Chapter 1
Introduction

The PLAAT family

Phospholipase A/acyltransferases (PLAATs) are a five-membered family of enzymes. They were first identified as H-Ras-like class II tumor suppressors (HRASLS).\(^1\) Numerous research studies in different scientific domains led to them being named in a variety of ways. These names are summarized in Table 1.1. The five members are a part of a larger, diverse superfamily NlpC/P60 thiol proteases or papain-like proteases because of their sequence similarities with lecithin: retinol acyltransferases (LRAT), a member of the NlpC/P60 family.\(^2\) PLAATs share a highly conserved sequence NCEHFV, which contains cysteine that acts as the nucleophile and is part of the catalytic triad.\(^2^-^5\) This triad consists next to the cysteine of two histidines, one acting as the base that deprotonates the sulfhydryl group of the nucleophile and the other stabilizes the imidazole ring of the basic histidine in PLAAT2-5. In PLAAT1, the latter histidine is replaced by asparagine. This cysteine-histidine-histidine catalytic triad is a hallmark signature of NlpC/P60 superfamily of proteins.\(^6\) Four of the PLAAT enzymes are membrane anchored proteins having a C-terminal trans-membrane domain, while PLAAT5, the largest of all PLAATs, lacks this anchor.\(^7\)

Table 1.1. Alternative names for PLAAT proteins found in literature

<table>
<thead>
<tr>
<th>PLAAT1</th>
<th>PLAAT2</th>
<th>PLAAT3</th>
<th>PLAAT4</th>
<th>PLAAT5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A-C1</td>
<td>PLA/AT-2</td>
<td>HREV107</td>
<td>TIG3</td>
<td>HRASLS-5</td>
</tr>
<tr>
<td>HRASLS-1</td>
<td>PLA(_{1/2})-2</td>
<td>HREV107-1</td>
<td>RIG1</td>
<td>RLP-1</td>
</tr>
<tr>
<td>PLA/AT-1</td>
<td>HRASLS-2</td>
<td>H-REV107</td>
<td>PLA/AT-4</td>
<td>HRSL5</td>
</tr>
<tr>
<td>HRSL1</td>
<td></td>
<td>PLA/AT-3</td>
<td>HRASLS-4</td>
<td>HRLP5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>PLA2G16</td>
<td>RARRES3</td>
<td>iNAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td>MCG118754</td>
<td>PL(_{1/2})-3</td>
<td>PL/AT-5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RLP-3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>adPLA</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>HRASLS-3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Only the three-dimensional structures of three PLAATs (PLAAT2, PLAAT3 and PLAAT4) are known. The PLAATs share similar secondary structure motifs and have an active site similar to those found in NlpC/P60 superfamily.\(^2^-^8,^9\) The crystal structures of PLAAT2 (PDB
entry 4DPZ) and PLAAT3 (4DOT) were first determined by Golczak et al.\textsuperscript{2} and the NMR structures of PLAAT3 (2KYT) and PLAAT4 (2MY9) were later solved by Xia et al. (Figure 1.1).\textsuperscript{8,9} The secondary structure comprises three $\alpha$-helices (crystal structures and NMR structure of PLAAT3) or four $\alpha$-helices (NMR structure of PLAAT4) and antiparallel $\beta$-sheet containing four strands (crystal structures) or six strands (NMR structures), similar to classic segregated $\alpha+\beta$-folds of papain-like proteases.\textsuperscript{2,6}

**Functions of PLAATs**

The PLAATs catalyze the first step of reactions leading to the formation of N-acylphosphatidylethanolamines (NAPE), which undergoes further steps to form N-acylethanolamines (NAEs), an important class of bioactive lipids that play roles in a variety of processes, such as anti-inflammation (N-palmitoyethanolamine),\textsuperscript{10–12} catabolism of fat (N-oleoylethanolamine),\textsuperscript{13} anti-apoptotic activity (N-stearoylethanolamine),\textsuperscript{13} and ligands for endocannabinoid receptors (anandamide), see Figure 1.2.\textsuperscript{14–16}

PLAATs, as the name suggests, demonstrate phospholipase A\textsubscript{1/2} (PLA\textsubscript{1/2}) activities, in which both phosphatidylcholine (PC) and phosphatidylethanolamine (PE) act as substrates.\textsuperscript{3,7,17–20} All PLAATs except PLAAT3 show specificity for the PLA\textsubscript{1} position, whereas for PLAAT3 contradicting evidence from various studies suggest that the protein may prefer either the A1 or the A2 position, depending on the type of substrate and assay conditions.\textsuperscript{21–23} The second part of the name comes from their ability to transfer the acyl-chain of a phosphoglyceride (for example PC) to the amino group of a phosphatidylethanolamine (PE), leading to the formation of NAPEs, as mentioned above. Furthermore, apart from being an N-acyltransferase, PLAATs can also act as O-acyltransferases, transferring an acyl-chain to the sn-1 or sn-2 position of a lysophospholipid, for example lysophosphatidylcholine.\textsuperscript{2,3,17,19,24} See Figure 1.3 for all the reactions catalyzed by PLAATs.
Figure 1.1 Crystal\textsuperscript{2} and NMR\textsuperscript{8,9}, structures of PLAAT proteins published till date.
In human beings, mice and rats, the gene for PLAAT1 is expressed mostly in skeletal muscle, heart and testes, where its physiological role has not been investigated so far. PLAAT1 shows an N-acyltransferase activity that is higher than the PLA1/2 activity. PLAAT2 is only found in human beings and especially in the trachea, stomach, colon and kidneys. It has been known to suppress tumors (class II) in breast cells and cervical cancer cells. This protein also exhibits strong N-acyltransferase activity and less PLA1/2 activity and, interestingly, prefers the sn-1 position of PC for the former activity. PLAAT3 was also discovered as a tumor suppressor, however contradicting reports showed that PLAAT3 rather increases tumor progression. In contrast to PLAAT1 and PLAAT2, PLAAT3 shows a preference for PLA activity over N-acyltransferase activity. However, PLAAT1 and PLAAT3 both share common function in organelle degradation in lens. PLAAT3 also plays an important role in viral entry pathways by acting as host factor for enterovirus. PLAAT3 is found mostly in white adipose tissue and less in brown adipose tissue. In white adipose tissue, it modulates lipolysis and therefore is a critical factor in obesity, as was elegantly demonstrated in mice models by Jaworski et al. PLAAT3 inhibitors, therefore would be potential therapeutic
anti-obesity and anti-viral targets. Like PLAAT3, PLAAT4 was also discovered and shown to work as tumor suppressor.\textsuperscript{41–46} PLAAT4, like PLAAT2 is a human-specific ortholog. It is found in skin cells, where it interacts with and activates transglutaminase I (TG1), which in turn produces cornified envelope, necessary for keratinocyte proliferation and survival, and skin to function as a physical and water barrier.\textsuperscript{47–50} As PLAAT4 is a class II tumor suppressor, it was shown to be down-regulated in psoriasis and skin cancer in a study by Duvic \textit{et al.}\textsuperscript{51}

PLAAT4 is a homolog of PLAAT3\textsuperscript{41,42} and it also shows more PLA\textsubscript{1/2} activity than N-acyltransferase activity.\textsuperscript{17,19,24} However, the two enzymes exhibit a contrasting characteristic regarding the PLA\textsubscript{1/2} activity. The transmembrane C-terminal domain is crucial for PLAAT3 membrane-attachment as well as PLA\textsubscript{1/2} activity and truncation of this domain has been shown by Uyama \textit{et al.} to result in loss of phospholipase activity.\textsuperscript{52} In contrast, Golczak \textit{et al.}\textsuperscript{2} demonstrated with truncated N-terminal domains of PLAATs that, unlike truncated PLAAT3, truncated PLAAT4 is capable of phospholipase activity, suggesting that the transmembrane C-terminal domain is not critical for PLAAT4 PLA\textsubscript{1/2} activity. Furthermore, these authors demonstrated by studying the rate of breakdown of short-chain phospholipids and studying the protein-acyl intermediates that the truncated PLAAT4 is more active as phospholipase than PLAAT3. Although PLAAT3 and PLAAT4 are homologs, they clearly differ in activity and specificity toward substrates. Physiologically, the phospholipase activity of PLAAT3 plays an important role in adipose tissue as it regulates triglyceride metabolism.\textsuperscript{21} Phospholipase activity of PLAAT4, on the other hand, plays a crucial role in tumor suppression, particularly in metastasis and invasion.\textsuperscript{46} Moreover Wei \textit{et al.} demonstrated that the NTD of PLAAT4 was found to be enhancing the cell death effect of the CTD, whereas the NTD of PLAAT3 was found to be inhibitory.\textsuperscript{8} Therefore, the differences in activity between the two enzymes can be studied by studying the roles played by the N-terminal and C-terminal domains on a molecular level. The findings could hold the key to the understanding of the molecular mechanisms and physiological significance of these enzymes that are yet to be studied in detail. Studying the roles of N-terminal and C-terminal domains of the two enzymes at a molecular level can help to obtain an understanding of the molecular mechanisms and physiological significance of these enzymes which may advance the design and discovery of selective PLAAT inhibitors.

The aim of the research presented in this thesis, therefore, was to elucidate the reason for the difference in activities (especially phospholipase activity) between PLAAT3 and PLAAT4, present despite their similar sequences and structures. It was hypothesized that differences in the dynamic properties could be the cause of the differences in activity, so NMR spectroscopy
and MD simulations were used to characterize the enzymes. Both methods are briefly introduced below.

**Figure 1.3** Various biochemical reactions catalyzed PLAATs. A, Phospholipase A1 (PLA1) reaction cleaves a fatty acyl chain from the sn-1 position of PC, resulting in the formation of a free fatty acid and lyso PC. B, Phospholipase A2 (PLA2) cleaves at the sn-2 position of PC to form a free fatty acid and lyso PC. C, N-acyltransferase reaction transfers the sn-1 fatty acyl chain from PC to PE to form NAPE. D, O-acyltransferase reaction converts lyso PC to PC, where the acyl donor is the sn-1 acyl chain of PC. E, Hydrolysis of NAPE to N-acylethanolamine (NAE) by NAPE-phospholipase D.

**NMR Spectroscopy**

Studying protein dynamics with NMR spectroscopy has created a new perspective on what proteins are and how they function. This applies to enzymes in particular. Proteins are no longer seen as static entities but as dynamic ensembles occupying various positions in an free energy landscape. protein dynamics occur at many timescales and different sets of NMR
experiments allow us to probe biologically relevant phenomena occurring at those timescales (Figure 1.4). Proximity motions occurring at ps-ns timescale can be studied using longitudinal relaxation (R₁), transverse relaxation (R₂) and heteronuclear nuclear Overhauser effect/enhancement (NOE). Several dynamics parameter, such as the order parameter (S²), global rotational correlation time (τc), effective rotational correlation time (τe) and the exchange rate (Rex) can be quantified using the model free analysis developed by Lipari and Szabo. The order parameter ranges from 0 (highly dynamic) to 1 (rigid), τc is an indicator of molecular tumbling time and is highly dependent on the size of the protein. τe is in the range of ps-ns (τe < τc) and is a measure of local motions, such as loop flexibility. Rex indicates exchange contribution to the linewidth (apparent transverse relaxation rate), due either to motions in the μs-ms timescale that cause chemical shift changes or to chemical exchange phenomena. Generally, data are acquired at two magnetic fields to improve the statistical fitting of the data to one of five “model-free” models, because relaxation rates are field dependent. These five models fit the measured relaxation rates with an increasing number of parameters to obtain S², τe and Rex for individual nuclei in the protein, usually of backbone ¹⁵N atoms. The term model-free refers to the fact that these parameters are defined without a predefined notion of the type of motion in mind, such as rotation in a cone. The ps-ns motions are often associated with flexibility of loops and termini, and in many cases such motions may not be related to biological function but rather are a property of protein matter. However, a growing number of studies have shown that protein dynamics at this timescale can also affect function.

Figure 1.4 A graphical representation of different protein motions occurring at different timescales and NMR experiments to study such motions. In protein dynamics, motions occurring at timescales below μs are termed as “fast timescale motions” whereas those occurring at timescales of μs or above are termed as “slow timescale motions”.

12
Biologically relevant and important phenomena occur at µs-ms timescale such as catalysis, protein folding, protein-protein or protein-ligand interactions. NMR experiments such as $R_{1p}$ and CPMG relaxation dispersion (RD) and chemical exchange saturation transfer (CEST) are crucial in providing information about dynamics occurring at this timescale.

The general principle of all these experiments is that different states have different NMR properties (chemical shifts, dipolar couplings or relaxation rates) and that the exchange between states affects the NMR signals, by causing changes in peak intensity or linewidth. Thus, these experiment can extract information about the population, exchange rates and structures of minor conformers that are in exchange with the major ground state conformation. Since the minor conformation is sparsely populated, the already low intensity peaks are usually further broadened out due to exchange, which makes them ‘invisible’. Therefore, information about the minor conformers is obtained from the effects of exchange on the NMR resonance of the major conformer, or on the signal average in case of fast exchange between the two states. In the CPMG-RD experiment, the chemical exchange is quantified by obtaining the exchange contribution to the apparent transverse relaxation rate ($R_2$) of the nucleus in the major conformer. This is achieved by using a train of 180° pulses at a frequency of $\nu_c = 1/4\tau_{cp}$ (where $\nu_c$ is the CPMG frequency and $2\tau_{cp}$ is time between the 180° pulses in the train) during a period $T$ to reduce the exchange contribution to the linewidth. By performing a series of experiments with varying $\nu_c$, the effective apparent line width ($R_{2,\text{eff}}$) can be determined from the peak intensities in each of these experiment ($I_{\text{CPMG}}$) relative to a reference spectrum for which $T$ is set to zero ($I_0$), equation 1.1.

$$R_{2,\text{eff}} = \frac{1}{T} ln\left(\frac{I_{\text{CPMG}}}{I_0}\right)$$  \hspace{1cm} (Equation 1.1)

Nuclei that do not undergo chemical exchange, do not have the exchange contribution and hence, the $R_{2,\text{eff}}$ is same at all CPMG frequencies and equals the intrinsic $R_2$. A plot of $R_{2,\text{eff}}$ VS. CPMG frequency shows a dispersion profile (Figure 1.5). A nucleus undergoing exchange will have a profile showing a decreasing $R_{2,\text{eff}}$ with increasing CPMG frequency until it reaches a plateau, being the intrinsic $R_2$. The parameters describing the exchange process, such as the rate of exchange ($k_{ex}$), population of the minor conformer ($p_B$) and the absolute chemical shift difference between the major and the minor conformer ($|\Delta\omega|$) can be obtained from the dispersion curves by minimizing the following equation:

$$X^2(\zeta) = (R'_{2,\text{eff}}(\varphi) - R_{2,\text{eff}})/\Delta R_{2,\text{eff}}$$  \hspace{1cm} (Equation 1.2)
Where $R'_{2,\text{eff}}$ is the calculated rate of relaxation obtained by solving the Bloch-McConnel equations,\textsuperscript{103,104} $R_{2,\text{eff}}$ is the effective transverse relaxation rate, $\phi$ is a function of dynamics parameters such as $k_{\text{ex}}$, $p_{B}$ and $\Delta \omega$. The detailed analysis of the CPMG RD experiments to deduce the structure of the minor conformer as well as the experiments required for calculating the sign of the $\Delta \omega$ are outside the scope of this thesis and can be found in many excellent review papers, such as.\textsuperscript{105–110}

Figure 1.5 A typical relaxation dispersion plot. The dashed lines represent an imaginary amide with no exchange and the dispersion curves for the backbone $^{15}$N nucleus of residue R18 in PLAAT4 undergoing chemical exchange, as seen by its decreasing $R_{2,\text{eff}}$ with increasing CPMG frequency. The data are obtained at two magnetic fields, red: 14 T (600 MHz), blue: 20 T (850 MHz).

NMR spectroscopy, though a powerful tool to study protein dynamics, is not bereft of limitations.

It is inherently insensitive due to the small energy gaps involved in magnetic resonance transitions, so concentrated samples are required. Furthermore, depending on the rate of the chemical process, NMR will often yield an average observable, making it difficult to determine the properties of the individual components that cause the averaged observable.\textsuperscript{111} Exchange dynamics is rich in information but can also hinder observation of nuclei considerably, for example if the resonances broaden beyond detection or due to exchange with solvent hydrogens of which the signals are necessarily suppressed in protein NMR experiments. All these factors can make it impossible to interpret dynamic properties in terms of structural changes.
Molecular Dynamics Simulation

Molecular dynamics (MD) simulation is another powerful tool for studying protein dynamics and can serve to complement the information obtained from NMR spectroscopy. MD simulations allow us to quantify and visualize atomic motions occurring in a protein at fast timescale as well as, due to the growth in modern computational capabilities, slower timescale. An introduction to MD simulations is presented in Chapter 3. For the possibilities of combining NMR spectroscopy on proteins and MD simulations, the reader is referred to two excellent reviews, by Case\textsuperscript{112} and Fisette \textit{et al.}\textsuperscript{113}

Thesis Outline

At the start of the research in October 2014, NMR assignments of PLAAT3\textsuperscript{114} and PLAAT4,\textsuperscript{115} the crystal structures of PLAAT2 and PLAAT3\textsuperscript{2} and the NMR structure of PLAAT3\textsuperscript{9} were available. The structure of PLAAT4 was reported in January 2015.\textsuperscript{8} Here, the first protein dynamics data of $^{15}$N-labelled PLAAT3 and PLAAT4 obtained with NMR spectroscopy are presented in Chapter 2. To complement our understanding of the dynamics of the two proteins and to get more information on the residues that could not be studied with NMR spectroscopy due to solvent exchange and slow-timescale-exchange-related line broadening, MD simulations were performed on the two proteins, described in chapter 3. The combination of NMR data and MD results enabled us to formulate a hypothesis to explain the differences in phospholipase activity of two proteins. The hypothesis was then tested by performing mutations both \textit{in silico} and \textit{in vitro}. MD simulations as well as activity assays on the variants were performed, leading to the remarkable conclusion that loop exchange from PLAAT4 to PLAAT3 can introduce enzymatic activity in the latter protein, a gain-of-function mutation. These results are described in Chapter 4. In Chapter 5, a study on CPMG RD NMR experiments is described, in which an artifact observed during the slow pulsing regimes is analyzed and the efficacy of different decoupling sequences in its removal was tested. Chapter 6 contains a general discussion on PLAAT3 and PLAAT4 dynamics and activity and the scope for future research on the phospholipase A/acyltransferase family.
References


Chapter 1


Chapter 1


Chapter 1

Protein Sci. 11, 221–232.


Chapter 1


