1. Introduction

Palmitoylation is a reversible, post-translational modification of a protein through the addition of the 16-carbon fatty acid, palmitate, to a cysteine residue. There are two types of palmitoylation, one called thio- or S-palmitoylation in which palmitate is added to the thiol side chain of a cysteine residue via a labile thioester bond [1]. The other type, N-palmitoylation, is the addition of palmitate to an N-terminal cysteine via a stable amide bond [2]. The two forms of palmitoylation are regulated by different families of palmitoyl acyltransferases (PATs) — S-palmitoylation via a family of multi-pass transmembrane proteins called DHHC (Asp-His-His-Cys) proteins [3] and N-palmitoylation via a family of multi-pass transmembrane proteins termed membrane-bound O-acyltransferase [4]. S-palmitoylation, the focus of this chapter, is more common and because of the labile thioester bond, can dynamically regulate protein sorting and function.

Palmitoylation increases the lipophilicity of the modified protein often changing its subcellular distribution in both dramatic and subtle ways. The larger-scale changes occur when cytoplasmic proteins relocate from the cytoplasm to membrane and when integral membrane proteins move from one membrane system to another, such as from the endoplasmic reticulum (ER) to the plasma membrane (PM). The more subtle changes, in terms of distance, occur at the nanoscale level within a membrane. The increase in lipophilicity upon palmitoylation often results in an altered affinity for a particular lipid microenvironment within that membrane [5]. For example, lipid rafts are small islands in membranes with distinct lipid compositions that selectively attract or exclude both peripheral (often exclusively by virtue of palmitoylation) and integral membrane palmitoylated proteins. Palmitoylated proteins have affinity for lipid rafts that are rich in cholesterol, while prenylated proteins have little or no affinity for these rafts [5]. Such lipophilicity-driven changes in protein dis-
tribution may alter access of a palmitoylated protein to extracellular ligands (when the protein moves from the ER to the PM), protein-protein interactions, or the engagement of the palmitoyl-protein in multi-molecular signaling complexes. The role of palmitoylation as a versatile protein sorting signal, regulating intracellular protein trafficking and targeting to membrane microdomains has been reviewed recently [6]. Palmitate may be the most common lipid species to occupy cysteine residues, but it is not the only one. Marilyn Resh and colleagues identified the lipid moieties resident on the cysteine residue of the N-terminal tail of Src family kinases [7-9]. While for these proteins the cysteine residue near the N-terminus is most frequently palmitoylated, it is also modified by palmitoleate, stearate, or oleate with a frequency that is apparently related to the abundance of palmitate in cells [10]. The physiological differences that result from proteins being modified by these other lipids has not been explored extensively; however, given their different physical properties, it seems reasonable that their impact on a protein should be subtly different than palmitate.

Unlike other forms of lipidation such as myristoylation and prenylation, palmitoylation is reversible, by virtue of the labile thioester bond. This allows for dynamic regulation of the protein’s lipophilicity [11-13]. By contrast, prenyl groups are attached to cysteines by a stable thioether bond and myristate to glycines by a stable amide bond. It is now apparent that many instances of palmitoylation are enzymatically mediated by a family of palmitoyl acyltransferases (PATs), whereas the mechanisms for depalmitoylation are poorly understood. Nevertheless, it is known that palmitate cycles on and off of many proteins at variable rates ranging from minutes to days. Such dynamic regulation makes palmitoylation unique among post-translational protein lipid modifications and places it in a category similar to phosphorylation. Discovering the molecular identity of PATs was a pivotal event that dramatically accelerated the pace of discovery in the field. Likewise, there has been increased interest in palmitoylation partly because many of the genes encoding PATs have been linked to human diseases like cancer. With a greater understanding of how palmitate is enzymatically attached to proteins, some of the most interesting questions include: What are the substrate(s) of each PAT?; how does a PAT recognize and palmitoylated a substrate?; how are PATs regulated?; and how is depalmitoylation regulated? The answers to these questions are beginning to unfold due to the recent discovery of pharmacological modulators of palmitoylation as well as the development of novel assays and refinement of existing assays. Our ability to understand palmitoylation and its importance to human health and disease is only as good as the methods we use to test our hypotheses. Thus, the discovery of potent and selective inhibitors of palmitoylation as well as the continued development of assays with increased sensitivity and selectivity is critical to this venture.

2. Palmitoylation and DHHC proteins

2.1. Molecular identity of palmitoyl acyltransferases (PATs)

It has been known for many years that palmitoylation is a critical regulator of diverse and complex signaling networks, but the mechanism responsible for palmitoylation of most pro-
Proteins remained a mystery and somewhat controversial until only recently. The apparent absence of a consensus site for palmitoylation encoded by the sequence of amino acid residues surrounding palmitoyl cysteines, as well as the difficulty in purifying and identifying the enzymes capable of mediating the reaction, led many to believe that it was autocatalytic. Given these issues and the high reactivity of cysteines and palmitoyl-CoA, especially in *in vitro* protein palmitoylation assays, the possibility was not unreasonable [11, 14, 15]. Many of the arguments for and against autocatalytic palmitoylation have been reviewed recently [16]. Yet, given the prevalence of palmitoylated proteins in parts of cells where signaling events are so highly concentrated, complex, and regulated, such as the neuronal synapse, it seemed somewhat unreasonable that all regulation of palmitoylation could be left to diffusion—a nagging reality that kept the search for an enzymatic mechanism alive despite the arguments to the contrary. Additionally, there was evidence over the years in support of the idea that these enzymes existed because PAT activity in detergent solubilized protein fractions had been measured using viral glycoproteins [17], p59*fy* [18], and H-Ras [19] as substrates among others.

The experiments that conclusively provided the molecular identity of PATs were presented in a series of papers spanning almost a decade. The experimental model organism that ultimately provided the information was yeast. First, palmitoylation-dependent alleles of yeast *RAS2* were identified. A genetic screen designed to identify mutations that rendered cells non-viable if Ras2p was not palmitoylated was utilized to identify mutations in two genes—*ERF2* and *ERF4/SHR5* [20, 21]. These mutations resulted in diminished palmitoylation of Ras2p and mislocalization of GFP-Ras2p (respectively or it takes both mutations to cause both effects [20, 22]). However, it could not be decisively concluded if the mutations in *ERF2* and *ERF4* were affecting Ras2p palmitoylation directly or indirectly by altering Ras2p trafficking (which could have prevented an interaction between the palmitoyl acyltransferase and Ras2p).

In collaboration with Maurine Linder, Deschenes and colleagues used an *in vitro* palmitoylation assay to show that Erf2p and Erf4p together constituted a Ras2p PAT that used palmitoyl-CoA as a donor [23]. Erf2p is a ~42-kDa integral membrane protein that is expressed in the ER. The protein contains the DHHC-CRD (Asp-His-His-Cys-cysteine rich domain), also referred to as the NEW1 or zf-DHHC domain (PF01529), which is found in an extensive family of membrane proteins ranging from unicellular eukaryotes to humans [24, 25]. This domain is now recognized as the molecular signature for PATs that add palmitate to cysteines via a labile thioester bond.

At almost the same time that the Erf2p/Erf4p complex was identified as the Ras2p PAT, Akr1p was identified as a PAT with specificity for Yck2p [26]. An important clue leading to the relationship between these two proteins came from the fact that mutants in both Ras2p and Yck2P exhibited a reduced rate of pheromone receptor internalization [27, 28]. Akr1p contains a DHYC-CRD instead of a DHHC-CRD as well as ankyrin repeats not present in Erf2p. The DHYC motif present in three yeast proteins (Akr1p, Akr2p and Pfa5) does not appear to occur in the mammalian genome. Akr1p and Akr2p are most closely related to the mammalian HIP14 (DHHC17) and HIP14L (DHHC13) which contains the variant DQHC—the only observed mammalian deviation from DHHC [3].
2.2. The ZDHHC family of PATs

The mammalian genome contains at least 23 members of the ZDHHC PAT gene family identified by the presence of the signature DHHC-cysteine rich domain. Members of the family had been identified as being genes of interest (e.g., “REAM” in metastatic cancer [29]) prior to understanding their function. The genomic structure of ZDHHC genes varies widely, including the number and differential use of exons that are spliced together to generate the mRNA. EC gene analyses (http://genome.ewha.ac.kr/ECgene/) of the mRNAs that encode PATs suggest that all of the genes are alternatively spliced at various sites throughout the protein coding sequence as well as within untranslated regions. Many of the putative, alternatively-spliced exons are predicted to encode small peptides that change the structure of the protein in a way that may alter substrate specificity. Similarly, splicing may alter sites for other post-translational modifications, such as phosphorylation or glycosylation, all of which may regulate activity, substrate specificity, subcellular distribution, or interactions with non-substrate proteins. ZDHHC7, for example, alters the use of a 111 bp exon that is differentially and specifically expressed in tissues such as placenta, lung, liver, thymus, and small intestine [30]. This exon encodes a 37-residue peptide (EKSSDCRP-SACTVKTGLDPTLVGICGEGTESVQSLLL) within the intracellular loop between transmembrane domain 2 (TM2) and TM3 that contains a PKC phosphorylation site. It is conceivable that phosphorylation of this serine changes DHHC7 in such a way that substrate specificity or the rate of palmitate transfer activity is altered. In addition to alternative mRNA splicing, aberrant splicing induced by mutations or single nucleotide polymorphisms has been shown to occur in at least two ZDHHC genes. A splice-site mutation in highly conserved residues of ZDHHC9, a PAT that has been shown to palmitoylate H-Ras and N-Ras [31], has been described in families with X-linked mental retardation (XLMR) [32]. This mutation creates an additional, stronger splice-donor site 140 nt before (toward the 5’ end) the normal donor site. Usage of the new site results in a mRNA that is frameshifted and that encodes a truncated protein. Single nucleotide polymorphisms that affect splicing of ZDHHC8 have also been implicated in schizophrenia [33] (also see: [34-36]).

Hydropathy analyses predict that the PATs encoded by these genes all pass through a membrane multiple times (at least four) and are expressed predominantly in the ER and Golgi membranes [30, 37, 38]. Currently, there is little published data on the numbers of TM domains in any of the PATs with the exception of Akr1p in yeast [39]. Predictions using TopPred II 1.1 [40] as presented by Ohno and colleagues [30] show that most PATs have an even number of TM domains with the DHHC-CRD motif in the cytoplasm. However for DHHC13, -16, -11, and -22, the DHHC-CRD motif resides just C-terminal to the first or third TM domain. Assuming the N-terminus is cytoplasmic, this places the DHHC-CRD motif either in the lumen of the ER (the membrane compartment of residence reported for each by Ohno and colleagues 2006) or outside of the cell if expressed on the PM. Given that the environment in these two locations is oxidizing in nature [41, 42] and assuming this topological model is correct, it is possible that the cysteines of the DHHC-CRD motif could form inter- or intra-molecular disulfide bridges rather than being involved in the transfer of palmitate. However, while it is possible that PATs may assume duties in addition to palmitoylation, it
seems somewhat unlikely they would do so in this arrangement as it represents a state in which it would be difficult to perform these functions. The highest-scoring predictions of the membrane topology using TMpred show that the human protein sequence of DHHC11 and -16 should contain four TM domains, DHHC13 eight TM domains, and DHHC22 either four or five TM domains, with each model placing the DHHC-CRD motif inside the cells. There is clear disparity among the predictions generated by the algorithms available and ultimately, any of these predictions of topology must be confirmed or disproved by experimental data. In any case, for a member of the PAT family to function as a PAT, the DHHC-CRD motif should probably reside in the cytoplasm (Figure 1A). The regions of the PAT proteins that contain the greatest diversity at the amino-acid level are the N- and C-terminal cytoplasmic tails (Figure 1B).

Figure 1. Predicted general structure of PATs. A) The predicted topology of PATs places the DHHC-CRD domain in the cytoplasm but such predictions must be confirmed experimentally. B) Each PAT is predicted to span the membrane four to six times; each is ~40 kDa with the greatest degree of sequence variability residing in the cytoplasmic N- and C-termini. The DHHC-CRD motif defines PATs. Palmitoylation of the cysteine in the DHHC portion is required for transfer of palmitate to a substrate. Most PATs also have a conserved DPG (aspartate-proline-glycine) motif and TTxE (threonine-threonine-asparagine-glutamate) motif, but their role in the function of PATs is not yet known.

In addition to the importance of PAT membrane topology, their membrane system of residence is likely to be an important aspect of their function. PATs have been localized to ER, Golgi, plasma membrane, endosomes, and the yeast vacuole [30, 43-48]; yet, little is known about how these proteins achieve their respective localizations. Immunolocalization of epitope-tagged DHHC proteins has been somewhat inconsistent among various cell types, between laboratories (e.g. DHHC2 [30, 37], and even in our own laboratory (unpublished observations SLP) in terms of within which membrane system a protein resides. Such inconsistencies suggest that the cell type, cell cycle, health of the cells, or even the location of the epitope tag may affect subcellular distribution. An interesting exception is DHHC2. DHHC2 has recently been shown to traffic between the PM and intracellular membranes via recycling endosomes [47]. Importantly, the C-terminal 68 amino acids of DHHC2 was shown to play an important role in defining
its intracellular localization; however, a defined targeting signal present within this region of DHHC2 and in other DHHC proteins has yet to be defined.

2.3. PAT genes, palmitoylation and human disease

PATs have already been linked, in varying degrees, to human disease despite their recent discovery. At least 7 genes encoding PATs have been implicated in human disorders, including ZDHHC8 with schizophrenia [33], ZDHHC17 with Huntington’s disease [49], ZDHHC15 and ZDHHC9 with X-linked mental retardation [32, 50], and ZDHHC2, ZDHHC9, ZDHHC17, and ZDHHC11 with cancer [29, 51-53]; most of the demonstrated and putative connections are with cancer.

Overexpression of some PATs has also been shown to alter cancer-related signaling. DHHC17 (HIP14) is oncogenic. DHHC9 and DHHC11 display characteristics that strongly suggest they also are oncogenic. Overexpression of DHHC17 has the ability to induce colony formation and anchorage-independent growth in cell culture and tumors in mice [53]. It has been shown that these effects occur, at least in part, by DHHC17-mediated palmitoylation of H-, N-, and K2A- RAS proteins [53]. DHHC9 is strongly upregulated in some adenocarcinomas of the gastrointestinal tract at the transcript and protein levels [52] and has also been shown to palmitoylate H- and N- RAS proteins \textit{in vitro} [31]. ZDHHC11 has a high incidence of additional genomic copies in cases of non-small cell lung cancer and bladder cancer in which it is strongly linked to high-grade, advanced stage and disease progression [51].

Conversely to the behavior of the oncogenic PATs, a failure to express ZDHHC2 results in an increase in metastasis in an \textit{in vivo} model leading to the suggestion that ZDHHC2 is a tumor/metastasis suppressor [29]. This absence of expression suggests that substrates of DHHC2 are no longer palmitoylated, and that whatever role palmitoylation had in signaling downstream from that event has been disrupted. Such is the case of DHHC2, where due to a lack of palmitoylation, one of its substrates, CKAP4, is no longer normally palmitoylated. One consequence of this is that CKAP4 no longer traffics efficiently (or at all) to the cell surface where it acts as a receptor for antiproliferative factor (APF) [37] [or presumably its other two known ligands, tissue plasminogen activator [54] and surfactant protein A [55]]. Without surface expression of CKAP4, APF is unable to initiate a wide range of downstream effects, including halting cellular proliferation and altering the expression of genes related to the progression of cancer [44].

CD9 and CD151, both tetraspanin proteins, have also been identified as substrates of DHHC2 [56]. CD9, which has been suggested to be a tumor suppressor [57, 58], is palmitoylated on multiple cysteines, but which of these are palmitoylated by DHHC2 is not known. Nonetheless, it is clear that suppression of DHHC2-mediated palmitoylation of CD9 in A431 cells affects cell behaviors that are consistent with it playing a role in tumor suppression. In particular, the cells undergo what appears to be epithelial-mesenchymal transition (EMT) — a process in which epithelial cells lose epithelial morphology and markers and gain a fibroblastic morphology during tumor progression [59-61]. It is not yet clear whether this change in cellular behavior was mediated solely by the reduction in CD9 palmitoylation or through reduced palmitoylation of this and other DHHC2 substrates such as CKAP4. It will be inter-
esting to learn if a select subset of cysteines of CD9 is palmitoylated by DHHC2 and also how decreasing palmitoylation of specific cysteines results in the observed cellular behavior. Several other substrates of DHHC2 have been identified ranging from the neuronal adaptor/scaffold protein PSD95 [62], the SNARE proteins SNAP-23/25 [63], the non-receptor tyrosine kinase Lck [64], and the intracellular signaling proteins Gaq2 [65], GAP43 [62], R7BP [66], and eNOS [48]. Notably, there is no apparent structural similarity between the reported substrates of DHHC2, or even any sequence similarities surrounding the palmitoylated cysteine residues. Thus, DHHC2 can apparently palmitoylate cysteines located in the N-terminal regions (PSD-95, GAP-43, and Ga), internally in the protein sequence (SNAP-23/25), in the juxtamembrane region of transmembrane proteins (CD9, CD151, and CKAP4) and close to an N-terminal myristoylated glycine (Lck and eNOS).

From these examples, it is clear that upsetting the homeostatic balance of protein palmitoylation, in either direction, can have significant and deleterious effects on signaling networks. It is also clear that identification of PAT cognate substrates will provide important information concerning the molecular mechanisms underlying the oncogenic nature of the affiliated signaling systems as well as reveal important, novel targets for pharmacologic intervention. The development of specific DHHC protein inhibitors would provide vital reagents with which to study the physiological and pathophysiological importance of many palmitoylated proteins and may offer potential for therapeutic development.

2.4. PAT functions in addition to palmitoylation

It is not surprising that a disruption in the homeostatic balance of protein palmitoylation, in either direction, can have pathophysiological consequences. However, one must remain mindful that palmitoylation may not be the sole function of these proteins. Recently, two PATs—HIP14 (DHHC17) and HIP14L (DHHC13)—have been shown to mediate the transport of Mg$^{2+}$ [67]. The first indication that these PATs were involved in Mg$^{2+}$ regulation was that the abundance of their corresponding mRNAs was increased in cells grown in medium with reduced Mg$^{2+}$ concentration. The authors then showed that Mg$^{2+}$ (but not Ca$^{2+}$) transport was both electrogenic and voltage dependent, and that the transport required palmitoylation of the PAT. The authors concluded that these two PATs fall into a category of enzymes called “chanzymes” or ion channels that also have enzymatic activity; a type of protein previously represented only by the transient receptor potential melastatin (TRPM) family of transporters [68, 69]. The fact that GODZ (DHHC3) does not appear to mediate Mg$^{2+}$ transport [70] but can mediate the transport of Ca$^{2+}$ [71] suggests that this is not a general property of all PATs. The discovery that these PATs transport Mg$^{2+}$ was astonishing especially in light of the fact that the DHHC-CRD motif appears, by sequence and predicted structure, to be a Zn$^{2+}$-binding protein; (a divalent cation with an atomic radius similar to Mg$^{2+}$)—not Mg$^{2+}$. However, Goytain and colleagues also found that HIP14 and HIP14L transported Zn$^{2+}$ with approximately half the efficiency as Mg$^{2+}$. The role of these and other PATs in binding to and/or transporting Zn$^{2+}$ remains to be elucidated, but demonstrates the importance of not limiting ones view of PAT function (or many other proteins for that matter) only to palmitoylation.
2.5. Enzymatic mechanisms of palmitoylation

The physical and chemical mechanisms that result in enzymatic palmitoylation have yet to be defined clearly, but some progress has been made using purified proteins. It has been established that mutation of the cysteine in the DHH motif of all PATs studied to date abolishes autoacylation of the PAT and palmitoylation of the substrate [23, 56, 62, 72]. This literature as well as discussion of potential physical mechanisms for the reaction have been reviewed recently [3, 73].

3. Palmitoyl-Cysteine prediction

Prior to the discovery of PATs, attempts were made to define stretches of amino acids that were preferred for palmitoylation. Palmitoylation near the N-terminus, following myristoylation, is among the predictable places for palmitoylation to occur provided there is one or more nearby cysteines. Navarro-Lérida et al (2002) fused a myristoylation motif (MGCTLS) to GFP with a short intervening sequence containing cysteines at various locations. These authors found a preference for cysteine palmitoylation at positions 3, 9, 15 and (to a much lesser degree) 21 residues away from the N-terminal methionine, but intervening residues were not evaluated. Commonalities in the composition of amino acid residues surrounding palmitoylated cysteines have been noted among members of the family of yeast amino acid permeases [74].

As more palmitoylated proteins and specific palmitoyl-cysteines are discovered, the task of predicting which adjacent amino acids provide a favorable environment for palmitoylation becomes easier. Algorithms trained with data from identified palmitoyl cysteines and adjacent amino acid residues are now able to provide predictions of the statistical likelihood that a cysteine of interest may be palmitoylated [75-78]. CSS-Palm 2.0, which was designed to predict potential palmitoylation sites, has been published [75]. The algorithm was trained to recognize potential palmitoyl-cysteines using a dataset of 263 experimentally determined palmitoylation sites from 109 distinct proteins. Interestingly, CSS-Palm 2.0 also successfully predicted most (~75%) of the same novel palmitoyl-cysteines in yeast proteins previously identified by Roth. et al [74] as well as palmitoyl-cysteines predicted by Roth et al., to be palmitoylated but not experimentally determined. This rate of success in both cases suggests that CSS-PALM 2.0 is more conservative at calling a site, potentially resulting in a greater rate of false negative results but is reasonably accurate nonetheless. This algorithm should prove useful when prioritizing which cysteine(s), often among multiple potential cysteines of a candidate palmitoyl protein, to analyze experimentally.

Patterns of amino acid residues surrounding palmitoyl-cysteines have emerged from these analyses. A diagram of favored residues generated by an early version of CSS-Palm 2.0 (NBA-Palm) [76] shows that leucines and additional cysteines are more commonly observed around palmitoyl-cysteines. The subsequent versions of NBA-palm used significantly improved predictive tests, but the rough sequence of preferred residues remains. An important aspect that cannot yet be considered when attempting to predict cysteine palmitoylation
with these algorithms is the complexity of the PAT-substrate recognition that is encoded by residues outside of those that immediately surround the palmitoyl-cysteine; the higher order components of the recognition sites.

3.1. The physical properties of cysteines and thioester bonds

The unique physical and biochemical nature of the thioester bond that links palmitate to cysteine residues is the basis for the design of many recent assays for palmitoylation. The cysteine residue is among the most nucleophilic entities in a cell [79] and is the most common site of palmitoylation. Other residues can be modified by palmitate, but their occurrence is relatively rare and the bond chemistries are different [2, 80-83]. Palmitoylation can also occur in other ways, for example, on an amine of an N-terminal cysteine as is the case with Hedgehog [2, 83, 84], a secreted signaling protein. An example of palmitate modifying the weaker –OH nucleophile of threonine occurs on the carboxyl terminus of a spider toxin [81]. The ε-amino group of lysine can also be modified by palmitate linked by an amide bond. This occurs in several secreted proteins including a bacterial toxin [80].

The reactivity of the thiolate anion of cysteine residues makes it a key component in the structure and function of many proteins by stabilizing higher order structures via disulfide bridges and post-translational modifications like nitrosylation, prenylation, and acylation [85-87]. The high degree of reactivity has also provided a well-characterized, indispensable target for modification by synthetic, thiol-reactive ligands, allowing capture and characterization of proteins [88]. An exceptionally useful application of such thiol-specific chemistry is isotope-coded affinity tags (ICAT) for mass spectrometric determination of relative protein or peptide abundance among two or more samples [89-91]. With these probes, changes in abundance of identified proteins or peptides are determined by changes in the ratio of heavy to light-isotope-modified peptides from mixed samples. Combining ICAT technology with functional genomics methods like siRNA-mediated PAT-gene knockdown is one of several mechanisms that will allow us to identify substrates of PATs [37].

In healthy cells the cytoplasm is generally a reducing environment, meaning that solvent-exposed cysteine side chains are not typically disulfides and thus available to engage in reactions with other molecules [92]. The reactivity of a free cysteine depends on the pKa of the cysteine which is a function of the local environment surrounding the residue within the context of the whole protein. Unlike other residues with nucleophilic side chains (–OH or –NH₂), thiol side chains undergo conjugation, redox, and exchange reactions [85]. Conjugation reactions (in addition to fatty acylation) include nitric oxide (NO) or S-nitrosylation, reactive oxygen species (ROS), and reactive nitrogen species (RNS) forming bonds that are not susceptible to cleavage by hydroxylamine at neutral pH. Hydroxylamine is a reagent used to selectively remove thioester-linked palmitate [93]. Importantly, we know that hydroxylamine does not perturb disulfides [94], and that it efficiently cleaves thioesters in a quantitative manner [95].

In addition to the linkage of palmitate to cysteines, another thioester bond that occurs in cells is the transient association between ubiquitin and the E1, E2, and certain E3 ubiquitination enzymes [87, 96]. However, these thioester bonds are easily distinguished from the thio-
ester bond that links palmitate to cysteines by their pKa; the pKa in the case of palmitoylation is near neutral pH (~7.4) whereas, for the thioester in the ubiquitin system it is pH 10.5 or greater. This wide differential allows for a high degree of selectivity when using hydroxylamine to cleave palmitate from proteins on the physical characteristics of ubiquitin-related cysteines. It is highly unlikely that they are ever in a position to be palmitoylated [97, 98].

Retinoic acid (RA) and RA-CoA have also been shown to be enzymatically attached to cysteines via a thioester bond that can be cleaved by hydroxylamine and reducing reagents such as βME at neutral pH. The reaction can be inhibited, but not fully, by myristate and palmitate suggesting that RA competes for the same cysteines as palmitate [99-107]. There is some debate in the RA field about how it binds to proteins, particularly the nuclear RA receptors, to carry out its signaling functions. RA binding to a hydrophobic cleft is the favored mechanism; however, there are many effects of RA (e.g. [108, 109]) that are independent of RA-receptor binding suggesting that cysteine modification may also have a place in the molecular mechanism of RA action.

3.2. Mass spectrometric identification of acyl groups that modify cysteines via a thioester bond

Lipid-modified thiols have been successfully identified using MALDI-TOF (matrix-assisted laser desorption ionization time-of-flight) mass spectrometry [110]. Using this method, direct information on the nature of the endogenous lipids on proteins or peptides (revealing interesting variability) can be obtained, whereas most other methods rely on surrogate markers for palmitate including thiol-reactive probes or radiolabeled palmitate. Using MALDI-TOF mass spectrometry, Marilyn Resh and colleagues found that the cysteine in the N-terminal Met-Gly-Cys of Src family kinases and two cysteines near the N-terminus of GAP43 are modified not only by palmitate but also (and to a lesser degree) by palmitoleate, stearate, or oleate [7, 8]. While palmitate appears to be the most common acyl group that forms a thioester bond to modify internal, cytoplasmic cysteines, it is clearly not the only one. The 16-carbon palmitate acyl group represents the longest chain synthesized by mammalian fatty acid synthase and is apparently the most abundant chain length present in some tissue types [111]. This relatively greater abundance may underlie the dominance of palmitate as the main acyl group to modify free thiols by S-acylation. The functional implications of incorporating lipids with shorter or longer acyl chains and especially those with different degrees of saturation may be that the proteins have different affinities for various lipid microdomains present in membranes. The specificity of PATs for chain-lengths different than 16 carbons has not been rigorously defined. However, it is known that acyl groups with differing carbon chain lengths and degrees of saturation can also be incorporated [7, 112].

3.3. PAT/Substrate recognition

Determining the nature of PAT/substrate recognition remains one of the more important tasks to be undertaken. This is especially true for PATs encoded by genes that have been linked to disease. There are two general approaches to defining PAT/substrate relationships:
1. identification of the PAT with specificity for a known palmitoyl protein and

2. identification of an unknown substrate of an individual PAT.

The first of these has been the most common. With this forward approach, each one of the 23 PATs is independently co-overexpressed with a known palmitoyl protein; the cells expressing the pair are metabolically labeled with $^{3}H$-palmitate and the proteins analyzed by SDS-PAGE and fluorography. The incorporation of $^{3}H$-palmitate onto the substrate protein in one or more of the co-overexpressions at a level significantly above background suggests that a particular PAT is responsible for palmitoylating that known substrate. Similarly, the presumptive PAT and substrate proteins can be purified and combined with $^{3}H$-palmitoyl-CoA in a tube, allowed to react, and the incorporation of $^{3}H$-palmitate measured as above. The current level of understanding of PAT/substrate recognition makes it unreasonable to assume that the more closely two PATs are related by sequence homology, the more likely they should palmitoylate a particular substrate. For this reason, assigning substrate status of a protein to a single PAT among a select group of tested, more closely-homologous PATs, to the exclusion of others because they are less homologous, may lead to erroneous exclusions. Similarly, we cannot yet assume that homology among residues surrounding palmitoyl cysteines of different proteins is an indication that they are palmitoylated by a particular PAT. The mechanism for molecular recognition is likely to be defined in part by the higher order structure (even quaternary as is the case with ERF2p and AKR1p) of the PATs and substrates. The reverse approach, defining unknown substrates of a single PAT can occur without these same biases as has been demonstrated in yeast and in human cells [37, 74].

4. Novel methods to discover and identify PAT/substrate specificity

The chemistry supporting novel assays to study palmitoylation and the reagents that are being incorporated into them have, for the most part, been known and available for years [88]. Most of the methods that are now being developed to study palmitoylation capitalize on many years of knowledge and development of cysteine-specific chemistries, developed mainly as methods to purify and/or specifically target proteins and peptides with various reagents. Many of the reagents that specifically label cysteines have been created as both affinity and fluorescent tags, the former for purification and structure determinations [88] and the latter as cellular reporters of protein abundance, subcellular distribution, protein conformation changes, the formation of the Golgi, and even the concentration of cellular analytes in specific subcellular domains. The following references provide a short list of some of the most clever uses of thiol chemistry [113-119]. Given the wealth of information on the unique chemistry of the palmitoyl thioester bond and the tools for capturing and characterizing cysteines in proteins, it is somewhat surprising that we are only now developing innovative assays to increase our understanding of palmitoylation. This recent increase is most likely tied to the dramatic increase in the utility of mass spectrometry as a proteomic tool. To provide a general frame of reference for the recent shift in the types of assays that are being developed, we will briefly discuss other assays that have been used successfully for a longer peri-
od of time. These assays are by no means outdated and some continue to be the most appropriate way to answer specific questions.

4.1. Chemistry and physical properties of palmitoyl cysteines: Reactions and probes

Working with palmitoylated proteins is inherently difficult due to the labile nature of the thioester bond and the increased hydrophobicity of the protein or peptide due to palmitate. On the other hand, the unique physical and chemical properties of thiols, palmitoylated thiols, and the thioester bond make them particularly amenable to modification by highly specific chemistry and a wide variety of thiol reactive probes.

Reactions of free thiols in the cytoplasm

Thiol modification occurs most commonly in cells by one of two routes: disulfide exchange or alkylation. Many of the reactive groups that undergo these two reactions are relatively stable in aqueous environments; the reactions are rapid and provide high yields of thioether and disulfide bonds [88]. Thiols will also react with many amine reactive reagents including isothiocyanates and succinimidyl esters but lack a high degree of specificity, resulting in unstable bonds that are much less useful for routine modification of thiols in proteins. Thiol-specific reagents and chemistry figure strongly into the design and development of novel assays for palmitoylation. Most investigators are limited somewhat to reagents that are available from a catalog but, fortunately, there are already many useful reagents available. Among the most useful are thio-reactive chemicals that are linked to another moiety (reactive or reporter) by a spacer arm of variable length and physical characteristics. Such hetero- and homo-bifunctional crosslinking reagents have provided much of the foundation for recent developments in palmitoylation assays and provide a fairly rich toolbox for future assay development.

Chemical moieties that react with palmitoyl-cysteines

Iodoacetamide conjugates are among the most commonly used tools for modifying cysteine thiols. These undergo nucleophilic substitution to form stable thioether bonds at physiological pH in aqueous environments. When using iodoacetamide and its conjugates, one should remember that depending on the pH of the solution, they can also react with histidine, lysine, and methionine (at pH >1.7) residues and N-terminal amines. However, when used at slightly alkaline pH in the dark and in the absence of reducing reagents, cysteine modification will be the exclusive reaction [88]. A good example of iodoacetamide-based probes are the isotope-coded affinity tags or ICAT [120]. These have proved particularly useful in determining the substrates of DHHC2 [37].

Maleimides are also common constituents of heterobifunctional crosslinking reagents and blocking reagents that target cysteines. They are ~1000 times more specific for cysteine sulfhydryls at pH 6.5-7.5, but at higher pH some cross reactivity can occur with amines. Maleimides form stable thioether bonds by adding the sulfhydryl across the double bond of the maleimide.

Phenylmercury derivatives react with thiols, including nitrosothiols, under conditions similar to iodoacetamides and maleimides to form stable mercury-thiol bonds that can be re-
versed in 0.1N HCl and reducing reagents like dithiothreitol (DTT) but apparently not by TCEP. Phenylmercury derivatives also react faster with thiols than do the commonly used thiol-reactive N-ethylmaleimide (NEM).

Compounds containing disulfide bonds are able to undergo disulfide exchange reactions with another thiol by the free thiol attacking the disulfide bond and subsequent formation of a new mixed thiol. Two examples of useful compounds in this category are Methylmethane-thiosulfonate (MMTS) and pyridyl disulfide derivatives like biotin HPDP ((N-(6-(Biotinamido)hexyl)-3’-(2’-pyridylthio)-propionamide). MMTS can be used in some cases to block free thiols more effectively than NEM, as it is uncharged and thus more likely to modify all free reactive cysteines. MMTS has been shown not to react with nitrosothiols or existing disulfides [121].

4.2. Metabolic labeling with radiolabeled palmitate

The most common method to identify palmitoyl proteins and to determine the residence half-life of palmitate on a specific protein or palmitate turnover (e.g. [122] for a particularly interesting example) is to metabolically label cells with radiolabeled palmitate. $^{14}$C-, $^3$H- and $^{125}$I-labeled palmitate have all been used, but $^3$H-palmitate is most common because it is relatively inexpensive and widely available. However, using $^{125}$I-labeled palmitate provides some advantages. In practical terms, the time required for detection is considerably shorter - hours instead of (often) weeks with $^3$H-palmitate. The $\gamma$-irradiation from the $^{125}$I is also compatible with phosphorimaging technology which is much more rapid and quantitative than densitometric measurements from films generated by autoradiography (as is used with tritium). The principle downside of using $^{125}$I-labeled palmitate is that it is not commercially available, and the labeling must be done by the investigator. Reviews of the methods using radiolabeled palmitate and including technical details have been published recently [123-125].

4.3. Fluorescently-labeled peptide substrates for palmitoylation

Fluorescently-labeled peptides that mimic PAT substrates have been used to characterize PAT activity and for the discovery of inhibitors of palmitoylation [123, 126, 127]. The use of these peptides over the last several years was reviewed recently [123]. Peptide substrates for palmitoylation have also been genetically fused to fluorescent proteins and expressed in cells. This strategy has been used to determine how palmitoylation affects subcellular trafficking both between and within membranes [124]. Monomeric GFP-based reporters and fluorescence resonance energy transfer proved to be helpful in the identification of lipid rafts with an affinity for palmitate on the inner leaflet of the plasma membrane [5].

4.4. Acyl-biotin exchange: ABE

Most of the novel assays for palmitoylation utilize the same basic foundation first described for a palmitoyl protein by Schmidt and colleagues [128] and now most commonly known as acyl-exchange. First, free cysteines are blocked on proteins that have been extracted from
live cells or tissue. Next, palmitates are removed from cysteines by cleavage of the thioester bond with hydroxylamine (typically 1.0M) at neutral pH. This creates a new set of free thiols unique in that they were all formerly palmitoylated; ideally, no others should exist. Finally, this new set of formerly-palmitoylated cysteines is modified by any one of the many thiol-specific reagents. The uniqueness of the individual assays that incorporate these steps lies primarily in the choice of thiol-specific reagents, and this choice depends on what questions the investigator wants to answer. There are also variations in the reagents used to block free cysteines in the first step. Both NEM and MMTS have been used in the assays described below but NEM is used most commonly.

Cysteines that are palmitoylated can also be modified by fatty acids other than palmitate [7] including stearate and oleate. The acyl-exchange method cannot yet distinguish between palmitate and the other fatty acids modifying cysteines by a thioester bond. Two additional points that relate to the specificity of this method for palmitoylation are: 1) that it will not report modification of cysteines by prenyl groups (geranylgeranyl or farnesyl) because they are attached by a thioether bond that is not susceptible to cleavage by hydroxylamine and 2) it will not report myristoylated proteins because this 14-carbon acyl group is linked to an N-terminal glutamate by an amide bond which is also insensitive to cleavage by hydroxylamine.

The recent development of novel assays using the three-step acyl exchange method to study palmitoylation in a broader sense was invigorated by two publications describing a new twist on the method that incorporated the use of radiolabeled NEM assay [129, 130]. Work described in these papers showed that labeling palmitoyl cysteines with radiolabeled NEM resulted in a remarkable 5- to 12-fold increase in sensitivity to detect several known palmitoyl proteins, including PSD-95 and SNAP-25, when compared to labeling with 3H-palmitate. In addition, the authors demonstrated the utility of the biotinylated, heterobifunctional crosslinker, 4-[4′-(maleimidomethyl)cyclohexanecarboxamido] butane (Btn-BMCC), as an effective tool to capture and purify (using streptavidin-agarose) palmitoylated proteins. In doing so, they also demonstrated the general potential of using the wide variety of existing thiol-specific probes for the development of additional assays for palmitoylation that are beginning to materialize.

4.5. The palmitoyl proteome

The demonstration that one can effectively replace palmitate with a biotin group led to development of the first, large-scale, proteomic analysis of palmitoylation [74] in yeast, the model system in which the molecular identify of PATs was first determined [23, 26]. This method was dubbed “acyl biotin exchange” or ABE and used the same basic three-steps as described above. As the name implies, the proteins were labeled with a thiol-reactive, biotinylated heterobifunctional probe, [6-(Biotinamido)hexyl]-3′-(2′-pyridyldithio)propionamide (HPDP-biotin), with subsequent capture on streptavidin affinity matrix (for a detailed protocol see Wan et al. 2007 [131]). It is interesting to note the number of proteins that Roth and colleagues captured in the negative control samples (Figure 1a;[74]). The degree of overlap among proteins captured in the experimental and control samples suggests that the step in
which free thiols were blocked with NEM was not quantitative and/or that the wash steps following binding of biotinylated proteins to the streptavidin matrix were not sufficiently stringent (steps 7 and 16 respectively from Wan et al, 2007 [131]) thereby resulting in the potential for a higher number of false-positive hits. However, issues of signal to noise and limits of sensitivity are by no means unique to this work (avidin-biotin affinity purification is notoriously difficult); rather they are unavoidable issues faced by all developers of novel strategies and users of nascent technologies. Incremental improvements in important assays like this always follow.

One of the key features of all proteomic methods is the system used for detection of specifically-isolated proteins or peptides. Work by Roth et al. [74] identified proteins by multi-dimensional protein identification technology (MudPIT), a high-throughput, tandem mass spectrometry (MS/MS)-based proteomic technology [132] [see also [131, 133]]. Compared to other mass spectrometric methods, MudPIT has the potential to identify less abundant proteins with a higher degree of confidence, because multiple peptides of a single protein can be used to identify a protein of interest. One downside with MudPIT in this case is that the palmitoyl cysteine(s) cannot be pinpointed, as there may be many candidates among the individual peptides of a whole protein suspected as being a palmitoyl protein. After demonstrating the usefulness of this large-scale method for purification and identification of palmitoylated proteins, the authors used mutant strains of yeast lacking one or more of the seven yeast PAT proteins to identify substrates of individual PATs. Comparison of the degree of palmitoylation of individual proteins between wild type yeast (a full set of normally palmitoylated proteins) and those not expressing one or more of the yeast PATs (each with a specific set of hypo/depalmitoylated proteins) provided the identity of the substrates of individual PATs. Together, this work represents a very significant contribution to the identification and understanding of the yeast palmitoyl proteome and provided many important clues about potential homologous PAT-substrate pairs in other systems.

The complexity of palmitoylation is greater in a vertebrate system. With at least 23 genes encoding PATs identified in humans, the diversity at the most basic level is at least three-fold greater than in yeast. When one considers the additional variants encoded by alternative splicing of PAT mRNAs, the potential diversity increases even more. The greater number of PATs suggests (but does not prove) that there are also more palmitoylated proteins in mammals. The ability to genetically manipulate mammalian cells is improving but lags behind yeast. Nevertheless, defining the palmitoyl proteome or palmitoylosome and how it is regulated in mammals (humans in particular) is a task of significant importance and interest. Now that the enzymes capable of mediating palmitoylation have been identified, one of the most important questions that we face is which substrates are palmitoylated by each PAT—a question brought sharply into focus when one considers the known connections between mutations or deletions in PAT genes and human disease, in particular cancer. DHHC2 is deleted in many types of cancer (see above). Its absence is strongly correlated with an increase in the metastatic potential of cancer cells. The simplest inverse corollary in this case is that palmitoylated substrates of DHHC2 are responsible for keeping cells from metastasizing. Identification of these substrates and their associated signaling networks using novel assays
for palmitoylation has begun to provide supporting evidence for known mechanisms of cancer progression [56] as well as a novel signaling pathway for the regulation of cellular proliferation and metastasis [37].

**Figure 2.** Palmitoyl-cysteine Identification Capture and Analysis (PICA): Determining PAT-substrate specificity by differential labeling of palmitoylated proteins with Isotope-coded Affinity Tags (ICAT) (A) In one set of cultures, *ZDHHC2* expression is knocked down by transfecting HeLa cells with *ZDHHC2*-specific siRNA (Dharmacon). Proteins are extracted from experimental and control cells and treated with the thiol-specific blocking reagent MMTS. This step chemically modifies or protects all thiols (“X” on the proteins P1-P3) that are free at physiological pH and leaves the palmitoylated cysteines (P) undisturbed as depicted on P1-P3. (B) Following the protection or blocking of free thiols, palmitates are removed by selective cleavage of the thioester bond with hydroxylamine at pH 7.4, which generates a distinctive set of free, formerly-palmitoylated, reactive thiols (C) that can be selectively labeled with ICAT reagents. Iodoacetamide at one end of the ICAT reagent binds to the thiol sidechain of cysteines; on the other end, biotin pro-
vides a mechanism for affinity purification of thiol-captured peptides on an avidin column. (D) Proteins from knockdown and control conditions are mixed in equal amounts and digested in-gel with trypsin. ICAT labeled peptides are enriched by avidin affinity and analyzed by LC/MS. A pair of ICAT-labeled peptides is chemically identical and is easily visualized, because they essentially coelute and there is a 9 Da mass difference measured in a scanning mass spectrometer. Even if equal amounts of a single protein exist in two different samples, the quantity of protein that is captured depends directly on its degree of palmitoylation; if all of a single protein is palmitoylated under one condition, then all of it will be captured; if only half of this protein is palmitoylated under another condition then the capture rate of that protein will be half as much, relative to control, making it appear half as abundant. Proteins for which there has been no change in palmitoylation (ie, equal capture rates) will yield a heavy:light (H/L) ratio of 1. The degree to which palmitoylation is diminished will register as a decrease in the H/L ratio (ie, 50% reduction in palmitoylation will correspond to a H/L ratio of 0.5). A change in the capture rate that results in a change in the post-purification abundance is measured in the LC/MS phase. (E) Finally, the peptides are further fragmented into their constituent amino acids by MS/MS, enabling identification of the proteins corresponding to the captured peptides.

4.6. Palmitoyl-cysteine Identification Capture and Analysis (PICA): Identification of PAT substrates and palmitoyl proteins in vertebrates

With the aim of defining PAT-substrate specificity in a living vertebrate system, we developed a method to identify substrates of specific PATs in mammalian cells and tissues called Palmitoyl cysteine Isolation Capture and Analysis or PICA [37]. We used this method to identify CKAP4/p63, a known palmitoyl protein [134] as one substrate of DHHC2 in HeLa cells [37]. This method is similar to ABE as described by Roth et al (2006) but was inspired [135] in part by the work of Drisdel and Green (2004) and incorporated several novel features that will be discussed below.

The ability of PICA to identify PAT substrates is based on the principle that it quantifies the differential frequency of palmitoylation of individual proteins or peptides in control conditions versus conditions in which the function of a single PAT is reduced by siRNA-mediated gene knockdown. The process to identify substrates of DHHC2 consisted of four basic steps outlined in Figure 2. In the first part we generated two distinct pools of palmitoylated proteins, one from control HeLa cells and the other from HeLa cells in which the activity of one PAT (DHHC2) was reduced. These two distinct pools of palmitoylated proteins were then captured and compared directly to identify differences in the degree of palmitoylation of individual proteins between the two pools. To do this, we reduced the expression of ZDDHC2 mRNA (and consequently the abundance of the encoded enzyme, DHHC2) in HeLa cells using siRNA-mediated gene knockdown which resulted in a reduced level of palmitoylation of DHHC2 substrates. Total protein from knockdown and control cells was prepared by first blocking free thiols with MMTS in the presence of SDS. This was followed by selective exposure of all palmitoyl cysteines by cleavage of the palmitoyl-cysteine thioester bond with 1.0M hydroxylamine at neutral pH, thereby generating a unique population of formerly palmitoylated cysteines. Second, we selectively and differentially labeled the exposed, formerly-palmitoylated cysteines from knockdown and control cells with biotinylated, thiol-reactive heavy (H) and light (L) ICAT reagents, respectively. Third, we combined equal quantities of the ICAT-labeled protein from ZDHHC2 knockdown and control cells and digested the mixture with trypsin. The resulting H and L ICAT-labeled tryptic peptides were captured and purified on an avidin affinity column. Finally, ICAT-labeled, putative, formerly-palmitoylated peptides were analyzed by mass spectrometry. Peptides with a reduced
H/L ratio over four independent runs were analyzed further to confirm that the identified cysteine was indeed palmitoylated by DHHC2 under physiological conditions. Details of the protocol and reagents used and outlined in Figure 2 can be found in [37].

There are several unique aspects in the PICA method. First, we used MMTS to block the free thiols in the first step. NEM is used most commonly at this step, but MMTS is more reactive and smaller than NEM or iodoacetamide, enhancing its ability to modify all free reactive cysteines. Inefficient blocking of free thiols in the first step is one factor that could easily contribute to false-positive capture of proteins in the purification step. Qualitative evaluation (silver-stained SDS-PAGE) of protein capture in experimental and control (no-hydroxylamine) conditions [Figure 2, [37]] suggests that it may be more efficient than NEM (for comparison see Figure 1A [74]). However, it may also be that we captured very few proteins under control conditions because of a more stringent wash protocol than described by Roth et al [74]. The use of ICAT reagents in PICA allowed us to combine formerly-palmitoylated peptides purified from control and experimental cells in the same pool, and subsequently, a direct, simultaneous analysis of palmitoylation in the two pools in a single analytical sample. We defined a substrate of DHHC2 as one that had a consistently reduced H/L ratio over four independent PICA runs. This approach provided us with many (the vast majority), convenient internal control peptides which are peptides that were not substrates of DHHC2 that had unchanged H/L ratios. This approach significantly reduces the potential for identification of false-positive hits because, if a protein can be falsely labeled by an ICAT, it should do so with equal efficiency in both the control and experimental cells yielding a peptide with an H/L ratio of ~1. The greater risk with this approach is the failure to identify substrates that exist in low abundance. Using tandem mass spectrometry, we analyzed a sample of significantly reduced complexity including only ICAT-tagged peptides. As is inherent in such analyses, the most abundant peptides dominate the report. However, one advantage of this approach is that when a peptide is identified, whether it is a substrate of a single PAT or not, the palmitoylated cysteine(s) is also identified. In the case of CKAP4/p63 (and the majority of other peptides) there was only a single cysteine, and it was already known to be a site for palmitoylation [134]. Spectral counting has the potential to positively identify palmitoyl proteins of lower abundance because more than a single peptide from any given protein is factored into to the identification. There is greater overall coverage (identified peptide fragments of a protein) using this method thereby increasing the confidence level of identification. However, the disadvantage inherent in analyzing a complex mixture, including non-palmitoylated peptides by spectral counting, is that identification of the palmitoyl cysteine (in the cases where there are multiple candidate cysteines) must await subsequent and tedious analyses. The tradeoff between these two complementary approaches in mass spectrometric analysis is sensitivity versus specificity. Combining these analyses will provide a much greater depth of coverage.

4.7. Forward and reverse approaches to assigning PAT-substrate pairs

The first reports that identified PAT-substrate pairings took the reverse approach: start with a known palmitoylated protein then use metabolic labeling with radiolabeled palmitate and
co-overexpression of one PAT and the substrate (for a review see: [72]). Using this method, an increase in the incorporation of radiolabeled palmitate on the overexpressed substrate in the presence of an overexpressed PAT is used to claim specificity. This method is an important tool for increasing our understanding of palmitoylation-related phenomena including confirmation of putative PAT-substrate pairs identified by other methods. Likewise, when starting with a known palmitoyl protein and the intention of identifying the PAT responsible for its palmitoylation, it remains a useful method. However, we should remember that just because overexpression of a PAT can increase the incorporation of palmitate onto a specific protein does not necessarily mean that it does so in a live cell. Again, problems like this are not unique to this method and simply reflect our lack of knowledge about where and when PATs and their substrates are expressed, the degree of promiscuity among PATs and, how PAT function is regulated.

The potential for specific cysteines to be modified by both palmitate and RA via a thioester bond is an issue that deserves attention from those of us interested primarily in palmitoylation for at least two reasons. One is the potential that an exchange between the two modifications is a physiologically relevant means of regulating signaling and second, the possibility that proteins identified as being palmitoylated in assays utilizing some form of ABE chemistry are RA-modified instead.

4.8. Labeling palmitoyl proteins with bioorthogonal probes

This particularly interesting approach labels cysteines with isosteric, azido-derivitives of fatty acids that are able to substitute for fatty acids that occur naturally in cells (Figure 3) [112, 136]. Once bound to cysteines, the azido group on the fatty acid is reacted, with a high degree of selectivity, via the Staudinger reaction [137] with (triaryl)phosphines that are themselves derivatized with Myc, biotin, a fluorophore, or others. Using this method Hang and colleagues [112] found that ω-azido fatty acids with 12 and 15 carbons can be efficiently metabolized by mammalian cells and accurately report myristoylation and thio-palmitoylation, respectively.

Work by Kostiuk and colleagues identified palmitoyl proteins from mitochondria using azido-palmitate [136]. To accomplish this they purified proteins from cellular mitochondrial fractions, first by differential centrifugation, then by further purification based first on charge and subsequently by size using chromatographic separation. Labeling of proteins in this study was outside of a living system presumably leaving only the possibility of autocatalytic/non-enzymatic palmitoylation. Mass spectrometric analysis of selected bands identified 21 palmitoylated proteins, 19 of which were novel. The majority of the proteins labeled were metabolic-type proteins unique to the mitochondrion. This raises the interesting possibility that the principle mechanism of palmitoylation in this organelle is autocatalytic rather than enzymatic, and that, as suggested by the authors, the key role of palmitoylation in the mitochondrion is to inhibit enzymes by palmitoylation of cysteines in the vicinity of the active site.

These so called bio-orthogonal probes have been reported to be nontoxic and very stable under physiological conditions. Importantly, this two-step reaction is rapid and more sensitive
than labeling with $^{125}$I-palmitoyl-CoA. These features, especially their ability to effectively substitute for endogenous fatty acids, make them ideal for labeling palmitoyl proteins in live cells, providing a significantly more direct measure of protein palmitoylation than can be achieved in any other assay format. It is easy to imagine that use of such probes will come to dominate in experimental systems for studying palmitoylation.

**Figure 3.** Using Click chemistry and bio-orthogonal probes to label palmitoyl cysteines. A) A palmitoylated protein; the shaded box indicates the thioester bond. B) Azido-palmitate is transferred to a protein forming a thioester bond with a cysteine residue. The azide moiety of the azido-palmitate reacts, via the Staudinger reaction, with the tagged (in this case biotin) phosphine, forming an amide bond. The biotin-tagged proteins can then be affinity purified and analyzed in various ways including mass spectrometry. Tags and reporters other than biotin can be added to the phosphine providing a wide array of potential methods for subsequent analyses.

### 5. Pharmacological modulation of Palmitoylation

#### 5.1. Developing compounds that selectively target individual PATs

Existing chemicals used to inhibit palmitoylation are neither selective nor potent. The compounds used most commonly are 2-bromopalmitate (2BP), tunicamycin, and cerulenin. Each of these is a lipid-based molecule (Figure 4). 2BP has been used most frequently at a concentration of ~100 µM to block palmitoylation in spite of the fact that at least two studies have shown that the IC50 of 2-BP is ~10 µM [138, 139]. 2BP is not tolerated well by cultured cells.
and causes death even after a brief exposure to 100 µM. 2BP inhibits several enzymes involved in lipid metabolism, including carnitine palmitoyltransferase 1, fatty acid CoA ligase, glycerol-3-phosphate acyltransferase, and enzymes in the synthesis of triacylglycerol biosynthesis [140, 141]. This high degree of promiscuity as well as the toxicity of 2BP renders it nearly useless as a tool to determine anything specific about palmitoylation related signaling issues that equally plague cerulenin and tunicamycin. The uses and effects of these three inhibitors was reviewed recently [142].

Figure 4. Lipid-based inhibitors of palmitoylation.

Smith and colleagues [126] recently screened a compound library in an attempt to identify more selective and potent inhibitors of palmitoylation, in particular inhibitors of PATs. This screen identified single compounds from five chemical classes (Compounds I-V) that inhibited cellular processes associated with palmitoylation. The assays used in the screens included: measuring the in vivo and in vitro growth rate of an NIH/3T3 cell line that overexpressed DHHC17, displacement from the plasma membrane of myristoylated or farnesylated GFP, and in vitro palmitoylation of small, non-complex, myristoylated or farnesylated, synthetic, fluorescent peptides intended to mimic known palmitoylation substrates [123, 126, 127]. These assays could not discriminate a direct effect of any compound on any PAT. They could only report the activity of compounds that acted at some point (not excluding direct PAT inhibition) in any pathway that leads to or affects palmitoylation; compounds like 2BP, cerulenin, and tunicamycin. This assertion was borne out in follow-up studies on the same compounds [138] (see below). Perhaps the most intriguing finding in this report was that compounds I-IV were able to suppress the oncogenic behavior of human cells that overex-
pressed DHHC17. However, there is no direct evidence to indicate that compounds I-IV exert these actions via inhibition of DHHC17 or through the palmitoylation of Ras proteins by DHHC17 as was speculated [53, 126]. Nevertheless, since these compounds reduced the *in vivo* growth of tumors from cells overexpressing DHHC17 [126], it would be worth determining their exact mechanism of action.

Subsequent studies by Linder, Deschenes and colleagues [138] tested four of the five compounds identified by Smith and colleagues and found that they were not selective for DHHC proteins. This report also included a wealth of information defining the mechanisms by which 2BP inhibits palmitoylation. Briefly, only one of the four compounds re-tested, compound V and 2BP, inhibited the activity of any of the four DHHC proteins tested. Neither compound V nor 2BP was selective for any of the PATs tested, and 2BP was more potent. Both compound V and 2BP blocked autoacylation of the PATs; compound V was reversible, 2BP was not. Even though compound V was able to inhibit the activity of the four PATs tested and in the same manner as 2BP, these experiments could not determine whether compound V also blocks palmitoylation indiscriminately at steps prior to the actual palmitoylation event, as is the case with 2BP, cerulenin, and tunicamycin.

There would be no compelling reason to begin a drug discovery program to identify inhibitors of just any PAT. Rather one would choose to begin with a PAT that is linked to a disease state—a situation where misregulated expression or function of that PAT was clearly linked to a pathological state. As discussed earlier, links between PAT expression (but not yet altered function) have been demonstrated for both neurological disorders and cancer; thus, candidate PATs that would be appropriate targets for drug development exist. Both overexpression and absence of PAT expression have been implicated in the development of cancer. Dampening the activity of an existing PAT is a conceptually and mechanistically simpler task than accurately restoring the specific activity of a PAT that is not expressed or absent. This review is concerned with PAT assays and PAT inhibitors, so we will address the case of PAT overexpression in ideal terms as well as the technical issues that surround the development and implementation of the assays designed to discover PAT inhibitors.

### 5.2. Considerations for development of high-throughput screens to discover PAT inhibitors

The DHHC motif in PATs defines the active site and is highly conserved in all mammalian PATs [3]. The regions of highest diversity are primarily in the N- and C-termini of the PAT. Mutation of the cysteine in the DHHC motif abolishes PAT autoacylation and palmitoylation of the substrate, a property of all DHHC proteins studied so far. This high degree of homology in the active site sequence among PATs could give the impression that developing highly specific, active-site inhibitors for palmitoylation will be impossible. However, this same issue exists with kinases [143, 144], and yet the development of selective and potent active-site, ATP-competitive inhibitors has been successful (eg, [145]).

The specificity of palmitoylation must be derived in part from the unique physical interactions of individual PATs with their substrates. The sequence of amino acids surrounding a substrate cysteine partially defines the potential for that cysteine to be palmitoylated. How-
ever, the physical determinants for substrate recognition will likely extend throughout the accessible portions of the PAT and substrate as was elegantly demonstrated for DHHC17 [146]. Other factors that are likely to regulate palmitoylation are the temporal and spatial aspects of PAT and substrate expression.

There are many more palmitoylated proteins than there are PATs; therefore, modulating the activity of a single PAT, even with complete compound selectivity, will likely yield a change in the palmitoylation of multiple substrates. This conundrum is common to the development of highly selective and potent pharmacological modulators of all enzymes that mediate post-translational protein modifications, again kinases being a classic example.

Another challenge is that each PAT traverses the plasma membrane multiple times. A conservative guess would suggest that the membrane environment is important for determining PAT structure and substrate recognition. However, Jennings et al.,[138] demonstrated that at least four PATs can be purified from a membrane environment and remain enzymatically active. These findings are both remarkable and encouraging evidence that enzyme activity-based and drug-binding screens for selective PAT inhibitors can be accomplished with purified proteins.

5.3. Primary screen for PAT inhibitors

The discovery and refinement of drugs to modulate PAT activity will require the use of multiple assay types. The initial success of each can only be a matter of speculation at the beginning of the project, and the success of the primary screen will influence the choice of follow up assays. However, one unique aspect of palmitoylation suggests a logical starting point. The most dramatic visible change that can occur when a protein is palmitoylated is when it moves from the cytoplasm to the plasma (or other) membrane. The technology to measure such a translocation in living cells using high-throughput microscopy has been demonstrated [139, 147] and along with many other such morphometric analyses, has become well established in drug discovery programs and the basic life sciences [148-150]. This technology is often referred to as high-content screening (HCS). To develop an assay to identify inhibitors of a single PAT using HCS, it would be ideal to have identified the most clinically relevant, cytosolic substrate of the PAT of interest and to have determined that this substrate is palmitoylated exclusively by this one PAT or, alternatively, by no other PAT expressed in the cell type that will be used for the screen. However, biological systems rarely offer ideal situations, and accommodations will inevitably need to be made. The ideal substrate would then be fused to a monomeric fluorescent protein (FP) [5, 151] to generate a fluorescent reporter of palmitoylation that localizes primarily or exclusively to the PM. Cells stably expressing this reporter would then be grown in multi-well imaging plates and exposed to a chemical compound library, and the subcellular distribution of the FP-tagged palmitoylation substrate evaluated by HCS. Compounds that cause redistribution of the fluorescent reporter from the PM to the cytoplasm are candidates (or hits) for follow up analyses that will determine if they blocked palmitoylation of the reporter by directly inhibiting the PAT of interest or indirectly by some other mechanism. Typically, compounds in a large library (tens to hundreds of thousands of compounds) are tested at a single concentration and repli-
cated, often three times, to increase the confidence of selecting biologically active compounds. But, the relationship of replicates is solely statistical, not pharmacological. An alternative screening method for identifying hits is “titration-based screening” called qHTS [152]. This method, which has been used successfully by Jim Inglese, Doug Auld and Colleagues at the NIH Chemical Genomics Center, measures the assay system response to multiple (up to seven), different concentrations of a single compound. The increased density and accuracy of the data produced by this method can provide many benefits over screening at a single concentration (for a full description of the merits of qHTS see [152]). Among the most important benefits of screening at multiple concentrations is that it alleviates the problems of false-negatives and false-positives that plague screens run at a single concentration. The nominal, additional effort required at the front end of the process is generously compensated by a subsequent reduction in the effort required to choose which hits to pursue in follow-up assays.

Displacement or translocation of the fluorescent palmitoylation reporter from the PM to the cytoplasm in response to a compound cannot provide evidence that the compound has this effect by direct inhibition of a PAT. Secondary screens designed to determine which of the hits works by direct inhibition of PAT activity will be required. One option would be to determine the effects of each hit on the enzymatic activity of the PAT of interest. Jennings et al have demonstrated that a PAT can be purified from a membrane environment and retain its enzymatic function i.e., transfer of palmitate to a substrate. The metabolically active form of palmitate in a living system is palmitoyl-CoA. Transfer of palmitate to a substrate results in the liberation of CoA from palmitate, a chemical species that can be measured with accuracy and sensitivity in a high throughput manner (Figure 5) [153].

![Figure 5.](image-url)
Retesting hits from the primary screen in this secondary, enzymatic assay would determine if the compounds directly inhibited palmitoylation of the substrate in the reaction. It would also identify compounds that inhibit palmitoylation by competing with palmitoyl-CoA for access to the PAT active site, as well as allosteric inhibitors. Structural analysis of the compounds would provide information about how they inhibit palmitoylation. A binding assay in which compounds are tested for their ability to compete directly with palmitoyl-CoA for binding to the PAT would more conclusively determine the mechanism by which the compounds were inhibiting substrate palmitoylation.

It is likely that the inhibitors identified will represent multiple classes of compounds distinguished by their chemical structures. However, the chemicals identified in these screens are unlikely to represent the most potent or selective compounds that exist. Further refinement by probing and refining the chemical space of each compound in a medicinal chemistry effort will be required to achieve both objectives. Generating higher affinity analogs of the hit compounds for the PAT of interest will improve the selectivity and potency of the compounds for an individual PAT. However, the question of how selective any of these compounds is for a single PAT must be answered by determining their ability to inhibit other PATs—counterscreening. Since the family of PATs is relatively small (23 genes), it would not be unreasonable to measure the effects of the compounds on each of the other PATs in the enzymatic assay described above.

The assay development pathway proposed above is outlined only in very general and ideal terms, glossing over inevitable technical challenges that must be overcome for the project to be successful. However, even as brief as this description is, it exceeds by far the complexity of any published attempt to identify selective PAT inhibitors to date. A valuable, practical guide to choosing, developing, and validating assays including those proposed above is available at: http://ncgc.nih.gov/guidance/manual_toc.html.

6. Conclusion

The discovery of the molecular identity of PATs was a pivotal event that has fostered substantial progress in the field of lipidation, having a profoundly positive effect on many fields of biology. Many long-standing questions have been greeted with answers as well as a clearer direction in which new inquiries should proceed. While sometimes criticized as being stamp collecting, defining the palmitoyl proteome of specific cells and tissues would provide new and unforeseen insight into many cellular processes. The methods described here provide the technical foundation for defining the palmitoyl proteome. Defining the intrinsic and extrinsic mechanisms and factors that regulate PAT activity will also be crucial and challenging. Future assays to investigate such details will certainly benefit from the demonstrated usefulness of bioorthogonal probes that appear to be treated by cells as if they were palmitate. These probes may provide a more direct measure of palmitoylation than the exchange of cysteine-reactive probes for palmitate on purified proteins.
The number of signaling networks in which palmitoylation plays a pivotal role is large and growing. The relationship between PAT gene expression and cancer is perhaps the most evident as the fraction of PAT genes implicated in metastasis and tumorigenesis is notably high. There is also a remarkable confluence between our increasing understanding of palmitoylation and our increasing awareness of the importance of lipid rafts, one of the primary residences of palmitoylated proteins in cancer [154, 155]. In instances where there is a relationship between aberrant expression of PATs and cancer, the critical questions relate to the substrates of these PATs and their associated signaling networks. Identification of these signaling networks will potentially provide new therapeutic targets for the prevention or reversal of cancer progression. Given the preponderance of palmitoylated proteins resident in the neuronal synapse (which is itself a lipid raft of sorts), it is clearly another area of research that deserves (and has already received) a great deal of attention.

While there has been some progress made in identifying pharmacological modulators of palmitoylation [125, 126, 138], there is nothing yet known about how to specifically target individual PATs. From a practical standpoint, inhibiting specific PATs may be a simpler process than developing specific PAT agonists. The advent of assays with the ability to measure changes in the activity of a single PAT along with the identification of PAT-substrate associations should enable further development of new assays to identify specific pharmacological modulators of individual PATs as well as providing important information on the signaling networks associated with specific PATs.

Our ability to understand palmitoylation and its importance to human health and disease is only as good as the technological methods we use to make accurate and valid measurements. Our ability to investigate the basic mechanisms of how PATs work, of PAT/substrate relationships, and how palmitoylation affects signaling processes related to disease would be improved significantly by the development of selective and potent pharmacological tools.

Until such tools are available, we should be mindful that using compounds such as 2-BP, cerulenin, and tunicamycin may lead to erroneous conclusions. Developing non-lipid, selective inhibitors that target the PAT active site is feasible. The challenges that exist are conceptually similar in some aspects to those faced during the development of selective, small-molecule inhibitors of kinases that do not resemble ATP. Based on current knowledge, the most logical PATs to target first are those for which overexpression is oncogenic. However, the motivation to initiate drug discovery programs on a large scale will probably remain below the required threshold until more conclusive data are available from more sophisticated, whole-animal experiments that link PATs to oncogenesis.

Author details

Sonia Lobo Planey

Address all correspondence to: splaney@tcmedc.org

The Commonwealth Medical College, Scranton, PA, USA
References


[82] Kleuss, C., & Krause, E. Galpha(s) is palmitoylated at the N-terminal glycine. Embo J, 22(4), 826-32.


[150] Collins, M. A. Generating’omic knowledge: The Role of Informatics in High Content Screening. Combinatorial chemistry & high throughput screening.


