active site defined the substrate binding regions, but did not allow unambiguous fitting of the SA. To further understand ligand binding, we used NMR spectroscopy to map the binding sites and to show the dynamics of protein-substrate and protein-product interactions. A model for ligand binding at the catalytic site is proposed, showing a second binding site involved in ligand exit and entry. NMR chemical shift perturbations and NMR resonance line-width alterations (observed as changes of intensity in two-dimensional cross-peaks in [¹H, ¹⁵N]-transfer relaxation optimization spectroscopy) for residues at the Ω loop (A-B loop), E-F loop, and G-H loop besides the catalytic sites indicate involvement of these residues in ligand entry/egress.—Lim, S. M., D. Chen, H. Teo, A. Roos, A. E. Jansson, T. Nyman, L. Trésaugues, K. Pervushin, and P. Nordlund. **Structural and dynamic insights into substrate binding and catalysis of human lipocalin prostaglandin D synthase.** *J. Lipid Res.* 2013. 54: 1630–1643.

Supplementary key words lipid signaling • X-ray crystallography • nuclear magnetic resonance spectroscopy • lipophilic substrate • protein-lipid interaction • β trace protein

Prostaglandins are members of the eicosanoid family alongside leukotrienes, lipoxins, epoxy fatty acids, and

thromboxane (1, 2). They are involved in various biological responses such as inflammation, allergy, nociception, cell growth, cell proliferation, and tumorigenesis (3, 4). Hormonal or stress stimuli induce hydrolysis of glycerophospholipids to free arachidonic acid (AA) by cytoplasmic phospholipase A₂. The substrate AA enters the cyclooxygenase (COX) pathway to yield prostaglandin H₂ (PGH₂) (5), which acts as a substrate for various downstream cell-specific enzymes producing prostaglandin D₂ (PGD₂), prostaglandin I₂ (PGI₂), prostaglandin E₂, prostaglandin F_{2α}, and thromboxane. There are two types of prostaglandin D synthase (PGDS), namely hematopoietic PGDS (H-PGDS) and lipocalin prostaglandin D synthase (L-PGDS). They are distributed in different tissues and distinct in structure; nonetheless both perform the same catalytic reaction to produce PGD₂.

L-PGDS is a major PGD₂ synthase in the central nervous system, heart, and reproductive tissue, catalyzing the isomerization of PGH₂ to PGD₂ (6). This protein is involved in pain induction, nephropathy, immunomodulation, adipocyte differentiation, lipid and carbohydrate metabolism, as well as sleep-wake regulation (7–11). Similar

Abbreviations: AA, arachidonic acid; c.m.c, critical micelle concentration; COX, cyclooxygenase; DPC, dodecylphosphocholine; HSQC, heteronuclear single quantum correlation; H-PGDS, hematopoietic prostaglandin D synthase; IPTG, isopropyl-beta-D-l-thiogalactopyranoside; L-PGDS, lipocalin prostaglandin D synthase; MOX, methoxylamine; NOESY, nuclear Overhauser effect spectroscopy; OD, optical density; PA, product analog; PGD₂, prostaglandin D₂; PGDS, prostaglandin D synthase; PGES, prostaglandin E synthase; PGH₂, prostaglandin H₂; PGI₂, prostaglandin I₂; PGIS, prostaglandin I₂ synthase; PPAR, peroxisome proliferator-activated receptor; RA, retinoic acid; SA, substrate analog; TCEP, tris-(2-carboxyethyl)phosphine; TROSY, transfer relaxation optimization spectroscopy; TSA, thermal shift assay; TXAS, thromboxane A₂ synthase.

¹Present address of A. Roos: Department of Cell and Molecular Biology, Structural and Molecular Biology, Uppsala University, Uppsala, Sweden.

²To whom correspondence should be addressed.

e-mail: PNORDLUND@ntu.edu.sg (P.N.); KPERVUSHIN@ntu.edu.sg (K.P.)

^[S] The online version of this article (available at <http://www.jlr.org>) contains supplementary data in the form of four figures and two tables.

Copyright © 2013 by the American Society for Biochemistry and Molecular Biology, Inc.

This article is available online at <http://www.jlr.org>

This work was financially supported by the Singapore National Research Foundation Competitive Research Program (CRP) funding scheme, NRF2008NRF-CRP002-067.

Manuscript received 21 December 2012 and in revised form 26 February 2013.

Published, JLR Papers in Press, March 12, 2013

DOI 10.1194/jlr.M035410

to other prostaglandin synthases such as COX1, COX2, prostaglandin E synthase (PGES), and prostaglandin F synthase, L-PGDS is also associated with cancer progression (4). However, unlike PGES's prominent role in enhancing tumorigenesis and metastasis, PGDS expression is down-regulated in cerebral spinal fluid of patients suffering from a brain tumor (12). PGD₂ production by L-PGDS activates peroxisome proliferator-activated receptor (PPAR) γ to inhibit tumor cell growth (13). Apart from its anti-tumorigenic function, PGD₂ has also been identified as a potential diagnostic marker for hypertension-related renal injury and gentamicin-induced renal impairment (14, 15). Thus, L-PGDS knockout mice have been used as a model for diabetic studies (16). In addition, studies by various research groups have proposed the diagnostic application of L-PGDS in obstructive azoospermia (17), and neurological disorders (18). Recently L-PGDS and PGD₂ were linked to hair loss and were found to be overexpressed on the bald scalp of patients suffering from androgenetic alopecia (19).

L-PGDS, in addition to its enzymatic role, also functions as a transporter for lipophilic ligands. Interestingly, chicken and fish L-PGDS lack enzymatic activity and operate primarily as lipophilic ligand vehicles (20). Human L-PGDS binds to a range of all-*trans*-retinoic acids (RAs), biliverdin, bilirubin, and thyroid hormones with good affinity (21). Hence, it has been proposed to be a vital RA transporter and scavenger for bile pigments (22, 23). Lipocalin family members often display promiscuous binding to other lipophilic compounds beside their endogenous ligands (24). Its broad specificity in ligand binding suggests that L-PGDS could be exploited as a drug delivery system for lipophilic molecules (25).

Previous structural studies have been achieved using mutants whereby the catalytic cysteine residue (Cys 65 in human L-PGDS) has been substituted for alanine. Structures of mouse L-PGDS determined using both NMR and X-ray crystallography established the lipocalin family fold of the protein as well as the presence of a large cavity for substrate and transport-ligand binding (26–28). NMR titration experiments also predict the binding of substrate and hydrophobic transport-ligand to the large cavity (29). Mutation analysis has identified conserved Cys 65 as the critical residue that confers enzymatic activity (27). Subsequently, the crystal structure of mutant human L-PGDS (C65A) with bound fatty acids was determined, revealing an extended binding mode of two associated lipids in the L-PGDS pocket (30). Despite extensive efforts from several groups, detailed understanding of L-PGDS substrate binding and catalytic mechanism is still limited. No informative structural complex with substrate or product analogs has been obtained thus far.

We have attempted to determine the substrate binding modes of L-PGDS with substrate and product analogs using X-ray crystallography. Our data define the substrate/product binding regions in the active site and identify potential interactions of the substrate with the protein. Comparison of structures with and without analog revealed conformational changes in the protein, mediating substrate entry, enzyme catalysis, and product egress. This is

the first structure of wild-type human L-PGDS with substrate analog (SA) U44069, allowing a detailed description of the environment of the catalytic cysteine. Finally, NMR titration studies with substrate and product analogs as well as membrane mimetics further elucidated the dynamics of substrate binding and a potential mode of interaction with the membrane.

MATERIALS AND METHODS

Constructs were obtained from the Mammalian Genomic Consortium. All prostaglandin analogs and chemical assays were purchased from Cayman Chemicals.

Cultivation and expression

The sequence encoding human L-PGDS was subcloned into pNIC-CH2 vector with C-terminal 6xHis tag and expressed in Rosetta BL21-DE3 *Escherichia coli* in Terrific Broth media consisting of tryptone, yeast extracts, and glycerol. Cells were grown at 37°C until OD_{600 nm} = 1.0 and induced with 0.5 mM isopropyl-beta-D-1-thiogalactopyranoside (IPTG) at 18°C for 18 h. Cells were resuspended and sonicated in buffer containing 20 mM HEPES pH 8.0, 300 mM NaCl, 2 mM tris-(2-carboxyethyl)phosphine (TCEP), EDTA-free protease inhibitor cocktail (Merck), and 1 μ l of Benzonase (Merck) per 1 l culture. After centrifugation, the lysate was loaded onto a 1 ml Nickel-affinity column equilibrated with Buffer A (20 mM HEPES pH 8.0, 300 mM NaCl, 2 mM TCEP, 10 mM imidazole) and eluted with Buffer B (20 mM HEPES pH 8.0, 300 mM NaCl, 2 mM TCEP, 500 mM imidazole). Fractions containing L-PGDS were pooled and further purified by gel filtration using HiLoad 16/60 Superdex 75 equilibrated with Buffer C (20 mM HEPES pH 6.5, 150 mM NaCl, 2 mM TCEP). The color of protein fractions changed from yellow to colorless with increasing elution time. Protein elute later appeared less yellow and showed good dispersion in ¹⁵N-heteronuclear single quantum correlation (HSQC) measurement. Only colorless fractions were pooled and concentrated to 5.5 mg/ml for crystallization trials. Protein identity was confirmed by mass spectrometry and Western blot analysis.

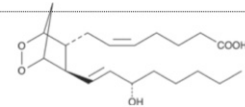
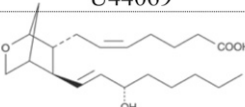
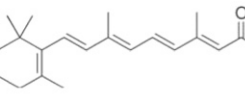
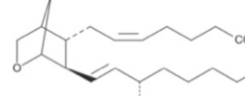
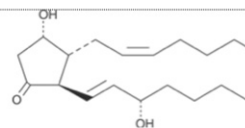
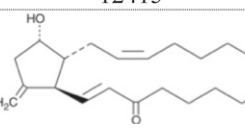
Crystallization

L-PGDS was cocrystallized with SA U44069 9,11-epoxymethano PGH₂ (Table 1) in condition A (0.1 M potassium thiocyanate and 30% PEG-MME 2000) in 1:1 protein-reservoir ratio. Crystals appeared after 5 days of incubation at 4°C by hanging drop vapor diffusion. Cocrystals were also obtained in condition B (1.4 M tri-sodium citrate pH 6.5) using a similar method except in 2:1 protein-reservoir ratio. Micro-crystals from condition A were used to seed crystallization of ligand-free L-PGDS in the same condition but in the absence of SA U44069. Crystals from condition A were cryo-protected using reservoir with 25% glycerol added while crystals from condition B were cryo-protected with 1.6 M tri-sodium citrate solution.

Data collection and processing

Native data sets were collected at beam line (BL)13C1 and BL13B1 at the National Synchrotron Radiation Research Center, Taiwan, Republic of China. Data sets were processed using HKL-2000 (31) and iMosflm (32), phases were generated by molecular replacement (MR, Phaser) (33) with mouse L-PGDS (PDB ID: 2CZT). Automatic building of the structure was carried out using ARP/wARP 7.3 (34), ligand fitting was performed in Coot 0.6.2

TABLE 1. Chemical representation of natural substrates, ligands, and analogs used in this study

Natural ligands	Analog
Substrate PGH ₂	Substrate analogs U44069
	
All-trans-retinoic acid	U46619
	
Product PGD ₂	Product analog 12415
	

(35), and refinement was performed using autoBUSTER (Global Phasing Limited) and REFMAC5 (36) in the CCP4 suite (37). Table 2 lists the final statistics for L-PGDS-ligand structure.

NMR protein sample preparation

¹⁵N single-labeled protein was prepared in M9 minimal media supplied with ¹⁵N-labeled ammonium chloride. Protein was induced at OD_{600 nm} = 0.7 with 0.5 mM IPTG at 18°C for 18 h. ¹³C and ¹⁵N double-labeled sample was prepared by supplementing ¹³C-labeled glucose at the beginning of cultivation and inducing at OD_{600 nm} = 0.6 with 0.5 mM IPTG at 18°C for 20 h. The final protein concentration was 0.3 mM for the ¹⁵N-labeled sample and 0.27 mM for the double-labeled sample. Both samples were in buffer consisting of 20 mM HEPES pH 6.5, 150 mM NaCl, 2 mM TCEP, 10% D₂O, and 0.5 mM 4,4-dimethyl-4-silapentane-1-sulfonic acid for NMR experiments. Selective labeling of ¹³C amino acids, leucine, and alanine in the background of [¹H-¹⁵N] were carried out as described in the study by Wagner et al. (38).

NMR spectroscopy

NMR spectroscopy experiments were measured using Bruker Avance 700 MHz with triple resonance z-axis gradient cryoprobe at 298 K. All experiments were recorded using a protein concentration of 0.3–0.5 mM. Data were processed by Topspin 2.2 (Bruker Corporation) and analyzed by CARRA (www.nmr.ch) (39). ¹⁵N-nuclear Overhauser effect spectroscopy (NOESY)-HSQC was measured at the Swedish NMR Center, Gothenburg using Varian 800 MHz. The backbone resonances were assigned using transfer relaxation optimization spectroscopy (TROSY)-HNCA, CBCA(CO)NH, and ¹⁵N-NOESY-HSQC experiments. TROSY-HN(CO) experiments were applied to further aid in residue-specific resonance assignment.

NMR titration

Unlabeled ligands SA U44069, SA U46619, and PA 12415 (Table 1) were titrated in ¹⁵N-labeled protein sample and a series of ¹⁵N-HSQC experiments were measured for analysis of ligand

TABLE 2. Data collection and refinement statistic

	SA Complex 1 (4IMO)	APO (4IMN)
Crystallization condition	0.1M potassium thiocyanate, 30% PEG MME 2000	0.1M potassium thiocyanate, 30% PEG MME 2000
Data collection		
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
Cell dimension		
a, b, c (Å)	36.24, 56.38, 72.92	36.41, 56.41, 72.99
α, β, γ	90, 90, 90	90, 90, 90
Wavelength	44.6–1.88 (1.92–1.88)	44.6–2.09 (2.09–2.16)
Resolution (Å)		
R _{sym}	0.062 (0.412)	0.060 (0.303)
I/σI	17.8 (3.6)	18.6 (3.23)
Completeness (%)	100 (99.2)	94.6 (95.3)
Redundancy	5.4 (5.5)	5.0 (4.7)
Refinement		
Resolution (Å)	1.88	2.09
No. reflections	13,902	4,675
R _{work} /R _{free}	0.184/0.240	0.174/0.243
No of atoms	1,386	1,303
Protein	1,241	1,221
Ligand	56	17
Water	89	65
B-factors	22.9	33.9
R.m.s deviations		
Bond length (Å)	0.020	0.018
Bond angles (°)	2.012	1.969
Ramachandran plot		
Most favored (%)	96.8	97.4
Allowed (%)	3.2	2.6
Disallowed (%)	0.0	0.0

Reported datasets are obtained from two different crystal structures. Values in parentheses are for highest-resolution shell. APO, apoenzyme.

binding interactions at concentrations of 2 mM, 3 mM, and 4 mM. Protein-ligand complex and detergent interactions were evaluated by adding dodecylphosphocholine (DPC) into the protein-ligand solution. DPC stock solution of 60 mM was prepared by dissolving the detergent in Millipore water. The samples contained 0.35 mM protein, 4 mM SA U44069 or PA 12415, and 3 mM DPC [critical micelle concentration (c.m.c.), 1.2 mM]. Chemical shift changes were mapped and the perturbations were calculated using the following equation,

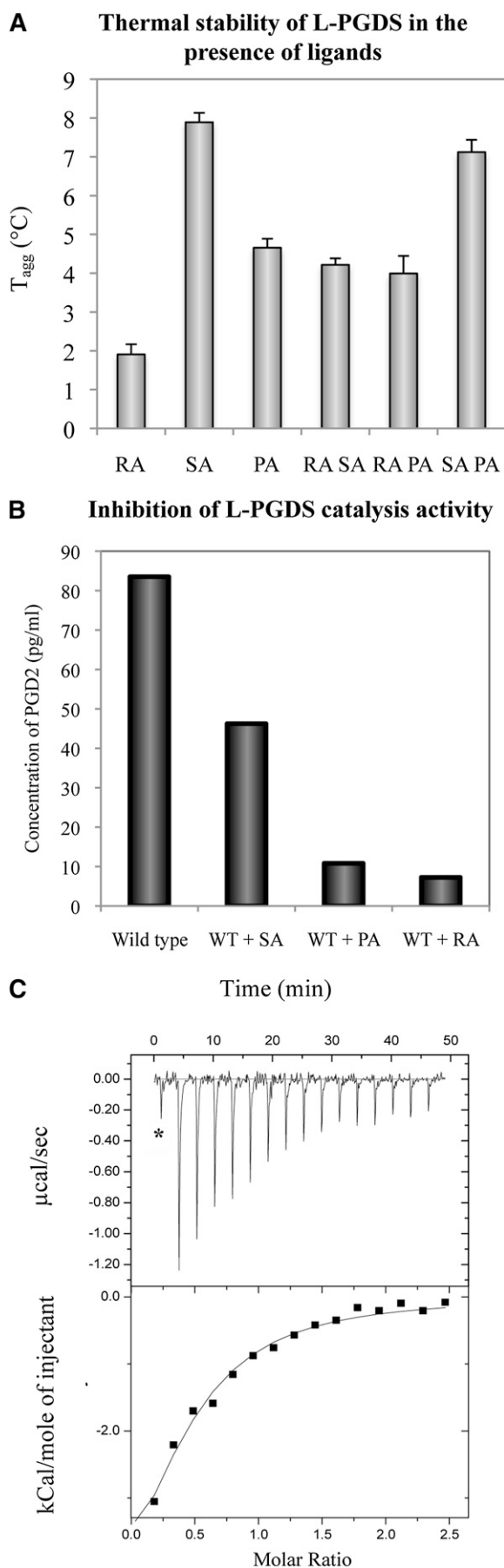
$$\Delta\omega = \{[\Delta\omega(^1\text{H})]^2 + [0.25 \times \Delta\omega(^{15}\text{N})]^2\}^{1/2}$$

Differential static light scattering assay

Protein thermal stability was measured using StarGazer-384 from Harbinger Biotechnology and Engineering. Under heat-induced denaturation, protein slowly unfolds and aggregates. The rate of protein aggregation can be measured by differential static light scattering and the intensity of scattering is plotted against temperature to obtain a temperature aggregation regression curve. This assay was carried out and analyzed according to the protocol stated in a previous study (40). Each well of a 384-well black assay plate (Nunc) had 45 µl of 0.3 mg/ml L-PGDS protein in buffer containing 20 mM HEPES pH 6.5, 150 mM NaCl, 2 mM TCEP, and 0.5 µl of ligand(s) dissolved in DMSO. The protein-ligand molar ratio was 1:10 for all ligands. To screen for stabilizing factors of L-PGDS, we used ligands such as RA, SA U44069, PA 12415, MgCl₂, CaCl₂, and in any combination of two to study if a mutually stabilizing effect could be observed.

Isothermal titration calorimetry

L-PGDS (200 µM) in Buffer C was titrated with 2.5 mM of product PGD₂ in the same buffer and concentration of DMSO to



measure the binding affinity of the protein with product using the MicroCal iTC200 system (Northampton, MA). The experiment was initiated with 0.5 μ l of titrant injection and followed by sixteen 2.4 μ l injections, each in 180 s intervals at 27°C. Stirring speed for the titration event was 1,000 rpm. Integration of heat peaks and regression analysis were performed with Origin software.

Enzymatic assay

Functional activity of L-PGDS was tested with Cayman Chemicals PGD₂-methoxylamine (MOX) ELISA kit assay (catalog number 512011). L-PGDS concentration was diluted according to the protocol for optimal enzymatic reaction. Detection of the PGD₂-MOX concentration was measured by the intensity of specific secondary fluorescence antibody provided.

RESULTS

Ligand binding to human L-PGDS

Because the biological substrate (PGH₂) and product (PGD₂) are chemically unstable, analogs were used for binding studies. To establish structural information of L-PGDS in complex with the substrate and product analog, we first wanted to characterize the binding of these ligands to L-PGDS and to confirm the catalytic activity of the recombinant protein. To identify suitable ligands/analog and conditions for optimal binding, we used a thermal shift assay (TSA) based on a differential static light scattering assay (DSLS, Stargazer, Harbinger Biotech). This assay monitors the rate of protein unfolding by measuring aggregation formed after thermally induced unfolding (41, 42). The thermal shifts for L-PGDS with SA U44069, PA 12415, and RA and any two combinations of the above were measured (Fig. 1A). SA U44069 binding to L-PGDS induced an 8°C shift of its aggregation temperature (T_{agg}), indicative of strong binding. PA 12415 and RA also increased the aggregation temperature of the protein, in the range of 2–3°C. It was previously reported that H-PGDS activity is Mg²⁺ dependent (43). However, neither Mg²⁺ nor other divalent cations like Ca²⁺ stabilized L-PGDS, so a specific binding site on the protein is not likely (data not shown). The combination of RA with SA U44069 or PA 12415 did not further increase the aggregation temperature suggesting overlapping binding sites (Fig. 1A).

The enzymatic activity of wild-type recombinant L-PGDS was measured based on detection of the product PGD₂ after incubation with substrate PGH₂ using Cayman Chemical's PGD₂-MOX ELISA kit. Recombinant L-PGDS was shown to be active (Fig. 1B) and a fixed time point assay of

Fig. 1. A: Thermal shift assay of L-PGDS in the presence of SA U44069, PA 12415, and RA shows various degrees of stability induced by these ligands. B: Recombinant protein is able to produce PGD₂ when supplemented with substrate PGH₂. Its catalytic activity is efficiently inhibited by SA U44069, PA 12415, and RA. WT, wild type. C: Human L-PGDS binds to its product PGD₂ at a K_d of 154 μ M, suggesting that the protein binds to its product even postcatalysis. Asterisk (*) indicates test injection; data was not included in integration.

L-PGDS measured a V_{\max} of 3.66 $\mu\text{mol}/\mu\text{g}/\text{s}$ and K_m of 4.15 μM . These values are in a similar range to those previously reported for recombinant mouse and human L-PGDS (27, 30). In addition, our data also show that SA U44069, PA 12415, and RA can inhibit the catalytic activity. This result agrees with Shimamoto et al. (29) who showed that RA inhibits mouse L-PGDS. In their study, they modeled two separate binding pockets for substrate and RA respectively. However the two sites were proposed to share one amino acid; it is not certain that the residue facilitates binding of both substrate and RA. Nonetheless, their Lineweaver-Burk analysis of a kinetic study claimed that the inhibition was noncompetitive (29). L-PGDS also inherently binds its product with a K_d of 154 μM (Fig. 1C). This is an interesting observation because most enzymes are designed to bind weakly to their products to facilitate release from the active site.

Cystal structure of L-PGDS in complex with substrate analog and in substrate analog free form

Human L-PGDS was cocrystallized with SA U44069 in a $P2_12_12_1$ unit cell with one monomer in the asymmetric unit. The overall structure had the classical lipocalin fold comprising an eight-stranded β -barrel with two α -helices. One of the helices, known as helix 2 on the Ω loop, lines the “entrance” of the large substrate and ligand binding cavity (Fig. 2A). The electron density map reveals two potential ligand binding sites in the cavity; one site near to the catalytic Cys 65 (site A, Fig. 2C) and one site beneath Trp 112 (site B, Fig. 2D). Crystals of L-PGDS (in the same crystal form) without SA U44069, were generated by micro-seeding with crystals grown in the presence of SA U44069 into a crystallization solution lacking the analog. The overall structure of L-PGDS without SA U44069 is very similar to the structure with it, except for distinct

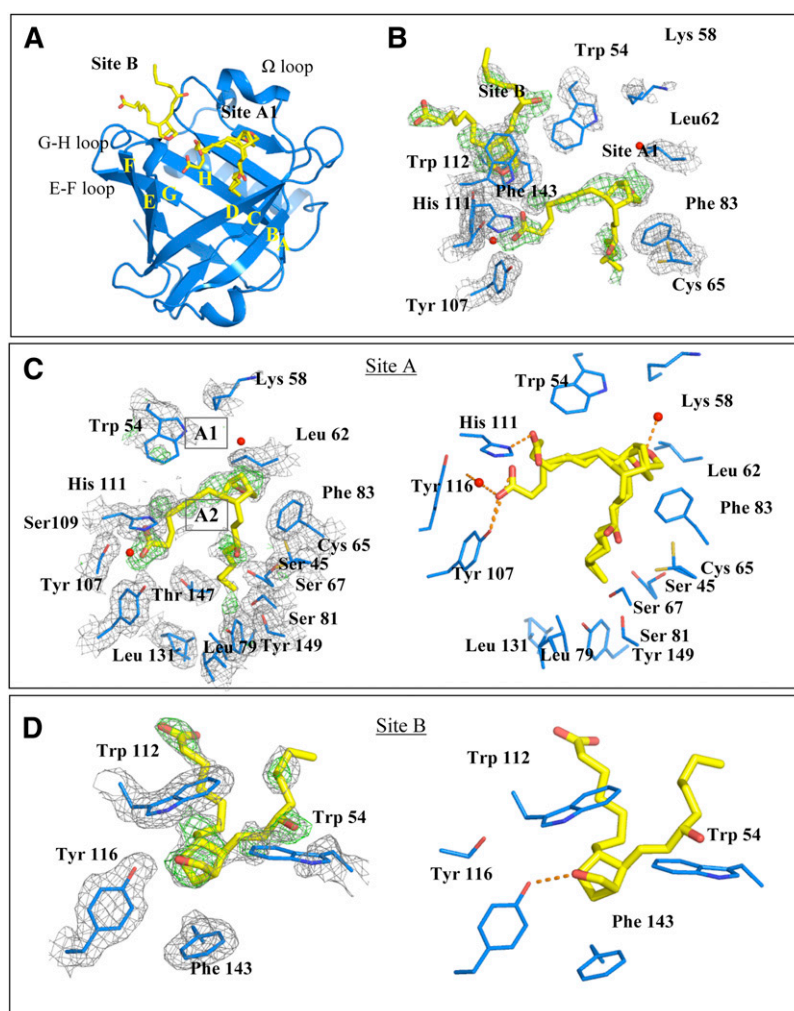


Fig. 2. A: Crystal structure of L-PGDS in complex with SA U44069 at site A and site B. β sheet A–H form the core barrel with four prominent inter-sheet loops namely the Ω loop (which consists of helix 2), C–D loop, E–F loop, G–H loop, and one long α helix H3. B: Detailed view of the binding site. Two SA U44069 were modeled into two regions of observed densities in site A and site B respectively. C: A zoomed in view of binding pocket A with two configurations that differ in the location of the carboxylate group at position A1. The head group of the bicyclopentane ring can also be positioned in A2, the Fo–Fc map at sigma level = 3.0 while the 2Fo–Fc map = 1.0. D: Binding pocket B with corresponding Fo–Fc map and 2Fo–Fc map at sigma levels 3.0 and 1.0, respectively.

conformational changes in the region of the Ω loop (Fig. 3). Density at site A is lacking in the SA U44069 free form, instead a long extended density is found stretching from site B into the cavity. We have interpreted this density as a bound polyethylene glycol (truncated PEG MME 2000, supplementary Fig. ID), a component found in the crystallization reservoir. The atomic coordinates and experimental data for the SA U44069 bound protein crystal, SAC (PDB ID: 4IMO) and ligand-free-form crystal, APO (PDB ID: 4IMN) have been deposited in the Protein Data Bank (www.wwpdb.org).

Detailed experimental data on substrate binding in L-PGDS has been elusive. Hence, we attempted to generate a molecular framework by solving the structure of human L-PGDS cocrystallized with SA U44069. The difference density map revealed a number of residual densities in the large hydrophobic cavity of L-PGDS that represent potential SA U44069 binding sites (Fig. 2C, D). Judging from the local interactions with the protein and in comparison with the analog free crystals in the same form, these density peaks are not likely to correspond to other buffer molecules but instead must originate from bound SA U44069. Even though the electron density is not sufficiently continuous for unambiguous modeling of SA U44069, it is very likely that the strongest density patches correspond to the SA head groups. Based on this hypothesis we have generated putative binding models for SA U44069 binding to L-PGDS at site A and site B.

Site A shows the major distinct prostaglandin head group density with a distance of 6 Å above the catalytic Cys 65 positioning its bicycloheptane ring in interaction with the Phe 83 side chain at site A1 (Fig. 2B). A second binding site, A2, is also possible; this model positions the SA U44069 further down in the pocket with less than 5 Å from the head group to Cys 65. The carboxylic group of SA U44069, bound in site A1, appears to extend into a polar patch formed by His 111 and Tyr 116. The hydroxyl group of the prostaglandin ω chain is in close proximity with Cys 65 and the aliphatic tail in this model is inserted into a hydrophobic pocket of Leu 79, Leu 131, and Tyr 149. The binding mode of SA U44069 in site A1 is not likely to be identical to the binding mode of the substrate during catalysis, as the bicycloheptane ring is too far from the

catalytic Cys 65. The bicycloheptane in site A2, however, is more representative of a productive substrate binding mode. With the head group in this second position, the aliphatic chain could extend deeper into the hydrophobic pocket, allowing the hydroxyl group of the prostaglandin ω chain to interact with Tyr 149 and Thr 147, equivalent to what has been proposed from substrate modeling in mouse L-PGDS (27). The carboxyl group of the substrate could still reach the polar patch containing His 111 and Tyr 116. Alternatively, it could reach Arg 92 and Arg 85 lining the groove of the cavity, as proposed for the mouse protein (27). In our L-PGDS structure, Arg 92 forms a salt bridge with Glu 35 of a neighboring molecule in the crystal lattice.

In site B, the prostaglandin head group is sandwiched between Trp 112 on the E-F loop and Phe143 of the H β -strand through stacking interactions. Modeling of the SA U44069 into this density enables its hydroxyl group to make a hydrogen bond with the hydroxyl group of Tyr 116 (Fig. 2C). Both the aliphatic and the carboxylate arms of the analog are exposed to the solvent region. In the SA U44069 free form, we could see a potential PEG MME 2000 molecule binding in site B, extending further into the large cavity. Similarly, in the structure of fatty acid-coordinated human L-PGDS, the fatty acid is clamped between Trp 112 and Phe 143 and extends into the large cavity. Previously, there was no direct evidence showing that the E-F loop containing Trp 112 and G-H loop containing Phe 143 are involved in the binding of substrate or product, although it has been proposed to facilitate the exit of product from the protein (27). Binding in site B reveals ligand stabilization of the E-F and G-H loops, and it is highly possible that the substrate and/or the product use this site as an intermediate site before exit/entrance.

The noncontinuous ligand density of the SA U44069 cocrystallized structure is most likely due to a combination of multiple binding modes and flexibility of the aliphatic moieties of the ligand. After a test refinement with models in sites A1 and B, the B factor values of the modeled ligands at both sites are in the range of 45–55 Å², while B factor values of neighboring residues are in the range of 20–25 Å². The most notable structural change upon SA U44069 binding is the rotation of helix 2 in the Ω loop, which unfolds outward, partially losing α helical rigidity. The Trp 54 side chain, as well as its main chain, showed high temperature B factor (110–120 Å²) in comparison with neighboring residues (B factor: 20–30 Å²). However, in the SA U44069 free structure it has a B factor of 30–44 Å². Hence, Trp 54 could serve as a gate for substrate upon entry/egress. Conformational change in the Ω loop is also observed when the SA U44069 bound structure is compared with the recently solved fatty acid bound C65A mutant structure (PDB ID: 3O22). Binding sites A and B of SA U44069 in our structure coincide with the binding sites for oleic acid and palmitic acid in the fatty acid bound structures.

Most residues such as Trp 54, Phe 83, Tyr 107, His 111, Trp 112, Tyr 116, and Phe 143 interacting with SA U44069 in our structure are also preserved in close homologs

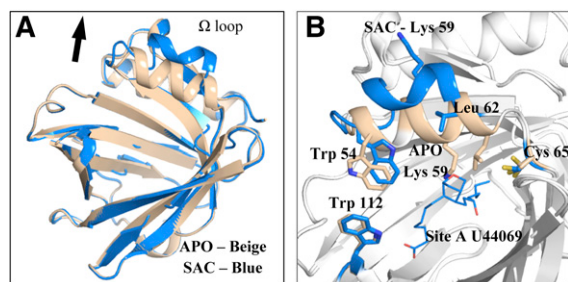


Fig. 3. A: Overlay of our substrate analog cocrystal (SAC) crystal structure in blue with apoenzyme (APO) structure in beige, root means square deviation values: 0.358. B: Ω loop of SAC structure, which is tilted outward in the presence of SA U44069; Trp 54 and Lys 59 on the Ω loop are observed to undergo major configuration changes.

(Fig. 4A). Even though Trp 54 is highly conserved across several species, mutation of Trp 54 in mouse L-PGDS only moderately decreases the catalytic activity of L-PGDS (27). Although Trp 112 is conserved in mammals and fish, it has been replaced by tyrosine in most other species, therefore retaining the basic aromatic structure required for base stacking interaction. Phe 83 is conserved in mammals, but substituted by tyrosine in *Xenopus laevis*. We agree with the prediction by Zhou et al. (30) that the aromatic side chain

is crucial in holding the bicyclopentane head group of the substrate as shown in our structure.

Structure of human L-PGDS catalytic site

The resolution of the structures allows for a detailed analysis of the environment of the catalytic cysteine. Previous crystal and NMR structures of mouse (27, 29) and human (30) L-PGDS were obtained from protein with the catalytic cysteine mutated to alanine. The structure of the

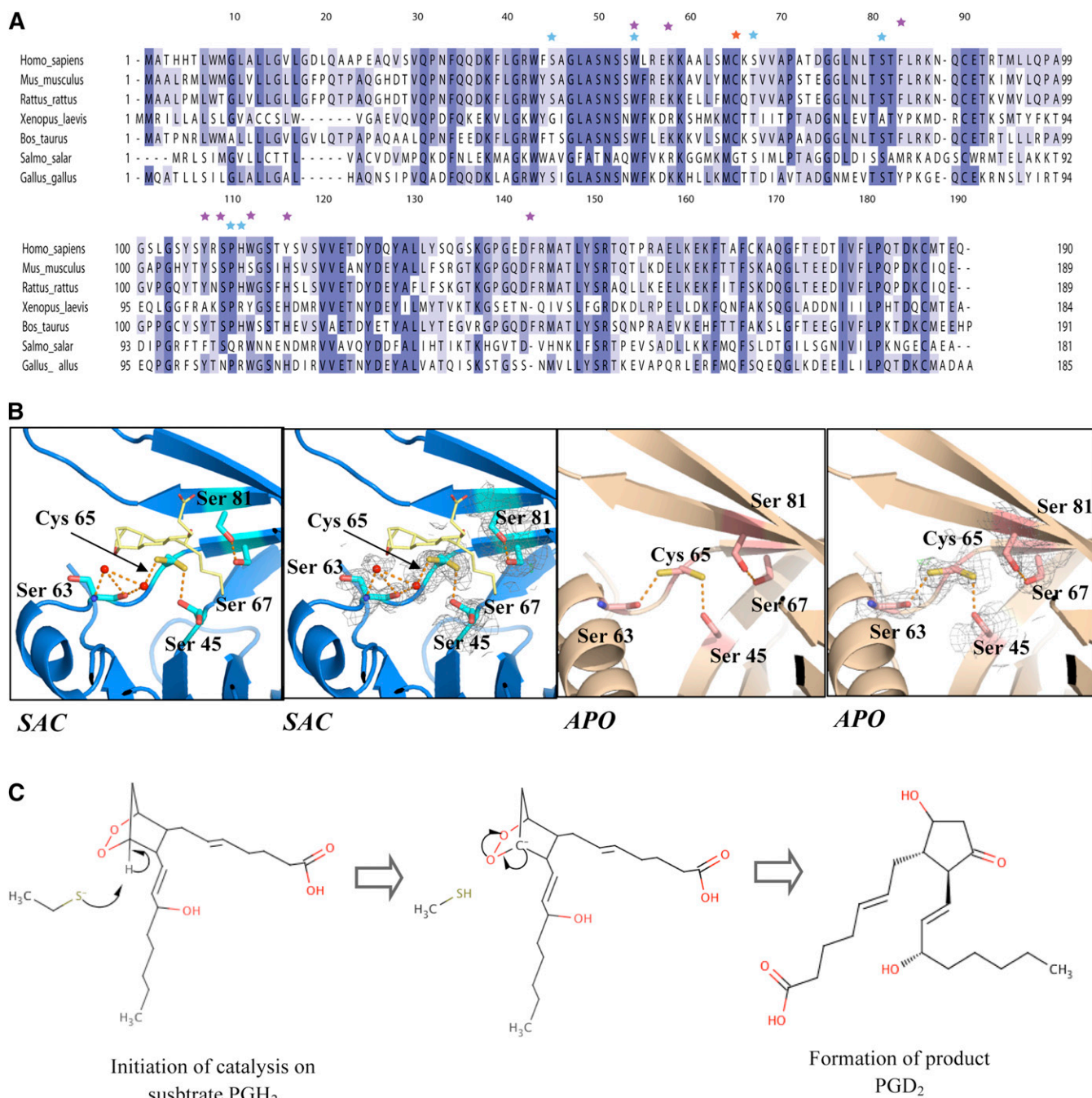


Fig. 4. A: Multiple sequence alignment of human L-PGDS with homologs. Residues proposed to interact with substrates from crystal structure results are shown as purple stars; mutations made by previous studies are highlighted as blue stars; and catalytic cysteine is indicated as a red star. B: Catalytic site comparison of Cys 65 environment in substrate analog cocrystal (SAC) and apoenzyme (APO) structure with respect to its neighboring serines. C: Schematic of L-PGDS catalytic reaction. Production of PGD₂ requires correct stereochemical positioning of SA U44069 in the catalytic site.

wild-type human L-PGDS without substrate reveals that Cys 65 has two potential side chain conformations. One of them has higher occupancy and is hydrogen bonded to the main chain carbonyl group of residue Ser 63. The thiol sulfur is not well exposed to the active site pocket but instead extends into a hydrophobic patch formed by Leu 62 and is exposed to the solvent outside the protein.

In the structure of SA U44069 bound L-PGDS, Cys 65 adopts two conformations and the dominant conformation is exposed into the pocket of L-PGDS, making a strong hydrogen bond to Ser 45 (2.8 Å) (Fig. 4B). Ser 45 also has two conformations. Interestingly, a distinct conformational switch is induced in the region of Lys 59 and Leu 62 upon substrate binding, resulting in a complete inversion of Lys 59 (Fig. 3B). This inversion is most likely due to the conformational changes observed in the Ω loop upon substrate binding, as discussed above. Mutation of K59A has been shown to increase enzymatic effect (23) by mimicking the effect of Lys 59 twisting outward to accommodate substrate binding. Shifting of the Leu 62 side chain results in the Cys 65 thiol group becoming more exposed to the substrate binding cavity with its side chain forming strong hydrogen bonding to Ser 45, thus activating Cys 65 for the subsequent reaction steps. Mutagenesis studies revealed the involvement of Ser 67 and Ser 81 in catalysis (23), but we did not observe any direct hydrogen bonding to the Cys thiol group in the structure presented here. However, we do not exclude the possibility that such hydrogen bonds could be formed upon productive substrate binding. In our structure, we identified a highly conserved residue, Leu 62, which was shown to play an important role in unraveling Cys 65 for substrate binding.

The picture that emerges from the structural data shows that substrate binding induces a distortion of helix 2 on the Ω loop and Leu 62, which exposes the Cys 65 thiol group into the substrate binding site. This allows Cys 65 to form a strong hydrogen bond to Ser 45, which helps in activating the cysteine nucleophile. The close juxtaposition of the nucleophile with substrate peroxide promotes the nucleophilic attack on C11 of the substrate (Fig. 4C). This might act in concert with the protonation of O9, to minimize net charge buildup along the reaction trajectory.

Mapping of L-PGDS interactions with substrate and PAs using NMR titration

NMR titration was used to further understand the dynamics of substrate binding to L-PGDS. L-PGDS was uniformly labeled with ¹⁵N isotope for [¹H, ¹⁵N]TROSY experiments and double-labeled with ¹⁵N and ¹³C for residue-specific assignment of backbone resonances. Reference spectra were acquired at 2 mM of SA U46619, an analog that has been used for NMR titration against the mouse protein (29). The analog saturation point was found to be at 2 mM. The spectra showed 93.3% of the expected number of cross-peaks (supplementary Table I and supplementary Fig. II). Subsequently, two resonance assignment experiments, namely HNCA and CBCA(CO)NH, together with a ¹⁵N-NOESY HSQC spectra were collected for residue-specific

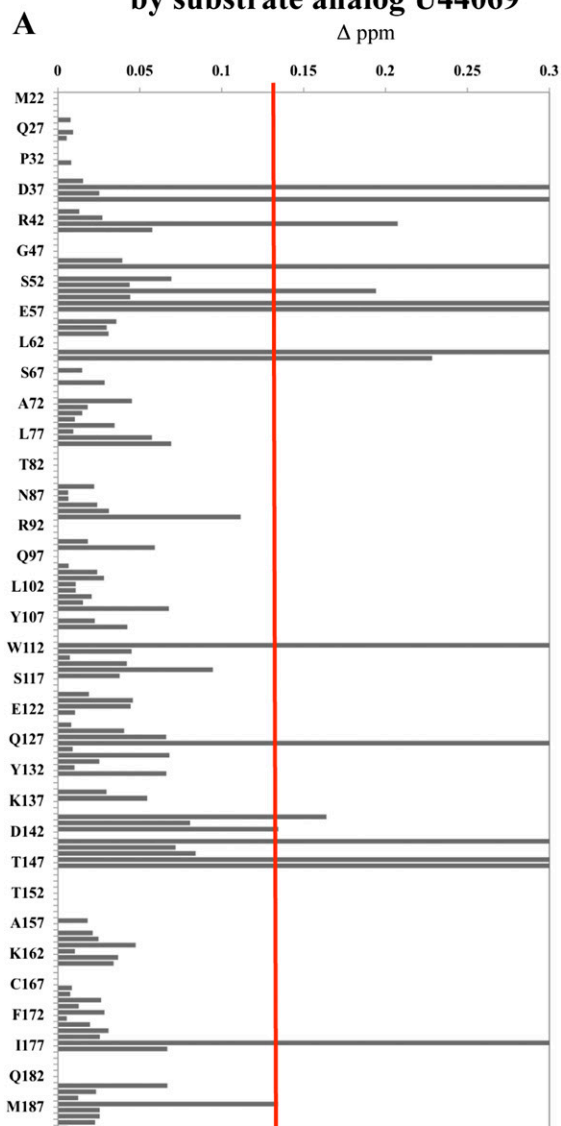
assignment. Overall, 75% of protein assignment was achieved for L-PGDS (supplementary Fig. II).

A series of titration studies were performed with SA U44069 and PA 12415 whereby similarly the saturation point was observed at ligand concentrations at or above 2 mM. Chemical shift differences above a threshold of 0.1 ppm were mapped onto the crystal structure (Fig. 5A, B). It was observed that both SA U44069 and PA 12415 induced similar chemical shift perturbations in L-PGDS. Resonance intensities for most cross-peaks were significantly increased when either SA U44069 or PA 12415 was added (supplementary Fig. III and supplementary Table II). This reflects a narrower line-width with the concomitant reduction of the conformational exchange-induced line broadening. In the presence of the SA U44069, residues Ala 49, Trp 54, Arg 56, and Glu 57 on the Ω loop as well as Trp 112 and Tyr 116 on the E-F loop showed large changes in chemical shifts (>0.1 ppm) consistent with the conformational changes seen in our crystal structures (Fig. 5A). This was accompanied by the lower chemical shift differences (>0.08 ppm) observed for Gly 140, Asp 142, and Arg 144 in the G-H loop (Fig. 5A). As expected, the chemical environment of the catalytic site residue, Cys 65, was altered and chemical shift changes of neighboring residues Met 64, Thr 147, and Leu 148 were also perturbed.

In the presence of PA 12415, changes in chemical shift at the Ω loop were only slightly less than that of SA (Fig. 5B). Two residues on the β-strand H, Met 145 and Ala 146, showed perturbations specific to PA. Chemical shifts of residues Asp 37, Trp 43, Phe 39, and Tyr 128 were consistently perturbed when both analogs were added. These four residues are situated at the narrow opening of the barrel structure, opposite to the Ω loop. It is likely that these residues alter their side-chain conformations to form a flexible plug during ligand entry/egress and binding.

Evaluation of the resonance peaks' intensity increments between bound and unbound L-PGDS are shown in supplementary Fig. III. We attribute this general increase of 2D cross-peak intensities caused by the addition of SA U44069 to the narrowing of the apparent resonance line-widths. This is due to the more restricted conformational space spawned by the protein's backbone and concomitant structural stabilization. Structural accretion due to ligand binding reduces conformational exchange-induced line broadening and this effect is also shown in the case of human protein disulphide isomerase (PDI) caused by peptide ligand binding (44). In L-PGDS approximately 24% of the peaks have intensity increases of more than 50%. These cross-peaks correspond to residues that undergo significant chemical shift perturbations. It is not known whether these residues are in direct or indirect contact with ligands because both intensity and chemical shifts were altered (45). Nonetheless, line-width reduction accompanied with increased intensity has been observed in the complex formation of the troponin C regulatory domain and troponin I (46). This suggests a stable and tight complex formation of the troponin protein. Likewise in the L-PGDS-analog complex, these NMR results are consistent

Chemical shifts differences induced by substrate analog U44069



Chemical shifts differences induced by product analog 12415

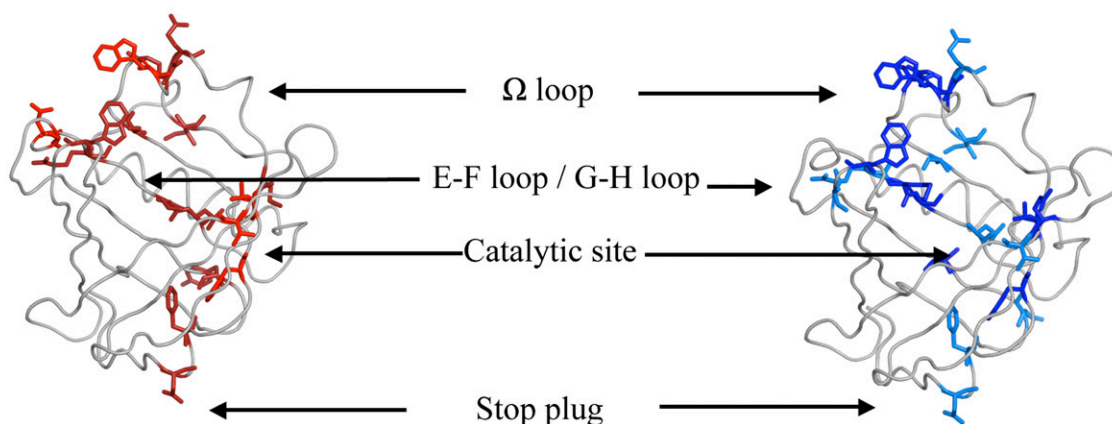
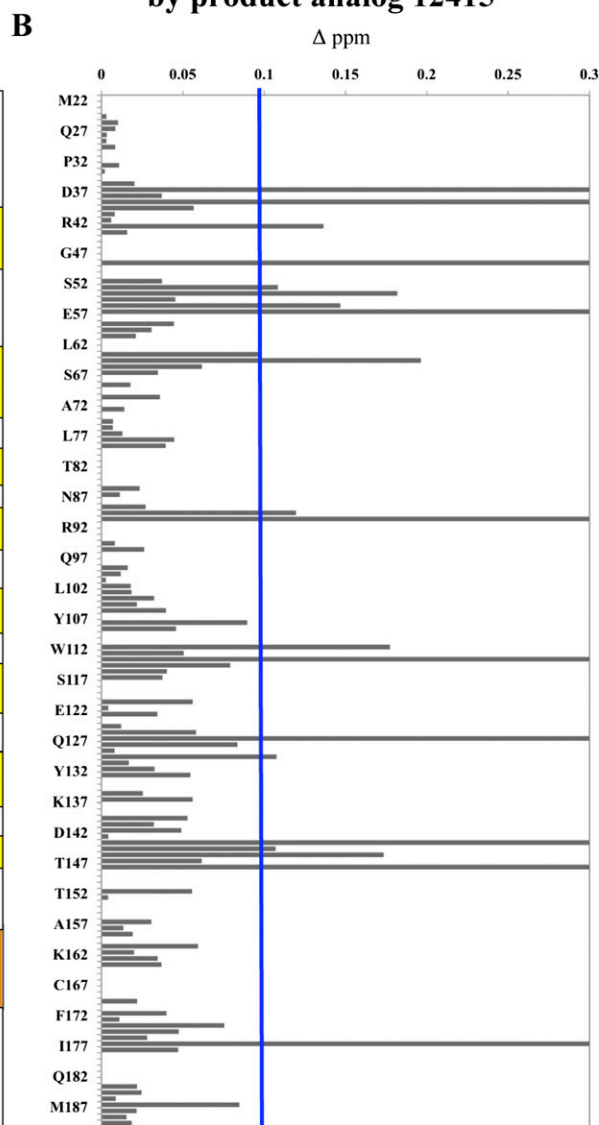


Fig. 5. A: Chemical shift perturbations induced by addition of SA U44069. In the crystal structure below, residues shifted 0.1–0.3 ppm are represented in red while residues peaks that appeared or disappeared are represented in dark red. B: Similar thresholds were set for perturbations observed in the presence of PA 12415. In the cartoon structure, chemical shift perturbations of 0.1–0.3 ppm are in dark blue while residues peaks with intensity attenuation are highlighted in blue. Secondary structure of L-PGDS was represented as follow: yellow blocks for the β sheet. orange block for the α helix; and white blocks for the inter-sheet loop.

with our TSA data and previous small angle X ray scattering data (28) showing that L-PGDS is greatly stabilized by SA U44069 and PA 12415.

By utilizing both X-ray crystallography and NMR spectroscopy data, we were able to provide a glimpse into L-PGDS substrate binding mechanics. Both methods suggested three key points of L-PGDS binding. First, substrate entry and exit was facilitated by both the E-F loop and the Ω loop. Second, residues in reaction center proximity, Met 64, Cys 65, and Ser 67, changed their rotamer configuration and backbone conformation upon ligand binding. β -sheet B (Fig. 2A), which is comprised of catalytic residues such as Met 64, Cys 65, and Ser 67, became disordered upon ligand addition. Third, N-terminal residues of the barrel cavity, usually described as the bottom of the calyx structure underwent “plugging” upon ligand binding. Phe 34 had been speculated to be involved in such an event due to poor electron density being observed (35). However our crystal data had clear density of Phe 34 in the presence and absence of SA U44069. In our NMR data, residues Asp 37, Phe 39, Trp 43, and Tyr 128 underwent chemical shift perturbations (Fig. 5A, B) while residues Thr 73, Gly 75, Gly 76, and Leu 77, which corresponded to the B-C loop, showed an increase in intensities upon ligand binding (supplementary Fig. III). We propose that the bulky side chains of Trp 43, Phe 39, and the hydrophobic side chain of Leu 77, together with the B-C loop, form a barrier at the barrel opening. **Table 3** summarizes the key residues involved in substrate recognition, proposed role in catalysis, product release, and plugging observed from both crystal data and NMR results.

L-PGDS interactions with dedecylphosphocholine to model membrane interactions

Besides protein-ligand interaction studies, protein-detergent interactions were studied in the absence and presence of ligand in order to clarify how L-PGDS interacts with a mimetic membrane and how this might regulate its activity. DPC has been commonly used as a mimetic membrane study model in solution NMR (47). Undeuterated DPC of 3 mM concentration was added to isotopically labeled protein premixed with 2 mM of SA U44069 or PA 12415. DPC micelles usually constitute a 70-80mer aggregation moiety at 25°C in 50 mM NaCl, which corresponds to a size of about 26 kDa (48). Signal disruptions of the labeled protein residues by undeuterated DPC can be detected in two-dimensional ^1H - ^{15}N -HSQC experiments. Perturbations resulting from DPC are distinct in the presence of SA U44069 and PA 12415, or in the absence of any ligands.

Figure 6A, B shows the magnitude of chemical shift differences when DPC was added in solution under two conditions, SA U44069 with DPC (SA-DPC) or PA 12415 with DPC (PA-DPC). ^{15}N -HSQC spectra were also acquired for all three samples with titrations of 3 mM and 6 mM DPC. Samples were measured for a total of 6 h with spectra being recorded hourly to observe protein state over time. There were no significant changes in chemical shifts of the protein within the period of 6 h or at higher concentrations of DPC. No degradation of protein was observed

TABLE 3. Summary of residues involve in the substrate binding, catalysis and exit

Residues	Conservation/Substitution in Orthologs	Proposed Function
Phe 83	Highly/tyrosine	Position bicyclopentane ring of substrate
Leu 62	Highly/methionine	Expose Cys 65 sulfur group
Lys 59	Conserved	Facilitate substrate entry
Cys 65	Conserved	Nucleophile
Ser 45	Partial/threonine	Deprotonate Cys 65
Ser 81	Highly/alanine ^a	Catalysis
Ser 67	Partial/threonine	Catalysis
Trp 54	Conserved	Gating/product exit
Trp 112	Highly/tyrosine or serine	Gating/product exit
His 111	Highly/arginine	Position carboxylic tail of substrate
Tyr 107	Highly/phenylalanine	Position carboxylic tail of substrate
Leu 79	Conserved	Positioning aliphatic tail of substrate
Leu 131	Conserved	Positioning aliphatic tail of substrate
Tyr 149	Conserved	Positioning aliphatic tail of substrate
Asp 37	Highly/glutamate	Plugging of clayx
Trp 43	Conserved	Plugging of clayx
Tyr 128	Conserved	Plugging of clayx
Thr 73	Highly/alanine	Plugging of clayx
Gly 75	Conserved	Plugging of clayx
Leu 77	Highly/methionine	Plugging of clayx

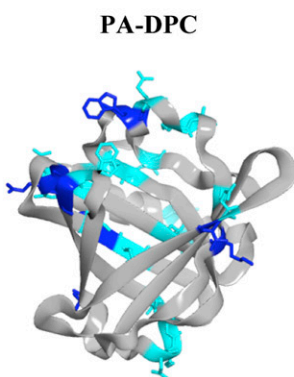
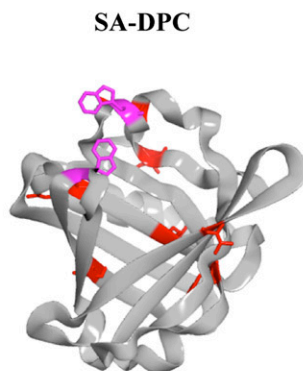
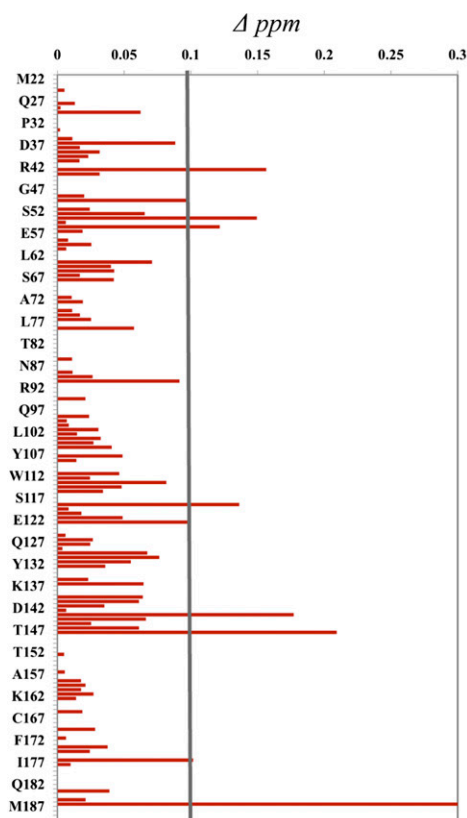
^aOnly in Xenopus.

throughout the experiments. Resonance intensity attenuation after addition of DPC was measured in both conditions (Fig. 6C, D).

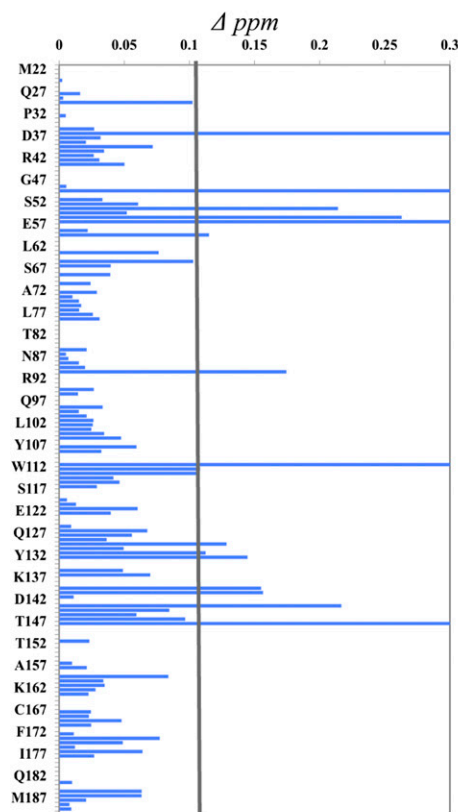
For the SA-DPC sample, residues Trp 54, Arg 56, and Ser 119 experienced alterations (>0.1 ppm) of chemical environment when compared with only SA U44069 titrated L-PGDS. Other perturbed residues included Asp 37, Ser 50, Thr 91, Ser 114, and Thr 123, which shifted between 0.08 and 0.1 ppm. Together with Ser 119, they corresponded to residues with a side chain facing the external environment of the protein. Hence, it is likely that the signal change at these residues is due to nonspecific interactions with the DPC micelles.

Overall, the effect of DPC was more pronounced in PA-DPC samples. Residues perturbed in PA-DPC (Fig. 6B) show similarity to residues perturbed in solely ligand titrated samples (Fig. 5A, B), with alterations concentrated at the E-F loop and the Ω loop. The cross-peak intensities which increased upon addition of PA 12415 (supplementary Fig. III) were attenuated to a larger extent when DPC was added (Fig. 6D). For example, residues Asp 37, Lys 59, Lys 60, Met 64, Leu 131, Asp 142, Leu 159, Glu 161, Phe 172, Ile 177, Asp 184, and Met 187 had signal reductions of more than 50% while signals of Lys 38, Ser 50, Glu 57, and Trp 112 were fully attenuated. In contrast, residues Ala 49, Glu 90, Thr 123, Ser 133, Arg 144, and Leu 148 showed more than a 50% increase in intensities (Fig. 6D). The overall decrease in intensities, which are coupled with the broadening of peaks, are due to the protein undergoing conformational exchange to release stabilization imposed by ligand binding. These findings suggest a plausible mode of product release from the protein.

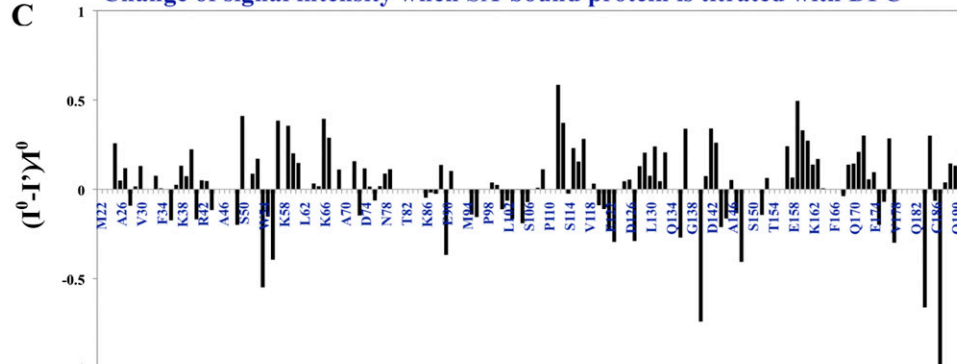
A Chemical Shift Perturbation induced by DPC in the presence of substrate analog



B Chemical Shift Perturbation induced by DPC in the presence of product analog



C Change of signal intensity when SA-bound protein is titrated with DPC



D Change of signal intensity when PA-bound protein is titrated with DPC

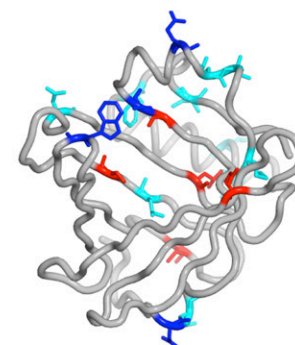
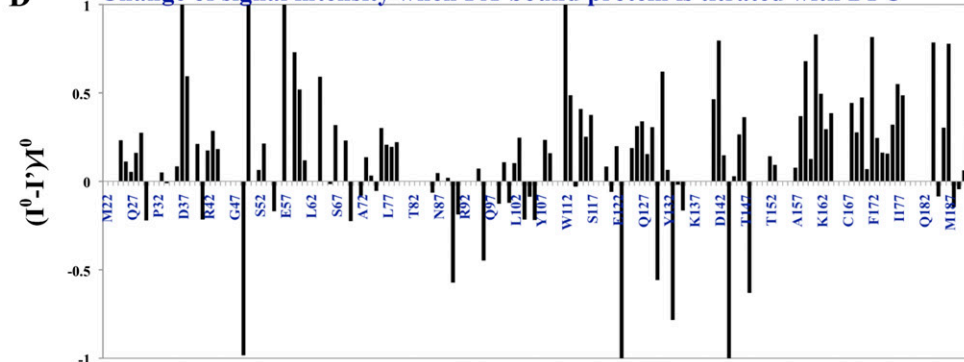


Fig. 6. Interactions of L-PGDS with DPC micelles in the presence of SA U44069 (SA-DPC) and PA 12415 (PA-DPC). Panel (A) shows the chemical shift perturbations for SA-DPC and (B) for PA-DPC. The residues involved are mapped to crystal structure in the middle. C: Intensity attenuation of L-PGDS cross-peaks in the event of DPC interaction either with SA U44069 or PA 12415 in (D). The residues involved were mapped to the crystal structure: blue, fully attenuated; dark blue, $\geq 50\%$ attenuation; red, $\geq 50\%$ intensity enhancement.

When the apoenzyme was titrated with DPC, amide cross-peaks became more resolved with a global increase in resonance intensities and chemical shift perturbations (supplementary Fig. IV). This pattern is similar to the observations made when SA U44069, SA U46619, or PA 12415 were added to the protein (Fig. 5B), which is a general stabilization of the protein. Judging from these titrations, it is likely that L-PGDS comprises intrinsic interactions with detergent micelles, and its ligand binding state could affect its interaction with membranes.

DISCUSSION

Analog bound L-PGDS cocrystals revealed putative binding sites for SA U44069 in the large cavity, where binding is accompanied by protein conformational change. Because both SA U44069-bound and analog free structures were determined in the same space group with similar crystallization buffer, this minimized the risk of artifacts, indicating that the observed structural changes of the protein are indeed due to substrate binding. The noncontinuous ligand densities are likely to be due to multiple binding modes and flexible interactions of aliphatic moieties of the SA U44069 with L-PGDS. Still, the strongest densities are most likely due to the coordinated head groups, which have allowed tentative modeling of the SA U44069 in the binding pocket. It should be noted that structures with well-defined binding of lipophilic ligands are relatively rare (49). Similar partial electron densities were observed for substrates in mouse urinary protein-pheromone complex (50), phospholipids in saposin B (53), and in 5-lipoxygenase-substrate complex (52).

Previous modeling of substrate binding in mouse L-PGDS and structural comparison with the catalytic site of H-PGDS have suggested that residues mPhe 83, mSer 81, mCys 65, mThr 67, and mSer 45 constitute the catalytic site in L-PGDS (27).³ In our crystal structure we observed the ligand density near Phe 83, Cys 65, and Ser 67. The aromatic ring of the conserved Phe 83 is positioned facing the cyclopentane ring of the modeled substrate analog in site A1 (as well as in site A2). This interaction appears crucial to position the head group of the substrate for catalysis. Mutant F83A has been shown to reduce L-PGDS enzymatic activity by more than 80% (30). Interestingly, we observed that binding of the SA U44069 induces a conformational change of Cys 65 and its neighboring residue. The side chain of Cys 65 was moved into the catalytic center and potentially assists in activating the thiol by establishing strong hydrogen bonding to Ser 67 for the attack of the cyclopentane peroxide bond of the substrate and subsequently produce PGD₂ (30). When no other suitable proton source is available in the active site, it is likely that Cys 65 also provides a proton to the substrate where nucleophilic attack and proton transfer might be concerted to minimize charge build-up.

³mPhe 83, mSer 81, mCys 65, mThr 67, and mSer 45 are the amino acids in mouse L-PGDS.

Apart from the expected binding at the catalytic site, our study reports a pseudo-binding site observed at the E-F loop of L-PGDS. This site binds polyethylene glycol in our apoenzyme structure and coincides with fatty acids binding in the structure from Zhou et al. (30). Residues of the E-F loop binding site are highly conserved across homologs of different organisms, making this a potential interaction point to facilitate the entry of ligand. The role of the E-F loop in ligand binding is highlighted in NMR titration studies in the presence of either SA U44069 or PA 12415. Chemical shift perturbations coupled with intensities augmentation upon ligand addition suggest ligand stabilization of the E-F loop. Because movement of E-F loop is coupled with conformational changes on the Ω loop and catalytic residues upon ligand titration, we speculate that the E-F loop either serves to usher the substrate into the catalytic site upon ligand binding or facilitate the release of product.

This then leads to the question of how PGD₂ is released into the right environment for its functional purpose. Because PGD₂ is produced in the lumen of ER by L-PGDS or H-PGDS, it has to be secreted into the extracellular compartment to reach its target, the receptor DP1. Even though de Waart et al. (53) suggested that PGD₂ transport through the plasma membrane might also be mediated by multidrug resistance associated protein that transport PGE₂, hitherto there had been no studies describing the transfer of PGD₂ from the lumen of endoplasmic reticulum (ER) to the cytoplasm side. Members of the prostaglandin synthases like PGI₂ synthase (PGIS) and thromboxane A₂ synthase (TXAS) are single-pass membrane proteins that bind their substrate on the luminal side of the ER membrane and release their product on the cytoplasmic side (54). However, as prostaglandin D synthase is a soluble protein, this model clearly does not apply.

Based on our data we hypothesized that L-PGDS could potentially regulate its product release through its ability to interact with membranes. Significant chemical shift perturbations are observed between the E-F loop and the Ω loop in the presence of DPC when L-PGDS is bound with PA 12415 but not with SA U44069. This is especially interesting as both the PA 12415 and SAU44069 are structurally similar (Table 1). Nonetheless, we cannot rule out the possibility that free detergent would interact with the hydrophobic core of the L-PGDS barrel especially in the apoenzyme state. Nevertheless, chemical shift perturbations and resonance intensity alterations in titration experiments with DPC revealed that the apoenzyme and the SA U44069- or PA 12415-bound protein clearly show distinct interaction modes with the micelles. In light of our data we propose that, when bound to its product, L-PGDS would be able to transiently bind the membrane of ER allowing the release of PGD₂ into the lipid bilayer. Similarly to PGIS and TXAS, L-PGDS would then be able to mediate both the synthesis and the transport of PGD₂. The present work of combining X-ray crystallography and NMR allows us to better understand the mechanisms and dynamics of ligand binding to L-PGDS. Both methods support a model whereby ligand binding is accompanied by protein

conformational changes mainly on the Ω loop and the E-F loop. Prostaglandin transport through the ER membrane into the cytoplasm was poorly understood, but our studies with membrane mimetics have provided evidence that differential protein-membrane interaction depends on ligand binding.

The authors would like to thank both the Protein Science Facility (PSF) in Karolinska Institutet and the Protein Production Platform (PPP) in Nanyang Technological University for doing initial cloning and screening of human L-PGDS constructs. The authors also thank Mr. Leo E. Wong who assisted in reviewing NMR resonance assignments.

REFERENCES

- Wymann, M. P., and R. Schneider. 2008. Lipid signalling in disease. *Nat. Rev. Mol. Cell Biol.* **9**: 162–176.
- Simmons, D. L., R. M. Botting, and T. Hla. 2004. Cyclooxygenase isozymes: the biology of prostaglandin synthesis and inhibition. *Pharmacol. Rev.* **56**: 387–437.
- Schneider, C., and A. Pozzi. 2011. Cyclooxygenases and lipoxygenases in cancer. *Cancer Metastasis Rev.* **30**: 277–294.
- Wang, D., and R. N. Dubois. 2010. Eicosanoids and cancer. *Nat. Rev. Cancer.* **10**: 181–193.
- Smith, W. L. 1989. The eicosanoids and their biochemical mechanisms of action. *Biochem. J.* **259**: 315–324.
- Urade, Y., and N. Eguchi. 2002. Lipocalin-type and hematopoietic prostaglandin D synthases as a novel example of functional convergence. *Prostaglandins Other Lipid Mediat.* **68–69**: 375–382.
- Hayaishi, O., and Y. Urade. 2002. Prostaglandin D₂ in sleep-wake regulation: recent progress and perspectives. *Neuroscientist.* **8**: 12–15.
- Eguchi, N., T. Minami, N. Shirafuji, Y. Kanaoka, T. Tanaka, A. Nagata, N. Yoshida, Y. Urade, S. Ito, and O. Hayaishi. 1999. Lack of tactile pain (allodynia) in lipocalin-type prostaglandin D synthase-deficient mice. *Proc. Natl. Acad. Sci. USA.* **96**: 726–730.
- Ragolia, L., T. Palaia, C. E. Hall, J. K. Maesaka, N. Eguchi, and Y. Urade. 2005. Accelerated glucose intolerance, nephropathy, and atherosclerosis in prostaglandin D₂ synthase knock-out mice. *J. Biol. Chem.* **280**: 29946–29955.
- Joo, M., M. Kwon, R. T. Sadikot, P. J. Kingsley, L. J. Marnett, T. S. Blackwell, R. S. Peebles, Jr., Y. Urade, and J. W. Christman. 2007. Induction and function of lipocalin prostaglandin D synthase in host immunity. *J. Immunol.* **179**: 2565–2575.
- Virtue, S., M. Masoodi, V. Velagapudi, C. Y. Tan, M. Dale, T. Suorti, M. Slawik, M. Blount, K. Burling, M. Campbell, et al. 2012. Lipocalin prostaglandin D synthase and PPARgamma2 coordinate to regulate carbohydrate and lipid metabolism in vivo. *PLoS ONE.* **7**: e39512.
- Saso, L., M. G. Leone, C. Sorrentino, S. Giacomelli, B. Silvestrini, J. Grima, J. C. Li, E. Samy, D. Mruk, and C. Y. Cheng. 1998. Quantification of prostaglandin D synthetase in cerebrospinal fluid: a potential marker for brain tumor. *Biochem. Mol. Biol. Int.* **46**: 643–656.
- Kim, J., P. Yang, M. Suraokar, A. L. Sabichi, N. D. Llansa, G. Mendoza, V. Subbarayan, C. J. Logothetis, R. A. Newman, S. M. Lippman, et al. 2005. Suppression of prostate tumor cell growth by stromal cell prostaglandin D synthase-derived products. *Cancer Res.* **65**: 6189–6198.
- Hirawa, N., Y. Uehara, M. Yamakado, Y. Toya, T. Gomi, T. Ikeda, Y. Eguchi, M. Takagi, H. Oda, K. Seiki, et al. 2002. Lipocalin-type prostaglandin d synthase in essential hypertension. *Hypertension.* **39**: 449–454.
- Nakayama, H., H. Echizen, T. Gomi, Y. Shibuya, Y. Nakamura, K. Nakano, H. Arashi, T. Itai, S. Ohnishi, M. Tanaka, et al. 2009. Urinary lipocalin-type prostaglandin D synthase: a potential marker for early gentamicin-induced renal damage? *Ther. Drug Monit.* **31**: 126–130.
- Evans, J. F., S. Islam, Y. Urade, N. Eguchi, and L. Ragolia. 2013. The lipocalin-type prostaglandin D₂ synthase knockout mouse model of insulin resistance and obesity demonstrates early hypothalamic-pituitary-adrenal axis hyperactivity. *J. Endocrinol.* **216**: 169–180.
- Heshmat, S. M., J. B. Mullen, K. A. Jarvi, A. Soosaipillai, E. P. Diamandis, R. J. Hamilton, and K. C. Lo. 2008. Seminal plasma lipocalin-type prostaglandin D synthase: a potential new marker for the diagnosis of obstructive azoospermia. *J. Urol.* **179**: 1077–1080.
- Lescuyer, P., A. Gandini, P. R. Burkhard, D. F. Hochstrasser, and J. C. Sanchez. 2005. Prostaglandin D₂ synthase and its post-translational modifications in neurological disorders. *Electrophoresis.* **26**: 4563–4570.
- Garza, L. A., Y. Liu, Z. Yang, B. Alagesan, J. A. Lawson, S. M. Norberg, D. E. Loy, T. Zhao, H. B. Blatt, D. C. Stanton, et al. 2012. Prostaglandin D₂ inhibits hair growth and is elevated in bald scalp of men with androgenetic alopecia. *Sci. Transl. Med.* **4**: 126–134.
- Fujimori, K., T. Inui, N. Uodome, K. Kadoyama, K. Aritake, and Y. Urade. 2006. Zebrafish and chicken lipocalin-type prostaglandin D synthase homologues: conservation of mammalian gene structure and binding ability for lipophilic molecules, and difference in expression profile and enzyme activity. *Gene.* **375**: 14–25.
- Kume, S., Y. H. Lee, Y. Miyamoto, H. Fukada, Y. Goto, and T. Inui. 2012. Systematic interaction analysis of human lipocalin-type prostaglandin D synthase with small lipophilic ligands. *Biochem. J.* **446**: 279–289.
- Beuckmann, C. T., M. Aoyagi, I. Okazaki, T. Hiroike, H. Toh, O. Hayaishi, and Y. Urade. 1999. Binding of biliverdin, bilirubin, and thyroid hormones to lipocalin-type prostaglandin D synthase. *Biochemistry.* **38**: 8006–8013.
- Tanaka, T., Y. Urade, H. Kimura, N. Eguchi, A. Nishikawa, and O. Hayaishi. 1997. Lipocalin-type prostaglandin D synthase (beta-trace) is a newly recognized type of retinoid transporter. *J. Biol. Chem.* **272**: 15789–15795.
- Flower, D. R. 1996. The lipocalin protein family: structure and function. *Biochem. J.* **318**: 1–14.
- Fukuhara, A., H. Nakajima, Y. Miyamoto, K. Inoue, S. Kume, Y. H. Lee, M. Noda, S. Uchiyama, S. Shimamoto, S. Nishimura, et al. 2012. Drug delivery system for poorly water-soluble compounds using lipocalin-type prostaglandin D synthase. *J. Control. Release.* **159**: 143–150.
- Urade, Y., and O. Hayaishi. 2000. Biochemical, structural, genetic, physiological, and pathophysiological features of lipocalin-type prostaglandin D synthase. *Biochim. Biophys. Acta.* **1482**: 259–271.
- Kumasaka, T., K. Aritake, H. Ago, D. Irikura, T. Tsurumura, M. Yamamoto, M. Miyano, Y. Urade, and O. Hayaishi. 2009. Structural basis of the catalytic mechanism operating in open-closed conformers of lipocalin type prostaglandin D synthase. *J. Biol. Chem.* **284**: 22344–22352.
- Miyamoto, Y., S. Nishimura, K. Inoue, S. Shimamoto, T. Yoshida, A. Fukuhara, M. Yamada, Y. Urade, N. Yagi, T. Ohkubo, et al. 2010. Structural analysis of lipocalin-type prostaglandin D synthase complexed with biliverdin by small-angle X-ray scattering and multi-dimensional NMR. *J. Struct. Biol.* **169**: 209–218.
- Shimamoto, S., T. Yoshida, T. Inui, K. Gohda, Y. Kobayashi, K. Fujimori, T. Tsurumura, K. Aritake, Y. Urade, and T. Ohkubo. 2007. NMR solution structure of lipocalin-type prostaglandin D synthase: evidence for partial overlapping of catalytic pocket and retinoic acid-binding pocket within the central cavity. *J. Biol. Chem.* **282**: 31373–31379.
- Zhou, Y., N. Shaw, Y. Li, Y. Zhao, R. Zhang, and Z. J. Liu. 2010. Structure-function analysis of human l-prostaglandin D synthase bound with fatty acid molecules. *FASEB J.* **24**: 4668–4677.
- Otwinowski, Z., and W. Minor. 1997. Processing of X-ray diffraction data collected in oscillation mode. In *Macromolecular Crystallography, Part A*. C. W. Carter, Jr. and R. M. Sweet, editors. Academic Press, New York. 307–326.
- Leslie, A. G. W., and H. R. Powell. 2007. Processing diffraction data with Mosflm. In *Evolving Methods for Macromolecular Crystallography*. R. J. Read and J. L. Sussman, editors. Springer, Dordrecht, The Netherlands. 41–51.
- McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, and R. J. Read. 2007. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**: 658–674.
- Langer, G., S. X. Cohen, V. S. Lamzin, and A. Perrakis. 2008. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**: 1171–1179.
- Emsley, P., and K. Cowtan. 2004. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**: 2126–2132.

36. Murshudov, G. N., A. A. Vagin, and E. J. Dodson. 1997. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**: 240–255.
37. Collaborative Computational Project Number 4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 760–763.
38. Takeuchi, K., E. Ng, T. J. Malia, and G. Wagner. 2007. 1–13C amino acid selective labeling in a 2H15N background for NMR studies of large proteins. *J. Biomol. NMR* **38**: 89–98.
39. Keller, R. L. J. 2004. The Computer Aided Resonance Assignment Tutorial ISBN 3-85600-112-3, first edition. Accessed September 30, 2011, at <http://cara.nmr-software.org/downloads/3-85600-112-3.pdf>.
40. Senisterra, G. A., E. Markin, K. Yamazaki, R. Hui, M. Vedadi, and D. E. Awrey. 2006. Screening for ligands using a generic and high-throughput light-scattering-based assay. *J. Biomol. Screen.* **11**: 940–948.
41. Vedadi, M., F. H. Niesen, A. Allali-Hassani, O. Y. Fedorov, P. J. Finerty, Jr., G. A. Wasney, R. Yeung, C. Arrowsmith, L. J. Ball, H. Berglund, et al. 2006. Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. USA* **103**: 15835–15840.
42. Ericsson, U. B., B. M. Hallberg, G. T. Detitta, N. Dekker, and P. Nordlund. 2006. Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* **357**: 289–298.
43. Inoue, T., D. Irikura, N. Okazaki, S. Kinugasa, H. Matsumura, N. Uodome, M. Yamamoto, T. Kumasaka, M. Miyano, Y. Kai, et al. 2003. Mechanism of metal activation of human hematopoietic prostaglandin D synthase. *Nat. Struct. Biol.* **10**: 291–296.
44. Byrne, L. J., A. Sidhu, A. K. Wallis, L. W. Ruddock, R. B. Freedman, M. J. Howard, and R. A. Williamson. 2009. Mapping of the ligand-binding site on the b' domain of human PDI: interaction with peptide ligands and the x-linker region. *Biochem. J.* **423**: 209–217.
45. Marintchev, A., D. Frueh, and G. Wagner. 2007. NMR methods for studying protein-protein interactions involved in translation initiation. *Methods Enzymol.* **430**: 283–331.
46. McKay, R. T., B. P. Tripet, R. S. Hodges, and B. D. Sykes. 1997. Interaction of the second binding region of troponin I with the regulatory domain of skeletal muscle troponin C as determined by NMR spectroscopy. *J. Biol. Chem.* **272**: 28494–28500.
47. Kallick, D. A., M. R. Tessmer, C. R. Watts, and C. Y. Li. 1995. The use of dodecylphosphocholine micelles in solution NMR. *J. Magn. Reson. B* **109**: 60–65.
48. Tham, L. K., editor. 2005. Protein-lipid interactions: from membrane domains to cellular networks. Wiley-VCH.
49. Forneris, F., and A. Mattevi. 2008. Enzymes without borders: mobilizing substrates, delivering products. *Science* **321**: 213–216.
50. Böcskei, Z., C. R. Groom, D. R. Flower, C. E. Wright, S. E. Phillips, A. Cavaggioni, J. B. Findlay, and A. C. North. 1992. Pheromone binding to two rodent urinary proteins revealed by X-ray crystallography. *Nature* **360**: 186–188.
51. Ahn, V. E., K. F. Faull, J. P. Whitelegge, A. L. Fluharty, and G. G. Prive. 2003. Crystal structure of saposin B reveals a dimeric shell for lipid binding. *Proc. Natl. Acad. Sci. USA* **100**: 38–43.
52. Gilbert, N. C., Z. Rui, D. B. Neau, M. T. Waight, S. G. Bartlett, W. E. Boeglin, A. R. Brash, and M. E. Newcomer. 2012. Conversion of human 5-lipoxygenase to a 15-lipoxygenase by a point mutation to mimic phosphorylation at Serine-663. *FASEB J.* **26**: 3222–3229.
53. de Waart, D. R., C. C. Paulusma, C. Kunne, and R. P. Oude Elferink. 2006. Multidrug resistance associated protein 2 mediates transport of prostaglandin E2. *Liver Int.* **26**: 362–368.
54. Deng, H., A. Huang, S. P. So, Y. Z. Lin, and K. H. Ruan. 2002. Substrate access channel topology in membrane-bound prostacyclin synthase. *Biochem. J.* **362**: 545–551.

Paper II

The structure and catalytic mechanism of human sphingomyelin phosphodiesterase like 3a – an acid sphingomyelinase homologue with a novel nucleotide hydrolase activity

Sing Mei Lim^{1,2}, Kit Yeung¹, Lionel Trésaugues², Teo Hsiang Ling¹ and Pär Nordlund^{1,2,3}

¹ Division of Biomedical Structural Biology, School of Biological Sciences, Nanyang Technological University, Singapore

² Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Stockholm, Sweden

³ Institute of Molecular and Cell Biology, A*STAR, Singapore city, Singapore

Keywords

acid sphingomyelinase like; binuclear metallophosphodiesterase; calcineurin-like phosphodiesterase; Niemann–Pick disease; nucleotide hydrolase

Correspondence

P. Nordlund, Division of Biomedical Structural Biology, School of Biological Sciences, Nanyang Technological University, Nanyang 637551, Singapore
Fax: +65 8180 5219
Tel: +46 70-433 6688
E-mail: Par.Nordlund@ki.se

(Received 21 October 2015, revised 29 December 2015, accepted 12 January 2016)

doi:10.1111/febs.13655

Human sphingomyelinase phosphodiesterase like 3a (SMPDL3a) is a secreted enzyme that shares a conserved catalytic domain with human acid sphingomyelinase (aSMase), the enzyme carrying mutations causative of Niemann–Pick disease. We have solved the structure of SMPDL3a revealing a calcineurin-like fold. A dimetal site, glycosylation pattern and a disulfide bond network are likely to be conserved also in human aSMase. We show that the binuclear site of SMPDL3a is occupied by two Zn^{2+} ions and that excess Zn^{2+} leads to inhibition of enzyme activity through binding to additional sites. As an extension of recent biochemical work we uncovered that SMPDL3a catalyses the hydrolysis of several modified nucleotides that include cytidine 5'-diphosphocholine, cytidine diphosphate ethanolamine and ADP-ribose, but not the aSMase substrate, sphingomyelin. We subsequently determined the structure of SMPDL3a in complex with the product 5'-cytidine monophosphate (CMP), a structure that is consistent with several distinct coordination modes of the substrate/product in the active site during the reaction cycle. Based on the structure of CMP complexes, we propose a phosphoryl transfer mechanism for SMPDL3a. Finally, a homology model of human aSMase was constructed to allow for the mapping of selected Niemann–Pick disease mutations on a three-dimensional framework to guide further characterization of their effects on aSMase function.

Database

Structural data are available in the PDB database under the accession numbers [5EBB](#) and [5EBE](#).

Introduction

Sphingolipids constitute a class of lipids that are characterized by their sphingosine backbone, including sphingosine, ceramide, sphingomyelin and glycosphingolipids. Ceramide, which is formed by the addition of a fatty acid chain to the sphingosine backbone, can be further decorated with either phosphocholine or

golipids. Ceramide, which is formed by the addition of a fatty acid chain to the sphingosine backbone, can be further decorated with either phosphocholine or

Abbreviations

ADPRM, Mn-dependent ADP-ribose/CDP-alcohol diphosphatase; alk-SMase, alkaline sphingomyelinase; AP, alkaline phosphatase; aSMase, acid sphingomyelinase; AU, asymmetric unit; CDP-choline, cytidine 5'-diphosphocholine; CMP, cytidine-5'-monophosphate; L-aSMase, lysosomal acid sphingomyelinase; NPD, Niemann–Pick disease; nSMase, neutral sphingomyelinase; PAP, purple acid phosphatase; S-aSMase, secretory acid sphingomyelinase; SMPDL3a, sphingolipid phosphodiesterase like 3a.

phosphoethanolamine to form sphingomyelin. Both ceramide and sphingomyelin form structural components of the membrane and act as signalling molecules in the sphingomyelin-ceramide signalling pathway [1,2]. Ceramide is known to be involved in the activation of the kinase suppressor Ras or c-jun kinase signalling pathway that leads to apoptosis [3,4]. Conversely, it can be phosphorylated into ceramide-1-phosphate to activate the cell survival signal through mitogen-activated protein kinase and protein kinase B pathways [5]. Therefore it is tightly regulated by various lipid transforming enzymes [6]; one such enzyme is sphingomyelinase (SMase). SMase hydrolyses the phosphodiester bond of sphingomyelin to release phosphocholine and ceramide. Since sphingomyelins are ubiquitous membrane lipids, it makes a good source for ceramide supply in the cells.

The human sphingomyelinase family is subdivided into five subgroups: lysosomal acid sphingomyelinase (L-aSMase), secretory acid sphingomyelinase (S-aSMase), alkaline sphingomyelinase (alk-SMase), Mg-dependent neutral sphingomyelinase (nSMase) and Mg-independent nSMase [7]. L-aSMase and S-aSMase are both Zn^{2+} -dependent sphingomyelinases that are encoded by a single acid sphingomyelinase (aSMase) gene. They are formed from differential precursor processing and alternate trafficking [8,9]. Mutation in L-aSMase leads to abnormal sphingomyelin accumulation that develops into Niemann–Pick type A (NPD-A) or type B (NPD-B) diseases, rare developmental disorders [10,11]. Patients with NPD-A suffer severe infantile neurological pathology and do not survive past early childhood while NPD-B patients have milder phenotypes and are still able to live until early adulthood. The secreted enzyme S-aSMase converts atherogenic lipoprotein associated sphingomyelin to ceramide that promotes lipoprotein aggregation, macrophage engulfment and subsequent foam cell formation [12], leading to impending atherosclerosis lesions.

In order to understand the structure and mechanism of L- and S-aSMase, we attempted to determine the structure of human aSMases or homologues thereof. Although we did not succeed in crystallizing aSMase or any of its domains, we were able to solve the structure of the catalytic domain of the less characterized secreted homologue of human aSMase, sphingomyelin phosphodiesterase like 3a (SMPDL3a). SMPDL3a shares 31% sequence identity (Fig. 1A) with aSMase and is most prominently expressed in liver, digestive and urinary tract tissue. It contains a signal peptide (residues 1–22) that facilitates targeting of SMPDL3a polypeptide to the

endoplasmic reticulum compartment for glycosylation and is relevant for the secretion of SMPDL3a to the extracellular space [13]. It is secreted by macrophages into human blood plasma whereas the non-secreted fraction is located either in the perinuclear region or in the Golgi body [13]. Increased expression of SMPDL3a is found in patients with bladder tumour [14]. The expression of SMPDL3a is under the control of the liver X receptor [15] whereby its secretion is stimulated by cyclic AMP activation and cholesterol loading onto macrophages. Our work on the structural characterization of SMPDL3a was well under way when it was shown to be unable to hydrolyse sphingomyelin; instead it appears to act as a nucleotide phosphodiesterase [13]. Recently, SMPDL3a was also shown to have phosphoramidase activity, probably being involved in activation of bisamidate prodrugs in the liver [16]. Another aSMase like protein known as SMPDL3b shares 30.8% sequence identity with aSMase and 41% sequence identity with SMPDL3a. It contains a C-terminal GPI anchor and has been found to affect macrophage lipid composition and membrane fluidity through regulation of Toll-like receptor signalling [17,18] (Fig. 1B).

Here we report the structure of the catalytic domain of SMPDL3a, the first crystal structure belonging to the acid sphingomyelinase family. The structure was first determined without added ligands and later we obtained a crystal in complex with a product, cytidine-5'-monophosphate (CMP) nucleotide. Based on the structure of SMPDL3a, we propose a homology model for aSMase to better understand the molecular consequences of the NPD-related mutations. We also extended the enzymatic studies of Traini *et al.* [13] to identify additional potential substrates for SMPDL3a and outline some principles for the catalytic mechanism of this family.

Results

Structure determination and overall architecture of SMPDL3a

SMPDL3a was cloned into a pFB-Sec-NH vector and tested for expression as secreted proteins in insect cells. A construct coding for residues 23–453 preceded by a signal peptide could be expressed and purified to homogeneity using affinity and size exclusion chromatography (Fig. 2A). Crystals of SMPDL3a were obtained and the structure was solved and refined to a resolution of 2.6 Å in the $P3_221$ space group with three protein molecules in the asymmetric unit (AU)

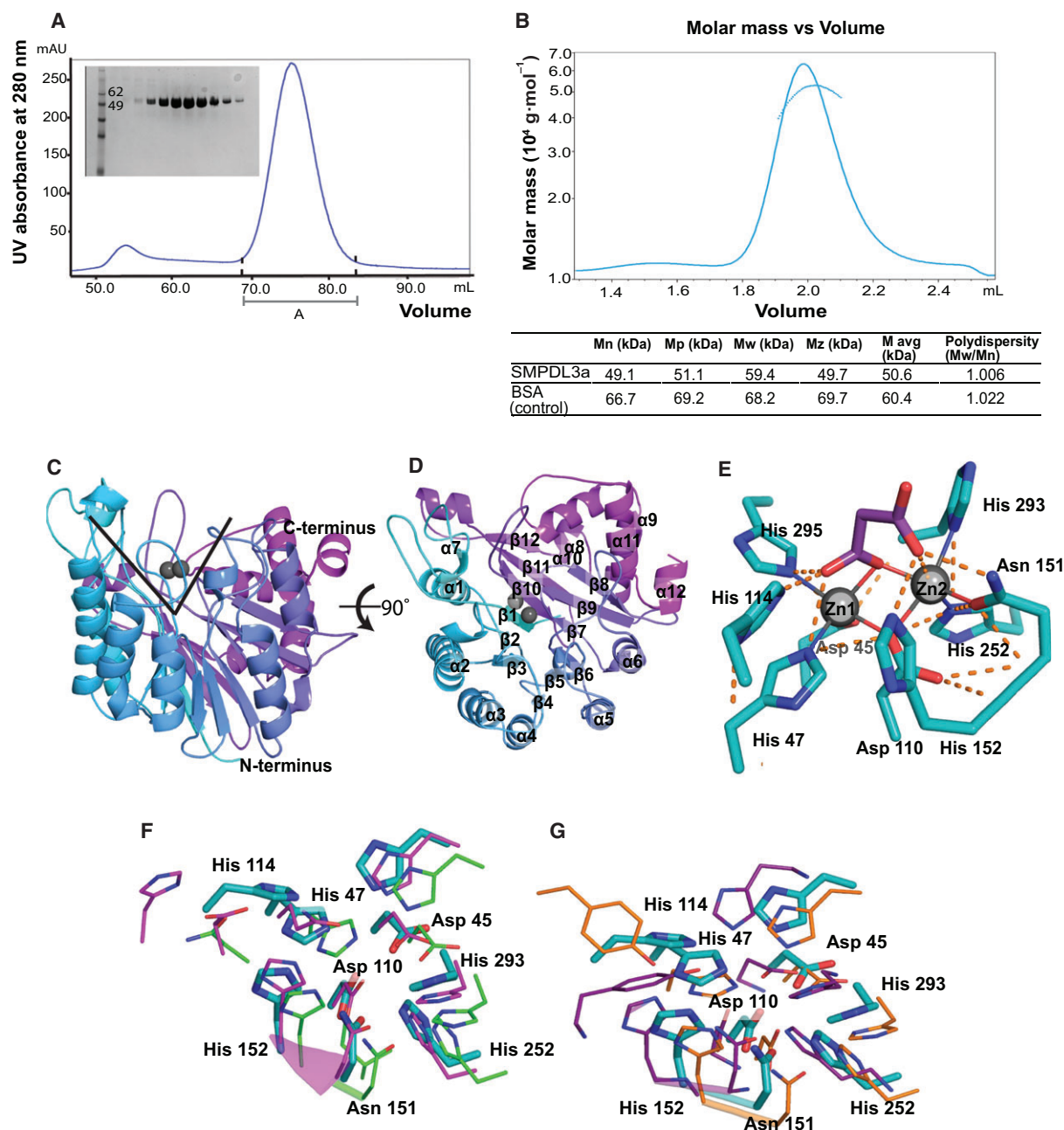


Fig. 2. (A) Elution profile of SMPDL3a from the HiLoad 16/60 Superdex 200 prep grade column and SDS/PAGE of the protein eluted corresponds to peak A. (B) Multi-angle laser light scattering elution profile of SMPDL3a using the Superdex 200 5/150 GL column and a table showing the molecular weight and dispersity of protein and control BSA. (C) The side view of the SMPDL3a structure coloured from the N terminus (cyan) to the C terminus (magenta) reveals a V-shaped binding pocket with two Zn^{2+} ions (grey). (D) Top view of SMPDL3a with all 12 α -helices and 12 β -sheets. (E) Zoomed-in view of the apoenzyme's active site showing a malonate, six conserved histidines, two aspartates and one asparagine interacting with Zn^{2+} (coordination bond shown in line). (F), (G) Comparison of SMPDL3a (teal) binuclear core with other members of the calcineurin-like family, including ADPRM (magenta; 2NXF), cyclic nucleotide phosphodiesterase, Lmo2642 (green; 2XMO), cyclic-AMP phosphodiesterase, Rv0805 (orange; 3IB7) and PAP (purple; 1WAR).

(Table 1, PDB: [5EBB](#)). The relatively small crystal packing interactions are consistent with multi-angle light scattering data supporting that the protein is

monomeric in solution (Fig. 2B). The overall structure of SMPDL3a consists of 12 α -helices wrapping around a β sandwich with six β -strands on each side (Fig. 2C,

Table 1. Data collection and refinement statistics. Values in parentheses represent the outermost shell. $R_{\text{merge}} = \sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle| / \sum_{hkl} \sum_j I_{hkl,j}$ where $\langle I_{hkl} \rangle$ represents the average of all symmetry related observations of a particular reflection. $R_{\text{cryst}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum F_{\text{obs}}$; it measures the agreement between the model and experimental structure factor data. R_{free} is computed like R_{cryst} but with 5% of reflection data removed from refinement.

	SPL3a		
	Native	Phasing Zn (SAD)	CMP-bound
PDB	5EBB		5EBE
Data collection			
Synchrotron	Diamond	BESSY	Australia Synchrotron
	I04-1	BL14.1	MX-2
Wavelength (Å)	0.9174	1.2826 (peak)	0.9537
Resolution range (Å)	48.46–2.60 (2.68–2.60)	48.44–2.89 (3.01–2.89)	47.56–3.0 (3.13–3.00)
Space group	P3 ₂ 21	P3 ₂ 21	P3 ₂ 21
Unit-cell dimensions (Å)	<i>a</i> = 147.79 <i>b</i> = 147.79 <i>c</i> = 139.90	<i>a</i> = 147.99 <i>b</i> = 147.99 <i>c</i> = 140.18	<i>a</i> = 147.65 <i>b</i> = 147.65 <i>c</i> = 142.30
Unit-cell angles (°)	90, 90, 120	90, 90, 120	90, 90, 120
Completeness (%)	99.5	99.1	99.9
Unique reflections	54349 (4446)	39760 (4438)	36234 (4376)
Mean <i>I</i> / <i>SD</i> (<i>I</i>)	9.8 (2.9)	14.6 (5.2)	6.3 (2.47)
Redundancy	5.7	19.7	3.7 (3.7)
<i>R</i> _{merge} (%)	13.2 (56.1)	14.6 (57.5)	15.6 (60.0)
Refinement			
Resolution range (Å)	48.46–2.60 (2.68–2.60)		47.56–3.0 (3.13–3.00)
<i>R</i> _{cryst} (%)	24.50		25.0
<i>R</i> _{free} (%)	26.20		29.5
Model content			
Protein atoms	9948		10074
Ligand atoms	165		262
Metal atoms	6		6
Water molecules	301		128
<i>B</i> factors			
Protein atoms	22.2		40.06
Ligand atoms	36.9		55.76
Metal atoms	17.52		34.85
Water molecules	22.0		20.76
rmsd bond (Å)	0.0216		0.0140
rmsd angles (°)	2.0987		1.7356
Ramachandran plot			
Favoured	1125 (94.14%)		1109 (90.83%)
Allowed	69 (5.77%)		110 (9.01%)
Outliers	1 (0.08%)		2 (0.16%)

D). This β – α alternating fold is commonly known as the calcineurin fold defining the calcineurin-like metal-lophosphodiesterase domain. A DALI search shows that the most similar structures in the calcineurin-like family are cyclic nucleotide phosphodiesterases (PDB: [2XMO](#) and [3IB7](#)), with rmsd values of 4.8 Å and 3.1 Å respectively. The binuclear metal centre of SMPDL3a is found tightly coordinated in the $\beta\alpha\beta\alpha\beta$ core, in the same position where other members of the calcineurin-like family have their catalytic metal centres. Several loops and short helices surround the binuclear site forming a solvent-accessible V-shaped

pocket, which constitutes the active site of SMPDL3a (Fig. 2C).

The binuclear metal site

Based on the electron density maps, we identified one di-metal centre for each SMPDL3a molecule in the AU with the two metals ~ 3.4 Å apart. The presence of Zn in the crystals was confirmed by an X-ray fluorescence scan of the native crystals whereas no signals were present for Mn, Mg and Fe. We therefore assume that all metals in our structure

correspond to Zn^{2+} ions. In the structure, the two Zn ions are coordinated by highly conserved residues where Asp45, His47 and His295 serve as terminal ligands for Zn1 and Asn151, His252 and His293 as terminal ligands for Zn2, with Asp110 bridging as a monodentate ligand (Fig. 2E). There is also an additional malonate ion present coordinating the metal ions in a bridging mode in all molecules in the AU (Fig. 2E).

The binuclear centre is a hallmark of the calcineurin-like metallophosphodiesterase where it contributes to catalysis by coordinating the scissile substrate phosphate in the reaction cycle. Superposition of SMPDL3a with other calcineurin-like metallophosphodiesterase structures, such as Mn-dependent ADP-ribose/CDP-alcohol diphosphatase (ADPRM, PDB: [2NXF](#)), purple acid phosphatase (PAP, PDB: [1WAR](#)) and cyclic nucleotide phosphodiesterase (Lmo2642, PDB: [2XMO](#); Rv0805, PDB: [3IB7](#)), reveals high similarity in the environment of the dimetal centre despite sharing merely 11–14% overall sequence identity (Fig. 2F,G). All residues coordinating the metals are conserved except for residues corresponding to His114 and His47 in SMDPL3a. His114 is replaced by either an asparagine in Lmo2642 (Fig. 2F), a tyrosine in Rv0805 and PAP (Fig. 2G) or aspartate in ADPRM. In ADPRM, the direct metal coordinating His47 is also substituted with Gln15 (Fig. 2F).

Glycosylation sites and disulfide links

Human SMPDL3a is an N-linked glycosylated protein predicted to have seven sites for glycosylation, three of which, Asn222, Asn263 and Asn356, have been confirmed by a proteomics study [19]. In our crystal struc-

ture, N-acetylglucosamines were identified on Asn69, Asn131, Asn263 and Asn356 respectively (Fig. 3A). The four glycosylation sites are found well exposed on the surface of the protein, specifically on $\beta 1$ – $\alpha 1$ loop, $\alpha 2$ helix, $\beta 6$ – $\alpha 6$ loop and $\beta 11$ – $\alpha 7$ loop. A thermal stability assay revealed a 15 °C destabilization of the de-glycosylated protein supporting a stabilizing role of these modifications (data not shown).

Three cysteine pairs are found close to each other with two being engaged in disulfide bonds (Cys62–Cys81 and Cys430–Cys443), while an additional pair (Cys420–Cys424) is reduced in the crystal structure (Fig. 3B). These cysteine pairs are highly conserved in the aSMase family and most probably play a role in maintaining the structural integrity of the proteins. The Cys594–607 pair in L-aSMase, which corresponds to Cys430–443 in SMPDL3a, has been proposed to be important for secretion of L-aSMase from the endoplasmic reticulum compartment into the lysosome [20]. The structure of SMPDL3a provides the first molecular insights into the conserved network of covalent modifications of members of the aSMase family.

Metal specificity of SMPDL3a

Different calcineurin-like metallophosphodiesterases use different metals as their cofactor, most commonly Fe^{2+} , Fe^{3+} , Mn^{2+} , Ni^{2+} , Mg^{2+} and Zn^{2+} . Some members of the superfamily like ADPRM require only Mn^{2+} for optimal function [21], while others like glycerophosphodiesterase are able to utilize various metals like Fe^{2+} , Zn^{2+} , Mn^{2+} , Cd^{2+} or Co^{2+} [22] or, like PAP, adopt a heterovalent metal centre comprising Fe^{3+} – Mn^{2+} or Fe^{3+} – Zn^{2+} [23]. Human L- and S-aSMase have been shown to use Zn^{2+} as their catalytic metal [24]. In order to determine whether

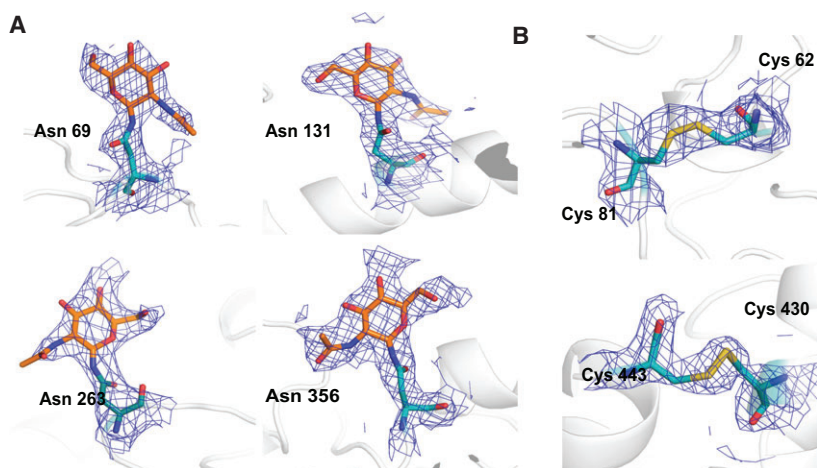


Fig. 3. (A) Four N-acetylglucosamine glycosylated asparagines and (B) two disulfide bridges observed in the crystal structure with $2F_o - F_c$ density featured in $\sigma 1.0$.

SMPDL3a has a restricted or broad metal specificity, we measured the thermal stability of the protein in the presence of different metals, which was indicated by the aggregation temperature T_{agg} . The thermal stability of SMPDL3a is significantly increased in the presence of 2 mM $ZnCl_2$ but not with other metals. SMPDL3a's

T_{agg} increased by almost 20 °C at pH 7.5 with $ZnCl_2$ (Fig. 4A). We have determined a crystal structure of Zn^{2+} enriched SMPDL3 revealing additional Zn^{2+} binding sites at Met300, Asp421 and Cys424 apart from the binuclear core (Fig. 4F). Asp421 and Cys424 binding with Zn^{2+} could stabilize the α_{10} – α_{11} loop;

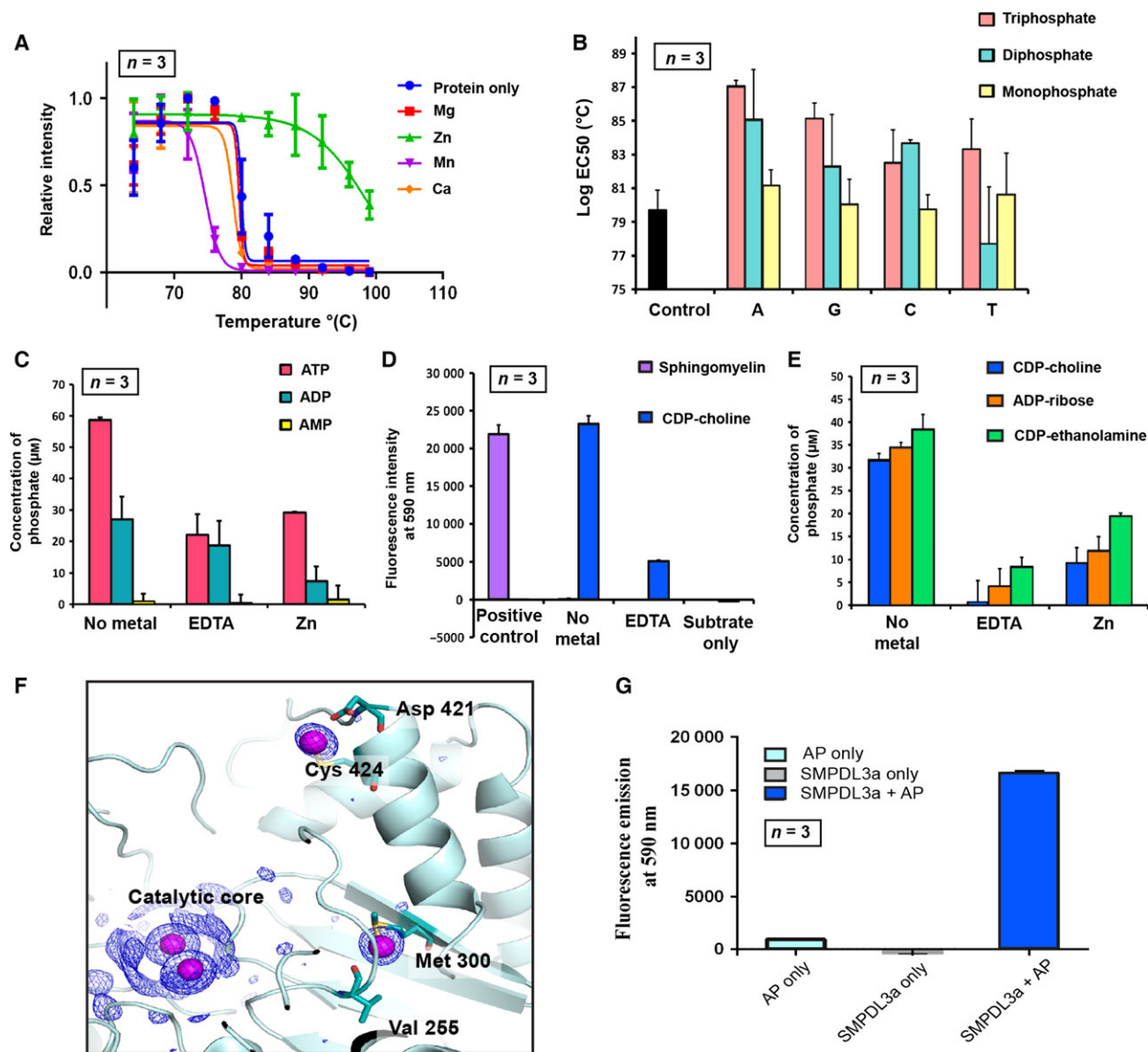


Fig. 4. (A) Thermal stabilization profiles of SMPDL3a with various metals. (B) Thermal stabilization of SMPDL3a with triphosphate, diphosphate and monophosphate nucleotides at pH 7.5. Log EC₅₀ is the temperature where 50% of the protein is aggregated. (C) Activity of SMPDL3a with substrate ATP, ADP and AMP in the presence and absence of additional Zn^{2+} at pH 7.5 measured by phosphate released using the malachite green assay. (D) Amplex Red detection of phosphocholine release when SMPDL3a was incubated with substrate CDP-choline and sphingomyelin. *Bacillus cereus* sphingomyelin was used as positive control. (E) Substrate CDP-choline, ADP-ribose and CDP-ethanolamine susceptibilities of hydrolysis by SMPDL3a were measured with the phosphate release assay. (F) Densities corresponding to Zn derivatives calculated from the anomalous map of Zn-soaked crystals revealed two additional Zn^{2+} sites near Asp421, Cys424 and Met300. (G) SMPDL3a activity with substrate CDP-choline in the presence and absence of AP. Error bars in (A) – (E) and (G) depict the SD of three independent experiments.

meanwhile, the Zn^{2+} –Met300 interaction, which is located near the ligands of the dimetal centre, i.e. His291 and His293, may help to strengthen the catalytic core from thermal denaturation and aggregation. These sites could contribute to the observed Zn^{2+} induced stabilization.

The anomalous Zn^{2+} signal observed in native crystals is likely to correspond only to the binuclear core metals. These metal bound native enzymes are catalytically active (Fig. 4C–E) and may have captured the metals during protein expression; they are presumed to be strongly bound since only small effects of EDTA treatment were observed on the activity (Fig. 4C–E). In addition, we also observed that SMPDL3a activity is repressed at a high millimolar Zn^{2+} concentration (Fig. 4C) as previously reported [13]. This inhibitory effect is less likely to be physiologically significant when the concentration of free Zn^{2+} ion concentration is typically much lower in cells [25] and extracellular space [26] but it reflects the structural similarities in the active site regions in these proteins. Thus, SMPDL3a most probably uses Zn^{2+} as its catalytic metal, which is similar to aSMase and consistent with the high conservation of residues at the Zn binding site between the two enzymes [24].

Substrate specificity of SMPDL3

SMPDL3a has been annotated as an enzyme with sphingomyelin phosphodiesterase activity (<http://www.uniprot.org/>). However, it was recently shown that recombinant human SMPDL3a could not utilize sphingomyelin as a substrate under standard *in vitro* assay conditions; instead it was demonstrated to have a nucleotide phosphodiesterase activity hydrolysing NTPs and NDPs but not NMPs. To further elucidate SMPDL3a substrate specificity, we used a thermal shift assay to investigate which nucleotides are potential ligands for SMPDL3. In a set of nucleotides, ATP is found to be the strongest stabilizer of SMPDL3a at pH 7.5 with an increase of 8 °C in T_{agg} , followed by GTP, CTP and TTP (Fig. 4B). Diphosphate nucleotides also stabilized the protein albeit to a lesser extent, while monophosphate nucleotides have no effect on protein stability. Since the thermal stability assay does not allow discrimination between substrate or product binding, we complemented these results with enzymatic assays. Our findings agreed with those of Traini *et al.*: SMPDL3a is found to hydrolyse ATP and, to a lesser extent, ADP but is inactive towards AMP [13] (Fig. 4C). Similarly, no activity against sphingomyelin could be detected (Fig. 4D). This coincides with another member of the aSMase like protein from plant

pathogen *Ralstonia solanacearum* where it was found to be a nucleotide hydrolase instead of a sphingomyelinase [27] (Fig. 1).

Apart from nucleotides, other substrates metabolized by calcineurin-like family members such as ADPRM include substituted NDPs – cytidine 5'-diphosphocholine (CDP-choline), ADP-ribose and CDP-ethanolamine [28] – and therefore we investigated the potential of these compounds to be substrates for SMPDL3a. The Amplex Red assay (supplemented with alkaline phosphatase, AP) detects the production of choline when CDP-choline is added as substrate for SMPDL3a (Fig. 4D). In order to investigate if SMPDL3a hydrolyses CDP-choline into CMP and phosphocholine or CDP and choline, AP is removed from the Amplex Red working solution mixture. In the absence of AP, phosphocholine produced from the reaction cannot be oxidized by choline oxidase to react with Amplex Red dye for fluorescence emission and detection. Our results showed that there is no increase in the fluorescence when AP is removed (Fig. 4G). This suggests that phosphocholine is the product of the reaction catalysed by SMPDL3a. A method based on the malachite green assay revealed that, in addition to CDP-choline, SMPDL3a also accepts ADP-ribose and CDP-ethanolamine as substrates (Fig. 4E, Table 2). SMPDL3a catalysed all four substrates at similar catalytic efficiency *in vitro* at pH 5 in the absence of supplementary metal with ADP-ribose and CDP-ethanolamine being slightly more efficient at a $k_{\text{cat}}/K_{\text{m}}$ of $5955 \text{ M}^{-1}\cdot\text{s}^{-1}$ and $5706 \text{ M}^{-1}\cdot\text{s}^{-1}$ (Table 2). Although these are relatively low efficiency rates, these substrate classes should be considered in order to investigate the physiological substrate for SMPDL3a.

The structure of SMPDL3a bound with cytidine-5'-monophosphate (CMP) nucleotide

We have attempted co-crystallization and soaking experiments extensively with different substrates and product analogues to understand the phosphodiesterase mechanism of SMPDL3a. Unfortunately, most structures did not reveal significant ligand binding, but after much effort we managed to obtain a complex of SMPDL3a and a tentative CMP ligand from soaking an apo-crystal with CDP-choline for 10 min at pH 5 (Fig. 5A,B) (Table 1, PDB: [5EBE](#)). A strong difference Fourier density in the active site in monomers A and C could be attributed to CMP (Fig. 5C,D). The electron density is smaller in monomer B, which we interpreted as the phosphoribose moiety of CMP (Fig. 5E). CMP density in molecule A and molecule C differs in the nucleoside orientation

Table 2. Kinetic parameters of SMPDL3a measured with various substrates at pH 5. Experiments were carried out in triplicate.

Substrate	K_m (μM)	V_{max} ($\mu\text{M}\cdot\text{s}^{-1}$)	k_{cat} (s^{-1})	k_{cat}/K_m ($\text{M}^{-1}\cdot\text{s}^{-1}$)
ATP	326.85 ± 33.87	0.07357 ± 0.003	0.7357 ± 0.033	2257.71 ± 131.86
CDP-choline	262.10 ± 63.78	0.129 ± 0.038	1.29 ± 0.38	4889.17 ± 283.37
ADP-ribose	348.15 ± 11.1	0.2072 ± 0.0009	2.07 ± 0.009	5955.5 ± 163.5
CDP-ethanolamine	390.7 ± 61.45	0.1864 ± 0.0083	1.864 ± 0.083	5706.79 ± 874.0
ADP	305.85 ± 18.6	0.0374 ± 0.02	0.374 ± 0.2	1227.57 ± 140.06

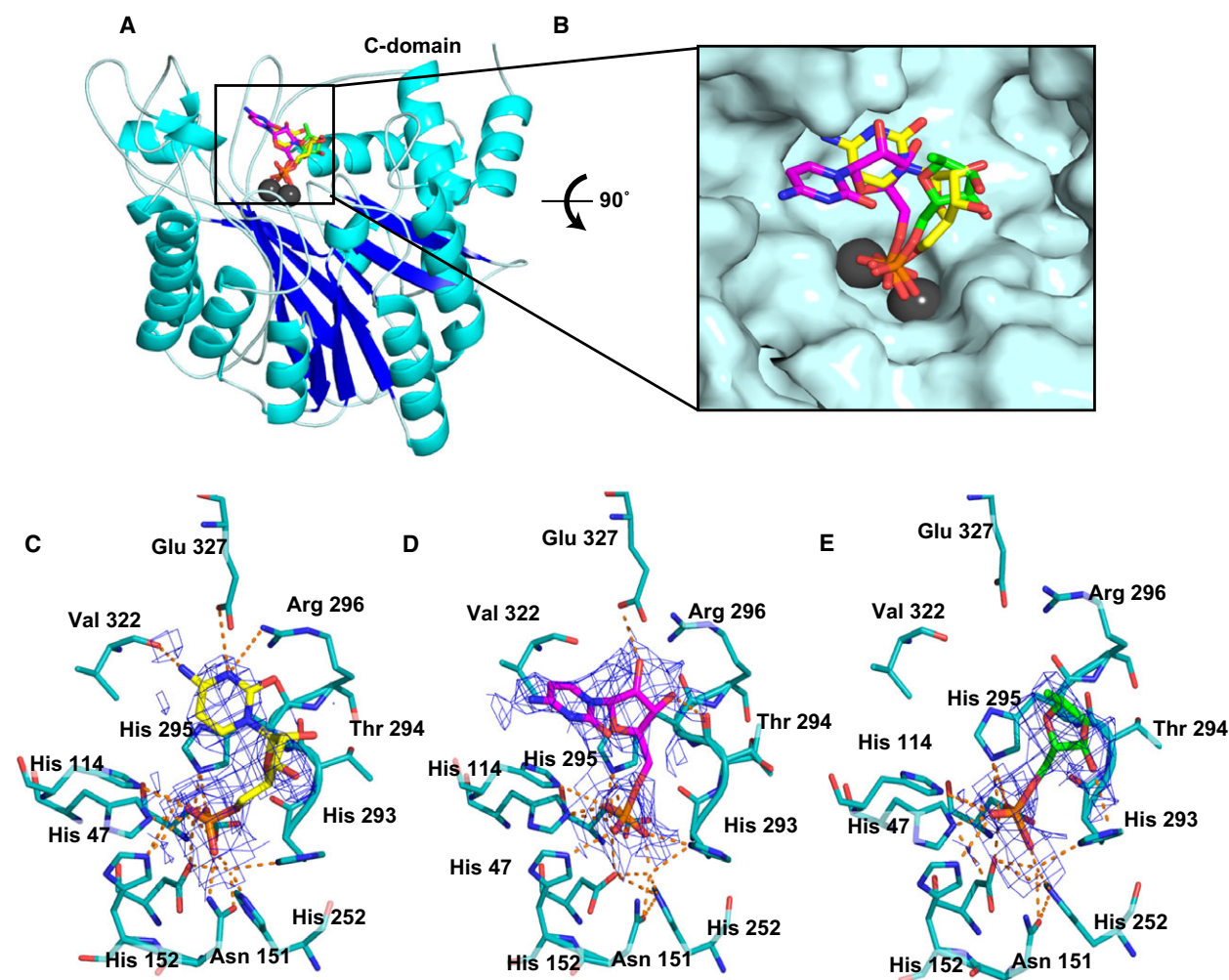


Fig. 5. (A) Crystal structure of SMPDL3a bound with different conformations of CMP from all three molecules superposed on each other. (B) View from the top of the SMPDL3a-ligand structure showing the cavity that accommodates the CMP (yellow in molecule A; magenta in molecule C) and phosphoribose (green in molecule B). (C) Key interactions of CMP with SMPDL3a residues in molecule A and (D) molecule C. (E) Only the phosphoribose moiety is observed in molecule B. The electron densities correspond to $2F_o - F_c$ at $\sigma 1.0$.

where the cytosine base is in an anti conformation in subunit A (Fig. 5C) but close to a syn conformation for C (Fig. 5D). These two conformations could possibly represent different stages of the release of the CMP product after the hydrolysis of CDP-choline. In subunit A, the cytosine base interacts strongly with the

backbone carbonyl of Val322 but weakly with the side chains of Glu327 and Arg296 via electrostatic interaction (Fig. 5C). Subsequently, upon releasing, the base of CMP in subunit C lost its interactions by shifting 3.5 Å away from its position in subunit A, leaving the ribose to make hydrogen bonds with the carbonyl of

Thr294 backbone and the side chain of Glu327 (Fig. 5D). The ribose group of phosphoribose ligand in subunit B forms only a hydrogen bond with His293 (Fig. 5E) since no pyrimidine density is observed. Nucleosides of the bound CMP have higher *B* factors (75–88) compared to the α -phosphates (50–55) and neighbouring residues (20–30) suggesting that it is more mobile, which hence explains the lack of cytidine density in subunit B.

The α -phosphates in all three subunits bind in a bidentate mode to the metal and interact via hydrogen bonding with His152, His114, Asn151, His293 and His295 (Fig. 5C–E). His152 and His114 are 4.5–5.5 Å away from metal Zn in our structure and do not contribute to the coordination bonding with the metal centre. It is likely that His152 plays a central role in substrate anchoring. The residue corresponding to His152 in ADPRM is His97 and this residue is proposed to be vital in hydrolysing CDP-choline, ADP-ribose, CDP-ethanolamine as well as CDP-glycerol [21]. His152 also corresponds to His319 in aSMase where the H319Y mutation in NPD-A patients results in residual aSMase activity [29]. Overlay of CMP-bound SMPDL3a molecule A with 5'AMP-bound metallophosphodiesterase Rv0805 (PDB: [3IB8](#)) showed that the nucleotide sits in similar positions. The two enzymes do not show direct interaction between the nucleoside and side chains of active site residues but are anchored via their phosphate moiety [30]. In addition, there are no major conformational changes or side chain shifts seen when comparing the apoenzyme structure and the ligand-bound structure of SMPDL3a.

Model of acid sphingomyelin phosphodiesterase, a target gene for Niemann–Pick disease

SMPDL3a shares 31% overall sequence identity with aSMase, including a conserved N-glycosylation site, disulfide bridges and catalytic residues in the binuclear centre (Fig. 1). Based on the SMPDL3a crystal structure we can now propose a more accurate model of the aSMase catalytic domain by using homology modelling with the SMPDL3a template to address the effect of these mutations on a structural framework.

The most challenging region for modelling is the β 10– β 11 loop (loop I) of SMPDL3a, due to the major difference in loop length between SMPDL3a and aSMase (Fig. 6A). A shorter loop in aSMase is likely to position Tyr488 towards the binuclear active site. The mutation Y488N in aSMase is found in NPD-B cases, suggesting that this residue could directly affect the activity [31]. The presence of tyrosine in this posi-

tion tightens the substrate binding site and this residue clashes with the SMPDL3a substrate when the aSMase model is superposed with the SMPDL3a–5'CMP complex. Two tryptophan residues (Trp435 and Trp437) on β 6– α 6 loop (loop II) of the aSMase model also narrow the active site opening compared to SMPDL3a where they are Ser261 and Leu403 respectively (Fig. 6A). However, it cannot be excluded that these loops will alter their conformation upon substrate binding to facilitate the entry of substrates. It is not clear from the model how the substrate sphingomyelin may bind. Moreover, unlike recently solved neutral ceramidase, there is no clear extended hydrophobic patch in the active site of aSMase, suggesting that the fatty acid chain of sphingomyelin may not be completely inserted into the active site [32] (Fig. 6B). Sphingomyelin is probably bound to the lipid binding saposin domain that acts as a chaperone for the substrate [33]. However, the role of the saposin domain in aSMase function is still relatively poorly characterized. Trp435 and Trp437 on loop II together with Trp285 and Trp391 of the aSMase model are protruding outward above the opening of the substrate binding pocket (Fig. 6A). These hydrophobic tryptophans as well as Arg387, Phe390 and Pro587 were predicted as potential membrane interaction sites by the PPM server (Fig. 6C) [34] suggesting the possibility for a direct physical interaction of the catalytic domain of aSMase with the membrane or its saposin domain. The orientation of the enzyme would then allow the substrate cavity to be aligned towards the membrane, or alternatively saposin, that displays the substrate. SMPDL3a on the other hand lacks these potential membrane interacting residues like aSMase.

Overview of NPD mutations on the aSMase model

The aSMase model allows mapping of the currently known NPD-A and NPD-B point mutation sites in the catalytic domain of aSMase (Fig. 6D). A significant fraction of these mutations probably do not directly modulate activity but rather affect protein expression and stability. Twenty-one aSMase's mutants were shown to be expressed in transfected COS-1 and COS-7 cells [35–37] out of which 19 were found with mutations in the aSMase catalytic domain (Fig. 6E). These point mutations may disrupt the enzyme's activity in three ways. First, the NPD mutations may interfere with substrate entry and membrane/saposin interaction. Five mutations on the β 1– α 1 loop reduced the enzymatic activity compared to wild-type even though they can be expressed in COS-1 cells. This β 1– α 1 loop

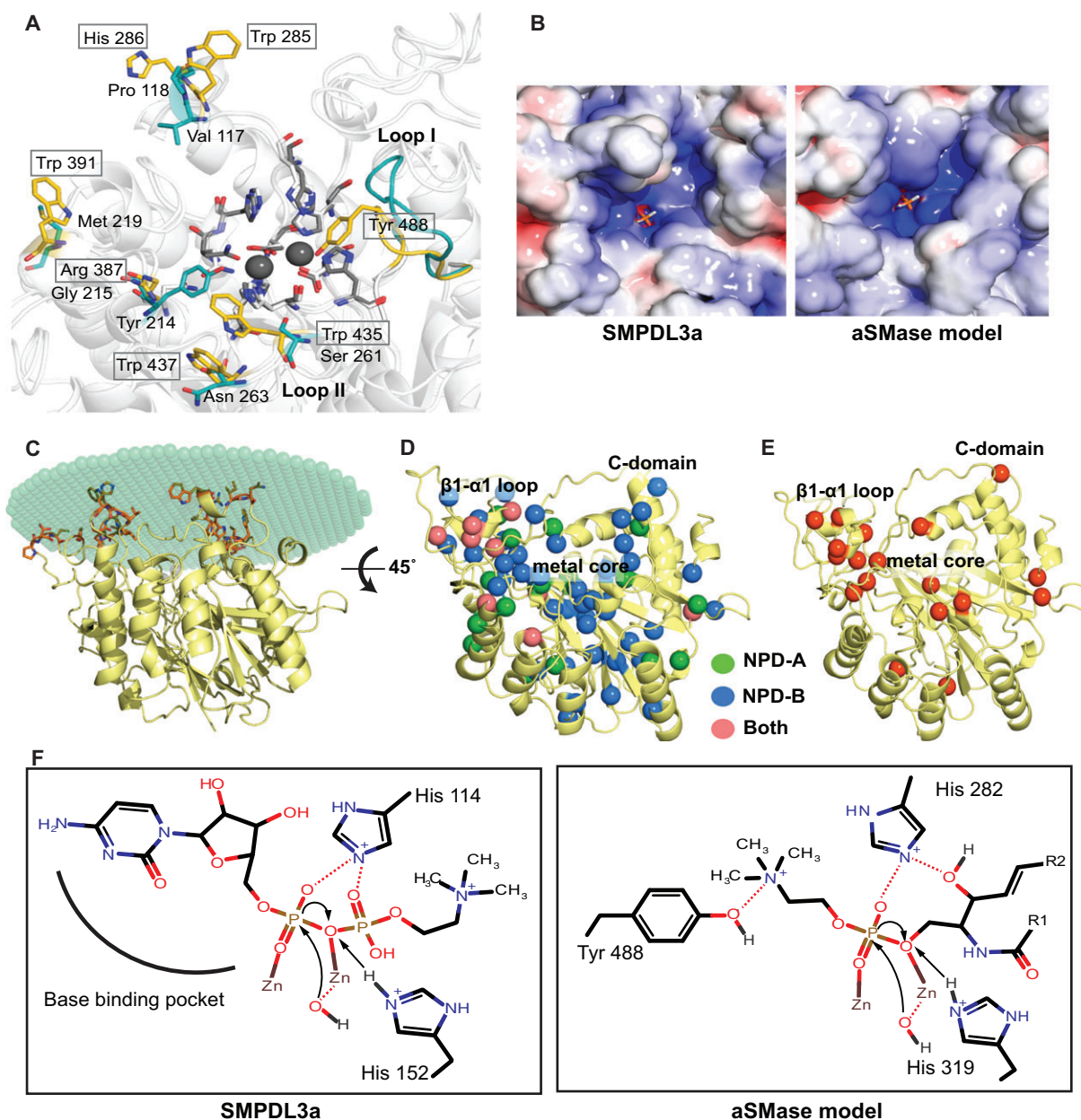


Fig. 6. (A) Key differences of residues surrounding the substrate binding pocket between the models of aSMase (yellow) and SMPDL3a (cyan). The residues at the binuclear core that remained conserved are shown in grey (SMDPL3a) and white (aSMase). Residues corresponding to aSMase are labelled in boxes. (B) Electrostatic potential surface of SMPDL3a and aSMase model with Zn and phosphate calculated with APBS at pH 6.0. (C) Membrane interacting plane of aSMase model predicted by the PPM server. (D) Mapping of known mutations found in NPD-A (green), NPD-B (blue) and both (pink) on the aSMase model. (E) Reported NPD mutations that do not affect protein expression in transfected COS-1 or COS-7 cells are shown in orange. (F) Schematic comparison of the substrate binding pocket of SMPDL3a and aSMase models with proposed Zn activated water as nucleophiles and His152/His319 as acid to protonate the leaving group.

sits next to the substrate entry site and hence it could be crucial for substrate entry and associated conformational changes needed for activity. Mutation of a predicted membrane interacting residue like W391G results in enzyme degradation despite being properly

folded and sorted into lysosomes [38]. Second, some NPD mutations can directly affect the binuclear core site; for example D251H, A281T and H425R mutants have low sphingomyelinase activity [35,39]. Third, some mutations could modulate structural integrity

and stability. For example, based on the aSMase model, Arg496 that is located directly below the binuclear site forms a salt bridge with Leu480 backbone carbonyl to keep the β -sandwich and binuclear site intact. Mutation of Arg496 is found in 32% of Ashkenazi Jewish patients with NPD-A [11]. Despite high mRNA expression observed in R496C transgenic mice, no sphingomyelinase activity is detected [40]. It is likely that the mutant is unstable to adopt a functional fold. Our model could now provide a glimpse of spatial understanding of these mutations and possibly pave the way for future molecular work.

Discussion

The crystal structure of SMPDL3a is the first structure of an acid sphingolipid phosphodiesterase family member. As expected, the structure shows a classical calcineurin fold with a catalytic dimetal site coordinated by primarily histidine and carboxylate residues. We show that the binuclear metal site is occupied by two strongly bound Zn^{2+} ions and that excess zinc can lead to the binding of Zn^{2+} ions at two other sites, with subsequent inhibition of the activity. This phenomenon of Zn inhibition is also observed in ADPRM [21] and S-aSMase [41]. In the case of SMPDL3a, it is possible that the additional Zn^{2+} ions could stabilize a catalytic inactive state allosterically. However, no major conformational changes were observed between the Zn-soaked, native enzyme and product-bound enzyme crystal structures. It is possible that the enzyme adopts a distinct conformation only in the presence of substrate and the effect of Zn allosteric inhibition will be pronounced in comparison with the enzyme–substrate complex.

The five subgroups of sphingomyelinase can be classified to three major types, namely nSMase, aSMase and alk-SMase. Currently there are only two bacterial nSMase structures available. Even though all three types of sphingomyelinases catalyse the same function, they do not share the same protein fold. nSMase belongs to the DNase-I like superfamily [42]; aSMase is a calcineurin-like phosphodiesterase; and alk-SMase is a type 1 nucleotide phosphodiesterase superfamily [43]. Apart from that, these enzymes also differ in their tissue/cellular localization, metal cofactor, pH of optimal function and catalytic site properties [7].

Nonetheless, they have one thing in common: nSMase has an N-terminal single-pass transmembrane domain [44], aSMase has an N-terminal saposin domain and alk-SMase has a C-terminal single-pass transmembrane domain [45]. We speculated that these membrane-associating domains are important in

imparting the sphingomyelinase activity by presenting the substrate to the catalytic site. Similarly SMPDL3b that contained a GPI anchor was reported with sphingomyelinase activity and SMPDL3b knockdown modulates global changes of lipid profile in RAW264.7 cells [17]. However, SMPDL3a overexpression did not seem to alter the ceramide and sphingomyelin composition in HEK cells [13]. Besides substrate presentation, it is possible that the SMPDL3a substrate binding site does not preferentially select sphingomyelin as substrate.

From our 5'CMP bound SMPDL3a structure, the α -phosphate moiety from the CMP product bridges the two metals to demark the site for the scissile phosphate of the substrate in the hydrolysis reaction. Two histidines are important in coordinating the substrate phosphate moiety, i.e. His114 and His152. His152 is also well suited to play the role as the catalytic acid donating a proton to the leaving group. The position of the general acid defines the geometry of the reaction and also the regions of the active site where different moieties of the substrate will bind. Based on our product complexes, we can now postulate the catalytic reaction of SMPDL3a, as well as aSMase, with tentative positions of the substrate and the catalytic His152, shown in Fig. 6F. In aSMase, the unique position of Tyr488 allows only phosphocholine head group to bind with the scissile phosphate that sits on the same position as SMPDL3a's substrate. In this case, His114 could probably interact with the hydroxyl group of the sphingosine backbone or the carbonyl moiety on the alky acid chain (Fig. 6F). Like some of the calcineurin family members, the nucleophilic water in the aSMase family proteins is likely to be coordinated to one of the metals and the metals will play a key role in compensating negative charge built up during the catalytic reaction, including charge concentration in the transition state.

Our homology model of SMase also allows for more detailed mapping of the 19 NPD mutations that are known to be expressed in COS cells, and this structural information might add further resolution to the understanding of their roles in NPD disease. Overall the mutations appear to have three major effects: they (a) modulate the substrate binding site environment, which might also disrupt membrane or saposin binding; (b) target the catalytic metal site; or (c) act as distant effectors of the metal site. The confirmation of their exact roles in modulating the activity, however, needs further studies but these can now be driven by an improved structural model.

In the light of recent work from Traini *et al.* together with our study, it is unlikely that SMPDL3a

functions in lipid metabolites [13]; it is more likely to function as a nucleotide hydrolase and this is consistent with the relatively polar surface in the active site in our structure. The binding of the base in our CMP product structure proposed the likely binding site for the nucleotide bases. SMPDL3a shares only 16% sequence identity with the human ADRPM that has been reported to catalyse CDP-choline, ADP-ribose and CDP-ethanolamine, but we show that SMPDL3a has activity towards these substrates. One possibility is that SMPDL3a acts as a housecleaning enzyme in the extracellular space to recycle different nucleotides towards the nucleoside for import into cells [46]. An additional possibility is that, as proposed, SMPDL3a works in a specific signalling context in the extracellular space, such as for example purinergic signalling [13].

Recently, an interesting study showed that SMPDL3a can act as phosphoramidase to activate a type II diabetes bisamidate prodrug, an inhibitor of fructose-1,6-bisphosphatase, in the liver. Furthermore, SMPDL3a might also be involved in the activation of anti-HIV nucleoside based prodrugs by catalysing the same reaction [16]. Now, many other physiological metabolites might serve as potential substrates for SMPDL3a with this newfound phosphoramidase ability. This also reflects that SMPDL3a may have evolved as an enzyme with broader specificity. Hence, further studies are clearly needed to pin down the exact physiological function of SMPDL3a, and that work can now be guided by detailed structural information.

Materials and methods

SMPDL3a protein expression and purification

The sequence encoding for human SMPDL3a (residues 23–453) (Genbank accession no. BC018999.2) was cloned into a pFB-Sec-NH and expressed as secreted protein in *Spodoptera frugiperda* Sf9 insect cells according to Invitrogen (Singapore) Bac-to-Bac[®] expression system protocol [47]. Cells were harvested by centrifugation at 4000 *g* for 20 min and supernatant was collected for protein extraction. The pH of the medium was adjusted to pH 7 with 10 × equilibration buffer (500 mM Tris pH 8.0, 1.5 M NaCl) followed by addition of 3 mL of 50% HisPur Ni-NTA resin (ThermoFisher Scientific, 88221) slurry in 1 × equilibration buffer to 1 L medium and mixing for 1 h at 4 °C. The protein–nickel mixture were loaded onto a gravity-flow column and washed with 20 mL of wash buffer (50 mM Tris pH 8.0, 150 mM NaCl, 2 mM imidazole) followed by elution with 3 mL of elution buffer (50 mM

Tris pH 8.0, 150 mM NaCl, 250 mM imidazole). Eluted proteins were concentrated and exchanged into storage buffer with a desalting column. Then the sample was incubated with tobacco etch virus protease in a ratio of 1 mg protease to 40 mg protein to remove the hexahistidine tag of the target protein. After overnight incubation at 4 °C, the sample was passed through 1 mL of Ni-NTA resin bed to separate the cleaved proteins from the uncleaved proteins and non-cleavable His-tagged tobacco etch virus protease. Finally, the cleaved proteins were subjected to size exclusion chromatography with a HiLoad 16/60 Superdex 200 prep grade column. SMPDL3a was concentrated to 21.6 mg·mL^{−1} in storage buffer (20 mM HEPES pH 7.0, 300 mM NaCl, 10% glycerol), flash-frozen in liquid nitrogen and stored at −80 °C until further use.

Crystallization and structural determination

A crystallization trial was set up using a protein concentration of 21.6 mg·mL^{−1} in a 96-well sitting drop vapour diffusion method at room temperature. The first crystal appeared in the form of a triangular prism after 7–14 days in the presence of 2.08 M disodium malonate pH 7.2, 0.23 M sodium thiocyanate and 0.01 M *tris*(2-carboxyethyl) phosphine (TCEP) with a protein to reservoir ratio of 0.2 µL : 0.1 µL. Meanwhile another condition that consists of 2.16 M disodium malonate pH 7.2 and 2,2,2-trifluoroethanol gave hexagonal crystals after 21 days with the same protein to reservoir ratio. Crystals of both conditions formed at room temperature. Native crystals were transferred to cryo-solution similar to the reservoir condition but with an additional 25% glycerol as cryo-protectant before being flash-frozen into liquid nitrogen. In order to solve the phase problem, native crystals were soaked with 4 mM ZnCl₂ and the dataset at peak wavelength of the Zn absorption edge (1.2826 Å) was collected and processed with XDS [48]. The phases were calculated with AUTOSHARP version 2.60 by using the single anomalous dispersion (SAD) method [49]. The complete polyalanine model was first manually built on the phased density map of 2.8 Å resolution and then used as molecular replacement template with PHASER-MR on a 2.6 Å native dataset. The phased model was rebuilt with ARP/WARP 7.5 and completed with 98.6% correctness in 15 cycles [50]. The model was manually optimized in COOT 0.8 and further refined with REFMAC in the CCP4 suite [51,52]. The CMP-bound crystal was obtained by soaking a native crystal in 2.08 M disodium malonate pH 5, 0.23 M sodium thiocyanate, 10 mM CDP-choline and 25% glycerol for 10 min at room temperature. Data were collected at beamline ID23.1 European Synchrotron Radiation Facility (ESRF), Grenoble, BL14.1 BESSY-Helmholtz-Zentrum Berlin and the National Synchrotron Radiation Research Centre, Taiwan. Statistics on data collection, processing and refinement are shown in Table 1. The PDB accession numbers for SMPDL3a

apoenzyme and SMPDL3a–5′CMP are [5EBB](#) and [5EBE](#) respectively.

Thermal precipitation assay

Five samples of 10 µg·mL^{−1} of SMPDL3a protein in storage buffer were added to 5 mM of MgCl₂, MnCl₂, CaCl₂ and ZnCl₂ respectively. Each protein–metal mixture was dispensed to 10 aliquots of 20 µL volume and subjected to heating with increments of 4 °C from 64 °C to 100 °C for 8 min. Next, the samples were transferred to a 0.65 µm filtration plate to remove the precipitate formed upon denaturation and aggregation. Soluble fractions were transferred to nitrocellulose membrane in a dot-blot manner and detected by western blotting with ThermoScientific His probe – HRP conjugate (catalog no. 15165). The temperature versus intensity curve was fitted with non-linear regression and T_{agg} is represented as log EC₅₀ where 50% of the proteins were aggregated.

Enzymatic assay of SMPDL3a with Amplex Red sphingomyelinase assay

SMPDL3a hydrolysis activities with substrate sphingomyelin and CDP-choline were measured with Amplex Red sphingomyelinase assay (Life Technologies, Singapore, catalog no. A1220) in the presence of 4 mM EDTA, 4 mM ZnCl₂ and in the native state. A two-step enzymatic assay was carried out according to the protocol with given protein concentration of 1 µM in 100 µL of 50 mM sodium acetate pH 5, 150 mM NaCl reaction buffer and 5 mM substrate. After 1 h incubation at 37 °C, the reactions were neutralized with an equal volume of 50 mM Tris pH 8.5, 150 mM NaCl buffer before development with Amplex Red fluorescence dye. A mixture of Amplex Red fluorescence development solution was prepared according to the assay protocol except for the experiment in Fig. 4G where AP was omitted. *Bacillus cereus* sphingomyelinase was used as a positive control.

Enzymatic assay of SMPDL3a with malachite green

SMPDL3a nucleotide phosphodiesterase activity was tested with 1 µM of enzyme and 5 mM of substrate ATP, ADP and AMP respectively in the presence of 5 mM EDTA, 10 mM ZnCl₂ or without supplement of cofactors or ligands in storage buffer. Reactions were incubated for 30 min at 37 °C. Enzymatic mixtures were aliquoted to a 384-Greiner flat bottom black plate, and then 10 µL of malachite green working solution prepared according to the protocol used by Baykov *et al.* was dispensed to every 20 µL reaction with a multichannel pipette [53]. Absorbance at 630 nm was measured immediately with a Tecan microplate reader.

Enzyme activity of SMPDL3a with CDP-choline, ADP-ribose, CDP-ethanolamine was carried out at pH 5 in a two-step reaction. First, 0.1 µM of enzyme and 2 mM of the respective substrate were mixed with (a) 5 mM EDTA, (b) 10 mM ZnCl₂ or (c) without both in 50 mM sodium acetate pH 5, 150 mM NaCl, 0.005% Tween 20 buffer. The reactions were incubated at 37 °C for 10 min followed by neutralization with 50 mM Tris pH 8.5, 150 mM NaCl buffer in a 1 : 1 ratio. Then, 1 U·mL^{−1} final concentration of calf intestinal phosphatase (CIP) was added to each reaction and the mixture was incubated for 30 min at 37 °C before malachite green development. CIP hydrolyses monoester bonds on SMPDL3a products (phosphocholine, AMP or CMP) to yield phosphate for detection. Initial reaction velocities were measured for 0.1 µM of enzyme with different substrate concentrations. The results were adjusted with non-linear regression based on the Michaelis–Menten equation. The CIP addition step was bypassed when nucleotide was the substrate.

Homology modelling and conservation analysis of human aSMase

The model of aSMase catalytic domain (residues 198–604) was built from SWISS-MODEL homology modelling server with SMPDL3a as model template (<http://swissmodel.expasy.org/>) [54].

Acknowledgements

We would like to acknowledge the staff of BESSY (Berlin, Germany), Diamond (Oxfordshire, UK), ESRF (Grenoble, France), SOLEIL (Gif-sur-Yvette, France), Australia Synchrotron (Melbourne, Australia) and the National Synchrotron Radiation Research Centre (HsinChu, Taiwan) for their support during our crystal screenings and data collection. We thank Ms Dina Darwis for her assistance and expertise in the insect cell culture work. We also thank Dr David James Gill who helped in the phasing process. This study is supported by grants from the Swedish Research Council, Swedish Cancer Society, National Research Foundation (NRF-CRP4-2008-02) and Nanyang Technological University. All research work was carried out at Karolinska Institutet, Sweden, and Nanyang Technological University.

Author contributions

S.M.L., H.L.T., L.T. designed the experiments; S.M.L. and K.Y. performed the experiments; S.M.L. analysed the data; S.M.L., L.T., P.N. wrote the paper. S.M.L.

and H.L.T. critically edited the manuscript before submission.

References

- Okazaki T, Bell RM & Hannun YA (1989) Sphingomyelin turnover induced by vitamin D3 in HL-60 cells. Role in cell differentiation. *J Biol Chem* **264**, 19076–19080.
- Kolesnick R & Golde DW (1994) The sphingomyelin pathway in tumor necrosis factor and interleukin-1 signaling. *Cell* **77**, 325–328.
- Zhang Y, Yao B, Delikat S, Bayoumy S, Lin XH, Basu S, McGinley M, Chan-Hui PY, Lichenstein H & Kolesnick R (1997) Kinase suppressor of Ras is ceramide-activated protein kinase. *Cell* **89**, 63–72.
- Haimovitz-Friedman A, Kolesnick RN & Fuks Z (1997) Ceramide signaling in apoptosis. *Br Med Bull* **53**, 539–553.
- Gangoiti P, Granado MH, Wang SW, Kong JY, Steinbrecher UP & Gomez-Munoz A (2008) Ceramide 1-phosphate stimulates macrophage proliferation through activation of the PI3-kinase/PKB, JNK and ERK1/2 pathways. *Cell Signal* **20**, 726–736.
- Kolesnick RN & Kronke M (1998) Regulation of ceramide production and apoptosis. *Annu Rev Physiol* **60**, 643–665.
- Goni FM & Alonso A (2002) Sphingomyelinases: enzymology and membrane activity. *FEBS Lett* **531**, 38–46.
- Samet D & Barenholz Y (1999) Characterization of acidic and neutral sphingomyelinase activities in crude extracts of HL-60 cells. *Chem Phys Lipids* **102**, 65–77.
- Barnholz Y, Roitman A & Gatt S (1966) Enzymatic hydrolysis of sphingolipids. II. Hydrolysis of sphingomyelin by an enzyme from rat brain. *J Biol Chem* **241**, 3731–3737.
- Ferlinz K, Hurwitz R & Sandhoff K (1991) Molecular basis of acid sphingomyelinase deficiency in a patient with Niemann-Pick disease type A. *Biochem Biophys Res Commun* **179**, 1187–1191.
- Levrin O, Desnick RJ & Schuchman EH (1991) Niemann-Pick disease: a frequent missense mutation in the acid sphingomyelinase gene of Ashkenazi Jewish type A and B patients. *Proc Natl Acad Sci U S A* **88**, 3748–3752.
- Marathe S, Schissel SL, Yellin MJ, Beatini N, Mintzer R, Williams KJ & Tabas I (1998) Human vascular endothelial cells are a rich and regulatable source of secretory sphingomyelinase. Implications for early atherogenesis and ceramide-mediated cell signaling. *J Biol Chem* **273**, 4081–4088.
- Traini M, Quinn CM, Sandoval C, Johansson E, Schroder K, Kockx M, Meikle PJ, Jessup W & Kritharides L (2014) Sphingomyelin phosphodiesterase acid-like 3A (SMPDL3A) is a novel nucleotide phosphodiesterase regulated by cholesterol in human macrophages. *J Biol Chem* **289**, 32895–32913.
- Wright KO, Messing EM & Reeder JE (2002) Increased expression of the acid sphingomyelinase-like protein ASML3a in bladder tumors. *J Urol* **168**, 2645–2649.
- Noto PB, Bukhtiyarov Y, Shi M, McKeever BM, McGeehan GM & Lala DS (2012) Regulation of sphingomyelin phosphodiesterase acid-like 3A gene (SMPDL3A) by liver X receptors. *Mol Pharmacol* **82**, 719–727.
- Kubota K, Inaba S, Nakano R, Watanabe M, Sakurai H, Fukushima Y, Ichikawa K, Takahashi T, Izumi T & Shinagawa A (2015) Identification of activating enzymes of a novel FBPase inhibitor prodrug, CS-917. *Pharmacol Res Perspect* **3**, e00138.
- Heinz LX, Baumann CL, Koberlin MS, Snijder B, Gawish R, Shui G, Sharif O, Aspalter IM, Muller AC, Kandasamy RK, *et al.* (2015) The Lipid-Modifying Enzyme SMPDL3B Negatively Regulates Innate Immunity. *Cell Rep* **11**, 1919–1928.
- Masuishi Y, Nomura A, Okayama A, Kimura Y, Arakawa N & Hirano H (2013) Mass spectrometric identification of glycosylphosphatidylinositol-anchored peptides. *J Proteome Res* **12**, 4617–4626.
- Chen R, Jiang X, Sun D, Han G, Wang F, Ye M, Wang L & Zou H (2009) Glycoproteomics analysis of human liver tissue by combination of multiple enzyme digestion and hydrazide chemistry. *J Proteome Res* **8**, 651–661.
- Lee CY, Tamura T, Rabah N, Lee DY, Ruel I, Hafiane A, Iatan I, Nyholt D, Laporte F, Lazure C, *et al.* (2007) Carboxyl-terminal disulfide bond of acid sphingomyelinase is critical for its secretion and enzymatic function. *Biochemistry* **46**, 14969–14978.
- Rodrigues JR, Fernandez A, Canales J, Cabezas A, Ribeiro JM, Costas MJ & Cameselle JC (2012) Characterization of Danio rerio Mn2+ -dependent ADP-ribose/CDP-alcohol diphosphatase, the structural prototype of the ADPRibase-Mn-like protein family. *PLoS One* **7**, e42249.
- Daumann LJ, McCarthy BY, Hadler KS, Murray TP, Gahan LR, Larrabee JA, Ollis DL & Schenk G (2013) Promiscuity comes at a price: catalytic versatility vs efficiency in different metal ion derivatives of the potential bioremediator GpdQ. *Biochim Biophys Acta* **1834**, 425–432.
- Schenk G, Ge Y, Carrington LE, Wynne CJ, Searle IR, Carroll BJ, Hamilton S & de Jersey J (1999) Binuclear metal centers in plant purple acid phosphatases: Fe-Mn in sweet potato and Fe-Zn in soybean. *Arch Biochem Biophys* **370**, 183–189.
- Schissel SL, Keesler GA, Schuchman EH, Williams KJ & Tabas I (1998) The cellular trafficking and zinc dependence of secretory and lysosomal

- sphingomyelinase, two products of the acid sphingomyelinase gene. *J Biol Chem* **273**, 18250–18259.
- 25 Simons TJ (1991) Intracellular free zinc and zinc buffering in human red blood cells. *J Membr Biol* **123**, 63–71.
 - 26 Frederickson CJ, Giblin LJ, Krezel A, McAdoo DJ, Mueller RN, Zeng Y, Balaji RV, Masalha R, Thompson RB, Fierke CA, *et al.* (2006) Concentrations of extracellular free zinc (pZn) in the central nervous system during simple anesthetization, ischemia and reperfusion. *Exp Neurol* **198**, 285–293.
 - 27 Airola MV, Tumolo JM, Snider J & Hannun YA (2014) Identification and biochemical characterization of an acid sphingomyelinase-like protein from the bacterial plant pathogen *Ralstonia solanacearum* that hydrolyzes ATP to AMP but not sphingomyelin to ceramide. *PLoS One* **9**, e105830.
 - 28 Canales J, Fernandez A, Ribeiro JM, Cabezas A, Rodrigues JR, Cameselle JC & Costas MJ (2008) Mn²⁺-dependent ADP-ribose/CDP-alcohol pyrophosphatase: a novel metallophosphoesterase family preferentially expressed in rodent immune cells. *Biochem J* **413**, 103–113.
 - 29 Sikora J, Pavlu-Pereira H, Elleder M, Roelofs H & Wevers RA (2003) Seven novel acid sphingomyelinase gene mutations in Niemann-Pick type A and B patients. *Ann Hum Genet* **67**, 63–70.
 - 30 Podobnik M, Tyagi R, Matange N, Dermol U, Gupta AK, Mattoo R, Seshadri K & Visweswariah SS (2009) A mycobacterial cyclic AMP phosphodiesterase that moonlights as a modifier of cell wall permeability. *J Biol Chem* **284**, 32846–32857.
 - 31 Simonaro CM, Desnick RJ, McGovern MM, Wasserstein MP & Schuchman EH (2002) The demographics and distribution of type B Niemann-Pick disease: novel mutations lead to new genotype/phenotype correlations. *Am J Hum Genet* **71**, 1413–1419.
 - 32 Airola MV, Allen WJ, Pulkoski-Gross MJ, Obeid LM, Rizzo RC & Hannun YA (2015) Structural basis for ceramide recognition and hydrolysis by human neutral ceramidase. *Structure* **23**, 1482–1491.
 - 33 Ponting CP (1994) Acid sphingomyelinase possesses a domain homologous to its activator proteins: saposins B and D. *Protein Sci* **3**, 359–361.
 - 34 Lomize MA, Lomize AL, Pogozheva ID & Mosberg HI (2006) OPM: orientations of proteins in membranes database. *Bioinformatics* **22**, 623–625.
 - 35 Desnick JP, Kim J, He X, Wasserstein MP, Simonaro CM & Schuchman EH (2010) Identification and characterization of eight novel SMPD1 mutations causing types A and B Niemann-Pick disease. *Mol Med* **16**, 316–321.
 - 36 Rodriguez-Pascau L, Gort L, Schuchman EH, Vilageliu L, Grinberg D & Chabas A (2009) Identification and characterization of SMPD1 mutations causing Niemann-Pick types A and B in Spanish patients. *Hum Mutat* **30**, 1117–1122.
 - 37 Dardis A, Zampieri S, Filocamo M, Burlina A, Bembi B & Pittis MG (2005) Functional in vitro characterization of 14 SMPD1 mutations identified in Italian patients affected by Niemann Pick Type B disease. *Hum Mutat* **26**, 164.
 - 38 Ferlinz K, Hurwitz R, Weiler M, Suzuki K, Sandhoff K & Vanier MT (1995) Molecular analysis of the acid sphingomyelinase deficiency in a family with an intermediate form of Niemann-Pick disease. *Am J Hum Genet* **56**, 1343–1349.
 - 39 Pittis MG, Ricci V, Guerri VI, Marçais C, Ciana G, Dardis A, Gerin F, Stroppiano M, Vanier MT, Filocamo M, *et al.* (2004) Acid sphingomyelinase: identification of nine novel mutations among Italian Niemann Pick type B patients and characterization of in vivo functional in-frame start codon. *Hum Mutat* **24**, 186–187.
 - 40 Jones I, He X, Katouzian F, Darroch PI & Schuchman EH (2008) Characterization of common SMPD1 mutations causing types A and B Niemann-Pick disease and generation of mutation-specific mouse models. *Mol Genet Metab* **95**, 152–162.
 - 41 Jenkins RW, Canals D, Idkowiak-Baldys J, Simbari F, Roddy P, Perry DM, Kitatani K, Luberto C & Hannun YA (2010) Regulated secretion of acid sphingomyelinase: implications for selectivity of ceramide formation. *J Biol Chem* **285**, 35706–35718.
 - 42 Ago H, Oda M, Takahashi M, Tsuge H, Ochi S, Katunuma N, Miyano M & Sakurai J (2006) Structural basis of the sphingomyelin phosphodiesterase activity in neutral sphingomyelinase from *Bacillus cereus*. *J Biol Chem* **281**, 16157–16167.
 - 43 Duan RD (2006) Alkaline sphingomyelinase: an old enzyme with novel implications. *Biochim Biophys Acta* **1761**, 281–291.
 - 44 Chatterjee S, Han H, Rollins S & Cleveland T (1999) Molecular cloning, characterization, and expression of a novel human neutral sphingomyelinase. *J Biol Chem* **274**, 37407–37412.
 - 45 Stefan C, Jansen S & Bollen M (2005) NPP-type ectophosphodiesterases: unity in diversity. *Trends Biochem Sci* **30**, 542–550.
 - 46 McLennan AG (2006) The Nudix hydrolase superfamily. *Cell Mol Life Sci* **63**, 123–143.
 - 47 O'Reilly DR, Miller LK & Luckow VA (1992) Baculovirus Expression Vectors: A Laboratory Manual. W.H. Freeman and Company, New York, NY.
 - 48 Kabsch W (2010) Xds. *Acta Crystallogr D Biol Crystallogr* **66**, 125–132.
 - 49 Bricogne G, Vonrhein C, Flensburg C, Schiltz M & Paciorek W (2003) Generation, representation and flow of phase information in structure determination: recent

- developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr* **59**, 2023–2030.
- 50 Langer G, Cohen SX, Lamzin VS & Perrakis A (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat Protoc* **3**, 1171–1179.
- 51 Emsley P, Lohkamp B, Scott WG & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**, 486–501.
- 52 Murshudov GN, Vagin AA & Dodson EJ (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240–255.
- 53 Baykov AA, Evtushenko OA & Avaeva SM (1988) A malachite green procedure for orthophosphate determination and its use in alkaline phosphatase-based enzyme immunoassay. *Anal Biochem* **171**, 266–270.
- 54 Arnold K, Bordoli L, Kopp J & Schwede T (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201.
- 55 Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465–W469.