Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Chemical Biology

Covalent protein modification: the current landscape of residue-specific electrophiles

D Alexander Shannon and Eranthie Weerapana

Functional amino acids that play critical roles in catalysis and regulation are known to display elevated nucleophilicity and can be selectively targeted for covalent modification by reactive electrophiles. Chemical-proteomic platforms, such as activity-based protein profiling (ABPP), exploit this reactivity by utilizing chemical probes to covalently modify active-site residues to inform on the functional state of enzymes within complex proteomes. These and other applications rely on the availability of a diverse array of electrophiles and detailed knowledge of the reactivity and amino-acid selectivity of these groups. Here, we survey the current landscape of electrophiles that covalently target various nucleophilic amino acids in proteins and highlight proteomic applications that have benefited from the unique properties of these electrophiles.

Addresses
Department of Chemistry, Boston College, Chestnut Hill, MA 02467, United States

Corresponding author: Weerapana, Eranthie (eranthie@bc.edu)

Current Opinion in Chemical Biology 2015, **24**:18–26

This review comes from a themed issue on **Omics**

Edited by **Benjamin F Cravatt** and **Thomas Kodadek**

http://dx.doi.org/10.1016/j.cbpa.2014.10.021

Introduction

Covalent modification of proteins is critical for various applications such as the development of irreversible inhibitors and activity-based protein profiling (ABPP). ABPP typically relies on the selective targeting of a single enzyme, or a functionally related family of proteins, with a chemical probe comprised of a protein-reactive electrophile and a reporter group [1•]. Two of the earliest examples of ABPP were the use of fluorophosphonate (FP)-based probes for targeting serine hydrolases (SHs), and vinyl sulfone and epoxide-based probes for lysosomal cysteine proteases [2,3]. These early ABPP studies exploited existing knowledge of affinity labels specific for these enzyme classes. More recently, non-directed approaches that utilize electrophiles such as sulfonate esters and chloroacetamides have enabled the profiling of enzymes for which cognate affinity labels do not exist, thereby expanding the enzyme classes amenable to ABPP [4,5].

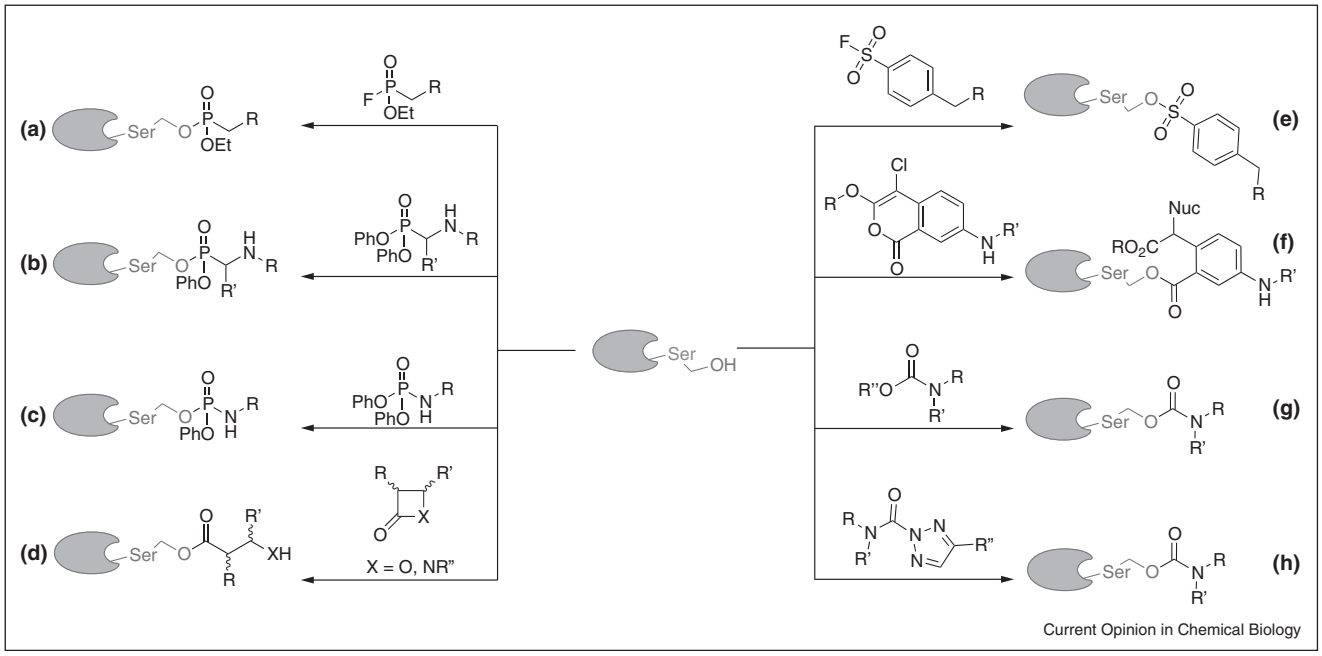
These non-directed probes were shown to bind to their respective targets in an activity-based manner, thereby selectively modifying functionally relevant amino acids despite the overabundance of non-functional residues in the proteome [6••]. This concept was intimately explored by measuring the reactivity of both functional and non-functional cysteine residues using a highly reactive iodoacetamide (IA) probe. Using a quantitative mass-spectrometry platform, termed isoTOP-ABPP, over a thousand cysteine residues in the proteome were ranked in order of nucleophilicity [7]. This analysis revealed that the subset of highly reactive cysteines was enriched in functional residues that were critical to catalysis and regulation. Therefore, using diverse electrophiles to covalently modify highly reactive amino acids in the proteome provides a means to identify novel functional loci across disparate protein families and extend the tools of ABPP to a larger subsection of the proteome.

Amino acids such as serine, cysteine, lysine, tyrosine, threonine, aspartate, and glutamate have the potential to be nucleophilic depending on the protein microenvironment. Covalent modification of these residues relies on the judicious selection of an electrophile with the appropriate affinity toward the activated amino-acid side chain. This affinity is mediated primarily by the relative hardness/softness of the nucleophile–electrophile pair. Diverse electrophiles have disparate amino-acid reactivity profiles, especially within the context of a proteome [6••]. Here, we survey the current landscape of protein-reactive electrophiles, with a particular focus on their amino-acid selectivity. Recent reviews have highlighted chemical transformations that are amenable to protein labeling [8•,9•]; here, we update these available reactions and emphasize their utility for chemical proteomics.

Covalent modification of serine

The SH enzyme family contains an activated serine within a catalytic dyad or triad and comprises approximately 1% of the human proteome [10]. The sensitivity of SHs to inhibition by fluorophosphates and FPs was well characterized, and the FP electrophile was adapted to generate ABPP tools to study the SH family [2]. FP probes have enabled the functional annotation of novel SHs, the discovery of selective inhibitors, and the characterization of dysregulated SH activities in diseases such as cancer [11]. The low reactivity of serine residues located

Figure 1



Serine-directed electrophiles for covalent protein modification.

outside of a prototypical catalytic triad/dyad, coupled with the high affinity of FP toward hydroxyl nucleophiles over other reactive groups such as thiols and amines, renders the serine-FP reaction (Figure 1a) highly effective for proteomic applications.

Diphenyl phosphonates (Figure 1b) have been used as covalent inhibitors and ABPP probes for the serine-protease subfamily of the SHs [12]. Most recently, a library of peptide-based diphenyl phosphonates contained members selective for many serine proteases, including chymotrypsin, cathepsin G and urokinase-type plasminogen activator (uPA) [13]. Additionally, diphenyl phophoramidate probes (Figure 1c) retain serine protease reactivity and selectivity while being amenable to strictly solid-phase synthesis techniques [14]. These electrophiles have significantly lower reactivity relative to FP, but incorporation of peptide-based binding groups direct these probes to protease active sites for covalent adduction.

β-Lactams and β-lactones (Figure 1d) constitute another important class of serine-reactive electrophiles due to the well-characterized covalent modification of the serine nucleophile in penicillin binding proteins (PBPs) by β-lactams, and the abundance of the β-lactone motif in a variety of electrophilic antibiotic natural products. β-Lactam and β-lactone probe libraries were synthesized and evaluated within bacterial proteomes [15,16], resulting in labeling of various SHs, including PBPs and the ATP-dependent caseinolytic protease Clp. However, labeling was also observed for proteins with active-site cysteines, such as β-ketoacyl-(acyl carrier protein) synthase III (KASIII), suggesting that these electrophiles are not selective for serine nucleophiles. More recently, it was shown that a related electrophile, the β-sultam, did not target serine-containing PBPs as expected, but instead reacted with an activated threonine residue in azoreductases [17].

The pan-serine protease inhibitor, 4-(2-aminoethyl) benzenesulfonyl fluoride (AEBSF), is widely utilized in protease cocktails, yet the proteome-wide reactivity of the sulfonyl-fluoride electrophile (Figure 1e) was poorly characterized. Sulfonyl-fluoride derivatives were shown to covalently label multiple serine protease sub-classes, demonstrating utility as a serine-reactive electrophile [18]. However, proteome-wide evaluation of protein labeling showed additional targeting of activated tyrosine residues, particularly within glutathione S-transferases (GSTs) [19•].

Coumarin and isocoumarin-based compounds (Figure 1f) inhibit serine proteases, driven by nucleophilic attack on the lactone carbonyl group by the active-site serine. 4-Chloro-isocoumarin probes were identified to be highly selective for cathepsin G, elastase, and, to a lesser extent, uPA [20,21]. Furthermore, 4-chloro-isocoumarins capable of enhancing *Toxoplasma gondii* invasion by inhibiting the SH, palmitoyl protein thioesterase-1 (PPT1) were identified [22]. Despite the observed reactivity with activated serine residues, these electrophiles have been shown to react with other nucleophiles such as the thiol group in